Document Title:	The Combination of Cranial Morphoscopic and Dental Morphological Methods to Improve the Forensic Estimation of Ancestry
Author(s):	Christopher A. Maier
Document Number:	306321
Date Received:	April 2023
Award Number:	2016-DN-BX-0003

Return Completed Form to:
Graduate School/0326
University of Nevada, Reno
Reno, NV 89557-0326
(775) 784-6869
Fax: (775) 784-6064
E-mail: gradschool@unr.edu
www.unr.edu/grad

RECEIVED
MAR 28 2017
GRADUATE SCHOOL

University of Nevada, Reno

DOCTORAL DEGREE – NOTICE OF COMPLETION

Please fill out on-line and print for signatures (No Handwritten forms will be accepted)

Name: Christopher Maier

NSHE#: 8000282711

Graduate Program: Anthropology

Degree: Doctor of Philosophy

Dissertation: ☑ Satisfactory ☐ Unsatisfactory

Title of Dissertation: The Combination of Cranial Morphoscopic and Dental Morphological Methods to Improve the Forensic Estimation of Ancestry

Dr. G. Richard Scott — G. Richard Scott — 3/27/17
Chair Name (Print) — **Chair Signature** — **Date**

Dr. Marin Pilloud (Co-Chair) — Marin Pilloud — 27 Mar 17
Committee Member Name (Print) — **Committee Member Signature** — **Date**

Dr. Christopher Morgan — 3/27/17
Committee Member Name (Print) — **Committee Member Signature** — **Date**

Dr. Joseph Hefner — 3/28/17
Committee Member Name (Print) — **Committee Member Signature** — **Date**

Dr. Guy Hoelzer — Guy Hoelzer — 3/27/17
Committee Member Name (Print) — **Committee Member Signature** — **Date**

Committee Member Name (Print) — **Committee Member Signature** — **Date**

Dr. Guy Hoelzer — Guy Hoelzer — 3/27/17
Graduate School Representative Name (Print) — **Graduate School Representative Signature** — **Date**

Dr. Geoffrey M. Smith — 3/28/17
Graduate Program Director Name (Print) — **Graduate Program Director Signature** — **Date**

APPROVED: Kara Cleveland — JUN 07 2017
Graduate Dean — **Date**

Return Completed Form to:
Graduate School/0326
University of Nevada, Reno
Reno, NV 89557-0326
(775) 784-6869
Fax: (775) 784-6064
www.unr.edu/grad

RECEIVED
APR 24 2017
GRADUATE SCHOOL

UNIVERSITY OF NEVADA, RENO GRADUATE SCHOOL
DISSERTATION FINAL REVIEW APPROVAL

This is to confirm that I have reviewed the Dissertation of Christopher A. Maier in its final form and have determined 1) that its format, citations, and content are consistent and acceptable 2) that all corrections, revisions, and modifications identified and required by the advisory committee at the defense have been completed.
The author will submit the dissertation (check one)

☐ In two paper copies

☑ Electronically via the Graduate School website

Any copies submitted prior to the date and time of signature below should be rejected.

24 April 2017 @ 12:00
Date/Time

Signature: Chair, Advisory Committee
G. Richard Scott Marin Pilloud

Drs. G. Richard Scott and Marin Pilloud
Print Chair's Name Here

Rev: 04/30/2014

University of Nevada, Reno

The Combination of Cranial Morphoscopic and Dental Morphological Methods to Improve the Forensic Estimation of Ancestry

A dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Anthropology

by

Christopher A. Maier

Dr. G. Richard Scott and Dr. Marin A. Pilloud /Dissertation Advisors

May, 2017

THE GRADUATE SCHOOL

We recommend that the dissertation
prepared under our supervision by

CHRISTOPHER A. MAIER

Entitled

The Combination of Cranial Morphoscopic and Dental Morphological Methods to Improve the Forensic Estimation of Ancestry

be accepted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

G. Richard Scott, Ph.D., Advisor

Marin A. Pilloud, Ph.D., Advisor

Joseph T. Hefner, Ph.D., Committee Member

Christopher Morgan, Ph.D., Committee Member

Guy Hoelzer, Ph.D., Graduate School Representative

David W. Zeh, Ph.D., Dean, Graduate School

May 2017

Abstract

The estimation of ancestry is central to the construction of a biological profile, and has traditionally been done through the analysis of cranial morphoscopic traits. Morphological traits of the dentition have also been used to assess population affinity, though largely outside of the forensic sphere. The Daubert standard of evidence requires that methods used in the forensic sciences be testable and have known rates of error. With few exceptions, the methods employed by anthropologists in the analysis of cranial morphoscopic and dental morphological traits do not comply with that standard. Furthermore, no method exists for reliably producing an ancestry estimate from multiple data sources.

This research examines 79 cranial and dental traits in a sample of 693 individuals from various ancestry groups, including several newly or recently standardized traits that have not been fully explored. The Asian/Native American sample was excluded due to sample size. For the remaining sample, the pool of 79 variables was reduced to 34, removing those that did not differ significantly among ancestry groups or were highly correlated with other variables. These 34 variables were used to build classificatory models using random forest modeling and naïve Bayes classification.

Overall these models correctly estimated ancestry in 67%-84% of cases. In general, the naïve Bayes classifier performed better than the random forest models. Also, models that combined cranial and dental data outperformed models based on a single data source. Although the improvement of the combined data models over the cranial data models was not statistically significant. Interestingly, the combined data model showed the most marked improvement in estimating the ancestry of Hispanic individuals. This suggests that the cranium and dentition provide different information with regard to ancestry, and more accurate ancestry estimates can be produced by combining them.

The methods used to produce ancestry estimates in this research comply with the Daubert standard of evidence, making them applicable to modern forensic casework. Additionally, the results highlight the potential improvement to ancestry estimation by combining data from different regions of the skeleton, and the utility of the dentition in forensic estimates of ancestry.

Acknowledgements

A project of this magnitude would be impossible to complete without the support and encouragement of a vast number of people and institutions. The skeletal data comprise the backbone of this research, and the collections and their curators deserve special thanks. I would like to thank: Dr. Bruce Anderson and the board of the Pima County Office of the Medical Examiner; Drs. Daniel Wescott, Kate Spradley, and Tim Gocha at Texas State University; Dr. Heather Edgar and Anna Rautman, of the University of New Mexico; Dr. Lyman Jellema at the Cleveland Museum of Natural History; Dr. David Hunt at the Smithsonian Institution; and Dr. Ginesse Listi of the Louisiana State University Forensic Anthropology and Computer Enhancement Services Lab. I would also like to thank the Peabody Archives for granting me access to some of the Hooton papers. I cannot thank you enough for allowing me access, answering my questions, and accommodating my often hectic travel schedule. Thank you.

As critically important as the data derived from these collections are, travel to those collections, and completion of this dissertation, would have been impossible without funding. I would like to thank the Graduate Student Association at the University of Nevada, Reno for providing a travel grant to fund my preliminary data collection. I would also like to thank Phi Kappa Phi for awarding the Love of Learning Award which was used toward funding preliminary data collection. The bulk of my data collection was supported by grant from the National Institute of Justice*. Thank you for seeing the importance of this work and choosing to fund its completion. Finally, I would like to thank the Bilinski Educational

Foundation for providing financial support while I wrote this dissertation. The financial support provided allowed me to focus on dissertation completion unencumbered by a teaching load or other employment.

I would like to thank my committee members, Drs. Christopher Morgan, Guy Hoelzer, and Joseph Hefner, for providing advice and critiques along the way as I formulated this project. Drs. Morgan and Hoelzer, I appreciate your serving on my committee even though my research is not directly related to your individual specializations. I have often found that an outside perspective is immensely helpful in identifying and fixing problems with research. That is no less true in this case, thank you. Dr. Hefner, I would like to thank you especially for agreeing to sit on my dissertation committee. I am not your student, nor am I a student at your university, yet when I approached you four years ago at the AAFS meetings in Seattle about this project you agreed without hesitation and with enthusiasm for my research. I also need to thank two UNR faculty who did not sit on my committee. Drs. Kyra Stull and Matthew Forrister, thank you for being available to answer my panicked questions when R would not cooperate, and helping me find solutions.

I owe the biggest debt of gratitude to my advisors, Dr. G. Richard Scott and Dr. Marin Pilloud. Richard, thank you for taking a chance in accepting me as a student. When I began, our research interests were not very closely aligned, but I have developed an appreciation for a whole new discipline, dental anthropology, because I have worked with you. You pushed me to explore topics with which I was unfamiliar, and for that I am a better anthropologist. Marin, you did not choose me as a student, and still I felt that I received as much attention and support from you as any of your students whom you did choose. I am prepared to enter the professional world because of the advice and the opportunities you have given me, and cannot stress how appreciative I am of the time you have invested in me and my research. You have both seen me through the end of this project with nothing but patience, advice, and encouragement.

To my friends and colleagues at UNR, thank you for helping me refine ideas and listening to any complaints I may have had along the way. I also need to thank the Heim family. Thank you for always making me feel like part of your family. Whether you offered a place to stay or sent care packages while I wrote my qualifying exams, you have always celebrated my achievements as proudly as my own family; the extra encouragement was always appreciated. Kelly, you, more than anyone else, have been through the ups and downs of this dissertation with me. Your constant confidence in my capability allowed me to keep working even when the process seemed insurmountable and frustrating. Finally, to my mom and dad, I cannot thank you enough. You have always encouraged me to be happy in what I am doing, and have continued that advice through three degrees. The things I have accomplished in pursuit of them degree are only possible because of the work ethic you instilled in me. I have been in college of some form for ten years now, and you finally get to see the finished product. ‘Thank you’ is not big enough to express how much I owe you.

* This project was supported in part by Award No. 2016-DN-BX-0003 awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. The opinions, findings, and conclusions or recommendations expressed in this publication are those of the author and do not necessarily reflect those of the Department of Justice.

Table of Contents

List of Tables

List of Figures

Chapter 1:
Introduction

1.1 Background

Forensic anthropology is the application of the methods and theory of anthropology, especially biological anthropology, within a medico-legal context. Typically, forensic anthropologists are called in when remains are skeletonized and the identity of the decedent is unknown. Identification begins with the construction of a biological profile, which usually consists minimally of estimations of the individual's sex, age, ancestry, and stature. Ancestry is fundamental to the positive identification of unknown individuals, but can be difficult to accurately estimate because of constantly changing genetic substructures and cultural definitions and ideals. This difficulty is further compounded by the fact that ancestry as assessed from the skeleton and self-identified race may not be congruent (Gill 1995, Hinkes 1993).

The global degree of human variation is vast and clinally distributed (Lewontin 1972, 1995, Livingstone 1962). However, some skeletal variation shows geographic patterns (e.g. Cavalli-Sforza 1997, Cavalli-Sforza and Piazza 1994, Dobzhansky 1962, Howells 1989, Jorde and Wooding 2004, Ousley et al. 2009, Relethford 2009), which roughly correspond to the social concept of race, especially as conceptualized in the United States (Ousley et al. 2009, Sauer 1992). The correspondence between biology, geography, and culture when it comes to human variation has been termed ancestry, and it is on this aspect of human identity that forensic anthropologists can comment (AAPA 1996, Brues 1992, Edgar and Hunley

2009, Hefner 2003, Kennedy 1995, Konigsberg et al. 2009, Ousley et al. 2009, Sauer 1992).

Forensic anthropological investigations are intrinsically tied to the legal system. Forensic anthropologists are frequently called upon to present their findings in a court of law; however, the guidelines governing the admissibility of those findings have been subject to many revisions (Christensen 2004, Grivas and Komar 2008). The first ruling with direct relevance to expert testimony is *Frye v. United States* (1923). The "*Frye* Rule" essentially permitted expert testimony if the methods informing the testimony were generally accepted by the appropriate scientific community (Christensen 2004). However, by the latter half of the 20th century, new methods were constantly employed, such that the "*Frye* Rule" was frequently ignored or modified (Christensen 2004). This uneven application of the "*Frye* Rule" led to the eventual enactment of the *Federal Rules of Evidence* in 1975. These new rules stated that an individual qualified by, "knowledge, skill, experience, training, or education," may testify as an expert on appropriate scientific information (Federal Rules of Evidence 1975). However, these new rules only created confusion about the "*Frye* Rule" as the general acceptance criterion of the latter was not mentioned (Christensen 2004).

The ruling in *Daubert v. Merrell Dow Pharmaceuticals Inc.* (1993), handed down by the U.S. Supreme Court, provided an evidentiary standard to which all scientific testimony must comply. Importantly, the *Daubert* ruling made the test of general acceptance less important, "in order that a reasonable minority opinion may be admitted into evidence" (Christensen 2004:2). This allowed for the inclusion of new techniques, like the emerging field of DNA analysis, although they may not yet be widely accepted. The *Daubert* ruling also established standards that evidence must meet to be considered scientific. These included a stipulation that the method(s) informing the testimony be testable, subject to peer review, and that the implementation of these methods be standardized and have known rates of error (Christensen 2004).

Many methods, both metric and morphoscopic, have been devised to aid in the estimation of ancestry from the human skeleton (e.g. Angel and Kelly 1990, Baker et al. 1990, Byers et al. 1997, Christensen et al. 2014, Edgar 2009, 2015, Giles and Elliot 1962, Gill 1998, Hefner 2009, Hinkes 1990, Krogman and Işcan 1986, Ousley and Jantz 2006, Rhine 1990, 1993). However, existing methods are not applicable in all circumstances (Gill 1998, Klepinger 2006, Rhine 1990) or may not meet current evidentiary standards (*Daubert v Merrell Dow Pharmaceuticals Inc.* 1993). Adding to the "toolbox" of methods from which anthropologists can draw when making ancestry estimates is vital to continued success in the identification of unknown individuals.

1.2 Research Question and Hypotheses

This research will answer the question, "Can ancestry estimation methods in forensic anthropology be improved through increased standardization, the inclusion of dental variables, and the infusion of more robust statistical methods?" The question above will be answered by testing several interrelated hypotheses:

1) <u>Morphoscopic traits of the cranium will separate groups by ancestry in the forensic context.</u>

    Previous work with morphoscopic traits (Hefner and Ousley 2014) has suggested they can be used to accurately estimate ancestry in over 85% of cases. This research will try to replicate those results, and explore what effect, if any, that the addition of more cranial variables has on the rate of successful classification.

2) <u>Variation in dental crown morphology will separate groups by ancestry in the forensic context.</u>

    In the United States, forensic anthropologists typically estimate ancestry as white, black, Asian/Native American, or Hispanic (Bass 2005, Burns 2007, Byers 2011, Kennedy 1995, Klepinger 2006, Sauer 1992, 1993, Spradley et al. 2008, Spradley and Weisensee 2013). If dental morphology alone is useful in predicting ancestry, then it is expected that a classification model based on dental morphological traits will be accurate more often than from chance; in this case, accurate more than 25% of the time. Since there are four categories in which an individual can be placed, features that have no predictive value would be expected to classify individuals randomly.

3) <u>The combination of cranial morphoscopic and dental traits will improve the accuracy of ancestry estimates.</u>

Assuming the above hypotheses are supported, the combination of cranial morphoscopic and dental morphological traits, will more accurately estimate ancestry than either method independently. The inclusion of more traits will allow the forensic anthropologist to make accurate assessments from remains that may have damaged or missing elements that preclude the use of all traits in a single method.

1.3 Importance

This project will provide a method of ancestry estimation that is accurate, validated, and reliable, including probabilities and known rates of error. Previous researchers (Hefner et al. 2012) have acknowledged that the defensible use of morphoscopic data requires large datasets, standard protocols for recordation, and statistics that are appropriate to the data being analyzed. This research meets those goals. The addition of dental morphology allows for more data to be used in the forensic estimation of ancestry. Teeth preserve well and are frequently recovered in forensic contexts, but there are a limited number of methods for their application in forensic anthropology that meet the evidentiary standard of *Daubert*. Additionally, dental and cranial morphology develop differently and likely capture different information with respect to ancestry (Martini et al. 2009, Scheuer and Black 2004). Therefore, the inclusion of dental morphology in the forensic estimation of ancestry may provide a broader picture of an individual's ancestry. Furthermore, the research proposed here focuses on standardization. Many of the cranial and dental traits employed in this research have been standardized (Hefner 2009, Turner II et al. 1991); however, a few are newly standardized and have not been fully explored; namely, palate shape, molar crenulations, diastema, and dental crowding. Therefore, not only does this research contribute data on the frequency of these traits in different populations, it also tests a standard method for defining and recording these traits that is currently lacking.

These contributions are important to forensic anthropology because the proposed methods comply with the established standards of admissibility in court (*Daubert v. Merrell Dow Pharmaceuticals Inc.* 1993) and address the shortcomings of forensic science in the United States enumerated by the National Academy of Science report (NAS/NRC 2009). A thorough investigation of forensic science conducted by a joint committee of the National Academy of Sciences and the National Research Council revealed a state of fragmentation and inconsistency that is incompatible with standards of evidence outlined in *Daubert* (NAS/NRC 2009). Among the most significant problems addressed by this report is a lack of standardization across forensic science disciplines; even in disciplines where standards exist, they are frequently unclear and unenforced (NAS/NRC 2009). The lack of consistency "pose[s] a continuing and serious threat to the quality and credibility of forensic science practice" (NAS/NRC 2009:6). Following these discoveries, the National Academy of Sciences called for more standardized methods, research that addresses the accuracy and reliability of existing methods, and the support of organizations that encourage such work (NAS/NRC 2009). In response to this call, many forensic sciences set out to make their own methods more robust. It is in part as a response to this call that this research has been conducted.

Chapter 2:
The History of Physical Anthropology and Race

"Physical anthropologists derive their professional existence in part from the fact that human variability exists" (Rhine 1993:54). This variability has been used to classify individuals into discrete groups termed 'races,' and the study of race has wavered in and out of focus in physical anthropology. However, the arbitrary nature of biologically-based population boundaries and political implications of studying race have led some to abandon its study entirely, resulting in two diametrically opposed views concerning race in physical anthropology (Gill 1990, Sauer 1993). There are those who believe biological races do not exist, have been proven to not exist, and should be abandoned as a subject of study in physical anthropology (e.g. Armelagos and Goodman 1998, Armelagos and van Gerven 2003, Lewontin 1972, 1995, Livingstone 1962, Smay and Armelagos 2000). There are also anthropologists who believe that human variation is patterned, that distinct populations exist, and that ignoring their existence is essentially bad science (e.g. Garn 1965, Gill 1990, Ousley et al. 2009, Sauer 1992, 1993).

Whether or not race is a biologically valid concept, or even an appropriate term to apply to human variation, treating it as a taboo subject in physical anthropology only serves to prevent the advancement of our understanding of it (Gill 1990). Rather than a point of contention among anthropologists, and within physical anthropology, an understanding of race, "…should be viewed as tantamount to an effective understanding of [the] evolutionary process itself" (Gill 1990: viii). Understanding the variation that has been used to classify groups into races, the mechanisms underlying that variation, and how that variation is truly distributed can only strengthen future anthropological research, and understandings of human variation.

To better understand what is meant when the term 'race' is applied, it must be considered historically. An evaluation of the evolving meaning and accumulating connotations of race helps pinpoint the variations used in delineating races, and assess if continued study of those variations is a fruitful avenue for physical anthropology.

2.1 The Origins of Human Races

An argument could be made that the idea of clearly divided races of human beings was a product of long-distance nautical exploration during the 15th-17th centuries (Brace 1995). Medieval explorers, like Marco Polo and Ibn Battuta, traveled widely and observed first-hand the biological diversity of extant human populations. However, the limitations of the available technology necessitated that long-distance trips be broken into several shorter segments. Since human variation is clinally distributed, the shorter segments of travel meant that the explorers were exposed to human variation at multiple points along that continuum. Consequently, the populations encountered did not appear to demonstrate any point at which a stark change in phenotype occurred, rather these changes occurred gradually as travel progressed (Brace 1995). The introduction of long-distance nautical travel, allowed for the European discovery of the Western Hemisphere, and fundamentally changed how human variation was viewed. Suddenly, European explorers were covering long distances at one time, and not observing any of the human variation occurring between the point of departure and their destination. Since populations more distant from one another in the clinal distribution were juxtaposed, the dramatic differences observed suggested inherent differences in groups of humanity.

Scientists during the Enlightenment (17th-19th centuries) were enthralled by the emerging fields of taxonomy and systematics. This new focus is perhaps best characterized by the work of Carolus Linnaeus. In his attempt to classify the physical world in *Systema Naturae* (1759), Linnaeus treats human races as taxonomic units equivalent to subspecies, identifying four distinct groups: *Homo sapiens afer, Homo sapiens americanus, Homo sapiens asiaticus,* and *Homo sapiens europaeus*. Although the Enlightenment is generally characterized by a decreased influence of religion on science, several concepts rooted in religious ideology remained influential. Among these, the Great Chain of Being had a notable effect on Linnaean taxonomy (Brace 2005, Hefner 2007, Ta'ala 2015). The Great Chain of Being posits that the world is hierarchically structured, with the most sophisticated and complex beings at the top. Once the concept of separate groups of humanity was introduced, the Great Chain of Being was expanded to include these groups in the hierarchy; classifying them on the basis of biology, behavior, and culture (Ta'ala 2015). The Linnaean system of human taxonomy, the first systematic classification of human variation, influenced subsequent attempts to define human races (Hefner 2007, Sauer 1993, Ta'ala 2015).

Roughly contemporary with Linnaeus, Johann Blumenbach, proposed a different racial classification scheme of humanity. Bearing in mind the considerable variation in human cranial form, Blumenbach (1775) proposed five races: Caucasian, Mongolian, Malayan, Ethiopian, and American. The primary difference between this view and that proposed by Linnaeus is the mechanism of variation. Whereas the Linnaean taxonomy is based on the idea of natural subdivisions in man, and their inherent hierarchy, Blumenbach employs an idea akin to clinal variation:

> "The variations of skin color, stature, body proportions, etc. which we have been able to observe, considerable though they may appear at first sight, have no absolute value; they all merge gradually one into another and, accordingly, classification into human races is arbitrary." (Bendyshe 1865 translation of Blumenbach in Comas 1960:16).

This approach to human variation, and his focus on skeletal morphology as a source of data have led some to call Blumenbach the first physical anthropologist (Brace 1982, Cook 2006).

Though progressive for the time, the classification system of Blumenbach is not without problems. Blumenbach was a monogenist; thus, the varieties of man were a product of varying degrees of degeneration from the ideal human form represented by an initial divine creation. Degeneration, as used by Blumenbach, is not an inherently hierarchical concept (Brace 2005). In this context, degeneration represented an absolute level of change from the hypothetical ideal, without a value judgment on those changes (Brace 2005). For Blumenbach, degeneration occurred through a combination of environmental factors including migration, diet, disease, and isolation (variables of striking resemblance to those we now recognize as key to human adaptability) acting on a common ancestral population (Brace 2005, Hefner 2007). In fact, Blumenbach himself spoke to the accomplishments of non-Caucasian races, and the equality of their intellects (Blumenbach in Bendyshe 1865). However, since Blumenbach identified the Caucasian race as the ideal ancestral population from which other races degenerated to varying degrees, future scholars would adopt his work as hierarchical, asserting that those races that had deviated furthest from the Caucasian ideal, were the most inferior. This idea has been wrongly ascribed to Blumenbach himself (e.g. Gould 1994, 1996).

The study of race in the United States during the middle of the 19th century was different than it had been in Europe, due largely to the prevailing political climate. The science of race had serious political implications as the U.S. was undergoing an unprecedented amount of immigration from Europe, and had more frequent dealings with Native Americans as westward expansion continued. In this socio-political climate, it was desirable to draw hard boundaries between groups of people, and the idea of races as separate creations, or polygenesis, took hold (Hefner 2007). One prominent scholar to adopt this point of view was Samuel Morton. Morton divided humanity into five races, the same as Blumenbach; however, these races were not the product of the unique pressures of the local environment, but separate lineages that arose in isolation from one another (Morton 1839). Morton's position in *Crania Americana*, and later reinforced by the examination of a series of

Egyptian mummies, is that distinct types existed too near in time to the Biblical flood for the sons of Noah hypothesis, often espoused by monogenists, to be accurate. Therefore, races must have been separate from an earlier point in time (Morton 1839, 1844, Thomas 2001). Although Morton is often credited with its inception (e.g. Thomas 2001), the more radical ideas regarding polygenesis and race are more likely the work of Morton's disciples, Josiah Nott and George Gliddon (Cook 2006). Their book, *Types of Mankind* (Nott and Gliddon 1854), focused on distinct creation events as the basis for human variation. They argue for a concept of 'zoological province' from which each race of man originated, and to which each race was ideally suited.

Though typological, Morton's concept of race was not based on preconceived notions of intellectual and cultural capacity. As an example, the monumental architecture present in the Americas was largely considered to be the work of some extinct population of 'Caucasian peoples,' an idea referred to as the "Moundbuilder Hypothesis," as it was commonly believed that complex architecture was beyond the capabilities of Native American groups (Cook 2006). Morton's work in *Crania Americana* (1839) demonstrated that individuals buried at the sites of this architecture shared racial affinity to extant Native American populations (Cook 2006), thus supporting his own view that the mounds were of Native American creation. Though his polygenist position on the nature of race has since been abandoned, Morton had an immeasurable impact on American physical anthropology. Like Blumenbach before him, Samuel Morton demonstrated the relationship between race and the features of the human skeleton, particularly craniometric measurements (Cook 2006, Hefner 2007). This relationship and the research it inspired form the foundations of physical anthropology.

2.2 The Founding Fathers of Physical Anthropology

Although the biological concept of race has roots in the Enlightenment and early naturalists, its position in American anthropology can be traced to three founding-fathers of the discipline: Aleš Hrdlička, Earnest Hooton, and Franz Boas (Caspari 2009, Marks 2010). Each of these men contributed incalculably to the development of physical anthropology in the United States, as well as to anthropology more broadly. The biological data on race and our understanding of genetics have advanced the conversation on race by monumental strides; however, the general theoretical perspectives on race can be traced to the views of these three men (Caspari 2009).

The American conversation concerning race in physical anthropology was largely influenced by the political climate in the early 1900s and the increasing interactions between the government and Native Americans as expansion continued to push west. Where other sub-fields of anthropology were heavily engaged in salvage anthropology, physical anthropology focused on typological studies of Native Americans and Alaskan Natives. A fair amount of research that looks back on this early work is justly focused on the racist overtones, but tends to obscure the fact that scholars were asking significant anthropological questions, particularly about Native American origins (Cook 2006). As this research was published, it became clear that there was no consensus regarding the definition of race. Race was simultaneously employed to describe the biological variation of human populations, the sociocultural meanings associated with that variation, and the supposedly stable essence of a population, or a type (Caspari 2009). It was into this confused academic sphere that Aleš Hrdlička, Earnest Hooton, and Franz Boas submitted their views on race that individually shaped the research of the field for the century to follow.

2.2.1 Aleš Hrdlička (1869-1943)

Aleš Hrdlička is known primarily for his work at the Smithsonian Institution. He is also well-known as the founder of the *American Journal of Physical Anthropology*, and germane to this research, he proposed an early classification system for shovel-shaped incisors and other traits of the incisors (Hrdlička 1920, 1921). However, his work concerning race has had a clear and lasting impact on the study of race in physical anthropology. In his opinion, understanding human variation is uniquely within the purview of the physical anthropologist, and as such, his career was devoted to determining the range of normal human variation through extensive description (Buikstra 2006, Caspari 2009). However, some have labeled his views on race, and its centrality to applied anthropology, as deterministic and unfounded (e.g. Blakely 1987, Caspari 2009).

To understand Hrdliçka's point of view regarding the relationship between race and physical anthropology, it is important to first consider his primary influence, Paul Broca. Though now more famous for his contributions to neuroscience, Broca was instrumental in shaping the trajectory of physical anthropology in Europe (Cook 2006, Ta'ala 2015). Broca was a polygenist, believing that the multiple origins, and fixity, of races was, "more scientifically sound," than the religious underpinnings of monogenesis (Ta'ala 2015:4). This belief was in direct contrast to evolution by natural selection as proposed by Darwin (1859), which was largely dismissed by contemporary European physical anthropologists (Brace 1982, Ta'ala 2015). Eventually, the point was ceded that species change over time, as evidenced by Broca's own research with craniometrics; however, natural selection as the primary mechanism of that change remained in doubt (Brace 1982, Ta'ala 2015).

Hrdlička was profoundly influenced by the French school of physical anthropology (Brace 1982, Cook 2006). Using Morton and Broca as guides, Hrdlička sought to apply craniometrics to the questions of race that contemporary anthropologists were facing. Hrdlička was a typologist, and used the power of craniometry to discretely subdivide humanity, establish the superiority of certain groups, and undertake a quest for the discovery so-called pure races (Blakey 1987, Caspari 2009, Ta'ala 2015). In the political climate of the early 1900s, Hrdlička sought to make physical anthropology an applied discipline, the methods and conclusions of which could be used in support of the eugenics movement and biological determinism (Caspari 2009, Ta'ala 2015).

Though the more deterministic leanings of Hrdlička's work have largely disappeared, the spirit of his methods persists in some studies of human variation. Hrdlička's work was based on a minimal understanding of human genetics and practically no statistical analyses (Ta'ala 2015), even going so far as to refer to statistics as, "the invention of the devil," and, "the ruin of the discipline" (Brues 1990). This emphasis on description and fear of statistical analyses persisted in physical anthropological studies of race until the latter half of the 20th century, when it was alleviated by the advent of computers and the modern evolutionary synthesis (Huxley 1942). For Hrdlička, the focus of physical anthropology in the United States was primarily racial description and categorization (Blakely 1987, Caspari 2009, Cook 2006). The goal of his work was to use science to demonstrate an inherent link between the biological and social concepts of race; a link that was not substantiated (Caspari 2009). Though his impact on the sub-discipline of physical anthropology is immense, his is a legacy of description with respect to the anthropological study of race.

2.2.2 Earnest Hooton (1887-1954)

Earnest Hooton had perhaps the most profound influence on modern American physical anthropology (Caspari 2009, Cook 2006, Hefner 2007, Spencer 1981, Ta'ala 2015). If not for his contributions to the discipline—which were many—then for the voluminous number of students he mentored in physical anthropology (Caspari 2009, Garn and Giles 1995). Therefore, examination of his perspective on race is vital to understanding the perspective taken by later members of his academic lineage. Though his impact on the field is widely agreed upon, his perspective on race is somewhat less clear. Some have argued he was a racist and a polygenist seeking to support the biological separation of races (e.g. Wolpoff and Caspari 1997), while others (e.g. Brace 1981) view him as taking a polyphyletic approach, viewing race as convergent toward similar forms due to environmental pressure.

Hooton was different from contemporary anthropologists (e.g. Hrdlička and Boas) in several important ways, and those differences influenced how he perceived the question of race. Firstly, Hooton was not trained as an anthropologist or even an anatomist, as many of his contemporaries were (Cook 2006). His interest in physical anthropology can be traced to his work with Sir Arthur Keith, an association likely responsible for Hooton's interest in races and his conception of them as discrete populations (Cook 2006). This typological approach allied him with others, such as Frederic Putnam, who viewed the degree of human variation within a population as evidence for multiple origins; this is especially evident in his seminal 1930 work *Indians of Pecos Pueblo* (Cook 2006, Hooton 1930).

For Hooton, the races were not inherently tied to different creation events, but rather to different environmental pressures acting on different populations of hominins (Brace 1981, Caspari 2009, Hooton 1931b, Wolpoff and Caspari 1997).

Race, therefore, could be viewed as a taxonomic unit, useful in the classification of the species (Caspari 2009). The traits that defined these groups were not simply inherited, but were indicative of the population's evolutionary history. For Hooton, there was a limited number of primary races, taxonomically equivalent to subspecies. These races could undergo varying degrees of mixture to produce secondary and tertiary races, thus explaining extant human variation (Caspari 2009). In spite of this approach to race, Hooton was not interested in human variation as a whole, but rather as it applied to specific questions founded in archaeology (Buikstra 2006). Early in his analyses of race, Hooton makes clear that those skeletal traits most informative to this question would be non-adaptive (Hooton 1918, 1926), a surprisingly modern position. He rightly assumed that traits with adaptive significance would not be informative with respect to common descent, a position reflected in his contribution to the first issue of the *American Journal of Physical Anthropology* (Caspari 2009, Cook 2006, Hefner 2007, Hooton 1918).

There are two problems with correctly fitting the views of Earnest Hooton into an examination of the race concept in physical anthropology. The first is that Hooton applies the term 'race' to multiple phenomena, so it is difficult to ascertain which aspect of human variation he is referring when 'race' is used. For example, Caspari (2009) explores four different ways that Hooton employs the term 'race' in Hooton's (1918) "On certain Eskimoid characters in Icelandic skull." First, Hooton used race to describe continental groups, such as European and American, a scheme that would later be echoed by the geographical races of Garn and modern forensic anthropology (Hooton 1918:71). Second, race described divisions within a continental group, such as the European taxonomy proposed by Ripley (1899) (Hooton 1918:55). Third, Hooton refers to his samples, Icelandic and Eskimo individuals, as separate races (Hooton 1918:57). Finally, race is used as a synonym for culture in the context of describing the behaviors of “civilized” and “primitive” races (Hooton 1918: 54). Therefore, Hooton’s use of ‘race’ is difficult to understand beyond the meaning of a group with some morphological or cultural commonality. This ambiguity is cleared in a later publication in which Hooton (1936) describes races as sharing morphological and anatomical features due to shared ancestry.

Despite some outmoded terminology and misguided beliefs concerning race, many of Hooton’s contributions to the field are surprisingly enduring. Though certainly not unique to Hooton, he was a proponent of the estimation of race from the characteristics of the bones, even hinting at the fact that some of these traits are correlated with one another (Caspari 2009, Hooton 1936), and that there is a certain amount of variation in human crania that cannot be attributed to differences in culture or environment (Cook 2006). Additionally, the methods devised by Hooton were an early attempt at standardization, and many of them persist in modern forensic anthropology (Hefner 2007), which will be further addressed in subsequent chapters. Standardization of non-metric observations was not common, and Hooton himself recognized the need for standardized data in making large comparisons (Hooton 1946).

Hooton’s mechanism for the emergence of human races is reminiscent of later multi-regional models, but with a much earlier last common ancestor (Caspari 2009, Hooton 1931b). While this was a clear influence on Carleton Coon (discussed below), Hooton did not believe that racial differences necessitated cultural or mental inequalities, or at least that such inequalities had yet to be scientifically proven (Hooton 1936, Ta’ala 2015). He was a proponent of the objective truth of science, being quoted as saying, “[A] physical anthropologist…desires emphatically to dissociate the finding of his science from the acts of human injustice which masquerade as ‘racial measures’ or ‘racial movements’ or even ‘racial hygiene’” (Hooton 1936).

Before moving on to the final of the early anthropologists under examination here, two of Hooton’s students, Carleton Coon and Stanley Garn, deserve mention. Their takes on the position of physical anthropology with regard to race strongly influenced the trajectory of race studies in later decades. Coon and Garn took drastically opposed views on human racial variation (Hefner 2007). Garn, more closely followed Hooton’s approach, with an emphasis on geography. Garn (1965) identified nine geographically discrete races, with the postulate that similarity is a function of distance between the large groups, and within smaller micro-races. This idea is similar to the modern concept of ancestry that is addressed later in this chapter. Carleton Coon, on the other hand, asserted that human races arose from five different populations of *Homo erectus* that achieved sapience at different points in their history (Coon 1962:658). Because of this deep time connection, and differing rates of achieving *Homo sapiens* status, Coon equated racial differences to differences in cultural development. The general reaction to Coon's work within the anthropological community largely steered physical anthropologists away from questions of race, and toward a focus on human variation (Marks 1995, 2010).

2.2.3 Franz Boas (1858-1942)

The name Franz Boas is practically synonymous with the American school of anthropology. Due to his background in both physics and geography, it is unsurprising that his vision of anthropology in the United States was highly scientific. Therefore, his views on race are fundamentally different from his contemporaries. Rather than viewing human variation in terms of separate groups with stable traits, Boas acknowledged human variation as continuous (Boas 1918, Caspari 2009, Marks 2010). In his view, the differences in so-called racial groups are better explained by geography than anything inherently biological (Boas 1912, Caspari 2009, Marks 2010). Furthermore, an understanding of human variation requires consideration of the evolutionary mechanisms that cause it. For Boas, the effects of environment and culture on biology are more substantial than the effect of biology, in the form of race, on culture (Boas 1912, Caspari 2009). Franz Boas became one of the most prominent opponents of racial science, and the concept of racial typology.

Though known today predominantly for his work in cultural and linguistic anthropology, Franz Boas was one of few American anthropologists whose research embodied the four-field approach. Perhaps Boas's most obvious contribution to physical anthropology was *Changes in Bodily Form of Descendants of Immigrants* (1912). Boas's work on race and human variation is rooted in the observation that there is no consistent relationship between assumed racial traits and human populations. In other words, no single racial feature can be used to divide human populations (Boas 1912, Caspari 2009). Prevailing ideas concerning race at the time identified cranial types, that were considered to be stable through time, such individuals of one type would be largely similar to individuals of that same type from later in a temporal sequence. This stability was the basis for the typological perspective on race, that is types existed and could be used to consistently separate human groups throughout time.

Using an immense set of craniometric data derived from Ellis Island immigrants, Boas (1912) investigated whether the cranial types of immigrants were passed on from generation to generation, and stable based on the population from which they were derived. For example, would the children of Italian immigrants display the same cranial type as their parents. Analysis of the cranial data, such as comparisons of cranial indices, revealed considerable differences between European-born parents and their American-born children. These results demonstrated not only the plasticity of the human skull, but also the capacity for cranial form to change within a single generation due to environmental factors (Boas 1912). More than any other, this study undermined the racial typology of the early 20th century (Gravlee et al. 2003). If cranial type can change within a single lineage over the course of a single generation, then there is no basis for the use of types to characterize entire populations, especially in series comprising multiple temporal elements (Boas 1912).

For decades, Boas's results were taken as proof that racial types do not exist; however, the analytical methods available to Boas in 1912 were limited. Therefore, subsequent studies have used more advanced statistics to re-evaluate Boas's data and conclusions (e.g. Gravlee et al. 2003, Sparks and Jantz 2002). His original conclusion was that environment played a significant role in cranial morphology. Gravlee and colleagues (2003) reanalyzed the data and agreed with Boas's overall conclusions: there can be generational changes in cranial form without change to the underlying genetics. The human cranium is plastic, and the differences observed between populations can be explained by different environments. However, Sparks and Jantz (2002), came to a slightly different conclusion. Although they agreed with the idea of cranial plasticity as used previously (Boas 1912, Gravlee et al. 2003), they found that the environmental response alone does not produce sufficient variation to account for between group differences. In fact, a principal component analysis of the Boas data revealed an East-West pattern that more clearly reflects differences among European ethnic groups than it does differences between American and foreign-born individuals (Hefner 2007, Sparks and Jantz 2002). The non-racial approach to human variation advocated by Boas is a hallmark of the "New Physical Anthropology" as described below (Washburn 1951). However, it is important to consider that Boas did not argue that human variation is not patterned.

He argued that racial typology and the concept of race cannot and should not be used as the basis for prejudice (Ta'ala 2015).

The three perspectives described above illustrate the varying climate regarding race as American physical anthropology approached the mid-twentieth century. Aleš Hrdlička fought for an applied anthropology based in the typology concept, and rife with racial description. Earnest Hooton also took a typological approach to race, though in a more standardized fashion. Importantly, he began to consider the mechanisms behind human skeletal variation and tried to account for the effects of evolutionary forces like natural selection. However, the work of some of his students (e.g. Carleton Coon) precipitated a shift away from race studies in American anthropology as a whole. Franz Boas, took an entirely different approach. He understood the interplay of biology, culture, and the environment, becoming perhaps the most vocal opponent of racial typology. This thread, looking a human variation as opposed to racial types, has continued through to the present day. Each of these early anthropologists influenced both one another, and the trajectory of physical anthropology.

2.3 Race and the "New Physical Anthropology"

Following World War II there was a change in science in general. There was a shift away from description and measurement, and toward research based in modelling and experimentation (Stini 2010). Advances in genetics and biology necessitated changes in physical anthropology. Most notably anthropologists began to re-examine long-held notions on the validity of human races. This change came in the wake of Nazi racial science, and was furthered by the American Civil Rights Movement. If before the war anthropologists argued for some inherent, but intangible, connection between biology, intellect, and culture, the war crimes of the Holocaust left many physical anthropologists, especially those clinging to typology, defending their interests in racial variation (Marks 2010).

For scholars interested in race after the war, population genetics and biodistance became the more acceptable options (Cook 2006). However, American physical anthropology faced an ideological dilemma. On one hand, the sub-discipline was essentially founded on studies of human races, and it would be unrealistic to abandon their study entirely. On the other hand, given the sensitive nature of the topic, it became increasingly difficult to disentangle scientific positions on race from social ones (Relethford 2010). Therefore, at this point, Marks (2010) identifies four directions in which anthropological considerations of race divert. First, some anthropologists sought to remove race entirely from the purview of physical anthropology, focusing instead on evolutionary dynamics, paleoanthropology, and non-human primates. The second approach was to retain a formal idea of race, but downplay the implied behavioral and cultural differences between races. The third and fourth approaches are somewhat tied together. Some anthropologists sought to re-conceptualize race in terms of global human variation and the effects of adaptation. In conjunction, many population geneticists were stepping in and redefining race in terms of genetics, instead of anatomy. Near the end of the 20th century, the relationship between physical anthropologists and the study of race had completely changed.

The removal of race from the sub-discipline of physical anthropology is best characterized by the "New Physical Anthropology" of Sherwood Washburn (1951). Washburn's academic career was multi-disciplinary in many ways, which carried over into his vision for American physical anthropology. His personal approach to research was "an amalgamation of functional anatomy, population genetics, and behavioral biology" (Stini 2010:177). This approach to studying humanity was focused on the interplay between biology and culture, and how modern variation was the result of generations of that relationship (Washburn 1951).

In 1950, Washburn organized the Cold Spring Harbor Symposium, along with Theodore Dobzhansky, renowned population geneticist, where the ideas that eventually became the "New Physical Anthropology" were first suggested. The principal themes of the symposium were the origins and races of humankind, with an emphasis on population variation, population genetics theory, examination of the evolution of whole populations, and a re-evaluation of race through the lens of genetic and population variation (Bates 1952, Marks 2010, Relethford 2010, Stini 2010). The Cold Spring Harbor Symposium did not redefine physical anthropology outright, but rather was a gathering of some of the most prominent physical anthropologists of the day, including Lewontin, Boyd, Hooton, Garn, Coon, Montagu, Howells, and Angel, and a coalescence of ideas that had been circulating since the end of World War II (Stini 2010).

One year later, when the "New Physical Anthropology" was published it took an unsurprisingly evolutionist bend. The "New Physical Anthropology" was characterized by the study of the process of evolution in both human and non-human primates. Classic evolutionary theory was used, underpinned by population genetics. Most importantly for the question of human variation, races were no longer considered appropriate units of analysis. Physical anthropologists were to study populations and the relationship of body and skeletal form to function, but study of the 'types' of man was considered anachronistic (Washburn 1951). The "New Physical Anthropology" does not seem like a radical departure from modern physical anthropology because it was the "New Physical Anthropology" that survived. However, in a sub-discipline that built its reputation on the scientific study of race, redefining it to exclude race studies was a risky move, and one that likely saved the discipline from the fate of more overtly biased fields of study like criminal anthropology, Bertillonage, and eugenics.

Following the "New Physical Anthropology," many organizations felt pressure to pronounce an official position on race and the future of race studies. The United Nations Education, Scientific, and Cultural Organization (UNESCO) drafted a team of scholars, including anthropologist Ashley Montagu, to pen their position (Marks 2010, Ta'ala 2015). Montagu was in attendance at the Cold Spring Harbor Symposium that spawned the "New Physical Anthropology" and was a vocal opponent of the biological basis of race (Witkowski 2017). Although written by the entire committee, the original UNESCO statement has come to be known as the

'Montagu statement' because his influence is so readily apparent (Brattain 2007, Marks 2010, Stewart 1961).

The original UNESCO statement on race was not as radical as the backlash it provoked may suggest. The official stance was that all human populations were biologically, intellectually, and culturally equal, and that race was more a social construct than a biological fact (Brattain 2007, UNESCO 1951). However, this statement was not well-received by many for two reasons. First, of the seven-member committee that wrote the statement, only two (including Montagu) were anthropologists. Though not inherently problematic, anthropologists had been the preeminent scholars of race, and their exclusion from the committee was viewed by many as an oversight resulting in an overly-idealistic and unfounded view of race (Brattain 2007, Marks 2010). In addition, many felt that the absence of those more familiar with the scholarship on race may have led to inaccuracies in the final product (Brattain 2007, Marks 2010, Müller-Wille 2007). Second, the language of the statement is unscientific in many places, and the language used became a tool for proponents of racial science to undermine the statement for lack of scientific grounding (Marks 2010).

Eventually, UNESCO caved to the criticism and convened a second committee to rewrite the statement on race. This time the panel was comprised mainly of biologists, many of whom had known racial agendas based on their vehement opposition to the original statement (Brattain 2007, Marks 2010). The revised statement maintained the spirit of the original in less absolute terms. For example, rather than stating the equality of all races with regard to intellect and cultural capacity, the second statement held that differences in these arenas between races had yet to be scientifically supported (Brattain 2007, Marks 2010, Relethford 2010). Following both the "New Physical Anthropology" and the UNESCO Statement in the 1950s, physical anthropology was essentially divided between those who had abandoned race and those that were determined to keep the study of race at the heart of physical anthropology.

Though popular, the "New Physical Anthropology" was not universally well-received. Many of the opponents of the so-called Montagu statement were equally vocal in their opposition of the "New Physical Anthropology", arguing that physical anthropology remained the most appropriate domain for the scientific study of human variation, and by extension, race. The less extreme members of this group (e.g. Stanley Garn) sought to retain the idea of races as a biological reality, influenced by geography and evolution, and divorce the concept from intellectual and cultural hierarchies (Garn 1965). This more moderate position was quickly eclipsed by those who sought to use physical anthropology as a justification for racism and segregation in the United States.

Many of the criticisms of physical anthropology at this time were precipitated by the publication of *The Origin of Races* (Coon 1962). In this work, Coon essentially equates races with sub-species of humanity, tracing their origins to five geographically distinct populations of *Homo erectus*. He argues that each population progressed to the status of *Homo sapiens* through parallel evolution with gene flow, although this far less important than in the polycentric model of Weidenreich, on which Coon's model is based (Trigger 1965, Weidenreich 1946). Furthermore, Coon (1962:658) proposed that each race became "sapient" at a different point in their history at that those differences equated to differences in cultural development. For example, he argued that the European race was more advanced than the African race because they transitioned to *Homo sapiens* earlier, conclusions that are completely unsupported by the paleontological record. The book was divisively received. Some praised the work for being a thorough review of paleoanthropology to date, but disagreed with Coon's conclusions (e.g. Hulse 1963, Mayr 1962), others strongly objected to the likelihood of his evolutionary mechanism and its implications for human races (e.g. Birdsell 1963, Dobzhansky 1963, Montagu 1963, Oschinsky 1963, Trigger 1965), while still others proffered it as evidence for the need to keep the races separate (Jackson 2001, Marks 2010, Relethford 2010).

The political climate in the United States when *The Origin of Races* (Coon 1962) was published necessitated that a decision be made. The uptake of anthropological literature by racists and segregationists, like Carleton Putnam, proved problematic for physical anthropology at large. Among all the anthropological sub-disciplines, physical anthropology was already viewed as legitimizing biological differences. Silence with regard to the adoption of their work by segregationist groups would be construed as apathy at best and complicity at worst, a positioned many critics levied against Coon himself (e.g. Dobzhansky 1963, Jackson 2001).

The combination of the vision of the "New Physical Anthropology" and the controversy surrounding *The Origin of Races* led many physical anthropologists to abandon the study of race with even more finality than following the "New Physical Anthropology." However, physical anthropologists had been the pre-eminent scholars of race, and their abandonment of the topic left an academic vacuum. Geneticists stepped in to investigate the biological differences in human populations. The basis of these traits in genetics and not anatomy somehow made the differences more objective, and their analysis more rigorous (e.g. Boyd 1940, 1950, 1963, Dunn and Dobzhansky 1952). Cultural anthropologists stepped in to explore differences in ethnicity or ethnic groups, which effectively became a substitution for studies of race, ultimately concluding that those differences perceived as race were of social not biological origins (Marks 2010, Visweswaran 1998).

Finally, although race as a concept was relegated to the background, questions of human variation persisted in physical anthropology. The analysis of human variation needed to be re-conceptualized. Initial attempts with genetic markers (e.g. Boyd's 1950, 1963 serological analyses) have been deemed overly arbitrary by some (e.g. Marks 2010), but demonstrated the potential for genetic variation to provide a promising new avenue of research to physical anthropologists. Based on a large sample of frequencies of the sickle-cell allele Livingstone (1962:279) made his lasting pronouncement, "there are no races, only clines." Acknowledging that a central question of physical anthropology is the existence of population variability, Livingstone (1962) proposed, like Boas, a non-racial approach to human variation with a focus on the clinal distribution of human traits. This shift in perspective marks the beginning of the transition to the modern view on race, with an emphasis on the forces of evolution over static racial types. Most work on human variation for years after this point either dealt with clines (e.g. Lewontin 1972, 1995), and therefore the non-existence of races, or looked to biodistances between similar locally-adapted populations to answer questions of human variation (Cook 2006). However, some physical anthropologists continued to hold onto the idea that there are geographically patterned differences between human populations, and some form of the race concept persisted in applied areas of physical anthropology, like forensic anthropology, (Relethford 2010, Smay and Armelagos 2000, Ta'ala 2015).

The dichotomous view on the appropriate treatment of race in physical anthropology that began with the "New Physical Anthropology" (Washburn 1951) persisted for decades ultimately necessitating pronouncements by professional anthropological bodies. In 1996, the American Association of Physical Anthropologists saw fit to make an official statement for the field. In summary the AAPA has stated that all humanity is a single species and cannot be effectively differentiated at a sub-specific level (AAPA 1996). They further acknowledged that differences between populations exist, but are clinally distributed across the globe and are the result of the interaction of genes and environment, and that no single trait, or even small subset of traits, can be used to effectively delineate populations from one another (AAPA 1996). These points are directly related to the emphasis on clinal variation proposed by Livingstone (1962) and the backlash against *The Origin of Races* (Coon 1962). However, research into the geographic distribution of human traits, and the ability of forensic anthropologists to accurately assess ancestry indicated that human variation is not random. Therefore, the AAPA has also acknowledged that human variation does seem to be patterned somewhat and that "human beings who speak the same language and share the same culture tend to select each other as mates" (AAPA 1996).

Physical anthropology in the latter half of the 20th century arrived at two important conclusions regarding the place of race and race studies in the subfield. First, the concept of race used by typologists breaks down in many cases. Furthermore, the focus on description by early scholars served only to obscure the mechanisms underlying human variation (Relethford 2010). Therefore, emphasis should be placed on understanding what forces cause variation, not on creating categories into which all of humanity must fit. Second, human variation can be meaningfully and successfully analyzed without using race as a unit of analysis. Livingstone (1962) demonstrated the potential for clinal analysis of human variation, and to some extent, Garn (1965) demonstrated that geographic origin may be a more meaningful unit of analysis; there is no need to impose hard divisions on human variation (Lewontin 1972, 1995). Resolution of the seemingly incongruous positions against race, but for patterned human variation, are further explored in subsequent sections.

2.4 The Modern Concept of Race

Anthropology has rejected the biological validity of races as stable, definite divisions of the human species. Race is a social construct, based on a person's definitions of self. However, even as a social construct, race is in part the result of the interaction between objective biological differences among global populations, and the cultural perceptions of self and otherness (Marks 2010). Physical anthropology is uniquely positioned to examine that interaction by reintegrating biology with culture. Skeletal analysis provides no direct assessment of so-called racial features like skin color, but allows for an assessment of geographical origins on a broad scale, such as African, Asian, or European (Brace 1995: 171).

Physical anthropologists generally agree that human variation exists and is geographically patterned, and that non-biological factors shape that pattern (Edgar and Hunley 2009). However, argument continues over the extent of this pattern, and appropriate nomenclature for its description (Edgar and Hunley 2009). This is particularly important in the applied field of forensic anthropology. Social race is a critical part of a forensic identification, but the skeleton provides biological information only. Therefore, anthropologists estimate ancestry. Some argue that the use of 'ancestry' is a lexical difference only, and recognizing this as a valid concept only serves to reinforce ideas of racial typology in students of biological anthropology (Armelagos and Goodman 1998, Armelagos and van Gerven 2003, Smay and Armelagos 2000). However, 'race', as it has been used historically, and 'ancestry', as it is used in forensic anthropology, are different things.

Race as a biological concept was conceived of as a framework for the classification and description of inherent differences in groups of humanity, such that they were clearly and consistently distinguishable (Blumenbach 1775, Caspari 2009, Cook 2006, Linnaeus 1759). The typological approach was based on the idea that for every race there was an ideal type that was static and unchanging (Caspari 2009, Cook 2006). Although a specimen may not have matched exactly the suite of traits that defined the type, it could be classified based on a higher or lower degree of similarity to different types (Krogman 1962, Krogman and Işcan 1986). Inherently tied to the idea of race were ideas about not only biological differences, but differences in mental capacity and cultural sophistication (e.g. Coon 1962). As mentioned above, the "New Physical Anthropology" (Washburn 1951), and the demonstration that the majority of human variation is shared among groups (Livingstone 1962, Lewontin 1972, 1995), made racial typologies more difficult to accept. As a result, physical anthropology as a discipline abandoned the study of race for a time. In the absence of physical anthropologists, understanding the differences among what were perceived as races fell to other disciplines (Marks 2010).

Although anthropology has moved beyond linking race with cultural development, it has not really recovered from the impact of that connection (Smay and Armelagos 2000). Therefore, the commonly held position today is that "races do not exist”; however, this position is equally dangerous as it, “…fails to advance our understanding of the subject [and] tends to create a social taboo on a subject greatly in need of rational, dispassionate examination” (Gill 1990: viii). The more appropriate position is that races do exist, but as socially defined means of self-identifying and claiming membership in groups; therefore, races have little biological meaning (Konigsberg et al. 2009:77).

“[H]uman evolution has been, and is, characterized by many locally differentiated populations coexisting at any given time, but with sufficient genetic contact to make all of humanity a single lineage sharing a common evolutionary fate” (Templeton 1998:632). This idea is the basis of the now preferred concept of ancestry in physical anthropology. Ancestry is, “a scientifically derived descriptor of the biological component of population variation” (Konigsberg et al. 2009:78). Through examination of the forces of microevolution (i.e. natural selection, gene flow, mutation, and genetic drift), human variation can be studied without a focus on race, but instead on a geographic origin along a continuum of human variation, a position that echoes the sentiments of anthropologists like Stanley Garn (1965).

Rather than assuming inherent differences in populations, the concept of ‘ancestry’ is based on the idea that different human populations share heritable traits that make them more similar to one another and more dissimilar to individuals in other populations (Brace 1995, Brues 1977, Ousley et al. 2009). This current understanding of ancestry is based on the biological reality that microevolutionary forces have affected populations differently, and resulted in recognizable differences. Ancestry considers the population history and the resultant effect on the genetic structure of a population (Klepinger 2006, Konigsberg et al. 2009, Ross et al. 2004, Spradley et al. 2008, Spradley and Weisensee 2013). For example, ancestry intersects with social race by recognizing that social barriers (and geographic barriers) frequently serve to magnify biological differences between populations (Brace 1995, Ousley et al. 2009). Although the majority of variation occurs within populations (Lewontin 1972, 1995, Livingstone 1962), the remaining 10-15% of variation is geographically structured such that clusters of individuals correlate closely with self-assigned ancestry (Bamshad et al. 2004, Edwards 2003, Hubbe and Neves 2007, King and Motulsky 2002, Novembre et al. 2008, Rosenberg et al. 2002).

2.5 Concluding Remarks

The culmination of recent genetic, craniometric, and morphoscopic studies have reawakened old conversations regarding human variation. However, that conversation is no longer focused on separating extant groups of humanity, but rather on understanding the global pattern of human variation. Concurrent with those explorations, anthropologists continue to grapple with the appropriate use of the term race within biological anthropology. While it is true that the historical concept of race is invalid in a biological sense, it is equally incorrect to claim that the social concept of race has no bearing on biology. The concepts of race, patterned human variation, and the use of ancestry in the forensic setting cannot be effectively understood from a biological or a cultural vantage point alone. The most appropriate course forward is to recognize that humanity is a biocultural species, and that especially with respect to race or ancestry, biology and culture must be considered together (Brues 1993, Buikstra and Beck 2006).

Over 250 years of investigations into global patterns of human variation have led to several conclusions. First, scholars recognize human populations show certain physical differences, and that populations that are geographically closer to one another tend to be more similar (e.g. Cavalli-Sforza et al. 1994, Eller 1990, Relethford 2004). Second, physical and biological differences between human populations are a function of different environmental pressures and have no bearing on culture or intelligence (e.g. Armelagos and van Gerven 2003, Washburn 1951). Third, human variation is continuously distributed; that is, no boundaries can be drawn between human populations that consistently delineate groups (e.g. Lewontin 1972, 1995, Livingstone 1962). However, although continuous, human variation is geographically patterned (Bamshad et al. 2004, Marks 2010, Novembre et al. 2008, Relethford 2010). Therefore, it is reasonable to assume that the trained anthropologist can identify an individuals' general areas of origin along that continuum (Sauer 1993).

Chapter 3:
Estimating Ancestry in the Forensic Context

Race and ancestry can be difficult to reconcile, and the two concepts are often conflated in critiques of forensic anthropology. Race is socially defined, and is therefore beyond the scope of information estimable from the skeleton. Ancestry, on the other hand, is firmly rooted in human variation and microevolution, and the human skeleton provides many indicators of ancestry. Correctly estimating ancestry is critical in the analysis of unknown human remains, for purposes of both identification and in some cases repatriation. Forensic anthropologists have developed a variety of methods to estimate ancestry from the skeleton. The methods rely on the geographic pattern of human variation and focus on both metric and morphoscopic data. Although ancestry estimation has a long history in forensic anthropology, the stricter standards of evidence imposed by the *Daubert* ruling have necessitated the re-evaluation of some methods. A thorough understanding of the tools currently available, and their limitations, is necessary to continue developing new methods for estimating ancestry in the forensic context.

3.1 What Is "Ancestry" and Why Do Forensic Anthropologists Estimate It?

As defined in the previous chapter, ancestry is "a scientifically derived descriptor of the biological component of population variation" (Konigsberg et al. 2009:78). It is for that reason that forensic anthropologists estimate ancestry and not race when constructing a biological profile (Konigsberg et al. 2009). The problem arises in that the social label, 'race', is what is reported when an individual goes missing; therefore, forensic anthropologists must relate the biological reality of ancestry to the social reality of race (Kennedy 1995, Sauer 1992, 1993, Spradley and Weisensee 2013). The practice of estimating ancestry, but aligning it with social race in individual identification is seemingly incongruous. Therefore, some have proclaimed that the practice of ancestry estimation in forensic anthropology is tantamount to neo-racism and a rebirth of the typological approach to human variation (e.g. Armelagos and Goodman 1998. Armelagos and van Gerven 2003, Smay and Armelagos 2000). The primary argument is that there is no necessary congruence between the social and biological labels. Meanwhile, practicing forensic anthropologists defend ancestry estimation, acknowledging that it is imperfect, but necessary (Hinkes 1993, Kennedy 1995, Sauer 1993, Ta'ala 2015). These anthropologists argue that general patterns can be gleaned from the continuous distribution of human variation, and that looking at those broad patterns allows the anthropologist to appreciate human variation as a tool for identification (Gill 1990, Sauer 1993). What then is the justification for aligning ancestry with the concept of race? It is here that a biocultural perspective on human variation, such as is currently used within forensic anthropology, is particularly informative.

Race is a social construct not directly reflected in the biology of the individual. However, as a social force, race influences biology in profound ways (Ousley et al. 2009). Particularly in the United States, institutionalized racism and anti-miscegenation laws of the past have created a social structure where black individuals can trace their ancestry to West Africa (Ousley et al. 2009, Parra et al.

1995, 1998), white individuals trace their ancestry to Europe (Spradley and Weisensee 2013), and Hispanic individuals claim Native American, European, and/or African influence depending on their geographic origin (Bonilla et al. 2005, Cerda-Flores et al. 2002, Lisker et al. 1986, 1990, 1996, Rangel-Villalobos 2008, Rubi-Castellanos et al. 2009).

In the prehistoric past, continental populations (i.e. African, Asian, and European) were separated by vast geographic expanses, such that the forces of evolution, particularly genetic drift and natural selection, served to differentiate these populations from one another (Ousley et al. 2009). Without the technology to efficiently cross these expanses, the homogenizing effects of gene flow were negligible, causing further accentuation of differences among populations.

As the world becomes more interconnected, environmental deterrents to gene flow are less pronounced (Templeton 1998). Therefore, the primary mechanisms dictating the exchange of genes between populations are cultural (Edgar 2007, Hefner 2007, Ousley et al. 2009, Parra et al. 1998, Sauer 1992). The AAPA has recognized that “people who speak the same language and share the same culture tend to choose each other as mates” (AAPA 1996). While mating structure may not be an inherently biological process, in this case positive assortative mating, it has a significant effect on biology (Manly 1985, Wright 1921). Positive assortative mating is structured such that phenotypically similar individuals are more likely to mate with one another, producing offspring that are phenotypically similar to both parents. Although legal prohibitions against interracial marriage were abolished following the case of *Loving v. Virginia* in 1967, the sociocultural impact of centuries of institutionally enforced isolation has had a profound impact on biology (Brace 1995). In the 50 years since *Loving v. Virginia*, the rate of interracial marriage in the United States has increased only to 7% (US Census 2010). This relatively small proportion suggests that even in the absence of legal barriers, the longer-lasting, socio-cultural implications of these institutions have reinforced a positive assortative mating structure on human populations, particularly in the United States.

It is the combined effect of ancient differentiation and a history that has created a sociocultural positive assortative mating structure in the United States that allow forensic anthropologists to effectively align the concepts of ancestry and race. For example, in the United States individuals who can trace their origins to Africa are likely to have been identified, and self-identified, as "black" (Ousley et al. 2009, Parra et al. 1995, 1998, Sauer 1992). Although forensic anthropologists agree that human variation is vast and continuously distributed, it is not unreasonable to expect that the anthropologist can estimate a "general area of origin along that continuum" that corresponds to the features of the remains being examined (Sauer 1993:82). This is not to say, however, that the labels applied to ancestry groups in the United States are the most elegant solution to the race-ancestry disparity. Relethford (2010) effectively describes the problem by comparing the labels assigned to ancestry groups to height. Height is a continuous variable, but it is convenient when describing an individual to reduce their height to a broad category, such as “short”, “average”, or “tall” (Relethford 2009, 2010). Based on this categorization, few (if any) would assume that humanity exists at only three heights, but rather that these labels are a convenient, if imprecise, means of discussing the variation in human heights; the labels applied to ancestry groups (i.e. African, Asian, European), and their subsequent alignment with social race as reported for missing persons (e.g. white, black) accomplish the same objective (Relethford 2009, 2010).

The use of somewhat imprecise labels, and the biological realities on which they are based, should not be interpreted as a validation of the outdated typological approach to race (Hubbe and Neves 2007, Sauer 1992). Methods designed to aid in the estimation of ancestry from the skeleton are not meant to “subdivide the entire human species into discrete…categories” (Ta’ala 2015:10), but rather to estimate a broad geographical origin of that individual, and/or the most likely race label assigned to that individual during life (Sauer 1992, Ta’ala 2015). In the words of Madeline Hinkes (1993:49), “we cannot know where a deceased individual placed himself on the racial continuum, but we can determine from his remains the preponderance of the various characteristics more often associated with one race over another.” These estimates may not be correct. Individuals can identify with multiple social races, the race with which an individual identifies may not be concordant with biological ancestry (e.g. Klimentidis et al. 2009), and race labels like “Hispanic” can represent several different geographic ancestries (Bonilla et al. 2005, Cerda-Flores et al. 2002, Lisker et al. 1986, 1990, 1996, Rangel-Villalobos 2008, Rubi-Castellanos et al. 2009). However, estimates of ancestry in the forensic context are based on the most-likely area of ultimate geographic origin of an individual, and are used effectively in the identification of unknown human remains (Hinkes 1993, Sauer 1992, 1993, Ta'ala 2015).

3.2 Methods for Estimating Ancestry

The human skeleton has long been observed to be highly variable, and this variation seems to exhibit a geographic pattern (e.g. Bamshad et al. 2004, Cartmill 1998, Edwards 2003, Howells 1989, Hubbe and Neves 2007, King and Motulsky 2002, Novembre et al. 2008, Relethford 2009, Rosenberg et al. 2002) that is tied to the effects of millennia of microevolution (e.g. Angel and Kelly 1990, Beals et al. 1983, 1984, Byers 2011, Garn 1965, Gill 1998, Klepinger 2006, Konigsberg et al. 2009, Relethford 2010, Ross et al. 2004, Spradley et al. 2008, Spradley and Weisensee 2013, Washburn 1951). The geographic pattern of human variation and the effects of microevolution on the human skeleton are most often focused on the skull, and as such, the preponderance of ancestry estimation methods are derived from intensive study of the human skull (e.g. Angel 1944, 1946, Angel and Kelly 1990, Boas 1912, Brues 1990, Giles and Elliot 1962, Gill 1971, 1995, 1998, Gravlee et al. 2003, Hefner 2003, 2007, 2009, Hooton 1931a, 1946, Howells 1970, 1973, 1989, 1995, Hughes et al. 2011, Jantz and Ousley 2006, Rhine 1990, Sparks and Jantz 2002). Cranial and dental methods are the focus of this chapter and the remainder of this dissertation. However, in the forensic context, the skull is not always recovered, and a discussion of ancestry estimation methods would be remiss to not mention the many methods that have been developed for the estimation of ancestry from the post-cranial skeleton (e.g. Baker at al. 1990, Bass 2005, Christensen et al. 2014, Craig 1995, DiBennardo and Taylor 1983, Gilbert and Gill 1990, Gill 2001, Holliday and Falsetti 1999, Stewart 1962, Wescott 2005). These methods are largely based on population differences in body proportions (Holliday and Falsetti 1999) that have long been known in biology and physical anthropology (Allen 1877, Bergmann 1847). Subsequent sections of this chapter discuss ancestry estimation as it relates to four larger groups of methods: trait lists, craniometrics, statistical analysis of morphoscopic traits, and dental analysis.

3.2.1 Trait Lists

Phenotypic and skeletal differences between populations have been at the root of the race concept since its inception. However, the traits most commonly used in the forensic estimation of ancestry can largely be traced to Earnest Hooton. Hooton preferred morphoscopic traits to measurements and indices because, "1) they spring to the eye, [and are] qualitative as well as quantitative; 2) they are dependent upon form differences rather than size/proportions; [and they are] more certainly heritable" (Hooton 1931a: no page #, in Hefner 2003:15). Through years of research, he eventually compiled a list of traits that came to be known as the "Harvard List" (Brues 1990, Hefner 2007). The traits on this list reflected Hooton's desire to study non-adaptive traits, which therefore spoke to biological relationships between populations (Hooton 1926, 1946). These traits focused on four categories: bone shape, feature morphology, suture shape, and feature prominence (Brues 1990, Hefner 2007,2009), and were therefore inherently different from the nonmetric traits identified by broader physical anthropology. Nonmetric traits can be considered non-pathological skeletal variation, characterized by the absence or proliferation of skeletal material (Buikstra and Ubelaker 1994, Pink et al. 2016). The traits on which Hooton focused are traits of the skeleton that closely reflect soft-tissue differences in the living (Pink et al. 2016). Rather than focusing on discrete expressions of skeletal variation, as with nonmetric traits, the Hooton traits focus on bone shape and feature morphology (Pink et al. 2016).

One of the biggest criticisms of cranial morphoscopic traits has been lack of standardization, and the subjectivity of the method. Hooton himself was cognizant of this problem (Hooton 1946). To test the replicability of his proposed traits, Hooton enlisted several of his students, including Carleton Coon and J. Lawrence Angel, to make observations on twenty skulls (Hooton 1946). The low degree of consistency observed among his students led him to conclude that, "…the observational techniques as practiced by these four men could do with a goodly measure of standardization." (Hooton 1946: no page #). Following these results, Hooton's vision for the future of the field included the creation of scales for the standardization of observations, directions for making observations and measurements, and the application of statistics to the analysis of these traits (Hooton 1946: no page #). In spite of these early conclusions, the list of traits proposed by Hooton was passed down, largely unchanged, through Hooton's academic lineage (Brues 1990, Hefner 2003). At least two of Hooton's students, J. Lawrence Angel and Alice Brues, inherited the "Harvard List", and its influence is clear in their later work. The career of Angel ran the gamut of physical anthropological research, but his approach to ancestry and ancestry estimation was based on racial types, and reliant on 'race traits' throughout (e.g. Angel 1944, 1946; Angel and Kelly 1990). This is also true of the career of Brues, though she was considerably more focused on race in physical anthropology. A student of Hooton in the late 1930s, Brues continued to contribute to the trait list approach by adding new traits (e.g. nasal root contour) fifty years into her career (Brues 1990). Perhaps her most lasting impact on the trait list in forensic anthropology was in the role of advisor. As an advisor, she passed on the list of race traits that she had inherited from Hooton to her students, among them Stanley Rhine.

The trait list method in the 'Modern Period' (Warren et al. 2011) of forensic anthropology has become almost synonymous with the work of George Gill and Stanley Rhine, predominantly the latter. Gill (1971:152) uses many of the features typically associated with forensic ancestry estimation (e.g. orbit form, palate shape, presence of a nasal sill) in his dissertation under the title of "continuous non-metrical observations." These traits persist throughout his work on ancestry estimation in forensic anthropology (e.g. Gill 1995, 1998, Gill et al. 1988). In 1990, *Skeletal Attribution of Race* was published, and the trait list method of ancestry estimation in forensic anthropology was codified. *Skeletal Attribution of Race* was compiled by Gill and Rhine (1990) from the proceedings of the first meeting of the Mountain, Desert, and Coastal (MD&C) Forensic Anthropologists. At this meeting, a host of forensic anthropologists, including many Hootonian descendants, met to discuss the issue of "racial identification from the skeleton" (Gill 1990: xi). The resulting volume represents their collective knowledge, based on experience in identifying human skeletons in the southwest United States. Although an important volume in its own right, perhaps the most influential paper contained within it is, "Non-metric Skull Racing" by Stanley Rhine (1990).

"Non-metric Skull Racing" has, in the 25 years since its publication, become nearly the sole methodological reference for the application of cranial morphoscopic traits to forensic matters (Hefner 2007, Hefner et al. 2012, Hughes et al. 2011). The study seeks to take the trait lists of Hooton one step further, by establishing lists that are effective in estimating ancestry, as compiled by a group of experienced forensic anthropologists who explicitly defined those traits. Furthermore, Rhine (1990) provides the researcher with illustrations of the most 'typical' individual from each ancestry group. These lists and illustrations are consistently, and nearly ubiquitously, used to teach the practice of estimating ancestry from the morphoscopic traits of the skull (e.g. Burns 2007, Byers 2011, Warren et al. 2011). However, there are problems with this study and the trait list method in general. First, the sample sizes on which the trait lists presented are based are small, and representative of only of the New Mexico population at the time of publication, a fact acknowledged by Rhine himself (Hefner 2007, 2009, Rhine, 1990). Though he acknowledges that "this sample represents only a fraction of the variability to be seen worldwide" (Rhine 1990:13), the information presented is rarely treated as applicable to a limited sample (e.g. Hefner 2007, 2009). The second major problem is in the assignment of traits to particular ancestry groups. First, no acknowledgement is made of the occurrence of these traits in 'unexpected' ancestry groups. Therefore, the trait lists are treated in a somewhat exclusive fashion. Second, in several instances, traits are assigned to ancestry groups for which little or no support was observed. For example, post-bregmatic depression and molar crenulations are listed under traits typical of American Black individuals (Rhine 1990). However, a post-bregmatic depression was observed in a single individual, and molar crenulations were not even recorded for any of the three American Black individuals in this sample (Rhine 1990). Similarly, sagittal keeling is in the trait list for the Southwest Mongoloid population, but was not observed in any Hispanic or American Indian individuals in this sample (Rhine 1990). This suggests that at least some component of the trait lists presented by Rhine (1990) is based on anecdotal evidence and the experience of the attendees of the MD&C, not on direct observations. These problems do not invalidate the traits included in trait list approaches in and of themselves; however, the limitations and oversights of this approach have not fully been addressed until recently (e.g. Hefner 2007, 2009, Hughes et al. 2011).

The methodological problems with the Rhine (1990) study should not be read as a condemnation of cranial morphoscopic traits. The problems raised with

Rhine's (1990) study have been primarily methodological, decrying the typological nature in which traits are employed, and the small samples on which large generalizations are based (e.g. Hefner 2007, 2009). The traits themselves, however, have been successfully used in estimating ancestry for decades. Recent work has investigated the efficacy of 'trait list' traits in estimating ancestry. Hughes and colleagues (2011) devised a mathematical simulation to test the traits on Rhine's (1990) list, varying the list of traits used, number of traits on the list, and the point at which one classification is made over another. Through all scenarios, cranial morphoscopic traits were used to correctly estimate ancestry in over 90% of cases (Hughes et al. 2011). It is clear from these results that it is not the concept of cranial morphoscopic traits that is flawed, but rather the ways in which they have been conceptualized and analyzed. A move away from lists and toward trait frequencies has opened up new avenues of research, and allowed the forensic anthropologist to refine the analysis of cranial morphoscopic traits.

3.2.2 Craniometrics

The use of metric data in the analysis of population affinity has a long history in anthropology, especially in American anthropology. That there were metric differences in the crania of global populations was long known (e.g. Blumenbach 1776, Broca 1862, Camper 1791, Daubenton 1784, Hilaire 1795, Huxley 1863, Retzius 1843), and the perceived objectiveness of metric data is an appealing feature, especially under today's stricter standards of evidence. However, the perceived objectivity of craniometric measurements does not ensure the objectivity of the conclusions drawn from them. Both Hooton (1930) and Morton (1839) used craniometric analyses to bolster the paradigm of racial types, conclusions which have since been re-examined and largely debunked (e.g. Gould 1981, Kaplan et al. 2015). Even researchers who used craniometrics to argue against racial typology (e.g. Boas 1912) have had their work re-examined in the context of our modern understanding of human variation (e.g. Gravlee et al. 2003, Sparks and Jantz 2002).

On the one hand, researchers like Hooton were using craniometrics to quantify real, observable phenotypic differences among world populations (Hooton 1930). On the other hand, Boas's work effectively showed that craniometric data could be used to demonstrate the plasticity of the human skeleton and debunk the typological approach to race (Boas 1912). The re-examinations mentioned above (Gravlee et al. 2003, Sparks and Jantz 2002) did little to resolve this debate. Gravlee and colleagues (2003) are opposed to the use of craniometry to estimate population affinity from the human skeleton, and their reanalysis of Boas's data supported his conclusions: that cranial measurements are too sensitive to environmental conditions to be used in assessing population affinity. Sparks and Jantz (2002), however, support the use of craniometrics in estimating individual ancestry. Their reanalysis acknowledged the plasticity of the human cranium, but also found a component of variation that cannot be explained simply by environment. These two groups of researchers came to opposing conclusions using the same data, illustrating that questions asked of the data are as important, if not more so, as the data themselves (Kaplan et al. 2015).

In some ways, the conclusions of both Gravlee and colleagues (2003) and Sparks and Jantz (2002) are correct. It is at the intersection of biology and environment that the application of craniometric data to the forensic context enters the discussion. The skeleton is plastic, responding and adapting to the environment. At the same time, different environments result in different adaptations and different skeletal morphology, producing distinct geographic patterns of variation. Although not focused on crania, the classic example of this principle is the effect on body size and proportions demonstrated by Bergman's (1847) and Allen's (1877) rules. Further research has demonstrated that these effects are not limited to the post-cranial skeleton, and that a certain amount of cranial variation is correlated with climate (Beals et al. 1983, 1984, Byers 2011, Gill 1998). It is reasonable to assume that centuries of evolutionary processes affect populations in contrasting environments differently, resulting in a genetically coded signature of past environmental adaptations. That relationship is the basis of craniometric, and indeed all, ancestry estimation methods. The differences in cranial form highlighted by craniometric analysis are not the result of inherent differences between distinct racial types, nor are they random and uninformative with respect to ancestry. Craniometric differences among groups are the cumulative effect of centuries, or millennia, of microevolution acting on local populations.

While studies of population affinity on a global scale generally dwindled in physical anthropology until recently, that never really happened within the forensic context. Although researchers were recognizing flaws in ancestry estimation, their ultimate goal was the identification of unknown human remains; therefore, research focused on methods that helped achieve that goal (Relethford 2010, Smay and Armelagos 2000, Sauer 1993, Ta'ala 2015). Even to this day, craniometric methods for the estimation of ancestry remain the most popular (Dirkmaat et al. 2008, Jantz and Ousley 2006, Spradley and Weisensee 2013). Craniometric data are continuous, and are perceived as a better representation of human variability, and consequently more appropriate for the estimation of ancestry outside of a typological paradigm (Algee-Hewitt 2011, Giles and Elliot 1962, Gill et al. 1988, Howells 1970, 1973, 1989, 1995, Konigsberg et al. 2009, Krogman and İşcan 1986, Olivier 1969, Ousley et al. 2009, Ousley and Jantz 2006, 2012, Relethford 1994, 2009, Relethford and Harpending 1994, Roseman and Weaver 2004, Snow et al. 1979, Spradley et al. 2008, Wescott and Moore-Jansen 2001).

One of the earliest, statistically robust applications of craniometric data to the problem of ancestry estimation was presented by Giles and Elliot (1962). Noting that the cranium provides more information regarding ancestry than any other skeletal element, and the difficulty in estimating ancestry based on visual inspection alone, Giles and Elliot (1962) proposed the application of discriminant function analysis (DFA) to craniometric data. This new approach allowed them to correctly classify unknown individuals into broad ancestry groups in 82-88% of cases.

Eventually, due to repeated success, an interest in assessing affinity from cranial measurements was reawakened in physical anthropology.

The resurgence of using craniometric data is best represented by the career of W.W. Howells (1973, 1989, 1995). Over a decades long career, Howells collected data from an astounding 2524 individuals from 28 populations (Auerbach 2014, Howells 1973, 1989, 1995). These data were used to answer questions of human origins (Howells 1973), explore the geographic pattern of human variation (Howells 1989), and eventually demonstrate the utility of craniometric data to estimate population membership at a scale more refined than the broad populations of Giles and Elliot (Howells 1995).

Modern techniques for the estimation of ancestry from craniometric data are more sophisticated but employ the same basic concept as the early work of Giles and Elliot (1962). The most famous, and most often used, example of modern DFA in the forensic context is the software package *FORDISC* (Jantz and Ousley 2006). *FORDISC* is an expansion on the Giles and Elliot (1962) concept. This software relies on a massive database of known individuals from a variety of populations to make estimations of ancestry (Jantz and Ousley 2006). Furthermore, it utilizes a larger battery of measurements, thus capturing more of the important interpopulational variation. For the purposes of reliability and replicability, *FORDISC* has two major advantages over previous discriminant function methods. First, ancestry (and sex) estimates provided by *FORDISC* are accompanied by a posterior probability, indicative of the likelihood of group membership. Second, a typicality probability is provided. This is perhaps the more important of the two statistics. The nature of DFA is that it classifies any data it is 'given' into a pre-defined group; the typicality probability is a gauge of the validity of that classification. These features have made *FORDISC* the most widely-used software package in forensic anthropology for craniometric analysis, and with it anthropologists can achieve correct classification in 80-90% of cases (Ousley and Jantz 2012)

Recently, craniometric data are being generated and analyzed for new populations (Spradley et al. 2008), combined with other datasets, such as cranial morphoscopic traits (Hefner et al. 2014), and analyzed in new, statistically robust ways (e.g. Algee-Hewitt 2011). However, craniometric methods for ancestry estimation are prone to problems, like traditional trait-list methods. The primary problem relates not to the data themselves, but rather to DFA, the statistical approach employed most often in their analysis. The mechanics of DFA necessitate that classification accuracy decreases when applied to populations not included in the DFA reference sample (Elliot and Collard 2009, Feldesman 2002, Gotelli and Ellison 2004, Krzanowski 2000). This has been a problem noted since Giles and Elliot's (1962) discriminant functions. Tests of those functions found that individuals from populations other than those on which the functions are based were incorrectly classified in 35-60% of cases (Birkby 1966). The problem of correct reference populations has been repeatedly raised as a concern with craniometric ancestry analysis (Elliot and Collard 2009, Feldesman 2002, İşcan and Steyn 1999, Ramsthaler et al. 2007, Snow et al. 1979, Williams et al. 2005), to the point that some have argued that it should not be used (e.g. Kosiba 2000, Williams et al. 2005). A secondary problem with craniometric methods, and one frequently cited by proponents of morphoscopic traits, is that craniometric data are difficult to generate from fragmented remains (Gill 1998, Hefner 2007, Hefner et al. 2012, Rhine 1990). In the case of fragmented human remains, as are often encountered by the forensic anthropologist, a craniometric analysis may not be possible.

Ancestry estimation methods employing craniometric data are effective and extremely popular, especially under current strict standards of evidence. Many have advocated against their use on the grounds of inaccuracy (e.g. Kosiba 2000, Williams et al. 2005), claiming that programs such as *FORDISC* will provide an ancestry estimate for a soccer ball if you let it (Freid et al. 2005, Ousley et al. 2007, 2009). These claims represent a lack of understanding of the statistics underlying craniometric ancestry estimation, more than a failure of the methods. The more pressing problem with craniometric methods is their inability to be applied to fragmented or distorted remains. Skeletons in less than pristine condition are frequently encountered by the forensic anthropologist; a circumstance that necessitates that non-metric, morphoscopic methods for estimating ancestry be available to supplement the repertoire of available techniques.

3.2.3 Morphoscopic Trait Frequencies

The problems associated with craniometric methods of ancestry estimation and the traditional trait list method, have left a hole in the methodology of ancestry estimation that needs to be filled. Hughes and colleagues (2011) demonstrated the utility of traditional morphoscopic traits in the estimation of ancestry. Therefore, it is not the traits themselves that are problematic, but rather the ways in which they have been applied. The 're-introduction' of cranial morphoscopic traits under the standards of evidence outlined by *Daubert* necessitated that their use be moved away from typology and total reliance on researcher experience. This shift required three major changes to trait-list methods. First, cranial morphoscopic traits should be reported in terms of population frequencies. It is not useful for an accurate understanding of human variation to report traits as indicative of a particular population, when their actual distributions reflect more variation. Second, trait lists are full of descriptions of traits and character states that are not conducive to consistent recordation. To bring the analysis of cranial morphoscopic traits in line with legal standards, a standard system based on anatomy and degrees of expression needed to be implemented. Finally, the analysis of cranial morphoscopic traits needed an infusion of statistical power. Traditional methods were based, at best, on a percent match to typical trait-list traits, and at worst on the experience and opinion of the observer (Hefner 2003, 2007, Hughes et al. 2011).

In response to these problems, Hefner (2003, 2009) presents definitions of cranial morphoscopic traits. The term "macromorphoscopic" was coined to distinguish the skeletal features used by forensic anthropologists from the nonmetric traits employed in physical anthropology more broadly (Hefner and Ousley 2006). This term was eventually shortened to "morphoscopic" (Hefner

2009), which will be used in this research. Definitions of the cranial morphoscopic traits and their character states are rooted in anatomy, and avoid the confusing descriptive labels that had traditionally been applied to them. One of the best examples of this change can be seen in his categorization of nasal bone contour. Where Brues (1990) describes the states of this trait as "Quonset hut", "tented", and "steepled", the scale proposed by Hefner (2009) consists of numeric stages defined by the angularity of the nasal bones in the defined region. This scale codifies essentially the same variation as that proposed by Brues (1990), without the descriptive labels that are somewhat arbitrary. Furthermore, standard recordation practices allow rates of error for these methods to be known, a shortcoming of the descriptive trait-list method. Finally, this work moves the analysis of cranial morphoscopic traits away from the 'trait-list' method. Scores for morphoscopic traits are not presented as characteristic of a particular ancestry group, but rather the frequency of all expressions of a trait across populations are presented (Hefner 2009). This allows future researchers to examine in which populations a given character state is most common, without ignoring the interpopulational variability of that trait.

Recent work has focused on the statistical analysis of cranial morphoscopic traits, and bringing ancestry estimation from cranial morphoscopic traits into compliance with the standards defined by *Daubert* (e.g. Hefner 2007, Hefner and Ousley 2014, Hefner et al. 2012, 2014, 2015, 2016). Primarily this work has focused on applying non-parametric statistics, not typically used in physical anthropology, to the analysis of cranial morphoscopic traits. There has been considerable success in analyzing morphoscopic traits using k-nearest neighbor modeling, random forest modeling (RFM), support vector machines (SVM), artificial neural networks (aNN), canonical analysis of principal coordinates (CAP), and a new method called the optimized summed scores attributes (OSSA) (Hefner 2007, Hefner 2015, Hefner and Ousley 2014, Hefner et al. 2014). In each of these cases, ancestry estimates have achieved 78-88% accuracy, and demonstrate the future utility of cranial morphoscopic traits in the estimations of ancestry. The use of cranial morphoscopic traits is on the rise having been tested for reliability and replicability (e.g. Klales and Kenyhercz 2015), proven to produce ancestry estimates comparable to those derived from metric methods (Hefner et al. 2014), and expanded to include trait list traits that have yet to be standardized (Maier et al. 2015, Maier 2016).

In spite of the recent interest in cranial morphoscopic methods, they are not without problems. One of the biggest problems is the inherently subjective nature of morphoscopic trait analysis (Hefner 2007). The utility of these traits lies in the recognition of various character states, necessitating a judgment call by the observer. For that reason, morphoscopic traits will never be completely objective, but the subjectivity can be mitigated through improved standards, improved methods of recordation, and more rigorous analysis (Hefner 2007). These problems are being actively addressed (e.g. Hefner and Ousley 2014, Hefner et al. 2012, 2014, 2015, Maier et al. 2015, Maier 2016), but the problem of subjectivity in morphoscopic trait analysis continues to be the most prevalent argument against their use (e.g. Algee-Hewitt 2011, Hefner 2007, 2009, Hughes et al. 2011, Konigsberg et al. 2009). One of the other limitations of ancestry estimation from cranial morphoscopic traits is in their analysis. An appropriate statistical framework exists, and there are many applicable methods (Hefner 2007, Hefner and Ousley 2014, Hefner et al. 2012, 2014, 2015, 2016); however, most of those methods are not widely employed in physical anthropology and are somewhat inaccessible to those less familiar with statistical software packages like R (R Core Team 2016). Furthermore, the comparative data needed to draw robust conclusions from the results of these analyses are not yet publically available. Both of the problems are being addressed and are soon to be resolved. On a grant from the National Institute of Justice (Award No. 2015-DN-BX-K012) the Macromorphoscopic Databank is being developed by Dr. Hefner (Hefner 2017, Pink et al. 2016). Upon its completion in 2017/2018, the databank will provide researchers with a large comparative sample, and a user-friendly interface, similar to FORDISC (Jantz and Ousley 2006), with which to analyze cranial morphoscopic trait data. These advances, including the research contained in this dissertation, represent a significant contribution to the pool of legally defensible methods from which the forensic anthropologist may draw when making an estimate of ancestry.

3.2.4 Dentition

Finally, this discussion turns to a grossly under-used dataset: the dentition. Traits of the dentition are highly heritable (Scott and Turner II 1997) and teeth preserve more readily than bone (Edgar 2005). Of particular interest to the forensic anthropologist, teeth are often protected from trauma inside the oral cavity and are therefore more likely to be recovered intact than cranial indicators of ancestry (Kimminau 1993). Furthermore, dental traits have already been standardized and tested for reliable recordation, helping to reduce the error inherent in ancestry estimation (Edgar 2015). Despite acknowledgement of the forensic utility of the dentition more than two decades ago, relatively little has been done to incorporate teeth into the fold of forensic techniques (e.g. Edgar 2005, 2013, 2015, Irish 2015, Pilloud et al. 2014). The inclusion of dental morphology in forensic anthropology has largely been hindered by two factors. First, forensic anthropologists are concerned with individual identification, while dental anthropologists have traditionally focused on population-level variation. This difference in scope may lead to a certain degree of data incompatibility, which will be touched on at the end of this chapter. Second, on more than one occasion preeminent dental anthropologists have voiced the opinion that dental morphology may be of little use in a forensic setting (Lasker and Lee 1957: 418, Turner II personal communication in Irish 2015). These two factors have been difficult obstacles to surmount. However, nearly a century's worth of research has led to three inter-related conclusions. First, global variation in dental morphology exists. Second, that variation corresponds closely to broad, geographically-defined populations. Finally, understanding these broad populations can allow forensic anthropologist to accurately estimate ancestry based on the dentition.

Population differences in the dentition have nearly as long a history in physical anthropology as craniometric and cranial morphoscopic traits (e.g. Dahlberg 1945, 1951, Hellman 1928; Hrdlička 1920, 1921, Lasker 1950, Lasker and Lee 1957). At around the same time that Boas (1912) was using craniometric data to explore questions of race in American immigrants, and Hooton (e.g. 1918, 1931, 1946) was exploring the efficacy of cranial morphoscopic traits in assessing population affinity, questions of origins, population movement, and race were being answered with the dentition. Although these early works were encumbered by the typological views of race that persisted at the time, they explored variation in the human dentition and how that variation coincided with group membership. Arguably the best known of these early attempts is "Shovel-Shaped Teeth" (Hrdlička 1920). In this article, Hrdlička explores in depth the trait that we now recognize as incisor shoveling, and codifies a means for recording it. This paper made two major contributions to the study of dental morphology and its global distribution. First, this paper represents one of the first attempts to standardize the recordation of dental morphology (Hrdlička 1920). Second, it established shoveled incisors as being a distinct form that is common in Native American and Asian populations and uncommon in African and European populations (Hrdlička 1920). Based on this trait, 'racial' differences in the dentition of man began to be documented. The following year, Hrdlička (1921) observed several more incisor traits, documenting their degrees of expression, and where appropriate, noting if the trait had any 'racial' significance (e.g. congenital absence of teeth is more common in Europeans).

Other authors soon took up the mantle of describing ‘racial’ variation in the dentition. Hellman (1928) focused on the Y-5 molar groove pattern in apes and hominins, and its changing frequencies in modern populations. Hellman defined degrees of deviation from the ancestral Y pattern, and discussed the populations in which various forms of the trait occur (Hellman 1928). For example, he notes that European populations have the highest frequencies of non-Y pattern molars and that Native American populations can be difficult to classify because they frequently have cusp 6 (Hellman 1928). Hellman (1928) describes the Y-patterned molar as the ancestral condition and deviations away from that pattern as derived. Using this as a basis, he proposed a hierarchy of world populations in terms of evolutionary status, which is unsurprising for the time. Although the evolutionary hierarchy was an unfortunate step, Hellman’s (1928) work added significantly to our knowledge of global dental variation and our understanding of ancestral and derived dental traits.

More major work regarding the dentition as a tool for assessing ‘racial’ affinity comes from the latter half of the 20th century (e.g. Dahlberg 1945, 1951, Lasker 1950, Lasker and Lee 1957). Until this time, researchers had focused primarily on the global distribution of single (or small sets of) traits in narrowly defined samples of world populations (e.g. Hellman 1928, Hrdlička 1920). As these data began accumulate, the door was opened for more comprehensive studies of global dental variation. Dahlberg (1945) began to synthesize years of dental data to gain an appreciation for the global pattern. His approach was unique in that he felt it was important to understand the mechanisms that were driving differentiation of dental morphology in various populations. Though not explored in detail, he argues for the significance of isolation and genetic drift in shaping the dental complex of a given population, concluding that the trend in the human dentition is toward reduction and simplification of size, form, and number of teeth (Dahlberg 1945). This change occurs at different rates in different populations based upon the specific conditions experienced by that population; therefore, geographically distant populations will exhibit different suites of traits based upon their unique microevolutionary histories (Dahlberg 1945). This conclusion is at the foundation of the search for geographically based dental complexes that would continue his work in last decades of the 20th century and beyond (e.g. Irish 1993, 1997, Scott and Turner II 1997, Scott et al. 2013, Turner II 1983, 1985, 1987, 1990). Other focuses of the time were the heritability of dental morphological features (e.g. Lasker 1950) and the documentation of dental variation in more narrowly defined populations (e.g. Dahlberg 1951). Taken together, the work in dental anthropology in the middle of the 20th century would soon be synthesized into a comprehensive description of the global variation in the dentition of man.

Perhaps the seminal work of the time regarding 'racial' affinities and the dentition was Lasker and Lee's (1957) "Racial Traits of the Human Teeth." They argue that the hopes for the use of dental features to assess race "have not been fulfilled" (Lasker and Lee 1957: 401). The gross differences in teeth (i.e. presence, absence, impaction) are globally uniform, and it requires closer inspection of the cusps, grooves, and other features of the teeth to elucidate their potential utility

(Lasker and Lee 1957). This work addresses two short-comings of previous attempts to understand 'racial' variation in teeth. First, Lasker and Lee (1957) suggest that a large study comparing many groups from many different populations is needed to truly appreciate significant population differences. Second, they undertake a tooth-class by tooth-class description of the major morphological differences among world populations, lamenting the lack of standardization in trait description and recordation (Lasker and Lee 1957). They argue (correctly) that this lack of standardization makes data from different researchers difficult, if not impossible, to compare (Lasker and Lee 1957). Finally, echoing the sentiments of Dahlberg (1945), they conclude that there are major differences in morphology between European, African, and Asian populations but in light of intragroup differences and "the lack of distinct racial boundaries" the use of the dentition in forensic applications may be limited (Lasker and Lee 1957:418). These early studies anticipated the eventual codification of global patterns of dental morphological variation. Expansion of our understanding of these patterns, and refinement of the traits that define them would be the call of future dental anthropologists. The question of forensic utility, however, would be essentially dropped for several decades; its eventual re-emergence is discussed later in this chapter.

The work of early researchers, despite its faults, laid the foundations for American dental anthropology, especially as concerned questions of population variation. The subsequent decades saw two major advances in the field of dental anthropology: 1) the development of standardized systems of data recordation, and 2) the collection of prodigious amounts of data from populations the world over. Standardization in dental anthropology is often discussed in conjunction with the Arizona State University Dental Anthropology System (ASUDAS; Turner II et al. 1991); however, it would be remiss to jump into this discussion at that point without at least reviewing its forbears.

Early on, researchers recognized that dental morphology was expressed along a gradient, and that simply recording a trait as present or absent would not encompass the full range of its variation. In his 1920 work on shovel-shaped incisors, Hrdlička proposed a scale for the recordation of shoveling. Recognizing various degrees of shoveling, he proposed a four-point scale that encompassed no shoveling, trace shoveling, semi-shoveled, and shoveled (Hrdlička 1920). However, shoveling is but one of many traits that can be recorded. In 1956, Dahlberg proposed a new, more comprehensive, system for standardizing the recordation of dental morphological traits (Dahlberg 1956). This system comprised definitions and plaster plaques of trait expressions on several traits including shoveling, double shoveling, the hypocone, Carabelli's trait, and the protostylid (Dahlberg 1956). However, as Dahlberg (1956) himself suggested, the system left room for expansion and improvement. Finally, the most comprehensive system, and today's standard, the Arizona State University Dental Anthropology System (ASUDAS) was proposed by Christy Turner II and colleagues (1991). The ASUDAS comprises standard definitions for 38 dental crown and root traits (as well as several associated traits of the oral cavity), descriptions of the trait and its various character states, and accompanying plaques depicting degrees of expression for most traits (Turner II et al. 1991). The standardization allowed by the ASUDAS definitions, descriptions, and plaques has allowed for dental morphology to be one the most consistently recorded datasets available to physical anthropologists (≈85% concordance; Nichol and Turner II 1986, Scott 1973, Scott and Turner II 1997, Sofaer et al. 1972b). A new guidebook is devoted exclusively to detailed descriptions and illustrations of the Arizona State University Dental Anthropology System (Scott and Irish 2017).

As mentioned above, the second major change to this field was in the sheer volume of data that was being collected specifically for dental anthropological work. Like standardization, the majority of this work was done by Christy Turner II and his students, but the patterns that emerged and subsequent interpretations of those patterns are the product of the efforts of many researchers. The combination of a standard method to record dental morphology and the knowledge that population differences in dental morphology exist led to the codification of geographically based dental complexes. Dental morphological differences between spatially and temporally disparate groups, are presumed to reflect different microevolutionary histories (Scott and Turner II 1988); therefore, if broad geographically defined groups exist, then members of those groups would be more dentally similar to one another than to members of a different dental complex. The task of many dental anthropologists was to define those complexes, and test their validity through the examination of many populations.

With respect to global dental complexes, the two most widely known are Sinodonty and Sundadonty, as defined by Christy Turner II (1983, 1985, 1987, 1990). These complexes are derived from the proposed 'Mongoloid Dental Complex' discussed by Hanihara (1967). This complex was based on observations of the deciduous dentition, and found five characteristics to be common in Asian populations: shoveling, deflecting wrinkle, protostylid, metaconule, and cusp 7 (Hanihara 1967). Upon further examination, Turner II (1983, 1985, 1987, 1990) discovered that while the 'Mongoloid Dental Complex' sufficiently described dental variation in Northeast Asian populations and Native Americans, the dentition of Southeast Asian populations was markedly different. Therefore, two complexes (Sinodont and Sundadont) were proposed (Turner II 1983, 1985, 1987, 1990). These complexes are geographically separate and can be characterized by differing degrees of dental complexity. The Sinodont dentition is characterized as mass additive, while the Sundadont complex is more generalized (Turner II 1983, 1985, 1987 1990, Scott and Turner II 1997).

The Sinodont dental complex that characterizes Northeast Asian and Native American populations is the most morphologically complex in the world (Scott and Turner II 1997: 236). These populations exhibit high frequencies of shoveling, double shoveling, winging, interruption grooves, odontomes, enamel extensions, cusp six, and the deflecting wrinkle (Scott and Turner II 1997, Turner II 1990). Traits found in markedly low frequencies in Sinodont populations are relatively few, but include the mesial canine ridge, metaconule, cusp 7, four-cusped lower molars, and Y-shaped groove pattern on the lower molars (Scott and Turner II 1997). Several of these features are in disagreement with the 'Mongoloid Dental Complex' proposed by Hanihara (1967). This is likely a reflection of sample size, and not of error on the part of the researchers. Sinodont populations have low frequencies of traits like the metaconule and cusp 7 on a global scale (Scott and Turner II 1997), but when compared specifically to European populations, as in Hanihara (1967), the frequencies of these traits may appear elevated. Finally, although Sinodont crown traits are characterized by added complexity, root traits in Sinodont populations tend to be simple, showing higher frequencies of undivided roots (Scott and Turner II 1997). One notable exception to this generalization is the relatively high frequency of three-rooted lower first molars. This root feature has been effectively used in conjunction with genetic and linguistic data to support a three-wave model of the peopling of the New World (Turner II 1971, 1985).

The Sundadont dental complex is very different. Characteristic of Southeast Asian, Polynesian, and Micronesian populations (Scott and Turner II 1997), the Sundadont dentition is more generalized and is characterized by retention of traits rather than by intensification (Turner II 1983, 1987, 1990). On a global scale, Sundadont dental morphology falls in the middle range of variation. That is, Sundadont populations exhibit no crown or root traits that occur significantly more or less in Sundadonts than in other world populations (Scott and Turner II 1997). The high frequency traits (e.g. Carabelli's trait and cusp 6) and low frequency traits (e.g. mesial canine ridge, cusp 7, and four-cusped lower molars) that characterize the Sundadont dentition, are shared by other world complexes (Scott and Turner II 1997). As such there is no diagnostically Sundadont dental trait, and this complex is best understood as the general impression given by all recordable dental traits.

A little understood, and potentially related, dental complex that characterizes populations of Australia, New Guinea, and Melanesia is the 'Australian Dental Complex' (Townsend et al. 1990). This complex was described by Townsend and colleagues in 1990. They found that the frequencies of cusp 6 in aboriginal Australian populations were among the highest in the world, a result which corroborates the work of previous researchers (Smith et al. 1981, Townsend et al. 1990). The resulting 'Australian Dental Complex' includes high frequencies of cusp 6, metaconule, and Carabelli's trait, and low frequencies for cusp 7 and the protostylid (Townsend et al. 1990). In spite of these apparent commonalities, an 'Australian Dental Complex' complex should be defined with caution (Scott and Turner II 1997). There are many traits that group populations of the Sahul-Pacific region together, but there are also traits that separate them; in fact, only the frequency of the metaconule separates populations included in the 'Australian Dental Complex' from other populations in a statistically significant way (Scott and Turner II 1997). This complex is presented here not as an argument for or against its validity, but as a step toward understanding global dental variation, and the work that has been done to record it.

Behind Sinodonty and Sundadonty, the next most recognized world dental complex is the sub-Saharan African dentition, or the Afridont dental complex (Irish

1993, 1997, 2013). Like the Asian complexes, work with the Afridont dental complex has been done almost exclusively by a single author. Joel Irish recorded data from 32 populations across Africa and subjected them to a biodistance analysis (Irish 1993, 1997). He found a morphological divide between North African populations and those south of the Sahara Desert, leading him to define the sub-Saharan African dental complex, also called Afridonty (Irish 1997, 2013). There are 11 traits that characterize sub-Saharan African populations. These populations have globally the lowest or second lowest frequencies of double shoveling and enamel extensions and the highest or second highest frequencies of presence of the third molar, two-rooted lower second molars, three-rooted upper second molars, Y-shaped groove pattern on the lower molars, Carabelli's trait, two-rooted upper premolars, cusp 7, Tome's root, and the mesial canine ridge (Irish 1997). Furthermore, Scott and Turner II (1997) have noted a general lack of root fusion among Afridont populations, and that cusp 7, the mesial canine ridge, and a Y-shaped groove pattern on the lower second molar are considered "distinctive indicators" of the Afridont dental complex (Scott and Turner II 1997: 236).

The last major geographically-based dental complex is the European dentition. There have been several attempts to define a European dental complex. In the same article in which the 'Mongoloid Dental Complex" is defined, Hanihara (1967) also proposes a 'Caucasoid Dental Complex', consisting minimally of high frequencies of Carabelli's trait. Mayhall and colleagues (1982) made another attempt to define a European dental complex based on the analysis of American

White dental casts. This definition of the complex included a single high-frequency trait (Carabelli's trait) and six traits that were absent or found in very low frequencies (shoveling, winging, odontomes, protostylid, cusp 6, and cusp 7) (Mayhall et al. 1982: 256). Since then, the definition of this complex has been expanded, and the complex itself has been relabeled the 'Eurodont' dental complex (Scott et al. 2013). There are six traits found in high frequencies in the Eurodont dentition: Carabelli's trait, multiple lingual cusps on lower premolars, three-cusped upper second molars, four-cusped lower first molars, four-cusped lower second molars, and two-rooted lower canines (Scott et al. 2013). However, the hallmark of the Eurodont dental pattern is the high number of traits that are absent or found in low frequencies. These include: winging, shoveling, double shoveling, the mesial canine ridge, enamel extensions, Y-shaped groove pattern of the lower second molar, cusp 6, cusp 7, protostylid, deflecting wrinkle, and the three-rooted lower first molar (Scott et al. 2013). As can be seen from these lists of traits, the Eurodont dentition is more frequently characterized by the absence of traits (Scott and Turner II 1997).

Having established that dental morphology varies significantly across the globe, and that variation in morphology has a geographic pattern. Scott and Turner II (1997) set out to compare dental morphology on a global scale. In a massive undertaking involving data from 175 populations, Scott and Turner II (1997) discerned 12 traits that are useful in discriminating between broad geographic populations on the basis of dental morphology. These traits include shoveling, winging, Carabelli's trait, the metaconule, enamel extensions, two-rooted upper first (third) premolars, odontomes, two-rooted lower canine, four-cusped lower molars, cusp 6, cusp 7, and three-rooted lower first molars. It should be noted that all of the traits listed above are used in the definition of previously discussed global dental complexes. Therefore, it should not be surprising that based on these traits, Scott and Turner II (1997) were able to accurately distinguish among 5 broad geographic groups that roughly correspond to those complexes: Sino-America, Sahul-Pacific, Sunda-Pacific, Western Eurasia, and sub-Saharan Africa. This well-patterned distribution of dental morphological variants is the basis for the use of the dentition in the forensic estimation of ancestry discussed below.

The dentition has not been completely ignored in the forensic arena. Most forensic anthropologists are aware of the correlation between marked shoveling of the incisors and Asian ancestry or of the perceived correlation between Carabelli's trait and European ancestry. These traits frequently appear on non-metric trait lists for the estimation of ancestry (e.g. Burns 2007, Byers 2011, Kimminau 1993, Rhine 1990, Warren et al. 2011). In fact, Rhine (1990) includes at least seven dental traits (shovel-shaped incisors, winging, enamel extensions, buccal pits, Carabelli's trait, canine fossa, and molar crenulations) in the landmark "Non-metric Skull Racing", and those traits represent the extent of most forensic anthropologists use of the dentition, despite years of research demonstrating its utility. The problem with this approach, is that it too employs the trait-list method, characterizing certain traits as indicative of specific ancestry groups with little consideration of interpopulation frequencies, and no application of statistics. As an example, Birkby and colleagues (2008) provide a list of characteristics, both biological and cultural, that have aided in the identification of migrants in Pima County, Arizona. Included on this list are incisor shoveling, and enamel extensions, two of the traits proposed by Rhine (1990) as indicative of "Southwestern Mongoloid" ancestry. These traits are included based on the experience of the authors, but neither the frequency with which these traits are encountered, nor the means by which this association was confirmed are discussed. The use of the dentition, especially dental morphology, in the forensic context was not attempted in earnest until recently.

Dental morphology has received the majority of the attention with respect to ancestry estimation. The standard traits of the ASUDAS are easy to observe and record, once a researcher has been trained to recognize these features, and do not require any specialized equipment, making them ideal for a forensic setting (Edgar 2015). Although they are the more popular choice for ancestry estimation, relatively few authors have presented the analysis of dental morphology in a forensically applicable way (e.g. Edgar 2005, 2009, 2013, 2015, Irish 2015, Scott et al. 2016). However, dental morphology has been used to demonstrate dental morphological differences between individuals of European and African ancestry, and that those differences allow the anthropologist to accurately estimate ancestry in 90% of cases (Edgar 2005). The few dental traits included in Rhine's (1990) publication have largely been demonstrated to be ineffectual in the estimation of ancestry, especially in isolation from other features (Edgar 2009). Finally, dental morphology has been used to effectively estimate ancestry through the application of logistic discrimination equations (Edgar 2013, 2015). Although these efforts have varied in demonstrating the utility of the dentition in estimating ancestry (46-89% accurate depending on group comparison), they have provided a foundation on which future analyses of the dentition can build. However, these methods share a few drawbacks. Most notably, the logistic discriminant equations (Edgar 2013) require that specific teeth be present for the equations presented to be applicable; namely an unworn upper first molar to distinguish between Hispanic and non-Hispanic populations, and an unworn lower canine and first (third) premolar to distinguish between African and European populations (Edgar 2013). Furthermore, the earlier research (Edgar 2005) while useful, can be cumbersome to apply and, like OSSA, is limited only to estimating African or European ancestry.

More recently, the utility of the dentition in the forensic context as demonstrated by Edgar (2005, 2009, 2013) has pushed other dental anthropologists to apply their knowledge and vast datasets to forensic questions. Irish (2015) presents a method for estimating ancestry from 10 ASUDAS traits considered to be diagnostic of major world populations. In this method, traits are recorded relative to established breakpoints, effectively dichotomizing the ordinal degrees of expression inherent in the ASUDAS to simple presence/absence data (Irish 2015). These frequencies are then compared to frequencies in five world groups: East Asian, American Indian, White, Polynesian, and Black. An unknown individual is given a summed score based on presence or absence of the defined traits and the population that corresponds to the highest summed scored is given as the ancestry estimate (Irish 2015). Though this method seems logical, it is not yet appropriate for the forensic sphere. The use of a large reference sample, certainly makes the data on which this method is based robust; however, this method contains no known rates of error making it in admissible under strict post-*Daubert* (Daubert v. Merrell Dow Pharmaceuticals Inc. 1993) standards of evidence. A second method, in development Richard Scott and colleagues (2016), applies more appropriate statistical methods to the estimation of ancestry from dental morphology. Currently called *rASUDAS,* this method combines the standardization of the ASUDAS, the computing power of the program R (RCoreTeam 2016), and the classificatory power of Bayesian statistics to arrive at a relatively robust estimate of ancestry for an unknown individual. As this method is still under development, there are considerations that have yet to be addressed. For example, for forensic applications, *rASUDAS* should be based, on individual rather than population data to better account for the range of human variation and to avoid becoming typological. Work continues to be done in this area, and dental anthropology, specifically dental morphology, is beginning to make a foray into forensic anthropology.

Dental morphology is not the only useful type of data derived from the dentition when estimating ancestry. Recent work has explored odontometric ancestry estimation. Like craniometric data, odontometrics provide some advantages over dental morphology; namely, the data are continuous and quantitative making them subject to more rigorous statistical analyses, and thus more applicable to the forensic context. Furthermore, dental metrics have the additional advantage of being applicable to worn teeth, where dental morphology may not be. Although researchers (Harris and Foster 2015, Pilloud et al. 2014) have demonstrated success in estimating ancestry from dental measurements, this technique has not been widely adopted. This is due in part to the specialized equipment needed to take dental measurements and the specialized training needed to take them correctly.

Dental metrics have also been employed with great success (over 90% accuracy) in the estimation of ancestry from the juvenile dentition (Lease and Sciuli 2005). The importance of this study is two-fold. First, methods for ancestry estimation in juveniles are very few; therefore, an accurate method from the dentition expands the repertoire of the forensic anthropologist. Second, and particularly germane to this research, this study represents the power of combining data from different datasets to arrive at an ancestry estimate. Though metrics are used, the Lease and Sciuli (2005) study combined odontometric and morphological data to arrive at this accurate method.

The above described methods, especially those pertaining to dental morphology, are not without problems and require refinement and validation. Many of the data generated from an analysis of the dentition are non-continuous, and the computing power and statistical approaches needed to properly analyze these data have not been available until recently. Although methods, such as discriminant function analysis, have been applied with success, their results may not hold up under the strict rigors of the medico-legal system, as dental morphological data violate many inherent assumptions of these methods. Furthermore, the addition of dental data to forensic anthropology is somewhat hindered by data incompatibility. Where forensic anthropology is concerned with the individual, dental anthropology has traditionally been concerned with populations. As a result, traditional dental morphological data (e.g. Scott and Turner II 1997) has been dichotomized and then coalesced into population frequencies, effectively erasing the signal of individual variation important to the process of identification, and making the application of existing dental data to forensic questions difficult.

Despite these difficulties, the estimation of ancestry from the dentition is a promising avenue of future research, though still in its nascent stages. Forensic anthropologists and dental anthropologists alike are beginning to recognize the potential in drawing on the knowledge of these two fields, and research in this area has boomed in the last five years (e.g. Adams 2015, George 2015, Irish 2015, Pilloud et al 2014, Scott et al. 2016, Willermet et al. 2016). Research involving the dentition has expanded to include odontometric estimations of ancestry and population history (Adams 2015, George 2015, Pilloud et al 2014), user friendly applications for the analysis of dental morphology (Irish 2015, Scott et al 2016), and the creation and application of statistical methods more appropriate for the analysis of dental data (Willermet et al. 2016). The application of dental anthropology to the forensic context represents one of many future directions of expansion, especially in ancestry estimation techniques.

3.3 Concluding Remarks

Given the promise of dental morphology in ancestry estimation, and the established utility of cranial traits toward these same goals, it stands to reason that the combination of these two datasets may yield more accurate estimates of ancestry than are possible from either data type independently. The teeth and the skull develop differently, and are differently affected by the environment. Therefore, an ancestry estimation method that combines the information available from both sets of data may provide a clearer picture of individual ancestry. The environment of forensic anthropology is primed for methods that synthesize data from different sources to improve the methods available for individual identification, especially as the statistical computing power available to the individual researcher increases. The research outlined in the remainder of this dissertation describes one such approach to the combination of cranial and dental data in pursuit of estimating ancestry from the human skeleton.

Chapter 4:
Materials and Methods

4.1 Materials

To explore the utility of combining cranial and dental traits in the forensic estimation of ancestry, data were generated for a sample of 693 individuals of European, African, Asian/Native American, and Hispanic ancestry (Table 4.1). A maximum of 79 traits (12 cranial, and 67 dental) was collected for each individual, and subjected to statistical analyses appropriate to the categorical and ordinal data generated by morphological trait analysis. Previous research has indicated that both cranial morphoscopic and dental morphological traits show negligible sexual dimorphism because the traits analyses are dependent on shape, not overall size (Hefner 2003, Scott and Turner II 1997); therefore, sexes were pooled for these analyses.

Table 4.1: Sample composition for this research.

Collection	**European**			**African**			**Asian/Native American**			**Hispanic**			**Total**
	M	**F**	**U**	**M**	**F**	**U**	**M**	**F**	**U**	**M**	**F**	**U**	
PCOME	2	0	0	0	0	0	2	1	0	76	4	1	*86*
LSU	11	6	0	5	7	0	0	0	0	0	0	0	*29*
Texas State	56	38	0	3	2	0	0	0	0	5	1	0	*105*
Operation ID	0	0	0	0	0	0	2	0	0	31	27	2	*62*
Terry	12	7	0	54	42	0	5	0	0	0	0	0	*120*
Hamann-Todd	46	8	0	64	12	0	0	0	0	0	0	0	*130*
UNM	96	52	0	4	2	0	0	0	0	3	4	0	*161*
Total	***223***	***111***	***0***	***130***	***65***	***0***	***9***	***1***	***0***	***115***	***36***	***3***	***693***

4.1.1 Ancestry Groups

A brief note regarding the use of the terms "European", "African", and "Asian/Native American" in this research. These terms, as applied in this dissertation, are meant to denote ultimate, not proximate, geographic ancestry. Furthermore, the individuals housed in the sampled collections are catalogued, in most cases, by social race (e.g. white, black, etc.). Based on the concordance of ancestry and social race in the United States cited in previous chapters (e.g. Sauer 1992), ancestry assignments were made based on the reported race of the individual. Finally, the ancestry label "Hispanic" is used in this research for ease of recognition. Within the forensic community there is a debate regarding the appropriate nomenclature for the group of individuals to whom the label "Hispanic" has traditionally been applied. The composition of these populations varies by region within the United States, with Southwestern populations having European and genetic contributions, and Southeastern populations exhibiting both of these contributions in addition to an African component (Bertoni et al. 2003). Alternative labels, such as "Admixed Native American" have been proposed (Anderson personal communication 2015) but "Hispanic" is used here to avoid confusion. Finally, as addressed in Chapter 1, these groups represent the most relevant categories for forensic anthropology in the United States, and are not meant to be universally applicable (Bass 2005, Burns 2007, Byers 2011, Kennedy 1995, Klepinger 2006, Sauer 1992, 1993, Spradley et al. 2008, Spradley and Weisensee 2013).

4.1.2 Sampled Skeletal Collections

For a method to be forensically relevant, the populations on which it is based must reflect the composition of the extant population as accurately as possible. Several craniometric analyses have demonstrated the effects of secular change on the human skull (e.g. Jantz and Jantz 2000, Wescott and Jantz 2005). Although these changes affect the size and shape of the cranial vault more than the facial skeleton (Jantz and Jantz 2000), the individuals used in this research were limited to recent individuals where possible; to ensure applicability to contemporary populations, no archaeological specimens were used. Most individuals studied here were born in the 20th century; however, in the interest of a robust sample size and more balanced demographics, individuals from two older anatomical collections (i.e. Terry and Hamann-Todd) were included. Descriptions of each collection and their respective demographic profiles are below.

4.1.2.1 Robert J Terry Anatomical Skeletal Collection

In the early 1900s, Robert Terry collected skeletons from cadavers used by the medical school at Washington University in St. Louis. These skeletons largely represent individuals whose bodies were left unclaimed after death and became property of the state. As such, the earliest component of the Terry Collection represents individuals of low socio-economic status. However, in the mid-twentieth century, legislation was passed that required written permission from next of kin for a body to be used for scientific study. As a result, the later component of the Terry collection represents individuals of middle to upper-middle class economic status. Presently, the Terry Collection represents 1728 individuals of known sex, age, and "ethnic origin" (Hunt 2016). These individuals were born between 1822 and 1943, and range in age from 16 to 102 years, with most over 45 years of age

4.1.2.2 Hamann-Todd Human Osteological Collection

This collection contains the skeletons of over 3000 individuals of known sex, age, and 'race' collected by Drs. Carl Hamann and T. Wingate Todd between 1912 and 1938. As with the Terry Collection, most individuals represented by the Hamann-Todd Collection were unclaimed and became property of the state, and subsequently were meant for use as anatomy teaching tools. Similar to the Terry Collection, the status of these individuals as unclaimed, likely means that they were of low socioeconomic status. The adults range in age from 18-75 years of age, meaning that all individuals in this collection were likely born after 1837 but before 1921.

4.1.2.3 FACES Lab Donated Skeletal Collection

The Forensic Anthropology and Computer Enhancement Services (FACES) Lab is at the Louisiana State University in Baton Rouge, Louisiana. Legislation passed in 2006 made the FACES Lab the central hub for missing and unidentified persons throughout the state of Louisiana (Louisiana House Bill 1140 Act 227).

During its existence, in addition to the unidentified whom they curate by law, the lab has amassed a small collection of donated skeletal material, the earliest of which was donated in 1982. This collection comprises 31 individuals of African and European ancestry. These individuals are predominantly forensic cases that were ultimately donated by the families for research and teaching purposes.

4.1.2.4 Texas State University Donated Skeletal Collection

This collection is composed of skeletons that were donated to Texas State University for scientific study. As such, individuals in the collection have known sex, age, and ancestry, as well as a variety of other information including occupation and health status. At the time of data collection, Texas State University had received 306 donations, only a fraction of which had been completely processed and were available for study. These individuals are aged between 18 and 102 years. As a forensically driven, donation-based collection, individuals housed in the collection have died since the collection's establishment in 2008, and are born exclusively in the 20th or 21st centuries.

4.1.2.5 Maxwell Museum Documented Skeletal Collection

The Maxwell Museum Documented Skeletal Collection was established in 1984, and includes over 278 individuals of known sex, age, and ancestry. The collection is built through donation from the individual, the deceased's family, or in the case of unclaimed remains, donation from the Office of the Medical Investigator.

Given its date of establishment (1984), all individuals in the Maxwell Museum Collection died within the last 33 years, making it an appropriately modern sample for forensic work.

4.1.2.6 Pima County Office of the Medical Examiner

The Pima County Office of the Medical Examiner (PCOME), located in Tucson, Arizona, works with a high volume of unidentified human remains. Since 2001, the PCOME has dealt with over 2100 cases of migrant deaths. Part of the mission of the PCOME is to identify these individuals and return them to their countries of origin. A byproduct of that process is that the PCOME houses those individuals who have yet to be positively identified. Using a biocultural profile crafted by the anthropologists at the PCOME (Birkby et al. 2008), the migrants at the PCOME have largely been given an ancestry estimate of "Admixed Native American," which is used in this research to signify Hispanic ancestry. It is important to note that these are not known individuals, and in many cases no ancestry estimate could be given; therefore, an additional seven individuals for whom there is no ancestry information were observed but will not be included in these analyses.

4.1.2.7 Operation Identification

Recent changes to the United States policy on border control have funneled migrants into increasingly inhospitable environments; one of those routes is through the Sonoran Desert. This has resulted in high numbers of migrant deaths in the Tucson area. However, Texas has also seen a rise in migrant deaths, and in 2012 the number of migrant deaths in Texas surpassed that of Arizona (USBP 2012). To identify deceased migrants, many unknown burials were exhumed and brought to the Forensic Anthropology Center at Texas State (FACTS) for storage and processing. Presently, Texas State University curates over 100 individuals as part of Operation Identification. As with the individuals at the PCOME, ancestry is not known for many of these individuals; 23 individuals, for whom no ancestry estimate was available, were not included in the following analyses.

4.2 Methods

4.2.1 Cranial Morphoscopic Traits

Cranial morphoscopic traits form the basis of the trait list method described in the previous chapter. The use of these traits in the past has involved descriptions of bone shape, suture configuration, trait degree of expression, or trait presence/absence (Hefner 2009). Further research regarding these traits has standardized their application, thus mitigating some of the subjectivity in their use (e.g. Hefner 2007, 2009). Eleven of the twelve cranial morphoscopic traits used in this research have previously been standardized and have demonstrated utility in the estimation of ancestry from the human skeleton (Hefner 2009). These traits include: the anterior nasal spine, inferior nasal aperture, interorbital breadth, malar tubercle, nasal aperture width, nasal bone contour, nasal overgrowth, postbregmatic depression, supranasal suture, transverse palatine suture, and zygomaticomaxillary suture shape (Table 4.2). For a full description of these traits and their degrees of expression see Hefner 2009. These traits have been used effectively in the estimation of ancestry (e.g. Hefner 2007, Hefner and Ousley 2014, Hefner et al. 2014, 2015), and have demonstrated low inter- and intraobserver errors (Hefner 2009).

4.2.1.1 Palate Shape

Though these 11 traits have been successfully employed, the trait lists contain traits that have yet to be standardized, but may be useful in the forensic context. Among these traits is palate shape. The utility of palate shape as an indicator of ancestry can be traced to the Harvard List developed by Earnest Hooton (Brues 1990), a connection which has been explored in depth elsewhere (e.g. Hefner

Table 4.2: List of cranial morphoscopic traits observed for this research; scoring system after Hefner 2009 unless otherwise noted.

Cranial Morphoscopic Traits	
Trait	**References**
Anterior Nasal Spine (ANS)	
Inferior Nasal Aperture (INA)	
Interorbital Breadth (IOB)	
Malar Tubercle (MT)	
Nasal Aperture Width (NAW)	
Nasal Bone Contour (NBC)	
Nasal Overgrowth (NO)	
Post-bregmatic Depression (PBD)	
Supranasal Suture (SNS)	
Transverse Palatine Suture (TPS)	
Zygomaticomaxillary Suture Shape (ZSS)	
Palate Shape (PS)	*Maier 2016*

2003, 2007). Even given this seemingly deep history within physical anthropology, there is little agreement on defining palate shape, or on a uniform means of describing its variation. The original classification proposed by Hooton was that "pinched" palates were characteristic of European individuals, "narrow" palates were indicative of African populations, and Native American groups exhibited "wide" palates (Hooton 1931b). Other classifications have been developed for use in forensic anthropology. Krogman (1939) proposed that European populations were characterized by "narrow" palates, African populations by "wide" palates, and Asian or Native American palates as some "intermediate" form between the two extremes. Finally, the classification presented by Rhine (1990) and most commonly employed today, is that of Gill (1971, 1986, 1995) and Olivier (1969). Gill (1971, 1986, 1995) describes European and East Asian palates as "triangular" or "parabolic," African palates as "hyperbolic," and Native American palates as "elliptical" or sometimes "parabolic." Aside from sometimes contradicting previous classification schemes, these descriptive terms are highly subjective, and with respect to geometry, are inaccurate in their characterization of palate shape.

Previous research demonstrated the utility of palate shape as an indicator of ancestry; however, the methods employed in that research necessitate the use of a digitizer and computer, which may not be available in all forensic settings (Maier et al. 2015). The current research proposes an ordinal scale, akin to those presented in Hefner (2009), to standardize the recordation of palate shape, and eliminate the need for subjective and inaccurate descriptive terms.

Palate shape is here defined as the overall appearance of the maxillary alveolus and dentition. Shape, assessed on a five-point scale following Maier (2016), is depicted in Figure 4.1:

- **Grade 1**: the overall arcade has a rounded shape, and the posterior teeth, especially third molars, move toward the midline of the mouth. This curve can be considered ovular in shape, and has traditionally been described as "elliptical".
- **Grade 2**: is an intermediate shape between Grades 1 and 3. The palate has less angled sides than in Grade 3, but is overall less ovular and rounded than in Grade 1.
- **Grade 3**: the palate exhibits a highly curved, almost pointed, anterior segment. The sides of the arcade angle away from the midline as the arcade is followed posteriorly. This curve is triangular in shape, and most closely aligns with the classic "parabolic" shape.
- **Grade 4**: is another intermediate state. Palates in Grade 4 are intermediate between Grades 3 and 5. The anterior portion of the curve is somewhat flattened as compared to Grade 3, but not to the degree seen in Grade 5. Additionally, the angle of the sides is less pronounced than in Grade 3.
- **Grade 5**: the anterior segment of the palate shape curve is flattened, and the sides are roughly parallel to one another as the curve is

followed toward the back of the mouth. This curve could be described as rectangular, and is most similar to the “hyperbolic” shape.

This scale allows researchers to record a wider array of extant variation than the existing three-shape system. Furthermore, assessment of a character state is based on several factors that contribute to the overall appearance of the palate. That these definitions are rooted in the morphology of specific areas of the palate, not on an inaccurate description of overall palate shape, helps to make the recordation of palate shape more objective, and certainly more standard.

4.2.2 Dental Traits

The dentition has long been used to explore human variation on the population level, with recent applications to the forensic setting (e.g. Edgar 2009, 2013, Irish 2015, Scott and Turner II 1997, Scott et al. 2016). In this study, 27 traits were recorded for appropriate teeth on both the right and left sides. Of those traits, 24 are derived from the ASUDAS (Turner II et al. 1991). These traits include (Table 4.3): winging, shoveling, double shoveling, interruption grooves, tuberculum dentale, canine mesial ridge, distal accessory ridge, premolar accessory cusps, hypocone, Carabelli’s trait, parastyle, paramolar tubercle, enamel extensions, premolar lingual cusp number, odontomes, groove pattern, molar cusp number, deflecting wrinkle, anterior fovea, distal trigonid crest, protostylid, cusp 5, cusp 6, and cusp 7. Definitions of traits and depictions of character states can be found in

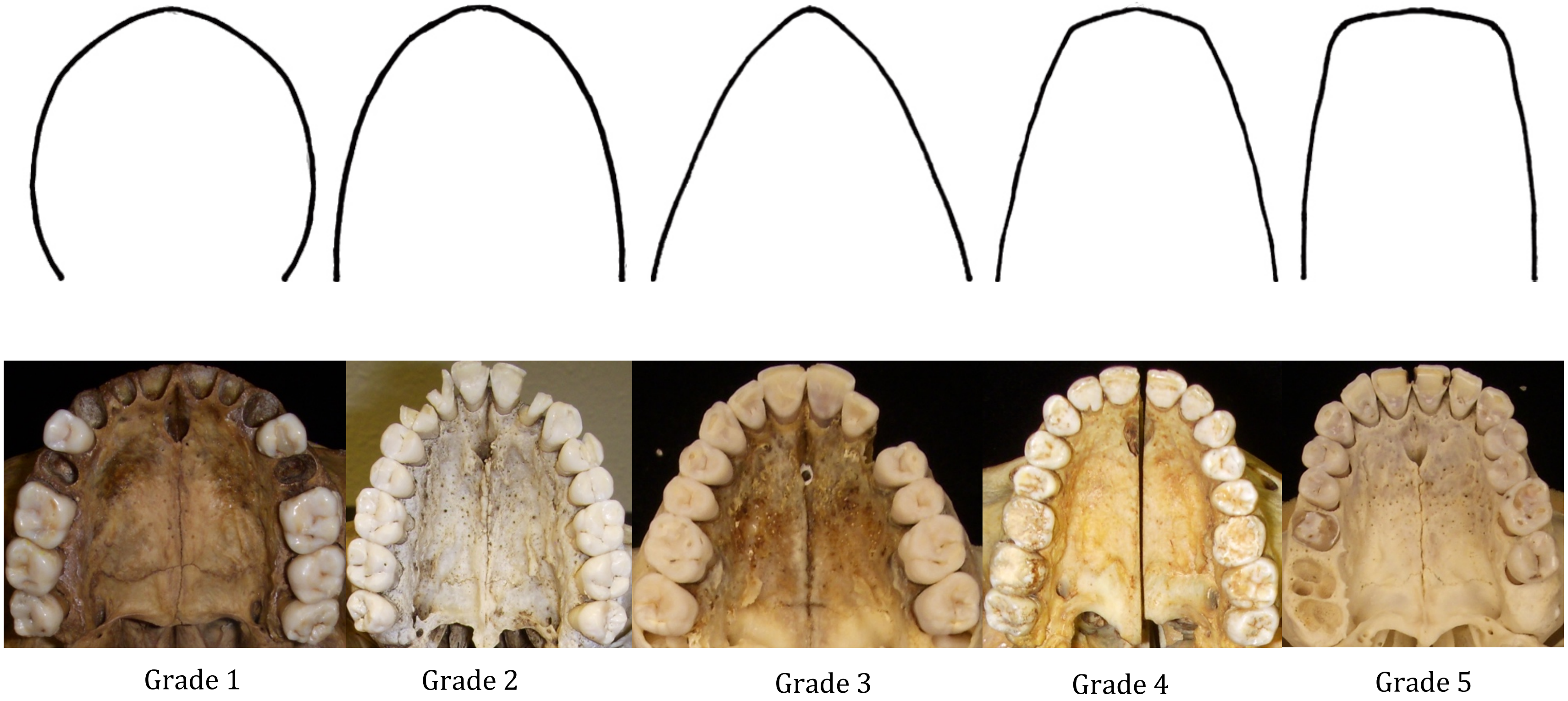


Figure 4.1: Recordation scale for palate shape.

the original publication (Turner II et al. 1991). As many dental morphological traits can be affected by occlusal wear, care was taken not to score overly worn teeth.

Table 4.3: List of dental morphological traits observed for this research. Trait definitions after Turner II et al. 1991 unless otherwise noted.

Dental Morphological Traits		
Trait	**Teeth**	**References**
Winging	Upper Central Incisors (WING)	
Shoveling	Upper Central Incisors (UI1_SHOV)	
	Upper Lateral Incisors (UI2_SHOV)	
	Lower Central Incisors (LI1.SHOV)	
	Lower Lateral Incisors (LI2.SHOV)	
Double Shoveling	Upper Central Incisors (UI1_DSHOV)	
	Upper Lateral Incisors (UI2_DSHOV)	
Interruption Groove	Upper Central Incisors (UI1_IG)	
	Upper Lateral Incisors (UI2_IG)	
Tuberculum dentale	Upper Central Incisors (UI1_TD)	
	Upper Lateral Incisors (UI2_TD)	
	Upper Canines (UC_TD)	
Canine Mesial Ridge (Bushman Canine)	Upper Canines (BUSH_C)	
Distal Accessory Ridge	Upper Canines (UC.DAR)	
	Lower Canines (LC.DAR)	
Premolar Accessory Cusps	Upper First (Third) Premolar (UP1.ACUSP)	
	Upper Second (Fourth) Premolar (UP2.ACUSP)	
Hypocone	Upper First Molar (UM1.HYPO)	

	Upper Second Molar (UM2.HYPO)	
	Upper Third Molar (UM3.HYPO)	
Carabelli's Trait	Upper First Molar (UM1.CARA)	
Parastyle	Upper First Molar (UM1.PARA)	
Paramolar Tubercle	Upper Second Molar (UM2.PARA)	
Enamel Extensions	Upper First Molar (UM1.ENEXT)	
	Upper Second Molar (UM2.ENEXT)	
	Upper Third Molar (UM#.ENEXT)	
	Lower First Molar (LM1.ENEXT)	
	Lower Second Molar (LM2.ENEXT)	
	Lower Third Molar (LM3.ENEXT)	
Premolar Lingual Cusp Number	Lower First (Third) Premolar (LP1.LCUSP)	
	Lower Second (Fourth) Premolar (LP2.LCUSP)	
Odontomes	Upper First (Third) Premolar (UP1.ODONT)	
	Upper Second (Fourth) Premolar (UP2.ODONT)	
	Lower First (Third) Premolar (LP1.ODONT)	
	Lower Second (Fourth) Premolar (LP2.ODONT)	
Groove Pattern	Lower First Molar (LM1.GPATT)	
	Lower Second Molar (LM2.GPATT)	
	Lower Third Molar (LM3.GPATT)	
Molar Cusp Number	Lower First Molar (LM1.CUSPNO)	

	Lower Second Molar (LM2.CUSPNO)	
	Lower Third Molar (LM3.CUSPNO)	
Deflecting Wrinkle	Lower First Molar (LM1.DWRIN)	
Anterior Fovea	Lower First Molar (LM1.ANTFOV)	
Distal Trigonid Crest	Lower First Molar (LM1.DTC)	
	Lower Second Molar (LM2.DTC)	
	Lower Third Molar (LM3.DTC)	
Protostylid	Lower First Molar (LM1.PROTO)	
	Lower Second Molar (LM2. PROTO)	
	Lower Third Molar (LM3.PROTO)	
Cusp 5 (Hypoconulid)	Lower First Molar (LM1.CUSP5)	
	Lower Second Molar (LM2.CUSP5)	
	Lower Third Molar (LM3.CUSP5)	
Cusp 6 (Entoconulid)	Lower First Molar (LM1.CUSP6)	
	Lower Second Molar (LM2.CUSP6)	
	Lower Third Molar (LM3.CUSP6)	
Cusp 7 (Metaconulid)	Lower First Molar (LM1.CUSP7)	
	Lower Second Molar (LM2.CUSP7)	
	Lower Third Molar (LM3.CUSP7)	
Molar Crenulations	Upper First Molar (UM1.CREN)	*Pilloud et al. 2017a*
	Upper Second Molar (UM2.CREN)	

	Upper Third Molar (UM3.CREN)	
	Lower First Molar (LM1.CREN)	
	Lower Second Molar (LM2.CREN)	
	Lower Third Molar (LM3.CREN)	
Diastema	Scored by Diastema Location (DIASTEMA)	*Pilloud personal communication 2015*
Dental Crowding	Scored for Maxilla (CROWDING_MAX) and Mandible (CROWDING_MAND)	*Pilloud personal communication 2015*

Several standard ASUDAS traits were not observed. In modern skeletal samples, teeth are often fixed in their sockets, and removal is discouraged as it may damage the surrounding alveolar bone. Therefore, root traits (e.g. two-rooted canine, 3-rooted lower first molar, Tome's root) were not recorded. Second, the ASUDAS trait "Pegged-Reduced-Missing M3" was not recorded because it is impossible to distinguish congenital absence from surgical removal in modern skeletal collections without taking dental radiographs of each individual. Finally, several other traits (e.g. labial convexity, distosagittal ridge, torsomolar angle) were not recorded because they do not demonstrate sufficient inter-population variation (Scott personal communication 2015). The remaining three traits, molar crenulations, diastema, and dental crowding, are recent developments on which relatively little data have been collected (Pilloud et al. 2017a, Pilloud personal communication). They are examined to aid in the further analysis and application of these traits to the forensic context.

4.2.2.1 Molar Crenulations

Molar crenulations are one of the few dental traits presented by Rhine (1990) as indicative of ancestry. However, there is no standard definition of molar crenulations, nor a standardized means of recording their presence available to forensic anthropologists. The presence of molar crenulations is widely acknowledged in the non-human primate and hominin evolution literature (e.g. Berger et al. 2015, Carbonell 1965, Dumont 1999, El-Najjar 1978, Halie-Selassie 2001, Leakey 1965, Swindler 1988, Vogel et al. 2008, von Koenigswald 1982, Weidenreich 1937, Williams 1985), but relatively little attention has been given to their presence in the human dentition. Lasker and Lee (1957: 410) make a passing mention of this trait, referring to it as “molar wrinkling,” and describe it as characteristic of “Mongoloid molars,” a conclusion reached by other researchers as well (e.g. Jacob 1967, Pedersen 1949). Others found this trait was more common in African populations (e.g. El-Najjar and McWilliams 1978, Herrick and Walsh-Haney 2010, Kiernberger 1955, Rhine 1990, 1993, Truesdell 2005).

The pattern here called ‘crenulations’ has long been recognized as potentially indicative of population affinity. However, little has been done with respect to the standardization and evaluation of the variation in this trait. There are potentially several reasons for this. First, objectively classifying the degree of crenulation is difficult (Grine 1981). Second, a minimal amount of wear obscures the expression of the trait (Grine 1981, Pilloud et al. 2017a). Finally, there is little agreement on how the trait should be defined. Previous studies (e.g. Truesdell 2005) have attempted to offer objective means of recording crenulations, but have done relatively little to clarify its definition.

Here, crenulations are defined as "curved fissures and ridges that surround the primary ridges (i.e. cristids) of each main molar cusp" following Pilloud and colleagues (2017a:164). Additionally, following that presentation, the scoring system advocated will be used in this research. Molar crenulations are recorded on a 3-point scale (Figure 4.2) as follows:

- **Grade 0**: No crenulations.
- **Grade 1**: Molar crenulations are shallow and do not involve all cusps.
- **Grade 2**: Molar crenulations are deep and involve all major cusps of the tooth.

4.2.2.2 Diastema

Anatomically speaking, a diastema is any gap between the teeth greater than 0.5mm (Keene 1963). Irish (1997) identified the midline diastema, a gap between the central incisors, as characteristic of the sub-Saharan African Dental Complex, but as this is not a standardly recorded dental trait, there are few comparative data. Others (Lavelle 1970, Pilloud personal communication2015) have suggested that the presence of a diastema to either side of the canine (i.e. between canines and lateral incisors, or canines and premolars), although common overall, may be more common in African populations. However, the work to date has had a population focus, and the utility of the diastema as an estimator of ancestry in forensic

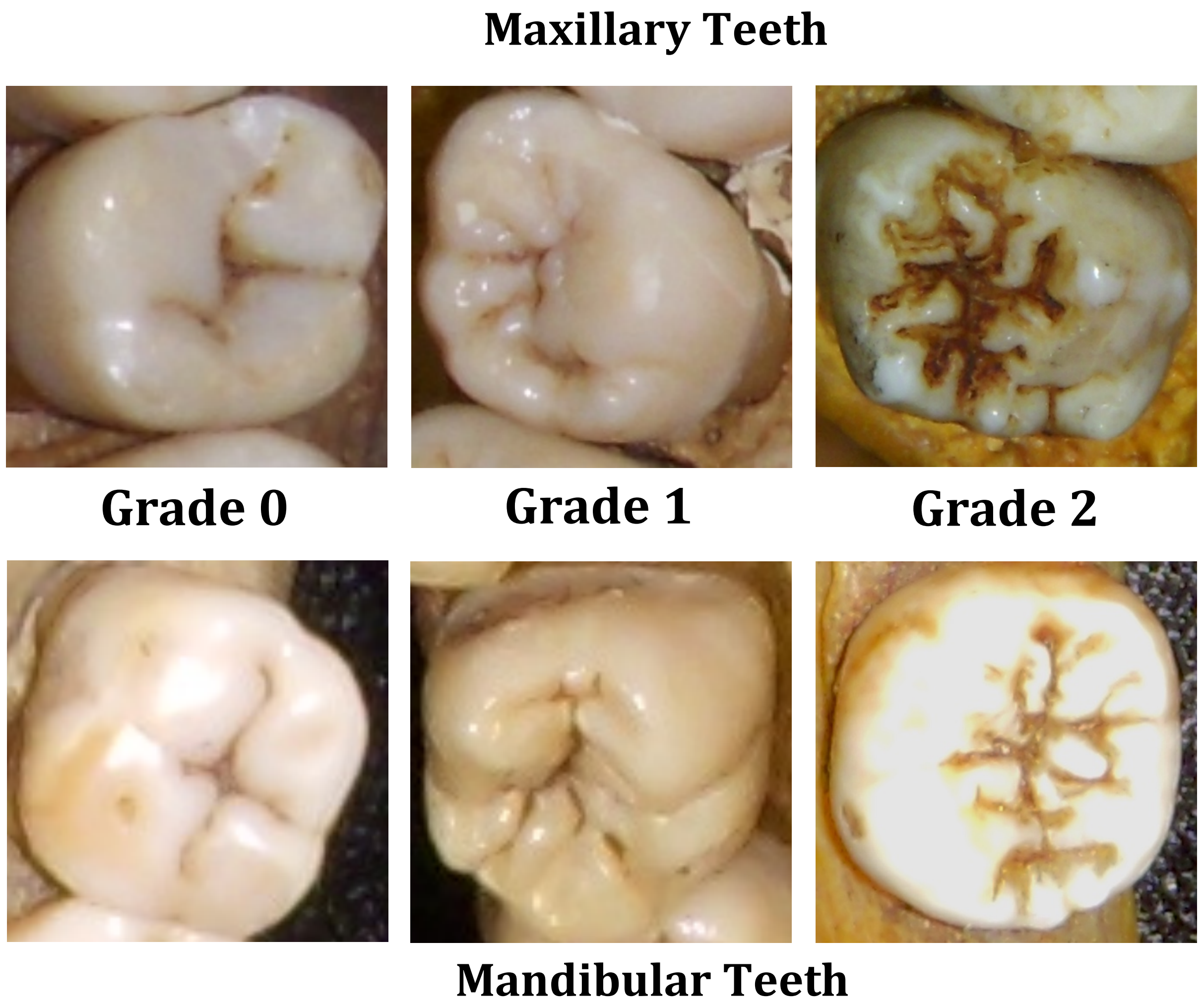


Figure 4.2: Recordation scale for molar crenulations; maxillary (top) and mandibular (bottom) teeth are presented.

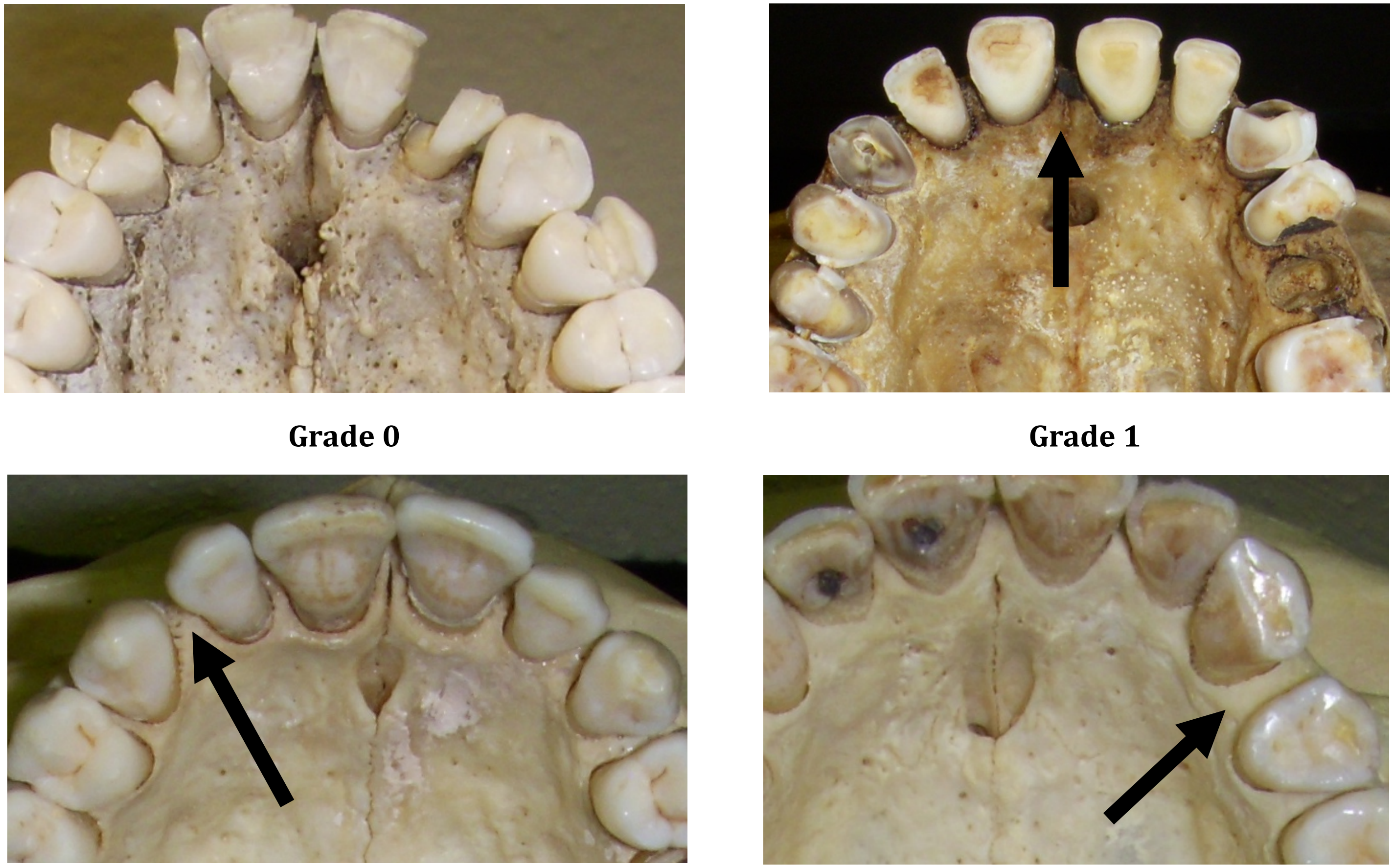
anthropology has not been explored.

Diastemata in this sample were recorded following a scale proposed by Pilloud (personal communication 2015) (Figure 4.3):

- **Grade 0**: No diastema
- **Grade 1**: Any gap between two central incisors
- **Grade 2**: Any gap between the canines and adjacent teeth

4.2.2.3 Dental Crowding

The inclusion of dental crowding for ancestry estimation is based on the disparity between tooth size and jaw size in some populations (Lavelle and Foster 1969, Moorees and Reed 1954), particularly Asian and Native American populations. Studies of tooth size have demonstrated that Asian and Native American populations fall in the intermediate range on the world scale of tooth size, while African populations tend to have large teeth, and European populations have small teeth (e.g. Pilloud et al. 2014, Schmidt 2008). However, studies of craniofacial size show a different pattern, with Asian populations exhibiting among the smallest jaws (Hanihara 1993, 1996, Howells 1989, Nicholson and Harvati 2006). Furthermore, studies of secular change in the face and jaws (e.g. Heim 2013, Jonke et al. 2007, Lavelle 1973, Smith et al. 1986, Truesdell 2005) indicate the degree of reduction in jaw size associated with modern diets differs among populations. Taken together, this evidence suggests that the incidence of dental crowding may be useful in the forensic estimation of ancestry. Although dental crowding differs from



Grade 0

Grade 1

Grade 2

Grade 2

Figure 4.3: Recordation scale for Diastemata. Grade 0-no diastema (top left); Grade 1-midline diastema (top right); Grade 2 – diastema to either side of the canine (bottom).

the other features in this study in that it has a more pronounced environmental component (i.e. it may be tied more strongly to socioeconomic status, and can be erased through the intervention of orthodontia), more recent studies (e.g. Normando et al. 2013) have confirmed a strong genetic component to dental crowding; therefore, it is included to explore what, if any, utility it may have in the forensic estimation of ancestry.

Dental crowding was recorded by quadrant of the mouth following the definition outlined by Pilloud (personal communication 2015) (Figure 4.4):

- **Grade 0**: No crowding
- **Grade 1**: Crowding present

4.2.3 Data Collection and Processing

Data were recorded using a custom-made UserForm Macro in *Microsoft Excel 2011* (Figure 4.5). This macro allows all data input to occur on a single screen, while the researcher selects a predetermined trait value from a drop down menu. The macro then uses researcher input to populate a single row of a spreadsheet. This method helps minimize user error while recording data, and produces a spreadsheet that is readily analyzed in statistical packages such as R.

Dental morphological data were recorded for both the right and left sides of the arcade. For analysis, the individual count method (Scott 1977, Scott 1980, Turner II 1967, Turner II and Scott 1977) was employed to reduce data redundancy. The underlying premise of the individual count method is that dental traits are

Grade 0-No Crowding

Grade 1-Crowding Present

Figure 4.4: Recordation scale proposed for dental crowding.

threshold traits (Falconer 1965, 1967, Scott and Turner II 1997), therefore the highest degree of expression for that trait in an individual is the most reflective of the individual's genotype (Scott 1980). In some cases, a dental trait's character states do not have an inherent hierarchical order (e.g. groove pattern). In those instances, in the case of asymmetry, the value for the left side was used in analysis. If the left tooth was absent, the antimere was used.

4.2.4 Statistical Analyses

4.2.4.1 Frequency Distributions

Scores for each of the traits were compiled into frequency distribution tables. These tables are a representation of the frequencies of each trait score in a given population. Presenting data in this format helps prevent typological associations of a given trait with a particular ancestry group, and more accurately presents the range of variation for a given trait. Furthermore, if a trait did not show any variation, it was removed from further analyses

4.2.4.2 Intraobserver Error

To assess intraobserver error approximately 10% of the individuals scored at each collection were scored a second time (n=73). The second observation was conducted at each collection 4 days after the initial observation, to maximize the period between observations. Intraobserver error was calculated using the

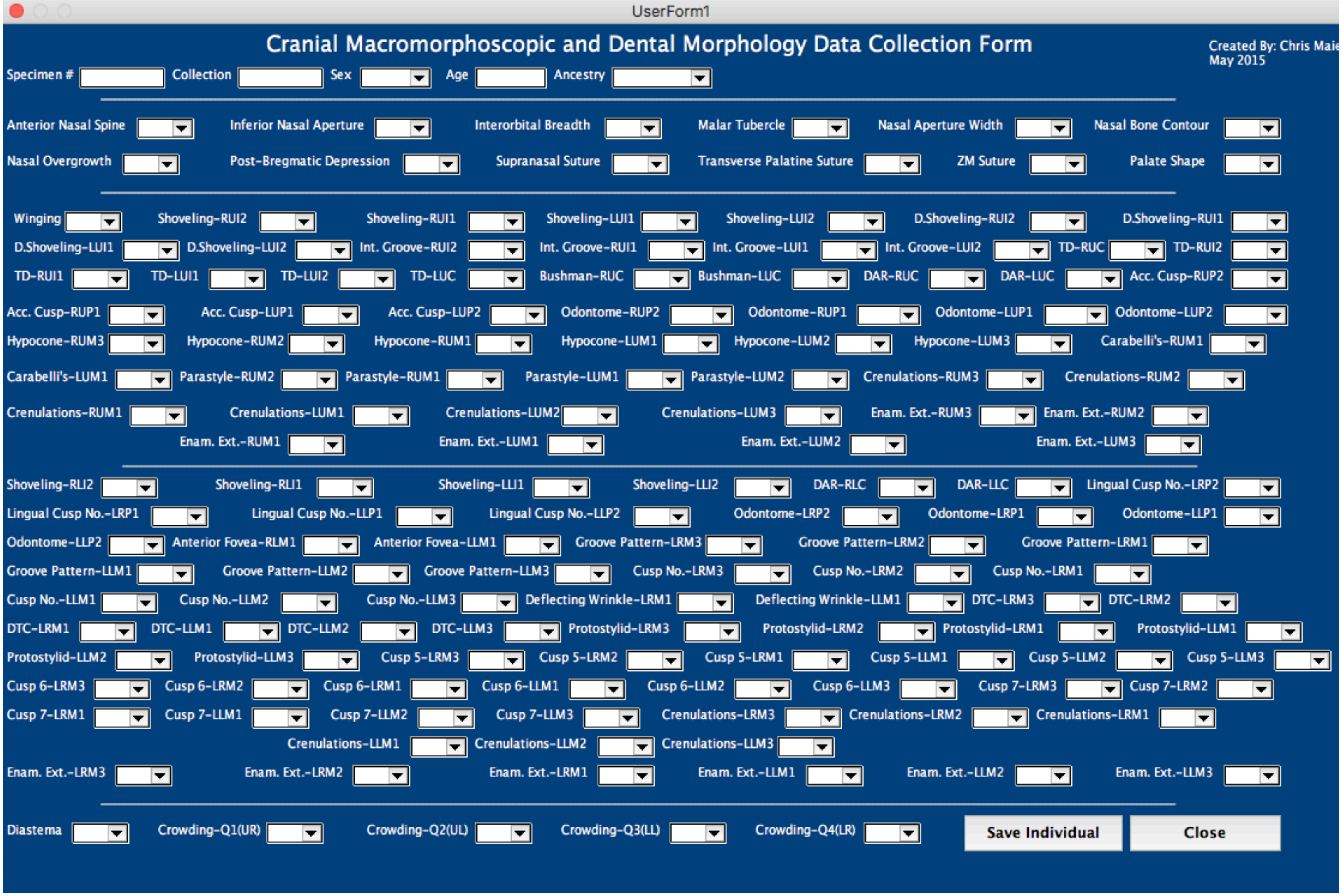


Figure 4.5: UserForm created for the collection of cranial morphoscopic and dental morphological data

weighted Cohen's Kappa statistic (Cohen 1968). The kappa statistic is a statistically robust method for determining the degree of agreement between two observers, or the same observer on different occasions. This statistic is more robust than percent agreement because it incorporates the possibility that observers agree by chance alone (Cohen 1960). The unmodified statistic can be overly conservative, in that it 'judges' any disagreement with equal 'wrongness'. However, a transformation of this statistic, weighted Cohen's Kappa is more appropriate for ordinal data, such as those generated in this research (Cohen 1968). The weighted statistic recognizes degrees of agreement, and weights them differently in the ultimate calculation of the level of agreement. For example, given a trait observed on a five-point scale, the weighted statistic recognizes that scores of 4 and 5 agree more than scores of 1 and 5. In the case of binary variables, the weighted statistic is equal to the unmodified statistic; therefore, equal-weighted Kappa will be used in the R package *irr* (Gamer et al. 2012) to generate rates of intraobserver error for all traits in this research.

4.2.4.3 Chi-Square Test

Data were examined for frequency differences among ancestry groups using the chi-square test in R (RCoreTeam 2016). The chi-square test can be used to assess differences between samples when the data are non-continuous and non-normal, making its use appropriate for the categorical and ordinal data generated in this research (Drennan 2009, Freund et al. 2010, Welkowitz et al. 1976, Krzanowski 2000).

As the chi-square test was used repeatedly to compare frequencies for many traits, the Bonferroni correction was used to adjust the p-values. As the number of statistical tests increases, in this case chi-square tests on 79 variables, the likelihood of a type I error, or incorrectly assessing a comparison as significantly different, also increases (Freund et al. 2010). Therefore, for a given number (m) of hypotheses tested and the desired level of significance (α), each individual hypothesis is required to meet a level of significance equal to α/m. For example, if individual chi-squares were calculated for all 79 recorded variables in this research, then the adjusted alpha for significance of each test would be α=0.00063291, to mitigate to possibility of a significant p-value due to chance.

The results of the chi-square tests are used as a univariate means to remove uninformative variables from further analysis. If a given trait does not show significant differences among ancestry groups, it is unlikely to be informative in the classification of unknown individuals.

4.2.4.4 Correlation

Every individual was scored for 79 variables. Some traits were observed on the same tooth (e.g. cusps 5, 6 and 7 are recorded on each of the mandibular molars), or the same trait was observed on the same class of tooth (e.g. shoveling was recorded on the central and lateral incisors). Previous researchers have demonstrated strong correlations between dental morphological traits, at least within a tooth class, (e.g. Kirveskari and Alvesalo 1982, Mizoguchi 1985, Scott and

Turner II 1997, Sofaer et al. 1972a), and among cranial morphoscopic traits (Hefner 2009). However, no measures of correlation between cranial and dental traits have yet been reported, and it is reasonable to assume these traits may be correlated with one another. Including redundant data in the classification model, even when the model can handle correlation, only serves to overinflate the accuracy of the model. To prevent that, polychoric correlations were calculated between all pairs of variables using the *hetcor()* function in the package *polycor* (Fox 2016). A polychoric correlation is appropriate for these data because it is designed to be used with ordinal data. The polychoric correlation assumes a latent continuous distribution underlying an ordinal recordation scale, and uses a polychoric series to account for that discrepancy by estimating the Pearson's correlation coefficient between two variables as if they were continuous (Pearson 1900, Ritchie-Scott 1918).

For these data, pairs of variables with correlations over an absolute value of 0.75 are examined for possible removal from model-building. The value of 0.75 represents an arbitrary lower boundary of correlation values considered to indicate a strong relationship between variables (e.g. Drennan 2009, Evans 1996, Hinkle et al. 2003, Shortell 2001, Wellkowitz et al. 1976). In each pair of strongly correlated variables, one variable will be removed to maximally reduce strong correlations in the dataset, but retain as many variables as possible for the generation of classificatory models.

4.2.5 Classificatory Models

The primary statistics used, random forest modeling and naïve Bayes classification, both require a training and a test set of data. A random 70% of the individuals (n=485) in the total sample are used as the training data, while the remaining 30% (n=208) are used to test the model. For the sake of reproducible results, a random seed is set in R before subsets are generated, and before any modelling takes place. Setting the random seed in R ensures that while individuals are chosen from the larger sample at random, every time the analysis is run, the same individuals are placed in the training and test samples. This step allows for reproducible point estimates of accuracy for both types of classification.

To fully appreciate the accuracy of each model, mean accuracies and ranges are calculated. After the point estimates are obtained, different randomly selected training and test samples (n=30) will be generated to be used for model validation. Models based on these randomly generated sets will be used to calculate a mean and confidence interval for model performance, illustrating the range of accurate classification for each model. This measure presents a more accurate picture of overall model performance than a point estimate alone.

Finally, this research addresses three hypotheses, concerning different subsets of the data (i.e. cranial data, dental data, and both datasets combined). Therefore, each type of modeling will be implemented on three subsets of variables: cranial variables only, dental variables only, and cranial and dental variables combined.

4.2.5.1 Random Forest Modeling

Random forest modeling (RFM) is an extension of simpler decision tree methods. The trees in both families of methods predict group membership on the level of a categorical variable, based on the values of a set of predefined predictor variables (Breiman et al. 1984, Feldesman 2002, Mercer 2013, Quinlan 1987). A subset of the data, termed the training set, is used to build a model and find sectioning points for each variable that best separate the data into the levels of the response variable, in this case ancestry (Breiman et al. 1984, Feldesman 2002, Quinlan 1987). Each variable is used as a node in a tree, such that a decision regarding class membership is made at each junction as the tree proceeds (Breiman et al. 1984). An unknown individual is run through the tree, following the decision rules at each node, until it reaches a terminal node where it is classified (Breiman et al. 1984). This method makes very few assumptions, and is very robust in that it can handle large sets of data (Breiman et al. 1984, Feldesman 2002, Quinlan 1987). The random forest extension of the decision tree method employs ensemble methods to add robusticity (Breiman 2001, Breiman and Cutler 2004). In a RFM many decision trees are generated from random subsamples of the training set, and each tree uses a random subset of the predictor variables at each node introducing two layers of randomness (Breiman 2001, Breiman and Cutler 2004, Liaw and Wiener 2002). The twofold randomness of the RFM ensures that every tree in the forest is unique, and helps to avoid the dangers of over-fitting the model to the data (Breiman 2001, Breiman and Cutler 2004, Liaw and Wiener 2002). The final classification of an unknown individual comprises the majority vote of the forest (Breiman 2001, Breiman and Cutler 2004, Strobl et al. 2008).

Furthermore, tree-based classification methods provide a variety of methods for dealing with missing data. Most often data imputation is used to replace missing data with a value based on similar individuals (Breiman et al. 1984); this is especially applicable when predictor variables are continuous. Surrogate variables can also be employed (Breiman et al. 1984). When a decision tree encounters an individual with missing data at a given node, the use of surrogate variables allows the tree to pass the individual to the next node based on another variable that results in a similar split as the variable for which the data are missing (Breiman et al. 1984, Rieger et al. 2010). The final option is to pass individuals with missing data on to the majority resultant node, and continue classifying from there. This method is the default in *cforest()* (Hothorn, personal communication 2016), requires the least data manipulation, and is the best representation of the actual data structure. Here, missing values were analyzed using the third option.

An extension of random forest modeling, using conditional inference trees, is used here because it employs a variable importance measure more appropriate for the structure of these data. Variable importance measures in traditional random forests, like the Gini Index used by the *randomForest* package (Liaw and Wiener 2002), can be biased toward variables with more levels (Deng et al. 2011, Hapfelmeier et al. 2014, Hothorn et al. 2006b, Jensen and Cohen 2000, Kass 1980, Shih 2004, White and Liu 1994), variables with a considerable amount of missing data (Hapfelmeier et al. 2014, Kim and Loh 2001), and spuriously correlated variables (Hapfelmeier et al. 2014, Strobl et al. 2008). These biases are the result of the criterion used to calculate variable importance, which in most cases is some measure of the decrease in impurity in the resultant nodes (Breiman et al. 1984, Kim and Loh 2001, Shih 2004). These biases are important to consider, as the variable importance measure is used in building the forest. At each node, a random subset of variables is selected, and the variable that best splits the data, is chosen for that node. Therefore, biased measures of variable importance, can bias the whole forest.

When comparing two predictor variables, measures of decrease in impurity are more likely to find a large decrease when variables can take more levels (Deng et al. 2011, Hothorn et al. 2006b, Jensen and Cohen 2000, Kass 1980, Shih 2004, White and Liu 1994). Consider two randomly generated variables that have no relation to the response. The first can be recorded on ten levels, and the second on only two levels. Simply because the first variable has more potential splits to examine, a split with a greater decrease in impurity is more likely to be found for that variable, even though both variables bear no relationship to the response (Jensen and Cohen 2000). This pattern extends even if the variable with fewer levels does correlate with the response (Dobra and Gehrke 2001). In other words, the relative importance of variables with different numbers of possible responses cannot be accurately assessed using impurity measures (Shih 2004).

A similar problem is encountered when a variable has a large amount of missing data. A splitting rule at a given node is based on the data that are present, ignoring the missing data (Breiman et al. 2004). Therefore, a variable with a large amount of missing data is less likely to be misclassified, and more likely to be considered important (Kim and Loh 2001). To borrow an extreme example from Kim and Loh (2001:596), consider a classification case with two possible responses (e.g. European and African ancestry), and a variable X that is missing for all but two individuals. If we further assume that these two individuals have different values for variable X, and belong to different classes, then any split that places these individuals in different nodes will result in an impurity score of 0, and elevate the relative importance of variable X in the overall classification model.

The final bias with regard to variable importance, as calculated by traditional random forests, relates to the problem of spurious correlation (Strobl et al. 2008). Consider two variables, one of which is highly correlated with the response, while the other is highly correlated with the first variable, but has little relationship to the response variable. Variable importance measures based on a decrease in impurity would likely indicate that both variables are important in predicting the response, though the second variable is not related to the response (Strobl et al. 2008). In other words, variable importance measures for traditional random forests are indicative of which variables effectively separate the data into the levels of the response variable, but do little to indicate which predictors bear a strong relationship to the response. Though this is not necessarily a problem for classification applications, it makes future variable selection difficult.

One solution to the above described biases is to use permutation importance measures of variable importance when growing forests (Hapfelmeier et al. 2014, Hothorn et al. 2006b, Rieger et al. 2010, Strobl et al. 2008). Permutation importance is measured by comparing the accuracy of a tree before and after the random imputation of a single predictor variable, and averaging that difference over all trees in the forest (Hapfelmeier et al. 2014, Strobl et al. 2008). The random permutation of the predictor variable destroys its relationship to the response; therefore, the difference in the two accuracies reflects the importance of that single variable to the overall accuracy of the forest (Hapfelmeier et al. 2014). Though effective, basic permutation importance cannot account for the effects of missing data. An alteration to the permutation importance formula that allows for missing data is provided by Hapfelmeier and colleagues (2014). In essence, the random permutation is performed on the individuals placed in the daughter nodes of a split rather than on the values of the split variable (Hapfelmeier et al. 2014). This allows the importance of the variable to be assessed, without the complications of missing data and surrogate variables.

Random forest models were generated using the *cforest()* command from the *party* package (Hothorn et al. 2006a, Strobl et al. 2007, Strobl et al. 2008). The *cforest()* command is most appropriate for these data because the predictor variables in this data set are expressed on many different levels, and data derived from human skeletal materials are likely to have a large number of missing values. Additionally, permutation importance, and the transformation by Hapfelmeier and colleagues (2014), is the default variable importance measure in *cforest().* Variable importance values were examined using the *varimp()* command from the same package in order to graphically represent the relative importance of the variables in the model.

In each of the three models, the pool of randomly selected predictors available at each node was roughly equal to the square root of total available predictors, and were selected with replacement, per the suggested parameters proposed by Breiman and Cutler (1984). Every forest is composed of 225 trees; this number was kept odd to avoid ties in forest voting.

4.2.5.2 Naïve Bayes Classification

Unlike the statistical methods discussed so far, naïve Bayesian classification draws on assumptions from a different school of statistical thought. Bayesian statistics are founded on the assumption that prior knowledge of a problem can inform the acceptance or rejection of a hypothesis. This has been formalized as Bayes' theorem:

$$p(c_j|d) = \frac{p(d|c_j)p(c_j)}{p(d)}$$

in which, $p(c_j/d)$ is the probability of some observation, d, occurring in class c_j. In the case of this research that would be the probability of a given individual being a member of a given ancestry group. The term $p(d/c_j)$ is the probability that members of class c_j have some variable *d.* This term will be explored in more depth in the explanation of naïve Bayes classification below, but it can be thought of as a measure of compatibility between the variable (*d*) and the class (c_j). For this research, this term represents the likelihood of observing a given variable (e.g. ANS score 5) in each ancestry group. The next term, $p(c_j)$, may be the most important to the Bayesian approach. This term accounts for the frequency of a given class in a sample. This amounts to the prior knowledge that informs a Bayesian analysis. As an example, if the majority of an analyzed sample belongs to the European ancestry group, the probability of an unknown individual from that population being European is higher. The final term, $p(d)$, is the probability of the variable occurring. This variable can effectively be ignored, as this probability is the same for all possible classes, and does not affect the relative probabilities of class membership (Domingos and Pazzani 1997, Langley et al. 1992, Rish et al. 2001).

The naïve Bayes classifier relies on an expansion of the second term of Bayes theorem. The probability of the evidence (*d)* given class (c_j) is calculated as the product of the probabilities of the individual predictor variables (Friedman et al. 1997, Rish et al. 2001). For this probability to be accurate, it is naïvely assumed that the predictor variables are independent of one another, hence the name (Domingos and Pazzani 1997, Friedman et al. 1997, Rish et al. 2001). This unrealistic assumption of the Bayes classifier was thought of as a problem for many years; however, despite this seeming handicap, the naïve Bayes classifier has consistently performed well on many data types (Domingos and Pazzani 1997, Friedman et al.

1997, Langley et al. 1992, Rish et al. 2001). Naïve Bayes classification has been shown to outperform more sophisticated classification methods, like decision trees (Domingos and Pazzani 1997), and has also been shown to perform optimally in the case of correlated predictor variables, in addition to the independent variables assumed by the method (Domingos and Pazzani 1997, Rish et al. 2001). These revelations regarding naïve Bayes classification, have reinvigorated interest in its application to classification problems.

For this research, naïve Bayes classification provides two advantages. First, although against methodological assumptions, the ability of the classifier to perform well on correlated predictor variables is an attractive feature. Although the number of highly correlated variables is reduced, a certain degree of correlation is expected. The second advantage of this method is that it deals with missing data by effectively ignoring it in probability calculations (Domingos and Pazzani 1997). The probability of each variable occurring in a class is calculated only from individuals who have data for that variable, without removing the other individuals from the data. This expands the size of the sample that can be used to inform a classification.

Furthermore, it is conceivable that a given variable score and ancestry designation never co-occur in the training set. This is problematic in that it introduces a zero probability element into the numerator of the classifier, in effect making the probability of an individual with that trait belonging to the given ancestry group also zero (Domingos and Pazzani 1997). The training set cannot be expected to completely represent the full range of extant human variation;

therefore, Laplace smoothing was used to avoid the zero probability problem. Laplace smoothing increases count data by one, to eliminate the possibility of a zero count for a trait value (Domingos and Pazzani 1997, Niblett 1987). The *naiveBayes()* function from the *e1701* package(Meyer et al. 2015) in R was used to conduct naïve Bayes classification.

4.2.6 Statistical Comparisons of Accuracy

Finally, the different models will be compared to one another to assess if differences exist in the rates of accurate classification, and if those rates ate significantly different from chance. The first comparisons will be between model accuracy and random allocation, using 25% as the benchmark for random assignment (four groups). Second, the different data types will be compared to one another to determine if there is a significant difference in accurate ancestry assessment if different data types are used. Finally, RFM and naïve Bayes will be compared to one another to assess if one statistic is the more powerful choice.

All of the comparisons will be assessed by calculating the z-score for proportions. The z-score is indicative of a standard distribution for which p-values are easily generated. Comparisons of proportions are calculated as follows:

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}$$

Where $\hat{p}$ is the target proportion, p_0 is the proportion to which it is being compared, and n is the size of the sample from which the proportions were calculated. Differences were assessed as significant if they reached the threshold of the appropriate adjusted p-value. In comparing data types for each model (i.e. Cranial vs. Dental for the OOB model, Cranial vs. Dental for the RFM test model, Cranial vs. Dental for the Bayes model, etc.), there are nine (9) potential comparisons; therefore, the adjusted p-value is 0.00556. There are six (6) possible comparisons when comparing the test set models to random chance (i.e. Dental RFM vs. chance, Dental Bayes vs. chance, Cranial RFM vs. Chance, etc.); therefore, the adjusted p is 0.00833. Finally, comparing RFM to naïve Bayes results in three (3) comparisons. Therefore, the adjusted p-value for these z tests will be 0.01667.

Chapter 5:
Results

5.1 Exploratory Data Analyses

5.1.1 Intraobserver Error

The results of the Cohen's weighted Kappa test for intraobserver error can be seen in Table 5.1. All variables were scored with "moderate agreement" or higher, following the designations of Landis and Koch (1977). The lowest intraobserver score (0.600) was seen in the recordation of Zygomaticomaxillary Suture Shape (ZSS), while the highest score (1.000) was observed in the recordation of several traits, including UI1 Interruption Groove, UI2 Interruption Groove, UP1 Odontome, LP2 Odontome, LM2 Distal Trigonid Crest, LM2 Cusp 7, and LM3 Cusp 7. None of the traits were scored inconsistently; therefore, no traits were excluded from classificatory modeling on the basis of inconsistent recordation.

5.1.2 Frequency Tables

Frequency distributions for the 79 variables collected in this research are shown in Appendix A. Two traits, upper second premolar odontome and lower first premolar odontome showed no variation, that is no individual in this sample had a recorded occurrence of odontome presence on these teeth. Therefore, these variables were excluded from subsequent analyses.

Additionally, the Asian/Native American sample in the research is very small (n=10), and for some traits contained as few as three recordable individuals. Therefore, this small sub-sample was removed from further analyses so measures of between group differences, and later classification model accuracies were not skewed by small sample size.

5.1.3 Chi-Square Tests

Once the Asian/Native American individuals, and the traits with no variation were removed from the sample, chi-square tests indicated that 40 of 77 variables exhibited significant differences among the 3 ancestry groups (Table 5.2). As 77 of the original 79 variables were used in chi-square tests, an adjusted α was used to assess statistical significance. In this case, tests needed to return a p-value below an adjusted $\alpha=0.0006494$, or have an adjusted p-value below 0.05. The 37 variables that exhibited insignificant differences among groups were removed from further analyses.

5.1.4 Polychoric Correlation

The matrix of polychoric correlation coefficients (Appendix B) revealed only 14 pairs of variables with correlations above the 0.75 cut off point. These pairs are shown in Table 5.3. To reduce collinearity in the data before building the classification models, UI1 Shoveling, LI1 Shoveling, UM2 Crenulations, LM2 Crenulations, UM3 Enamel Extensions, and LM1 Enamel Extensions were removed from the data set. These variables were chosen because they maximally reduce highly correlated data, with minimal reduction in the total number of variables.

Table 5.1: Results of the weighted Cohen's Kappa measure of intraobserver reliability.

Trait	Weighted Kappa Value	Landis and Koch (1977) Designation
UI1 Interruption Groove	1.000	Almost Perfect
UI2 Interruption Groove	1.000	Almost Perfect
UP1 Odontome	1.000	Almost Perfect
LP2 Odontome	1.000	Almost Perfect
LM2 Distal Trigonid Crest	1.000	Almost Perfect
LM2 Cusp 7	1.000	Almost Perfect
LM3 Cusp 7	1.000	Almost Perfect
UP2 Accessory Cusps	0.986	Almost Perfect
LM1 Enamel Extensions	0.986	Almost Perfect
LM1 Distal Trigonid Crest	0.986	Almost Perfect
LM3 Enamel Extensions	0.985	Almost Perfect
LM3 Cusp Number	0.985	Almost Perfect
LM2 Protostylid	0.985	Almost Perfect
LM1 Cusp Number	0.974	Almost Perfect
UM1 Parastyle	0.971	Almost Perfect
UM3 Enamel Extensions	0.971	Almost Perfect
LP1 Odontome	0.965	Almost Perfect
LM2 Cusp Number	0.963	Almost Perfect
UM1 Enamel Extensions	0.962	Almost Perfect
LM2 Crenulations	0.962	Almost Perfect
UP2 Odontome	0.958	Almost Perfect
UM1 Hypocone	0.957	Almost Perfect
LM3 Protostylid	0.956	Almost Perfect
LM3 Distal Trigonid Crest	0.953	Almost Perfect
Mandibular Dental Crowding	0.953	Almost Perfect
LM2 Cusp 6	0.948	Almost Perfect
UM2 Crenulations	0.948	Almost Perfect
LM1 Groove Pattern	0.946	Almost Perfect
UP1 Accessory Cusps	0.944	Almost Perfect
UM2 Paramolar Tubercle	0.942	Almost Perfect
LI2 Shoveling	0.941	Almost Perfect
LM1 Cusp 7	0.931	Almost Perfect
LM1 Protostylid	0.930	Almost Perfect
LM1 Deflecting Wrinkle	0.927	Almost Perfect

Maxillary Dental Crowding	0.926	Almost Perfect
LM3 Cusp 6	0.925	Almost Perfect
Nasal Overgrowth	0.922	Almost Perfect
Canine Mesial Ridge (Bushman Canine)	0.922	Almost Perfect
LM2 Enamel Extensions	0.920	Almost Perfect
LM1 Crenulations	0.919	Almost Perfect
Winging	0.917	Almost Perfect
LM2 Cusp 5	0.917	Almost Perfect
UM2 Enamel Extensions	0.910	Almost Perfect
LM1 Cusp 6	0.907	Almost Perfect
LI1 Shoveling	0.906	Almost Perfect
UM1 Crenulations	0.905	Almost Perfect
UM1 Carabelli's Trait	0.901	Almost Perfect
LM1 Cusp 5	0.882	Almost Perfect
UI1 Double Shoveling	0.878	Almost Perfect
UI2 Double Shoveling	0.878	Almost Perfect
UI1 Tuberculum Dentale	0.876	Almost Perfect
LM3 Cusp 5	0.861	Almost Perfect
Diastema	0.853	Almost Perfect
LC Distal Accessory Ridge	0.836	Almost Perfect
LM1 Anterior Fovea	0.835	Almost Perfect
LM3 Groove Pattern	0.819	Almost Perfect
UM2 Hypocone	0.810	Almost Perfect
UC Tuberculum Dentale	0.792	Substantial
UI1 Shoveling	0.788	Substantial
UI2 Tuberculum Dentale	0.786	Substantial
UM3 Crenulations	0.778	Substantial
LM3 Crenulations	0.773	Substantial
Post-Bregmatic Depression	0.768	Substantial
UM3 Hypocone	0.765	Substantial
Anterior Nasal Spine	0.759	Substantial
LM2 Groove Pattern	0.754	Substantial
UI2 Shoveling	0.724	Substantial
Inferior Nasal Aperture	0.715	Substantial
Transverse Palatine Suture	0.714	Substantial
LP1 Lingual Cusp Number	0.714	Substantial
Nasal Aperture Width	0.702	Substantial

Interorbital Breadth	0.666	Substantial
Malar Tubercle	0.658	Substantial
UC Distal Accessory Ridge	0.640	Substantial
Supranasal Suture	0.634	Substantial
Nasal Bone Contour	0.624	Substantial
LP2 Lingual Cusp Number	0.618	Substantial
Palate Shape	0.610	Substantial
Zygomaticomaxillary Suture Shape	0.600	Moderate

Table 5.2: Result of chi-square tests for differences between ancestry groups. Significant differences assessed if p-value<0.0006494 or adj-p<0.05. Traits above the bold line are significantly different among groups.

Trait	Chi-Square Value	p-value	adj-p
LM2 Crenulations	115.45	2.20E-16	0.0000
UM2 Crenulations	93.634	2.20E-16	0.0000
LI2 Shoveling	103.28	2.20E-16	0.0000
Nasal Overgrowth	102.49	2.20E-16	0.0000
UM2 Enamel Extensions	94.844	2.20E-16	0.0000
LI1 Shoveling	107.00	2.20E-16	0.0000
Anterior Nasal Spine	189.09	2.20E-16	0.0000
Inferior Nasal Aperture	327.48	2.20E-16	0.0000
Transverse Palatine Suture	121.32	2.20E-16	0.0000
Nasal Aperture Width	292.13	2.20E-16	0.0000
Interorbital Breadth	153.44	2.20E-16	0.0000
Nasal Bone Contour	134.64	2.20E-16	0.0000
Palate Shape	217.34	2.20E-16	0.0000
Zygomaticomaxillary Suture Shape	80.67	2.20E-16	0.0000
LC Distal Accessory Ridge	94.22	7.77E-16	0.0000
UI1 Shoveling	85.21	4.74E-14	0.0000
UC Distal Accessory Ridge	83.668	9.55E-14	0.0000
LM1 Enamel Extensions	70.602	3.08E-13	0.0000
LM1 Crenulations	59.765	3.25E-12	0.0000
LM2 Enamel Extensions	60.576	3.44E-11	0.0000
UM1 Crenulations	52.063	1.34E-10	0.0000
UI2 Shoveling	68.446	6.25E-10	0.0000
LM2 Groove Pattern	47.563	1.16E-09	0.0000
Supranasal Suture	44.717	4.55E-09	0.0000

LM3 Crenulations	44.459	5.15E-09	0.0000
UM3 Enamel Extensions	42.585	1.26E-08	0.0000
UM3 Crenulations	42.455	1.34E-08	0.0000
Canine Mesial Ridge (Bushman Canine)	46.814	2.04E-08	0.0000
Post-Bregmatic Depression	30.339	2.58E-07	0.0000
UM2 Hypocone	47.136	8.92E-07	0.0001
UM1 Enamel Extensions	33.522	9.34E-07	0.0001
LM1 Cusp 5	43.401	4.22E-06	0.0003
LM1 Cusp Number	30.014	4.86E-06	0.0004
Maxillary Dental Crowding	23.913	6.42E-06	0.0005
Winging	33.611	8.00E-06	0.0006
UI2 Double Shoveling	33.177	9.69E-06	0.0007
LM3 Enamel Extensions	23.362	1.07E-04	0.0083
LM1 Deflecting Wrinkle	27.63	1.10E-04	0.0085
LP1 Lingual Cusp Number	41.285	1.60E-04	0.0124
UM3 Hypocone	32.308	3.56E-04	0.0274
UI2 Interruption Groove	23.473	6.53E-04	0.0502
UC Tuberculum Dentale	33.297	8.69E-04	0.0669
LM1 Anterior Fovea	22.022	1.20E-03	0.0924
UI1 Double Shoveling	21.392	1.56E-03	0.1201
LM2 Cusp 5	28.144	1.71E-03	0.1318
Malar Tubercle	20.971	1.86E-03	0.1430
LM1 Protostylid	19.885	2.90E-03	0.2235
LM1 Cusp 7	21.169	6.71E-03	0.5168
LM2 Protostylid	23.779	8.21E-03	0.6322
LM3 Protostylid	26.314	9.69E-03	0.7460
UP1 Accessory Cusps	8.8302	1.21E-02	0.9309
UI2 Tuberculum Dentale	22.027	1.50E-02	1.0000
LM2 Cusp Number	12.287	1.53E-02	1.0000
UM1 Hypocone	16.721	3.32E-02	1.0000
Mandibular Dental Crowding	6.6488	3.60E-02	1.0000
LM3 Cusp 5	18.36	4.92E-02	1.0000
UM1 Carabelli's Trait	23.661	5.03E-02	1.0000
LP2 Lingual Cusp Number	23.539	5.21E-02	1.0000
LM1 Distal Trigonid Crest	5.8899	5.26E-02	1.0000
UI1 Interruption Groove	189.09	1.08E-01	1.0000
LM3 Cusp 7	12.916	1.15E-01	1.0000

LM1 Groove Pattern	4.2732	1.18E-01	1.0000
LP2 Odontome	4.0355	1.33E-01	1.0000
LM2 Distal Trigonid Crest	3.5097	1.73E-01	1.0000
LM3 Distal Trigonid Crest	2.6281	2.69E-01	1.0000
LM3 Groove Pattern	4.9429	2.93E-01	1.0000
LM2 Cusp 7	4.5651	3.35E-01	1.0000
LM3 Cusp 6	11.141	3.47E-01	1.0000
LM3 Cusp Number	4.3097	3.66E-01	1.0000
UM1 Parastyle	1.7264	4.22E-01	1.0000
UP2 Accessory Cusps	1.6147	4.46E-01	1.0000
UI1 Tuberculum Dentale	7.7364	4.60E-01	1.0000
UP1 Odontome	1.4259	4.90E-01	1.0000
UM2 Paramolar Tubercle	10.532	5.69E-01	1.0000
LM1 Cusp 6	6.4391	5.98E-01	1.0000
Diastema	2.1799	7.03E-01	1.0000
LM2 Cusp 6	3.3208	7.68E-01	1.0000

Table 5.3: Pairs of traits with polychoric correlation coefficients exceeding |0.75|. Traits removed from classification models are italicized.

Trait 1	Trait 2	Correlation Coefficient
UI1 Shoveling	UI2 Shoveling	0.86860787
	UI2 Double Shoveling	0.83466392
	LI1 Shoveling	0.86506956
	LI2 Shoveling	0.77344532
LI1 Shoveling	Post-Bregmatic Depression	-0.9457385
	UI2 Shoveling	0.76477853
	LI2 Shoveling	0.9892549
	UM1 Enamel Extensions	0.77942298
UM2 Crenulations	UM1 Crenulations	0.78671532
	UM3 Crenulations	0.7628298
LM2 Crenulations	LM1 Crenulations	0.80163239
	LM3 Crenulations	0.8134159
UM3 Enamel Extensions	UM2 Enamel Extensions	0.75523545
LM1 Enamel Extensions	UM1 Enamel Extensions	0.81926265
All correlations significant at p<0.01		

5.1.5 The Final Sample

After all exploratory data analyses were completed, 34 variables remained (Table 5.4), with which the classificatory models could be trained and tested. These variables represent those traits that are consistently recorded, exhibit significant differences among ancestry groups, and are not strongly correlated with one another.

Table 5.4: The final 34 variables used for classification modeling.

Cranial (n=11)	Maxillary Dental (n-12)	Mandibular Dental (n=11)
Inferior Nasal Aperture	UM2 Enamel Extensions	LI2 Shoveling
Nasal Aperture Width	UC Distal Accessory Ridge	LC Distal Accessory Ridge
Anterior Nasal Spine	UI2 Shoveling	LM2 Enamel Extensions
Palate Shape	UM1 Crenulations	LM2 Groove Pattern
Nasal Overgrowth	UM1 Enamel Extensions	LM3 Crenulations
Inter-Orbital Breadth	UM3 Crenulations	LM1 Crenulations
Transverse Palatine Suture	UM3 Hypocone	LM3 Enamel Extensions
Zygomaticomaxillary Suture Shape	Maxillary Dental Crowding	LM1 Cusp 5
Nasal Bone Contour	Mesial Canine Ridge (Bushman Canine)	LM1 Deflecting Wrinkle
Post-Bregmatic Depression	UM2 Hypocone	LP1 Lingual Cusp Number
Supranasal Suture	UI2 Double Shoveling	LM1 Cusp Number
	Winging	

5.2 Classification

5.2.1 Cranial Variables

After exploratory data analyses, 11 cranial variables remain for building classification models. The random forest model was built using no surrogate variables, and selecting from a pool of 3 random variables at each node, selected with replacement. The preliminary estimate of model accuracy was assessed using a measure of out-of-bag (OOB) accuracy. To calculate this measure of accuracy, the individuals not used in building a given tree are run through the tree to generate a measure of classification accuracy. This measure tends to over-fit the data, but provides an idea of model performance before further testing is attempted. The point estimate of OOB accuracy for cranial traits is 74.5%. The model was run 30 times, with every iteration producing a random training test set. The average OOB accuracy for the cranial trait model was 75.5%, with a 95% confidence interval of 74.3%-76.8% (Table 5.15). The point estimate confusion matrix for OOB accuracy can be seen below (Table 5.5). This matrix is useful in evaluating asymmetries in correct classification and biases in the model; however, it is important to remember this matrix represents one specific randomly generated training and test set. Row percentages (Row%) can be interpreted as the percent of correct classifications of the observed individuals. For example, the top left cell has a row percent of 83.2%, meaning that 83.2% of individuals of African ancestry were classified correctly. Column percentages (Col%) can be interpreted as the positive predictive value, that is, how often the model is correct when it predicts a specific ancestry. Looking at the same top left cell, when the model predicts African ancestry, it is correct 76.5% of the time.

Accuracy was also assessed using a holdout test sample. This measure of accuracy provides a more generalized measure of classification accuracy because it is calculated using individuals not included in the model, therefore, avoiding much

Table 5.5: Confusion matrix of OOB classification for the cranial trait RFM.

		Predicted Ancestry					
Observed Ancestry		**African**		**Hispanic**		**European**	
		n	***Row%/Col%***	**n**	***Row%/Col%***	**n**	***Row%/Col%***
	African	114	83.2%/76.5%	4	2.9%/6.2%	19	13.9%/7.2%
	Hispanic	11	10.2%/7.4%	47	43.5%/72.3%	50	46.3%/18.9%
	European	24	10.3%/16.1%	14	6.0%/21.5%	196	83.8%/74.0%

of the overfitting of the OOB test. The point estimate of classification using the test set is 71.6%. Mean accuracy for the test set is 75.2%, with a 95% confidence interval of 73.0%-77.3% (Table 5.15). The point estimate confusion matrix for test set accuracy is seen below (Table 5.6).

Table 5.6: Confusion matrix of test set classification for the cranial trait RFM.

		Predicted Ancestry					
Observed Ancestry		**African**		**Hispanic**		**European**	
		n	***Row%/Col%***	**n**	***Row%/Col%***	**n**	***Row%/Col%***
	African	47	81.0%/72.3%	0	0.0%/0.0%	11	19.0%/9.6%
	Hispanic	7	15.2%/10.8%	17	37.0%/70.8%	22	47.8%/19.1%
	European	11	11.0%/16.9%	7	7.0%/29.2%	82	82.0%/71.3%

The permutation importance measure calculated by the *party* (Hothorn et al. 2006a, Strobl et al. 2007, Strobl et al. 2008) package indicates which variables are most strongly related to ancestry. Following the recommendation of Strobl and colleagues (2009) variables are deemed unimportant to predicting ancestry if their importance value is less than or equal to the absolute value of the lowest importance value among all variables. Variable importance for cranial variables is shown in Figure 5.1; a dashed redline has been added to the figure to denote importance. All variables plotted to right of the red line are deemed important to the prediction of ancestry following the above cited recommendation (Strobl et al. 2009).

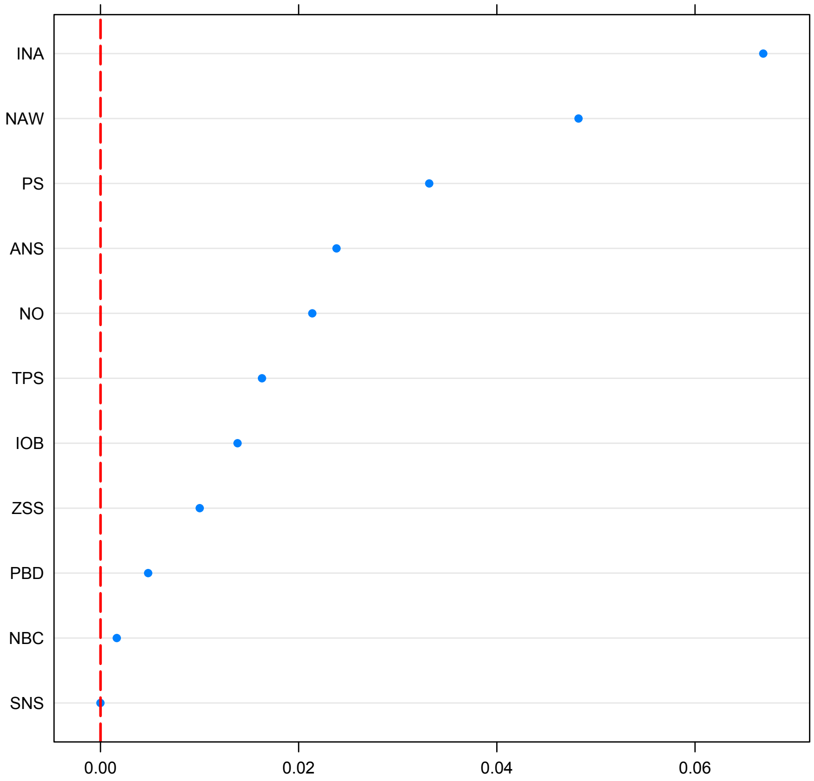


Figure 5.1: Variable importance plot of cranial variables (abbreviations in Table 4.2). Variables to the right of the red line are importantly associated with ancestry following the recommendation of Strobl and colleagues (2009).

The cranial data were also analyzed using a naïve Bayes classifier. Unlike many statistics, Bayesian statistics incorporate the distribution of the sample into the analysis. The prior probabilities for each ancestry group, given the fixed random seed training and test sets, are in Table 5.7. The same training set was used in the analysis of all sets of variables; therefore, the prior probabilities of the ancestry groups for the point estimates are the same in all analyses.

Table 5.7: Prior probabilities of group membership for the fixed seed training set.

Ancestry	N	Probability
African	137	0.28601252
Hispanic	108	0.22546973
European	234	0.48851775
Total	*479*	*1.00000000*

The conditional probabilities for each trait can be found in Appendix C. Conditional probabilities are essentially the probability of a given trait score occurring in an ancestry group. The combination of conditional and prior probabilities yields a posterior probability for an individual and ultimately a classification. For the fixed seed model based only on cranial traits, the point estimate of accuracy is 75.5% (Table 5.15). Thirty runs of the naïve Bayes classifier using different randomly generated training and test sets, yielded a mean accuracy of 79.6%, with a 95% confidence interval of 77.1%-82.1%. The confusion matrix for the fixed seed Bayes classifier is seen below (Table 5.8).

5.2.2 Dental Variables

Following the exploratory analyses outlined above, 23 dental variables remained for modelling. As the number of variables was increased, the number of variables available at each node in the random forest is also increased. In this case, every node drew from a pool of five potential variables when making a split. As before, a point estimate and range were calculated for both the OOB sample and the test sample. The point estimate of OOB accuracy for a model built using only dental variables is 64.5% (Table 5.15). The mean OOB accuracy for the dental variable model is 66.1%, with a 95% confidence interval of 64.7%-67.5% (Table 5.15). The confusion matrix for the dental variables OOB accuracy can be seen below (Table 5.9).

Table 5.8: Confusion matrix of test set classification for the cranial trait naïve Bayes classifier.

		Predicted Ancestry					
Observed Ancestry		**African**		**Hispanic**		**European**	
		n	***Row%/Col%***	**n**	***Row%/Col%***	**n**	***Row%/Col%***
	African	45	77.6%/76.3%	4	6.9%/10.0%	9	15.5%/8.6%
	Hispanic	3	6.5%/5.1%	28	60.9%/70.0%	15	32.6%/14.3%
	European	11	11.0%/18.7%	8	8.0%/20.0%	81	81.0%/77.1%

The 30% test sample was also analyzed using the dental trait only model. The point estimated for accuracy using the test set is 68.6%. Multiple iterations of training and test samples indicate an average test set accuracy of 67.4%, with a 95% confidence interval of 64.7%-70.1% (Table 5.15). The confusion matrix generated from the point estimate model of test set accuracy is seen below (Table 5.10).

Table 5.9: Confusion matrix of OOB classification for the dental trait RFM.

		Predicted Ancestry					
		African		**Hispanic**		**European**	
		n	***Row%/Col%***	**n**	***Row%/Col%***	**n**	***Row%/Col%***
Observed Ancestry	**African**	69	50.4%/79.3%	2	1.5%/7.4%	66	48.2%/18.1%
	Hispanic	7	6.5%/8.1%	21	19.4%/77.8%	80	74.1%/21.9%
	European	11	4.7%/12.6%	4	1.7%/14.8%	219	93.6%/60.0%

Table 5.10: Confusion matrix of test set classification for the dental trait RFM.

		Predicted Ancestry					
		African		**Hispanic**		**European**	
		n	***Row%/Col%***	**n**	***Row%/Col%***	**n**	***Row%/Col%***
Observed Ancestry	**African**	30	51.7%/81.1%	2	3.5%/15.4%	26	44.8%/16.9%
	Hispanic	6	13.0%/16.2%	11	23.9%/84.6%	29	63.1%/18.8%
	European	1	1.0%/2.7%	0	0.0%/0.0%	99	99.0%/64.3%

Permutation importance for the dental variables is shown in Figure 5.2; a dashed redline has been added to the figure to denote importance. All variables plotted to right of the red line are deemed important to the prediction of ancestry following the recommendation of Strobl and colleagues (2009).

Like the cranial data, the dental data were analyzed using a naïve Bayes classifier. The same training and test sets of data were used in calculating the estimate of the prior probabilities; therefore, these probabilities are the same as in the cranial traits only model (Table 5.7). For the fixed seed model based only on dental traits, the point estimate of test set accuracy is 72.6% (Table 5.15). The mean accuracy for this model is 70.5%, with a 95% confidence interval of 68.9%-72.1%.

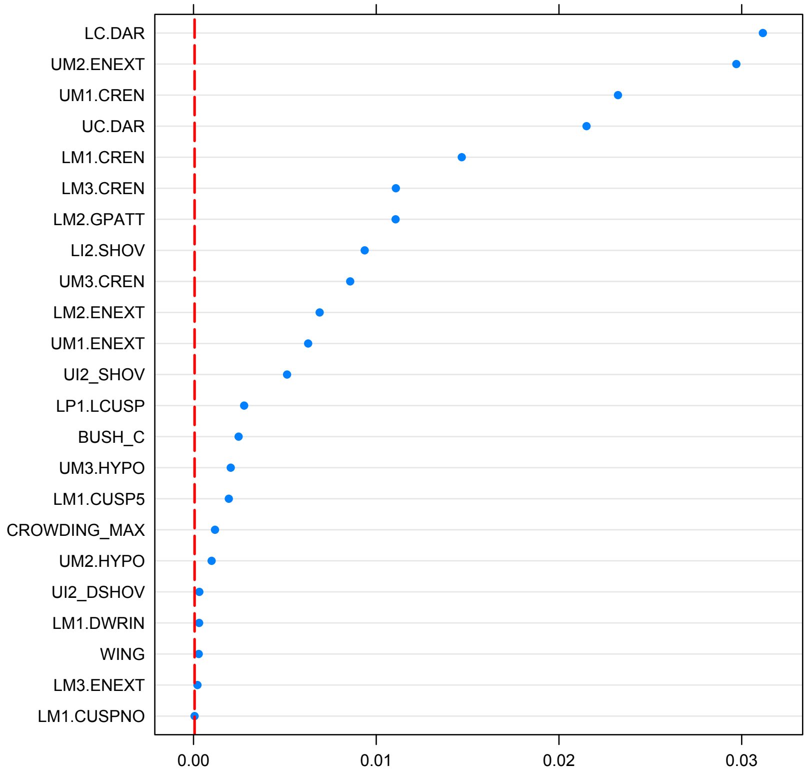


Figure 5.2: Variable importance plot of dental variables (abbreviations in Table 4.3). Variables to the right of the red line are importantly associated with ancestry following the recommendation of Strobl and colleagues (2009).

The confusion matrix for dental trait only naïve Bayes classifier is seen below (Table 5.11).

5.2.3 Cranial and Dental Variables Combined

Finally, the full 34 traits of the combined dataset were used to build both a

Table 5.11: Confusion matrix of test set classification for the dental trait naïve Bayes classifier.

		Predicted Ancestry					
Observed Ancestry		**African**		**Hispanic**		**European**	
		n	***Row%/Col%***	**n**	***Row%/Col%***	**n**	***Row%/Col%***
	African	38	65.5%/73.1%	2	3.5%/8.3%	18	31.0%/14.1%
	Hispanic	4	8.7%/7.7%	21	45.7%/87.5%	21	45.7%/16.4%
	European	10	10.0%/19.2%	1	1.0%/4.2%	89	89.0%/69.5%

Table 5.12: Confusion matrix of OOB classification for the combined data RFM.

		Predicted Ancestry					
Observed Ancestry		**African**		**Hispanic**		**European**	
		n	***Row%/Col%***	**n**	***Row%/Col%***	**n**	***Row%/Col%***
	African	116	84.7%/84.1%	2	1.5%/5.1%	19	13.9%/6.3%
	Hispanic	9	8.3%/6.5%	34	31.5%/87.2%	65	60.2%/21.5%
	European	13	5.6%/9.4%	3	1.3%/7.7%	218	93.2%/72.2

random forest model and a naïve Bayes classifier. The point estimate of OOB accuracy for the RFM built using the combined data set is 76.8%. The mean OOB accuracy for this model is 76.7%, with a 95% confidence interval of 75.5%-77.8% (Table 5.15). The confusion matrix for the combined data OOB accuracy is in Table 5.12.

The point estimate of test accuracy for the combined data set RFM is 74.0%. The mean test set accuracy for this model is 77.1%, with a 95% confidence interval of 75.5%-78.7% (Table 5.15). The confusion matrix for this model is seen below (Table 5.13).

Table 5.13: Confusion matrix of test set classification for the combined data RFM.

		Predicted Ancestry					
Observed Ancestry		**African**		**Hispanic**		**European**	
		n	***Row%/Col%***	**n**	***Row%/Col%***	**n**	***Row%/Col%***
	African	47	97.9%/73.4%	0	0.0%/0.0%	1	2.1%/0.8%
	Hispanic	7	16.3%/11.1%	11	25.6%/100.0%	25	58.1%/21.6%
	European	10	10.0%/15.6%	0	0.0%/0.0%	90	90.0%/77.6%

Overall variable importance, based on the permutation importance measure, is shown in Figure 5.3; a dashed, red, line has been added to the figure to denote importance. All variables plotted to right of the red line are deemed important to the prediction of ancestry (Strobl et al. 2009).

The naïve Bayes classifier based on all 34 cranial and dental variables has a point estimate accuracy of 83.3% (Table 5.15). The mean accuracy for this model is 83.8%, with a 95% confidence interval of 81.9%-85.7%. The confusion matrix for the Bayes classifier using all variables can be seen below (Table 5.14).

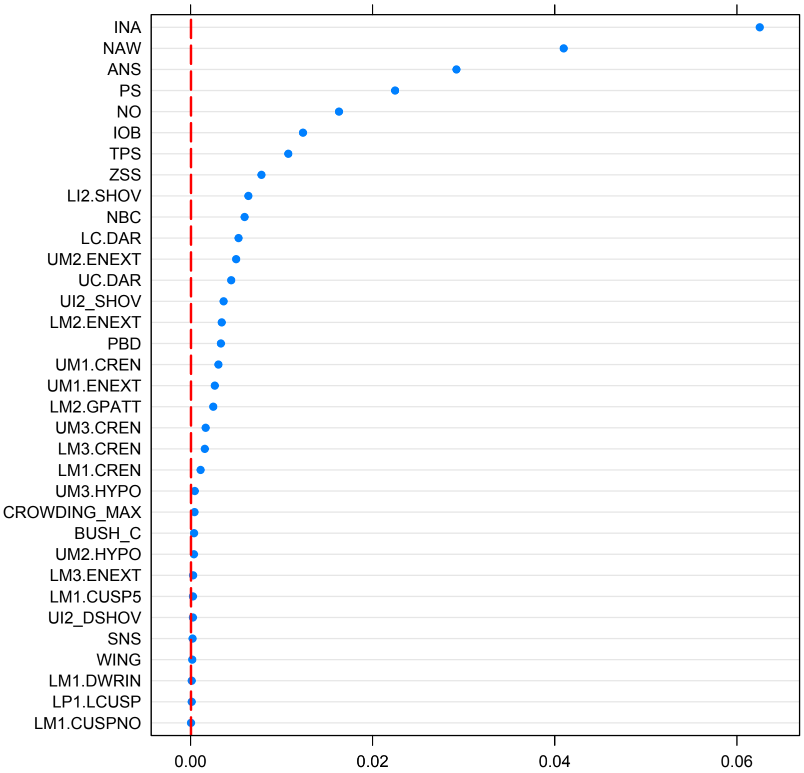


Figure 5.3: Variable importance plot of all variables used in these models. Variables to the right of the red line are importantly associated with ancestry following the recommendation of Strobl and colleagues (2009).

Point estimates of out-of-bag model accuracy, test sample accuracy, and the confusion matrices were calculated from the random training and test samples generated using a set random seed, allowing for reproducible results. Accuracy ranges were generated through multiple runs (n=30) of each model using different randomly generated training and test sets. All measures of model accuracy are in Table 5.15.

Table 5.14: Confusion matrix of test set classification for the combined data naïve Bayes classifier.

		Predicted Ancestry					
Observed Ancestry		**African**		**Hispanic**		**European**	
		n	***Row%/Col%***	**n**	***Row%/Col%***	**n**	***Row%/Col%***
	African	51	87.9%/80.9%	0	0.0%/0.0%	7	12.1%/6.5%
	Hispanic	2	4.3%/3.2%	31	67.4%/93.9%	13	28.3%/12.0%
	European	10	10.0%/15.9%	2	2.0%/6.1%	88	88.0%/81.5%

Table 5.15: Comparison of rates of accurate classification among models.

	RFM OOB Classification			**RFM Test Set Classification**			**Naïve Bayes Test Set Classification**		
	Point Estimate	**Mean Accuracy**	**95% Confidence Interval**	**Point Estimate**	**Mean Accuracy**	**95% Confidence Interval**	**Point Estimate**	**Mean Accuracy**	**95% Confidence Interval**
Cranial Variables	74.5%	75.5%	74.3%-76.8%	71.6%	75.2%	73.0%-77.3%	75.5%	79.6%	77.1%-82.1%
Dental Variables	64.5%	66.1%	64.7%-67.5%	68.6%	67.4%	64.7%-70.1%	72.6%	70.5%	68.9%-72.1%
Combined Variables	76.8%	76.7%	75.5%-77.8%	74.0%	77.1%	75.5%-78.7%	83.3%	83.8%	81.9%-85.7%

5.2.4 Statistical Comparisons of Accuracy

The z-scores and p-values for all comparisons can be seen in Table 5.16. All comparisons are significantly different than expected under random chance. In all tests of accuracy, the models based on dental data alone perform significantly worse than either the cranial variable model or the model with the combined data set

(Figures 5.4-5.6). Accuracy for the combined model is consistently higher than the cranial variable model, though these differences are not statistically significant (Figures 5.4-5.6). Finally, the naïve Bayes classifiers outperform the random forest models for all datasets; however, the difference in accuracy is only significant for the models based on the combined dataset (Figure 5.7).

Table 5.16: Statistical comparisons of accuracy rates.

Comparison	$\hat{p}$	p_0	Adj-p	z-score	p-value
Cranial RFM v. Chance	0.7515	0.3333	0.00833	12.8023	<0.00001
Dental RFM v Chance	0.6741	0.3333	0.00833	10.4506	<0.00001
Combined RFM v. Chance	0.7709	0.3333	0.00833	13.3927	<0.00001
Cranial Bayes v. Chance	0.7959	0.3333	0.00833	14.1523	<0.00001
Dental Bayes v. Chance	0.7049	0.3333	0.00833	11.3878	<0.00001
Combined Bayes v. Chance	0.8379	0.3333	0.00833	15.4279	<0.00001
Dental v. Cranial OOB	0.6610	0.7553	0.00556	-4.7970	<0.00001
Dental v. Combined OOB	0.6610	0.7665	0.00556	-5.4579	<0.00001
Cranial v. Combined OOB	0.7553	0.7665	0.00556	-0.5827	0.279947
Dental v. Cranial RFM	0.6741	0.7515	0.00556	-2.5587	0.005249
Dental v. Combined RFM	0.6741	0.7709	0.00556	-3.2918	0.000497

Cranial v. Combined RFM	0.7515	0.7709	0.00556	-0.6606	0.254306
Dental v. Cranial Bayes	0.7049	0.7959	0.00556	-3.2254	0.00063
Dental v. Combined Bayes	0.7049	0.8379	0.00556	-5.1548	<0.00001
Cranial v. Combined Bayes	0.7959	0.8379	0.00556	-1.6275	0.051869
Dental Bayes v. Dental RFM	0.7049	0.6741	0.01667	0.9401	0.173609
Cranial Bayes v. Cranial RFM	0.7959	0.7515	0.01667	1.4689	0.070916
Combined Bayes v. Combined RFM	0.8379	0.7709	0.01667	2.2772	0.011393

5.3 Summary

The results of this research indicate that all of the observed cranial morphoscopic and dental morphological traits can be reliably observed. About half (40 of 77) of the recorded traits exhibited significant differences between groups with the exception of the Asian/Native American sample, which was excluded due to small sample size. Several traits (n=6) were removed from further analyses because they were strongly correlated with other variables. The final analyses were conducted on 34 variables from a sample of 683 individuals. Classification into the correct ancestry group occurred in about 65-86% of individuals depending on the dataset examined and the method applied. All classifications performed

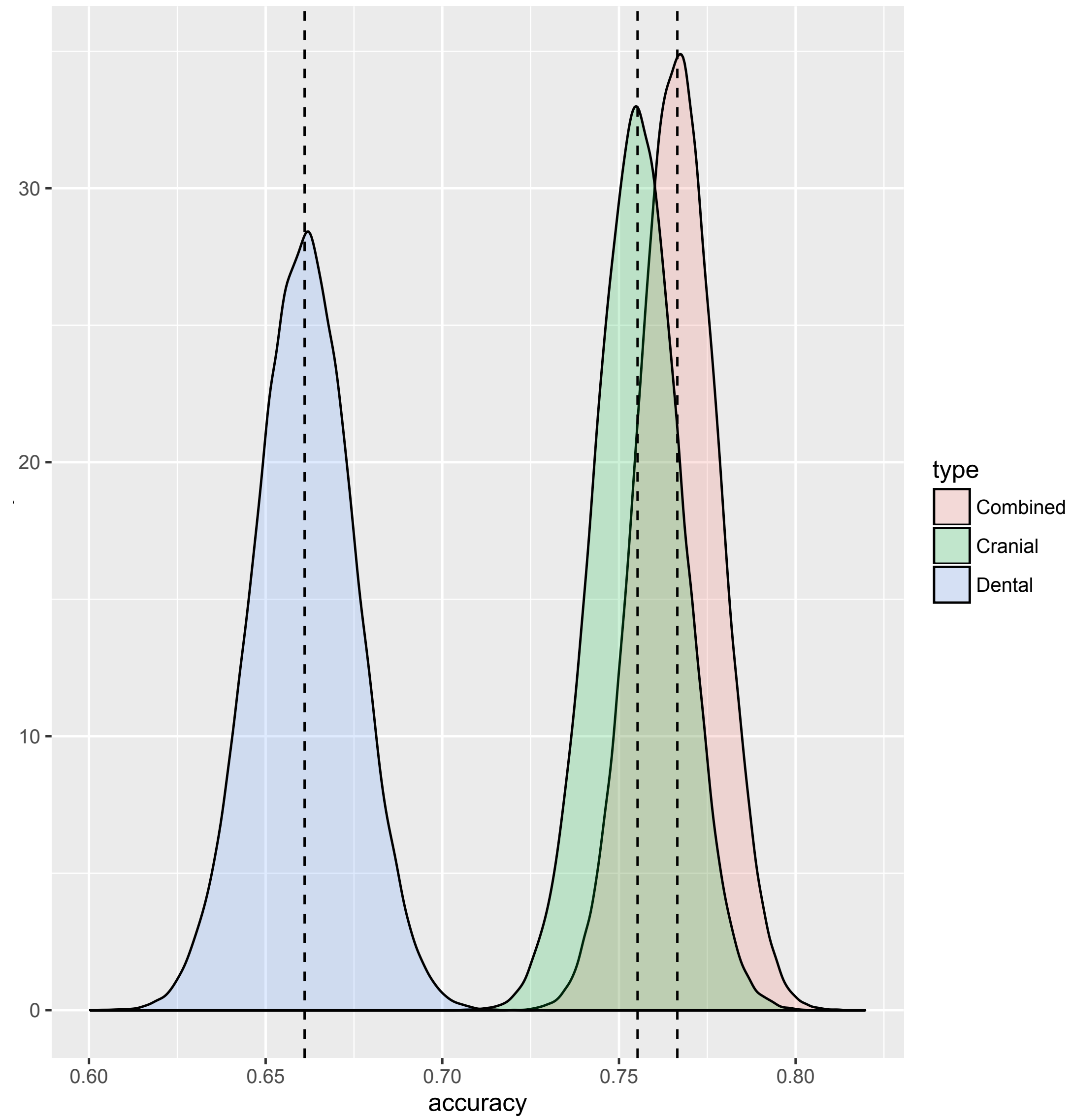


Figure 5.4: Plot of the mean out-of-bag (OOB) RFM accuracies and associated distributions for each of the three data sets (combined in pink, cranial in green, and dental in blue).

significantly better than chance. Finally, the models based on cranial data and combined data significantly outperformed those based on dental data. The combined data models also outperformed the cranial models, though this difference was not statistically significant.

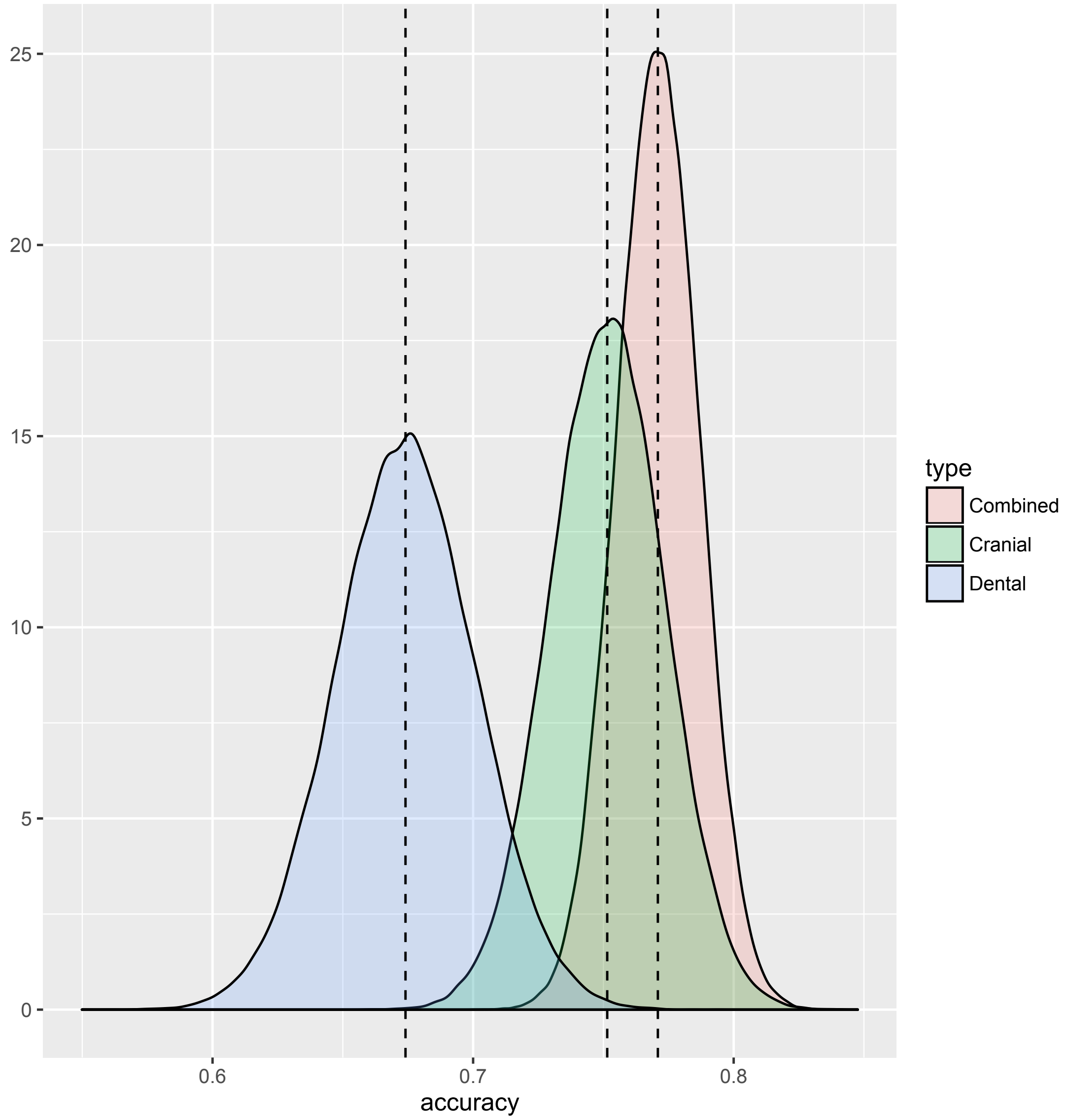


Figure 5.5: Plot of the mean test set accuracies from random forest modeling (RFM) and associated distributions for each of the three data sets (combined in pink, cranial in green, and dental in blue).

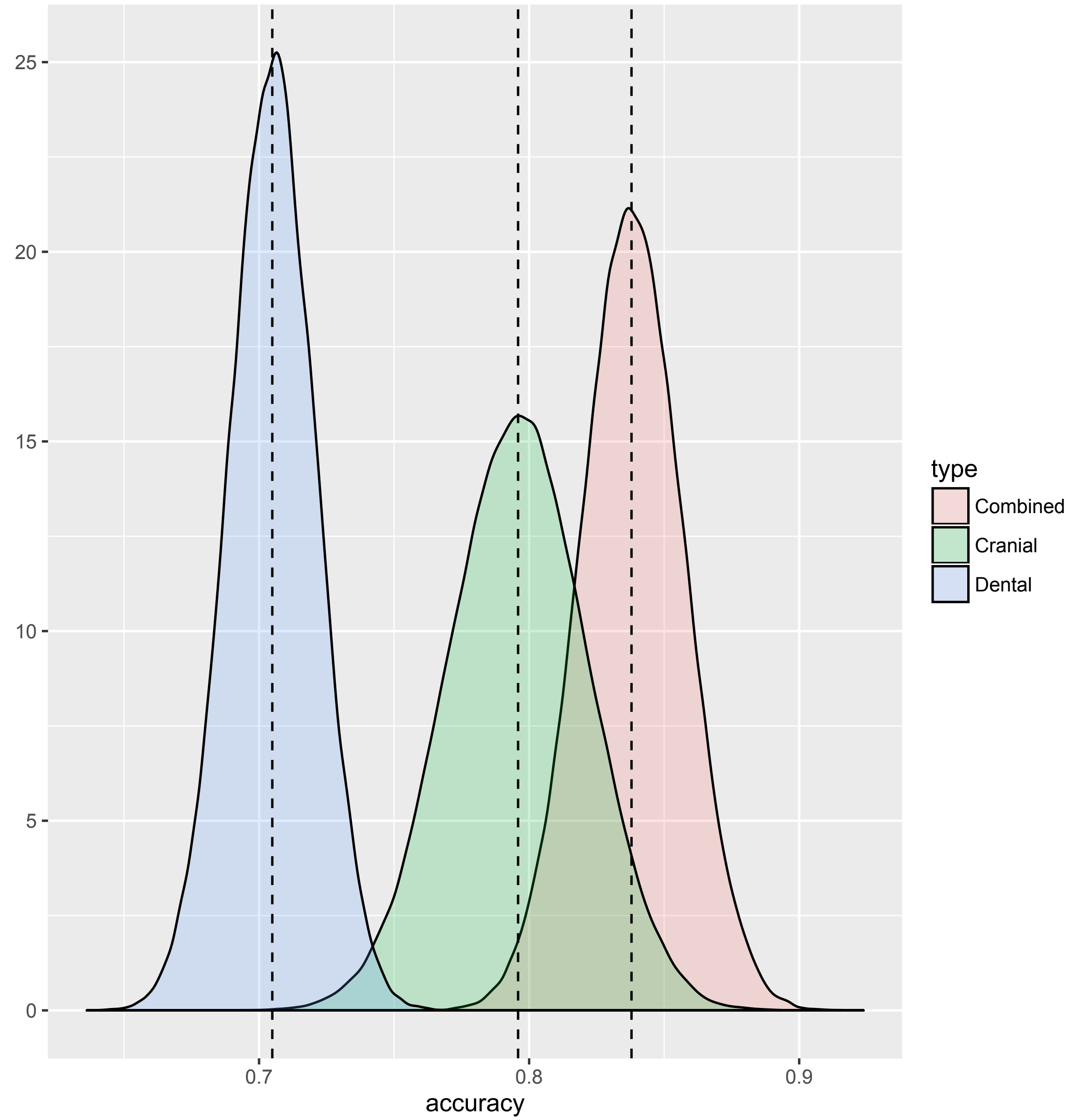


Figure 5.6: Plot of the mean test set accuracies from naïve Bayes classification and associated distributions for each of the three data sets (combined in pink, cranial in green, and dental in blue).

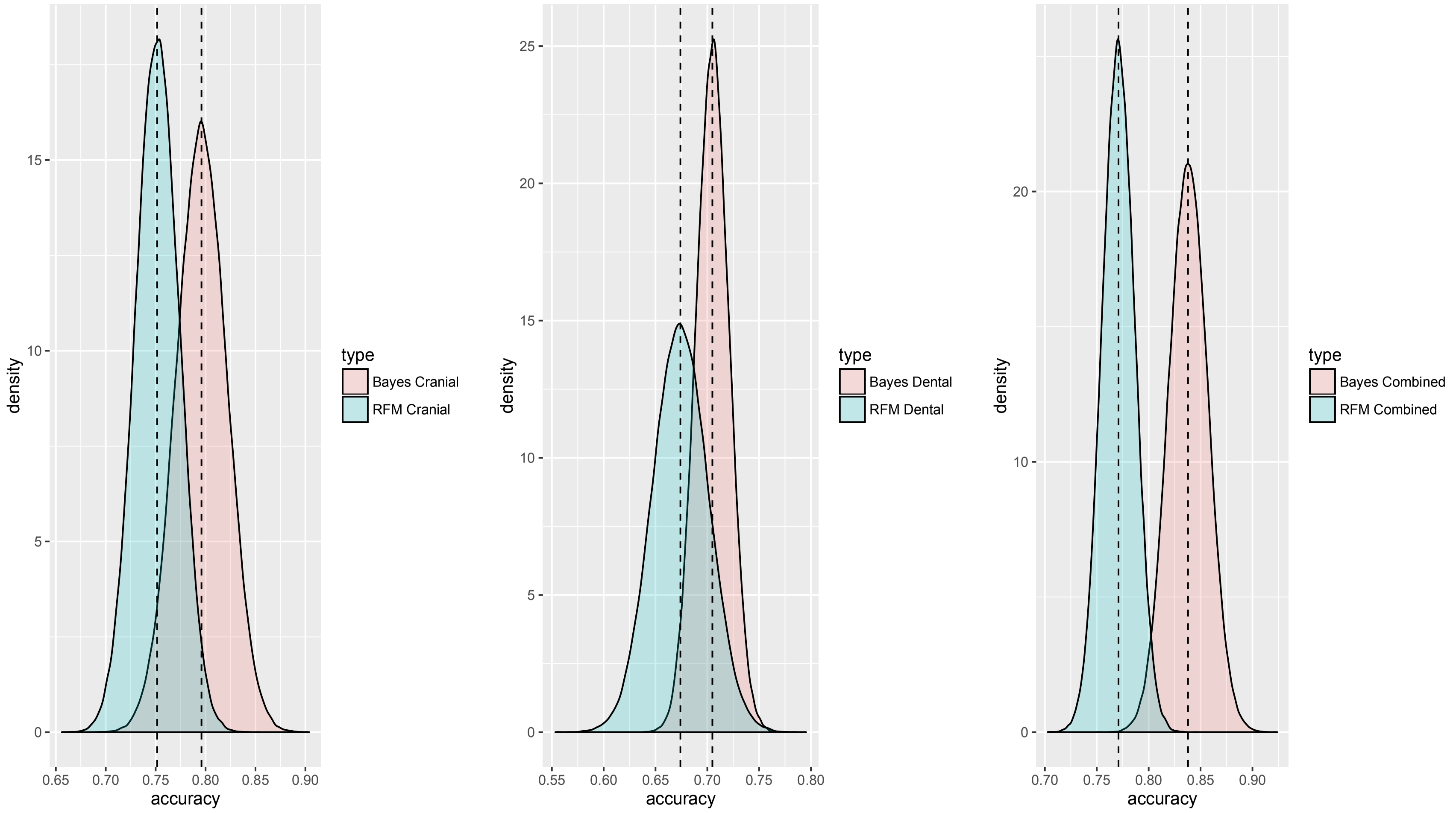


Figure 5.7: Comparison of random forest and naïve Bayes classification accuracy for each data set (left to right: cranial, dental, combined data).

Chapter 6: Discussion

This research is best interpreted through a comparison to previously published models. Cranial morphoscopic traits and dental morphology have each been used in the estimation of ancestry. Therefore, evaluation of the models presented here are considered relative to those benchmarks. These models must perform comparably well to those in existence to demonstrate the validity of the approach and to support the viability of increased classificatory potential through the combination of datasets.

6.1 Intraobserver Error

The Cohen's weighted Kappa measure of intraobserver error suggests that all of the traits recorded in this research were recorded consistently. The range of intraobserver error for the cranial morphoscopic traits is 0.600-0.922. This range is somewhat narrower, but overall lower, than that reported by Hefner (2009). A comparison of these two studies shows that some cranial morphoscopic traits are more consistently recorded than others, though all of the traits in both studies exhibit moderate agreement or better. Table 6.1, shows the 11 cranial morphoscopic traits that were examined in both studies.

Table 6.1. Intraobserver agreement of cranial morphoscopic traits in descending order.

Current Research		Hefner 2009	
Trait	**Weighted-Kappa**	**Trait**	**Kappa**
Nasal Overgrowth	0.922	Nasal Overgrowth	1.000
Post-Bregmatic Depression	0.768	Transverse Palatine Suture	1.000
Anterior Nasal Spine	0.759	Inferior Nasal Aperture	0.964

Inferior Nasal Aperture	0.715	Nasal Aperture Width	0.929
Transverse Palatine Suture	0.714	Malar Tubercle	0.929
Nasal Aperture Width	0.702	Interorbital Breadth	0.857
Interorbital Breadth	0.666	Zygomaticomaxillary Suture Shape	0.857
Malar Tubercle	0.658	Post-Bregmatic Depression	0.820
Supranasal Suture	0.634	Nasal Bone Contour	0.810
Nasal Bone Contour	0.624	Supranasal Suture	0.468
Zygomaticomaxillary Suture Shape	0.600	Anterior Nasal Spine	0.422

Although Kappa values cannot be directly compared because they were calculated differently, there are patterns in the overall consistency of trait recordation. Nasal Overgrowth, Inferior Nasal Aperture, and Transverse Palatine Suture are among the most consistently recorded traits in both studies. Observations were least consistent for Supranasal Suture and Nasal Bone Contour. The most consistently recorded traits are unsurprising for a number of reasons. First, morphoscopic features of the nasal region have been identified as some of the most useful in ancestry estimation (Brues 1990), and consequently exhibit some of the most distinctive morphological differences between groups. This distinctive morphology is easier to identify and repeatedly score than other less pronounced traits. Another potential source of low error, at least in the case of Nasal Overgrowth, is the recordation scale. Nasal overgrowth is scored as either present or absent, not on degrees of expression. This simple scale reduces the likelihood of the same observer scoring the trait differently on two occasions because the trait is either present or not. Finally, the consistency in recordation of the Transverse Palatine Suture may also be linked to the scale on which it is recorded. Unlike many of the other cranial morphoscopic traits recorded by Hefner (2009), the scale for TPS is not truly ordinal. The variations represented by the scale are more discrete. Therefore, the researcher is choosing among four different shapes, rather than scoring the observed trait on a degree of expression. In this way, the recordation scale for TPS may require less subjectivity on the part of the researcher than some other traits.

The traits that are the least consistently recorded are not surprising, but Nasal Bone Contour is exceptional. Hefner (2009) reports an intraobserver agreement of 0.810. This qualifies as “almost perfect” agreement by the standards of Landis and Koch (1977), but is third from the bottom when the scores for all traits are ranked (Table 6.1). In this study, NBC has an intraobserver agreement of 0.624, which is “substantial” agreement (Landis and Koch 1977). Reports of interobserver accuracy in Hefner (2009:993), also indicate difficulty in consistently recording Nasal Bone Contour (Fleiss’s *k*=0.231). As mentioned in Hefner (2009), part of this discrepancy may be in unfamiliarity with the contour gauge used to record NBC. An additional explanation may be difficulty in deciding on a shape category when the observed contour does not exactly match one of the ideal contours proposed on the recordation scale. It is in this second aspect that the creator of the scale may have an advantage, explaining the higher intraobserver error reported in this and other studies (e.g. Klales and Kenyhercz 2015). Less surprising are the relatively low levels of intraobserver agreement observed for the Supranasal Suture. Hefner (2009) argues that the difficulty in assessing this feature is due to the subjectivity inherent in determining the degree of suture closure and obliteration. That conclusion is supported by this research, though SNS was still recorded with relatively little intraobserver error. In summary, the results of this research support the conclusion that cranial morphoscopic traits can be consistently recorded and used in the forensic estimation of ancestry (Hefner 2009).

The morphological traits of the dentition exhibit lower levels of intraobserver error than the cranial traits. Though the range of values is comparable (0.618-1.000), the mean kappa value for the dental traits (0.904) is much higher than that for the cranial traits (0.698). The amount of agreement previously reported for dental morphology hovers around 85%, though it differs depending on the specific conditions of the study, with some observers reporting intraobserver agreement as high as 90% (Sofaer et al. 1972b) or 95% (Scott 1973). Although these measures are high, they are typically based on raw percent concordance, not a kappa statistic (Nichol and Turner II 1986, Scott 1973, Sofaer et al. 1972b), and therefore do not account for agreement by chance. Despite this difference in approach, the mean kappa value in this research is consistent with reported measures of agreement for dental morphology (e.g. Nichol and Turner II 1986, Scott 1973, Scott and Turner II 1997, Sofaer et al. 1972b). This demonstrates the high degree of replicability in observing dental morphology.

Of all 67 dental traits observed in this research, only two, distal accessory ridge on the upper canine (k=0.640) and lingual cusp number on the lower second premolar (k=0.618), were notably less consistently recorded than the others. In both cases, the feature in question is relatively small, and can be difficult to accurately score. The scale for the canine distal accessory ridge is based on prominence of the ridge (Turner II et al. 1991); therefore, lower degrees of expression are difficult to distinguish from one another. Premolar lingual cusps are also difficult to accurately score. Although the scale is based on a count of lingual cusp number (Turner II et al. 1991), faint cusps can be difficult to detect, and the presence or absence of a cusp may be affected by dental wear (Scott and Turner II 1997, Turner II et al. 1991). Dental wear can play a critical role in evaluating the level of intraobserver agreement in dental morphology, especially in deciding whether a trait should be recorded. The researcher must balance maximizing sample size with accurately recording the traits observed; therefore, it is common for some traits to be scored differently during different recording periods (Scott 1973). Therefore, dental wear may explain the inconsistent recordation of these two traits.

More common than low kappa values among the dental data are traits with kappa values indicating perfect agreement (k=1.000). In this sample, seven dental traits were found to have 100% intraobserver agreement. These traits included interruption grooves on the upper incisors, odontomes on the upper first and lower second premolars, the distal trigonid crest on the lower second molar, and cusp seven on the lower second and third molars. These high levels of agreement are mostly explained by the scales on which they are recorded and the rarity of the traits in question. Interruption grooves are recorded on a discrete, not an ordinal, scale. The trait is not scored by degree of expression, but rather location on the incisor, increasing the likelihood of identical scores on repeated observations. Odontomes and the Distal Trigonid Crest have a high rates of agreement, at least in part, for a similar reason. Both traits are recorded on the basis of presence or absence, making them easier to score consistently over repeated observations (Turner II et al. 1991). In addition, they are relatively rare traits; odontomes have a world frequency of 0.0-6.5% (Scott and Turner II 1997:191) and the distal trigonid crest has a world frequency of 0.0 18.7% (Scott and Turner II 1991:219). The rarity of these traits means that most individuals are given a score of zero, and those scores are easy to duplicate upon repeated observation (Scott 1973).

The final perfect agreement traits, cusp 7 on LM2 and LM3, are more difficult to explain. Cusp 7 on the first molar is rare outside of sub-Saharan African populations (5-10% of people; Scott and Turner II 1997); and rarer still on second and third molars (Scott personal communication 2017). Additionally, it is evaluated on a 5-point scale based on degree of expression, providing more opportunity for disagreement. Therefore, perfect intraobserver agreement is unexpected. The unusually high-level of intraobserver agreement seen in this study is likely a product of the rarity with which this trait was observed. On the lower second molar, cusp 7 exhibited any type of present expression in only 0.475% of the sample, and only 4.198% on the third molar. These frequencies are consistent with world frequencies reported by Scott and Turner II (1997), but those frequencies reflect only the prevalence of cusp 7 on the first molar. Though cusp 7 is expected to be common in African populations, the frequency of occurrence on second and third molars is much lower than on the first. As with odontomes and the distal trigonid crest, the rarity of this trait on these teeth helps account for the low degree of error as discussed above.

One of the major arguments in favor of the inclusion of dental morphology in the estimation of ancestry is that it has already been standardized and with minimal training can be consistently recorded (Edgar 2015). With few exceptions, the results seen here support those conclusions and the integration of dental morphological traits into forensic analyses.

Four recently developed trait scales were also tested in this research. With the exception of molar crenulations (k=0.503-0.909; Pilloud et al. 2017a), these traits have no previously reported rate of intraobserver agreement, but to have any forensic utility, must be demonstrated to be consistently recordable. The single new cranial trait, palate shape, exhibits the lowest intraobserver agreement of all the new traits (k=0.610). This is consistent with the overall pattern of intraobserver error between the cranial and dental traits, i.e., the cranial traits exhibit lower kappa scores than the dental traits. However, this number is relatively low even among cranial traits. This suggests that although this scale is an improvement on the purely subjective means of assessing palate shape, a certain degree of subjectivity remains when deciding between different grades on the scale, and further refinement may be possible.

Molar crenulations exhibited both the highest level of intraobserver agreement (LM2 crenulation, k=0.962) and the most variation in agreement (k=0.773-0.962) among the novel traits. The lowest agreement values for crenulations are on the third molars (LM3, k=0.773; UM3, k=0.778), which are still considered "substantial agreement" (Landis and Koch 1977). Third molars are often impacted or only partially erupted in some individuals (e.g. Carter and Worthington 2015, Farman 2007, Hattab and Abu Alhaijah 1999, Saker et al. 2008). Decisions regarding the appropriateness of scoring third molars in those cases are subject to the same inconsistencies described above for scoring dental morphology on worn teeth (Scott 1973). Excluding intraobserver scores for third molars, the level of agreement on molar crenulations is maintained at the "almost perfect" level (Landis and Koch 1977). These values are broadly consistent with those reported by Pilloud and colleagues (2017a).

Diastemata (k=0.853), maxillary dental crowding (k=0.926), and mandibular dental crowding (k=0.953) each exhibit "almost perfect" intraobserver agreement as defined by Landis and Koch (1977). Each of these traits are recorded on essentially a presence-absence scale, which makes them easily replicable. Importantly, for continued applications to the forensic context, each of the four new traits were recorded with "substantial agreement" or better (Landis and Koch 1977).

6.2 Associations with Ancestry

Based on chi-square analyses, 37 traits did not exhibit significant differences among ancestry groups. Of those 37, 36 non-significant traits were dental. Many of the non-significant traits are seen infrequently, even in groups where they are most common. Therefore, non-significant association with ancestry may be a product of the sample or the statistic used to measure significance, not indicative of a lack of relationship between a dental morphological feature and a population. Furthermore, the few Asian/Native American individuals observed were removed from analysis due to small sample size. Some of the non-significant traits, especially those associated with the Sinodont dental complex (e.g. double shoveling, interruption grooves, odontomes), may have exhibited significant associations with ancestry had the Asian/Native American sample been analyzed.

Of the novel traits, palate shape, maxillary dental crowding, and molar crenulations (on all recorded teeth) were significantly related to ancestry. Diastemata and mandibular dental crowding showed no significant differences among ancestry groups. The non-significance of mandibular crowding, may be associated with the lack of an Asian/Native American sample in these analyses. Though these individuals were not analyzed, they were observed, and the frequency distributions in Appendix A include the Asian/Native American sample. Examination of the table for mandibular crowding shows the

Asian/Native American sample exhibited a 100% frequency for mandibular crowding, compared to78.7% in Europeans, 76.5% in Hispanics, and 66.4% in Africans. These numbers align with the expected distribution of dental crowding due to the disparity between tooth and jaw size discussed in Chapter 4. However, when the Asian sample is removed, the remaining three frequencies are comparable. Similarly, diastemata are relatively evenly distributed among the groups analyzed. Diastema grade 1 is most common in individuals of African ancestry, and diastema grade 2 is slightly more common in Hispanic individuals, but the differences are minor and the overwhelming majority of individuals exhibit no diastemata.

Although the frequency distribution of palate shape (Appendix A) aligns closely with the classic 3-shape model, it also demonstrates a considerable amount of variation in palate shape that was not captured under the three-shape system. With one exception, all five grades of palate shape expression were observed in every ancestry group. A palate shape of grade 1 was not observed in any African individuals. This sample does not encompass the full range of human variation, and it is likely that some individuals of African ancestry would exhibit palates scored as grade 1. However, the classic understanding of palate shape would place African individuals at the higher end of the scale; therefore, lower frequencies of grades 1 and 2 are not surprising. In the European sample, over half (53.2%) of all individuals were scored as grade 3, which most closely approximates the classic parabolic arcade shape. When the two adjacent grades, 2 and 4, are considered as well, then the middle three grades of the scale account for 90.1% of all European palates. Similarly, the majority of African individuals (93.2%) are encompassed by grades 3-5, where grade 5 most closely resembles the hyperbolic shape usually associated with African ancestry.

The three-shape system for assessing palate shape is employed by ascribing the observation to one of three shapes, each of which is aligned with a particular ancestry group, i.e. Asian/Native American, European, or African. Individuals of Hispanic ancestry highlight one of the major flaws with this system, as they are not consistently thought of as members of any of the three ancestry groups used in the scale. Genetically, Hispanic populations are diverse and represent contributions from each of the ancestry groups represented by the old system to varying degrees depending on region of origin (e.g. Bertoni et al. 2003, Bonilla et al. 2005, Cerda-Flores et al. 2002, Lisker et al. 1986, 1990, 1996, Rangel-Villalobos 2008, Rubi-Castellanos et al. 2009). In these results, Hispanic individuals most frequently exhibit palate shapes of grade 2 (42.3%). If grades 1, 3, and 5 correspond to the elliptical, parabolic, and hyperbolic shapes, respectively, then the high frequency of grade 2 in the Hispanic sample suggests that the expression of palate shape in Hispanic populations was not being adequately evaluated with only three shapes, and justifies the addition of more shapes to the recordation of this trait.

Also somewhat surprising is the high frequency of grade 4 (44.6%) in the African sample. If the three-shape system was based on accurate associations between shape and ancestry, as previous research has suggested (e.g. Maier et al. 2015), then grade 5, which most closely resembles the hyperbolic shape, should have been most frequently expressed in African populations. Although grade 5 was present more often in the African sample (22.6%) than in the other samples, grade 4 was the most frequent expression. The high frequency of grade 4 has two potential explanations. First, it is possible that the original parabolic-hyperbolic-elliptical scheme, as proposed by Olivier (1969), and the current research are based on samples with different population histories. Although the effect of this potential difference cannot be known, it is reasonable to assume that different samples would express different proportions of the various levels of palate shape.

The second explanation seems more likely. The high frequency of grade 4 and lessened frequency of grade 5 can be attributed to differences in recordation scales. Under the three-shape paradigm, grades of 4 and 5 on the current scale would likely be classified as hyperbolic. In these results, those two grades together represent 67% of the African individuals observed, which aligns with previous characterizations of African palates as hyperbolic. However, when grades 4 and 5 are both available as potential scores, more precision is allowed in assessing palate shape. Under this system, scores that may previously have been classified as hyperbolic, but are not as extreme as the shape seen in grade 5, can be given their own category. Grade 5 palates exist and are observed more frequently in African populations. However, it represents an extreme of trait expression that is not seen as often as more moderate expressions. The differentiation between moderate and extreme forms of the "hyperbolic" shape cannot be captured by the three-shape system.

Whatever additional meanings may be imparted onto the frequency distribution of palate shape, the scale implemented in this research demonstrates a significant association with ancestry. Furthermore, clear definitions regarding different palate shapes and the elimination of subjective shape labels helps to mitigate the subjectivity inherent in observing this trait, and will allow its use to inform forensic estimations of ancestry.

The frequency distribution of maxillary dental crowding suggests it is most common in European individuals (52.8%) and least common in African individuals (26.6%). The Hispanic sample (29.6%) exhibits frequencies of maxillary dental crowding more comparable to the African sample, and the Asian/Native American sample, though not tested for associations, exhibits an intermediate frequency (40%). Dental crowding is caused by a disparity in tooth and jaw size (Lavelle and Foster 1969, Moorees and Reed 1954). It was hypothesized that Asian/Native American populations would exhibit the highest frequencies of crowding due to the disparity between intermediate tooth size (Pilloud et al. 2014, Schmidt 2008) and generally small cranial and jaw dimensions (Hanihara 1993, 1996, Howells 1989, Nicholson and Harvati 2006). This is not supported by these results. However, a high frequency of dental crowding in European populations is not surprising. Europeans, generally, trend toward small teeth on a global scale (e.g. Hanihara and Ishida 2005), and especially when compared to broadly defined African and Asian populations (e.g. Pilloud et al. 2014). Their cranial dimensions also tend to be smaller than many groups (Hanihara 1993, 1996, Howells 1989, Nicholson and Harvati 2006). Some research on secular changes in the human skull has suggested that cranial dimensions are changing faster than tooth size, and that this pattern is especially apparent in recent individuals of European ancestry (e.g. Heim 2013, Jonke et al. 2007, Lavelle 1973, Smith et al. 1986). The pattern of dental crowding observed here is consistent with those conclusions.

The association of dental crowding with the European sample highlights another interesting consideration. It has been argued that dental crowding is not useful in studies of population affinity because it is too affected by environmental variables like socioeconomic status to accurately reflect biological relationships (e.g. Corruccini et al. 1990, Hughes et al. 2001, Kawala et al. 2007, Townsend et al. 2009, von Cramon-Taubadel 2011). However, the results presented here appear to support the conclusions of Normando and colleagues (2013) that there is a considerable genetic component to the presence of dental crowding in an individual. The majority of the Hispanic sample (93.0%) represents migrant individuals. This population is typically of lower socioeconomic status, and has poor dental health (Birkby et al. 2008). Similarly, most of the African sample (88.2%) is derived from the Hamann-Todd and Robert J. Terry Collections, which were predominantly built from unclaimed individuals of presumed lower socio-economic status with more restricted access to dental care. Finally, 77.5% of the individuals in the European sample are curated in donated collections. In these collections, most individuals are of lower socioeconomic status; only individuals who donated themselves to the collections tend to be of middle class backgrounds (e.g. Komar and Grivas 2008, Wilson et al. 2007). Given the histories and demographics of the different collections, it is reasonable to assume that the individuals observed for this study had fairly comparable access to orthodontia. It should also be noted that many individuals (36.1%) in this sample are derived from older skeletal collections, such as the Terry and Hamann-Todd collections, and had died before orthodontic intervention was commonplace. Therefore, socioeconomic status and access to orthodontia should have a negligible effect on the expression of dental crowding, at least in this case, and the pattern observed can be attributed to differences among ancestry groups.

The presence of molar crenulations on any molar is significantly associated with ancestry ($p<0.0001$). In general, European individuals exhibit the highest frequencies of absence, or grade 0 crenulations (36.23%-87.29%). At the other extreme, individuals of African ancestry exhibit the highest frequencies of grade 2 crenulations (30.93%-68.33%). However, the distribution of different grades of crenulations varies by molar number. The third molar is the most variable in the expression of crenulations, as with other aspects of morphology and development (Butler 1939, Dahlberg 1945, Townsend et al. 2008). The general pattern of absence and extreme presence described above holds for the third molars, i.e. the highest frequency of grade 0 occurs in the European sample and the highest frequency of grade 2 occurs in the African sample. However, crenulations are more frequently present on third molars. For example, although grade 0 on third molars is most frequently expressed in the European sample, grade 1 rather than grade 0 is most common.

The second molar may be the most diagnostic for estimating ancestry using crenulations alone. For this tooth, grade 0 is most frequently observed in European individuals (86.2% maxillary, 80.6% mandibular), and is the most common score for those individuals. Furthermore, any expression of crenulations is most common in the African sample, with grade 2 practically absent in the non-African samples (< 6.0% in both jaws). The first molar exhibits a similar pattern to the second with one notable exception: absence of crenulations is most common for all groups except on the mandibular first molar in the African sample. The relative frequencies of presence are comparable to the second molar, meaning the African sample exhibits crenulations the most frequently and the European sample the least frequently, but the absolute presence of crenulations is least on the first molar.

The distribution of crenulations in the Hispanic sample is more difficult to characterize. Overall, the Hispanic sample expressed all grades of crenulations in more balanced proportions than either the African or European samples. There is a trend toward crenulation presence rather than absence, but it is very slight and dependent on which molar is being observed. In general, the frequency of crenulations on the first molar in Hispanic individuals is more similar to that of the European sample, while the frequency on the third molar is more similar to the African sample. Crenulations on the second molar are intermediate, falling between the frequencies for the European and African samples. A grade 0 score for crenulations is most often recorded on the first molars (66.7-77.3%) of either jaw for Hispanic individuals, and grade 2 score is most often recorded on the third molars (30.9-44.4%), a gradient in expression based on tooth position not seen as clearly in the non-Hispanic samples. Though not included in classificatory modeling, the small Asian/Native American sample shows an interesting distribution of crenulations. The overall trend is toward the presence of crenulations, either grade 1 or grade 2. The only teeth in which absence is more common than presence are the first and second molars of the maxillary dentition. However, the sample size of the Asian/ Native American population for this trait is ≤ 6 with an average sample size of 4 individuals; therefore, the patterns observed may not be characteristic of the larger population.

The trend of European individuals exhibiting absence of crenulations, and Hispanic and African individuals exhibiting some form of presence is consistent with the distribution of crenulations observed in other literature. The relatively few mentions of crenulations in human molars that exist suggest that this trait is most commonly observed in African (El-Najjar and McWilliams 1978, Herrick and Walsh-Haney 2010, Kiernberger 1955, Rhine 1990, 1993, Truesdell 2005) and Asian/Native American (Jacob 1967, Lasker and Lee 1957, Pedersen 1949) populations. These results support those assertions. The unique graduated pattern of the Hispanic molars may be tied to genetic information from both European and Native American populations being expressed in this population. Importantly these results confirm that molar crenulations are a trait of forensic interest, and that using the proposed scale allows for consistent recordation, encompassing the spectrum of existing variation.

6.3 Hypothesis 1-Cranial Morphoscopic Traits Perform Well in Ancestry Estimation

6.3.1 Comparison to Other Models

The mean accuracies of the models generated from cranial morphoscopic traits are generally comparable to those previously reported with respect to classification accuracy (Table 5.15). The naïve Bayes classifier generated from these data is an accurate classifier in 79.6% of cases, which is negligibly less accurate than the classifier reported by Hefner and Ousley (80.4%; 2014). This sample is imbalanced in favor of European individuals, which may increase the number of incorrect classifications returned with a Bayes classifier. The prior probability of an individual being European is higher than the other ancestry groups; therefore, it is more likely that individuals will be misclassified into this group since it is the most probable assignment. Even with the imbalanced sample, the classifier generated here performs comparably to one generated with a more balanced sample (Hefner and Ousley 2014), demonstrating the power of cranial morphoscopic traits in the estimation of ancestry.

The random forest model generated has an average correct classification rate of 75% (75.5% OOB, 75.2% Test set). Though an analysis of classification accuracy confirms that all models based on cranial traits alone classify correctly at a rate significantly greater than chance (p<0.00001), the accuracy rate of the models presented is about 10% less than that reported by Hefner and Ousley (2014). This difference may have more than one potential cause, one of which is the use of slightly different variables. This random forest model uses a newly standardized cranial variable, palate shape, and removes the malar tubercle since it was not significantly associated with ancestry. This change in variables could affect classification accuracy. However, palate shape demonstrated a strong association with ancestry, and, as discussed below, was found to be among the most important cranial classificatory variables; therefore, it's inclusion would not logically decrease model accuracy. The second, and more likely, explanation for the difference in accuracy between the two models is slightly different implementations of random forest modeling. The models in this study were generated using the *cforest()* command in the *party* package of R (Hothorn et al. 2006a, Strobl et al. 2007, Strobl et al. 2008). The primary difference between *cforest()* and many other implementations of random forest modeling, is the measure of variable importance. As variable importance measures are used at each node to select from the randomly generated pool of predictor variables, different measures can result in vastly different forests, and consequently different rates of accurate classification. As far as the author is aware, the specific method of building a random forest in Hefner and Ousley (2014) is not disclosed, but based on the reported measures of variable importance, mean decrease in accuracy and mean decrease in the Gini Index, in a related publication (Hefner et al. 2014), it does not seem likely that *cforest()* was used. The difference in forest construction likely accounts for the discrepancy between the two accuracy rates.

6.3.2 Variable Importance

The eleven cranial morphoscopic variables used in building these models are recorded on 2 (NO, PBD), 3(NAW, ANS, IOB, ZSS, SNS), 4(TPS), or 5(INA, PS, NBC) point scales. Previous studies (Deng et al. 2011, Hapfelmeier et al. 2014, Hothorn et al. 2006b, Jensen and Cohen 2000, Kass 1980, Shih 2004, White and Liu 1994) have found that variable importance measures that are based on a decrease in node impurity, such as the Gini Index, can be biased toward variables measured on more levels. Therefore, permutation importance is the most appropriate measure of variable importance for these data, and is the default variable importance measure in *cforest()*. The three most important cranial variables in this research are the inferior nasal aperture (INA), nasal aperture width (NAW), and the novel variable palate shape (PS). The three least important variables are post-bregmatic depression (PBD), nasal bone contour (NBC), and the supranasal suture (SNS). These traits are measured on 5, 3, 5, 2, 5, and 3-point scales respectively. As a point of comparison, the cranial morphoscopic traits that are both included here and reported as important by Hefner and colleagues (2014:6) are inferior nasal aperture (INA) and nasal bone contour (NBC), while the least important traits are nasal aperture width (NAW) and transverse palatine suture (TPS). The most striking differences are between nasal aperture width and nasal bone contour. The difference in the importance of these two variables between the two models may be a difference in recordation, or a difference in sample composition. However, given the different implementations of random forest modeling used, and the number of levels for each variable, 3 for nasal aperture width and 5 for nasal bone contour, the reversal of importance between models may be demonstrating the bias of decrease in impurity measures toward traits with more levels described above.

6.3.3 Accuracy

Although the model is slightly less accurate than those previously reported, the overall patterns of performance are very similar. As with Hefner and Ousley (2014), the RFM models are most often correct in classifying African individuals (76.5% OOB, 72.3% Test set), while Hispanic individuals are the least accurately classified (72.3% OOB, 70.8% Test set). This finding is slightly different for the Bayes classifier, with European individuals being most accurately classified (77.1%) instead of African individuals. This change in the pattern to the Bayes classifier is unsurprising since the European component of the sample has the highest prior probability of the groups included in modeling. The disparity between the highest rate of correct classification and the lowest is smaller for the RFM model (1.5-4.2%) than for the Bayes classifier (16.3%). Therefore, although the Bayes classifier is more accurate overall, the random forest model may be more generalizable, and performs more consistently across all ancestry groups assessed. The models based on cranial traits show a slight classification bias. When individuals are classified incorrectly, they are most often classified as European. This pattern is particularly pronounced for Hispanic individuals, with on average 17.5% of individuals classified as European being Hispanic. This bias may reflect a component of European ancestry in the Hispanic individuals analyzed in this sample, but is more likely an effect of the sample imbalance toward individuals of European ancestry. All the models classify based on the prevalence of traits in a given ancestry group. Since the European sample is larger, the variation expressed by a single individual is more likely to align with the variation in European sample, making a misclassification as European an expected outcome. In sum, results support hypothesis 1 that cranial morphoscopic traits can be used to effectively group individuals by ancestry.

6.4 Hypothesis 2- Dental Morphological Traits Perform Well in Ancestry Estimation.

6.4.1 Comparison to Other Models

Classification models based on dental crown morphology correctly classify individuals about 68% of the time (66.1% OOB; 67.4% Test set; 70.5% Bayes), a rate significantly different from chance ($p<0.00001$). These results are somewhat lower compared to previous attempts to use dental morphology in estimating ancestry (e.g. Edgar 2005, 2013, Scott et al 2016); however, the methods are not directly comparable. Edgar (2005) employs logistic regression and Bayesian classification and reports an accurate classification rate of 90%. This is considerably higher than found in this study, but the Edgar (2005) study only compares European and African American individuals. The models presented here also include Hispanic individuals. The presence of a third group increases the potential for an individual to be misclassified. Even if individuals were classified randomly, adding a third potential classification increases the chance of misclassification from 50% to 67%.

The later study by Edgar (2013) is more comparable in research design to this study. Using discriminant functions, Edgar (2013) reports accuracies ranging from 66.7-89.3% in differentiating European or African Americans from Hispanic Americans. Success in differentiating European and African Americans from each other ranged from 71.4-100% (Edgar 2013). Finally, two geographically distinct Hispanic populations, Southwest and Florida Hispanics, were differentiated from each other in 46.2-72.7% of cases. Edgar

(2013) examines two Hispanic samples, making the sample composition comparable to that of this study, but 3-way classification is not attempted; therefore, results cannot be directly compared. However, the results presented here are consistent with the results of both sets of equations including a Hispanic sample.

Perhaps the approach closest to that taken in this research is that of Scott and colleagues (2016). That presentation describes a method for assessing ancestry from dental morphology, including root traits, using naïve Bayes classification. In a comparison of three groups, Western Eurasia, Sino-American, and Sub-Saharan African, the program, called *rASUDAS,* correctly classified individuals in 78.53% of cases. This program still outperforms the models in this research, but the basis of the design is the most similar and the rate of accurate classification is the most comparable. The difference observed here is likely attributable to several differences in research design. First, the three group model presented by Scott and colleagues (2016) includes a sample of Sino-American, or Asian/Native American, individuals. That group was the most accurately classified (96.7%) by *rASUDAS*, which improves the classification accuracy overall. That population is absent in the analyses presented here, and the most closely related population, the Hispanic sample, has considerable influences from other populations, e.g. European and/or African populations (Bertoni et al. 2003, Bonilla et al. 2005, Cerda-Flores et al. 2002, Lisker et al. 1986, 1990, 1996, Rangel-Villalobos 2008, Rubi-Castellanos et al. 2009), and may not represent the Sino-American dental pattern. Interestingly, Hispanic individuals were the most accurately classified in the models used in this research, possibly supporting the classificatory power of Sinodont dental features observed by Scott and colleagues (2016). The second difference between the two studies is the traits included. The *rASUDAS* program incorporates data from root traits in addition to crown traits. Several of these root traits (e.g. 3-rooted lower first molar, Tome's root, 2-rooted lower canine) are indicative of ancestry, and information derived from those traits is not included in the current models. Finally, in its current form *rASUDAS*, is based on mean frequencies for each trait, and does not incorporate individual variation. Using mean frequencies creates an expected expression for each trait in each population. While this approach certainly reduces noise in the data and may lead to more accurate classifications, it does account for the degree of variability extant in human populations. This is a concern being addressed by further refinement of the *rASUDAS* program and the collection and analysis of individual dental morphological data (Pilloud et al. 2017b)

6.4.2 Variable Importance

Visual inspection of the variable importance plots for the dental morphology models reveal four roughly defined groups of traits (Figure 5.2). The most important traits are the distal accessory ridge on the mandibular canine (LC.DAR) and enamel extensions on the mandibular second molar (LM2.ENEXT). That group is closely followed in importance by crenulations on the maxillary first molar (UM1.CREN) and the distal accessory ridge on the maxillary canine (UC.DAR). The majority of the dental crown traits fall in the bottom two groups with respect to importance. The next most important group of traits includes crenulations on the mandibular first and third molars as well as the maxillary third molar (LM1.CREN, LM3.CREN, UM3.CREN), groove pattern on the mandibular second molar (LM2.GPATT), shoveling on the maxillary and mandibular lateral incisors (UI2.SHOV, LI2.SHOV), and enamel extensions on the maxillary first and mandibular second molars

(UM1.ENEXT, LM2.ENEXT). The final group of traits, which are significantly associated with ancestry but not powerful with respect to classification, includes lingual cusp number on the mandibular first (third) premolar (LP1.LCUSP), the mesial canine ridge or Bushman canine (BUSH_C), the hypocone of the maxillary second and third molars (UM2.HYPO, UM3.HYPO), the hypoconulid on the mandibular first molar (LM1.CUSP5), maxillary dental crowding (CROWDING_MAX), double shoveling on the maxillary lateral incisors (UI2.DSHOV), the deflecting wrinkle on the mandibular first molar (LM1.DWRIN), incisor winging (WING), enamel extensions on the mandibular third molar (LM3.ENEXT), and cusp number on the mandibular first molar (LM1.CUSPNO).

Although many of the observed dental morphological traits included are associated with geographically-based dental complexes, and are therefore expected to be indicative of ancestry, only one of the four most important traits, enamel extensions (Scott and Turner II 1997, Turner II 1990), meets that criterion. Molar crenulations have been suggested multiple times in the past to be indicative of ancestry (El-Najjar and McWilliams 1978, Herrick and Walsh-Haney 2010, Jacob 1967, Kiernberger 1955, Pedersen 1949, Rhine 1990, 1993, Truesdell 2005), but their standardized recordation is new to this study, and they are not included in previously defined dental complexes. However, the high placement of molar crenulations in terms of variable importance supports the utility of this trait in estimating ancestry.

Perhaps the most surprising finding with respect to variable importance is that of canine accessory ridges, both maxillary and mandibular. This trait is not associated with any geographically defined dental complex (Scott 1977, Scott and Turner II 1997). However, canine distal accessory ridges, at least on the mandibular canine, have been suggested to be correlated with tooth size; higher degrees of expression being more common in teeth with greater mesio-distal dimensions (Noss et al. 1983). Odontometric studies have suggested that African populations have among the largest teeth in the world, while European populations have some of the smallest, and Asian and Native American populations have teeth of intermediate size (e.g. Pilloud et al. 2014). Furthermore, in a comparison between Southwest American Indian and American White samples, the mandibular canine distal accessory ridge was seen more frequently in the Native American sample (Scott 1977). In these results distal accessory ridges were more pronounced in both African and Hispanic individuals. This trait has also been suggested to be among the more sexually dimorphic dental crown traits (Noss et al. 1983, Scott 1977), but there are no significant differences by sex for this trait in this sample (UC: X^2=15.55, df=10,p=0.1132; UC: X^2=12.954, df=10,p=0.2268). Therefore, the importance of the canine accessory ridge in these models, is likely highlighting the utility of tooth size in assessing ancestry, supporting previous conclusions (e.g. Pilloud et al. 2014).

Another interesting result is that many (8 of 9) of the traits in the fourth group described above, those least important to classification, are associated with geographically defined dental complexes. There are two likely reasons why these traits are important in defining a dental complex but not important in estimating ancestry. First, many of these traits (e.g. winging, shoveling, double shoveling, deflecting wrinkle, enamel extensions) are characteristic of the Sinodont dental complex, which is not well-represented in this sample. Individuals of Native American or Asian ancestry exhibit Sinodont dental features, and neither group is included in these analyses. Although there is a Hispanic sample, and a significant component of Hispanic ancestry is thought to be Native American (Bertoni et al.

2003), the dental morphology of Hispanic populations is not well understood, and based on the importance of the variables analyzed, does not seem to align well with the Sinodont dentition. The second reason for this disparity is simple prevalence of the traits. For example, the mesial canine ridge is strongly associated with African ancestry (Edgar 2009, Irish 1997, Scott and Turner II 1997). However, the trait itself is rare, occurring in only about 12-35% of individuals expected to exhibit the sub-Saharan African dental complex (Irish 2013 == in Scott and Irish edited volume). Though much more common in African individuals, it is so rarely present that it is not heavily weighted in classification.

6.4.3 Accuracy

In stark contrast to the models based on cranial traits, dental models most often correctly classify individuals as Hispanic, at least in two of three models (84.6% Test set; 87.5% Bayes). Although, the out-of-bag accuracy is highest for individuals of African ancestry (79.3%), it is closely followed by Hispanic individuals (77.8%) continuing a pattern of improved performance in classifying Hispanic individuals. This difference between the cranial and dental models likely reflects the influence of Sinodont dental traits on the Hispanic sample. In the work of Scott and colleagues (2016) the Sino-American sample was the most accurately classified, suggesting that Sinodont dental traits are useful in differentiating among groups. In this study, the only sample that would be expected to consistently exhibit Sinodont features is the Hispanic sample. The fact that it is the Hispanic sample that is most accurately classified, supports the classificatory potential of dental morphology, in particular the features associated with the Sinodont dentition. These results further highlight that the Hispanic dentition is not well understood.

The variable importance plots suggest that odontometric differences are driving the models; however, if that were the case, African individuals would be expected to have high rates of correct classification because of the higher prevalence of distal accessory ridges and large teeth. Although the dental models classify African individuals correctly, the degree to which Hispanic individuals are correctly classified suggests a greater influence of overall morphology than size-related traits in classification. Taken together, it is some combination of dental morphology and odontometric differences that seem to be driving the classification of these individuals based on the dentition.

The correct classification rate for the European sample is lowest across the board (60.00% OOB; 64.29% Test set; 69.53% Bayes). This result is similar to the pattern observed by Scott and colleagues (2016). The likely explanation for the poor performance of European individuals in both studies is the general lack of morphological features in the Eurodont dentition, and difficulty in modeling trait absence as important. Rarely is the positive expression of a trait indicative of European ancestry. Therefore, the presence of a so-called non-European trait (e.g. shoveling, cusp 7, enamel extensions), can cause that individual to be incorrectly classified into a group in which that trait is more common. This effect is further amplified when population means are used, as in *rASUDAS*. Models based on means are more likely to misclassify individuals whose morphology does not align with the most common expression of traits in a given population. In other words, European individuals with uncharacteristic positive expressions of traits are more likely to be misclassified in means based models that do not account for the possibility of that expression.

Like the models based on cranial data, individuals who were misclassified were most often misclassified as European; however, unlike the cranial models the rates of misclassification were more comparable between groups. That is, although European was the most common misclassification, it occurred only marginally more often than misclassification into a different group, and neither African nor Hispanic individuals were more frequently misclassified as European. For example, an average of 10.7% individuals classified as African were actually Hispanic, while 11.5% on average were actually European. This indicates that the models based on dental morphology, are not consistently misclassifying one ancestry group more than another.

Furthermore, there is a greater magnitude of difference between the highest and lowest correct classification rate in the random forests (19.3%-20.3%) than in the Bayes classifier (18.0%). This result runs counter to that observed in the cranial models, for which the random forests were more stable. The results for the dental models suggest that the Bayes classifier is more stable, in that the rate of accurate classification is more comparable across ancestry groups, possibly suggesting it is more appropriate in the analysis of dental morphology. Taken together, the various results derived from the dental morphological data support the effective use of dental morphology in the forensic context, and support the second hypothesis of this research.

6.5 Hypothesis 3- Combining Cranial Morphoscopic and Dental Morphological Traits Results in a Model that Outperforms Either Dataset Alone

6.5.1 Comparison to Other Models

Models based on the combination of cranial morphoscopic and dental morphological traits are on average accurate in 79% of cases (76.7% OOB; 77.1% Test set; 83.8% Bayes).

This combination of data sets has not been previously reported; therefore, there is no direct comparison appropriate to these models. The most similar study combined cranial morphoscopic traits with cranial measurements (Hefner et al. 2014). The techniques used in that study accurately classified individuals in 85.5-89.6% of cases when both data types were used (Hefner et al. 2014); a 4-10% improvement over the performance of cranial morphoscopic traits alone (Hefner et al. 2014, Hefner and Ousley 2014). Although the overall accuracy of the models presented here is lower, the magnitude of the differences between models is comparable with the combined models performing more accurately on average 2-11% of the time, as compared to the cranial variable models.

Similar to the models based on dental traits, the combined dataset models perform best when classifying Hispanic individuals. When classifying an individual as Hispanic, both the random forest and Bayesian classifier were correct in over 87% of cases (87.2% OOB; 100.0% Test set; 93.9% Bayes). In two of three cases, the models perform worst in classifying individuals of African ancestry (73.4% Test set; 81.0% Bayes), though still significantly better than chance ($p<0.00001$). In the remaining model, European individuals were least accurately classified (72.2% OOB). However, on average, the European sample is least often classified correctly (77.1%). These results are somewhat unusual because both datasets independently perform well at classifying individuals of African ancestry. Although European individuals are least accurately classified in the dental models, neither the cranial nor dental models performed poorly in classifying African individuals. The combined data models still perform well for African and European samples, with an average correct classification rate over 77%, but they are underperforming the classification of the Hispanic sample by roughly 15%. This discrepancy may highlight the value in combining dental and cranial data. In examining the accuracy rates across all models for the African sample, there is an increase in accuracy for the combined models compared to both the cranial and dental models, but the rates of accuracy are fairly consistent throughout. This pattern largely holds true for the European sample as well, although there is a considerable decrease in accuracy for the dental traits models. The relative consistency of accurate classification in European and African samples suggests concordance between the cranial and dental data. In other words, the cranial and dental data reveal similar information, so relatively little classificatory power is gained through their combination, at least in European and African samples. However, between the lowest and highest rate of classification for the Hispanic sample there is a 22% difference, with the models based on combined data performing best. This suggests that cranial and dental variables are capturing different information with respect to ancestry for Hispanic individuals, and their combination results in a more powerful classifier.

6.5.2 Variable Importance

The accuracy in classification for the models based only on cranial traits is significantly higher ($p<0.01$) than those based only on dental traits. Therefore, it is not surprising of the 34 traits used in the combined data models, only one dental trait, shoveling on the mandibular lateral incisor, is included in the top ten classificatory variables. However, it is interesting that the relative importance of many of the dental variables is different when combined with the cranial data. This is also true to a certain extent among the cranial variables, e.g. anterior nasal spine becomes more important than palate shape in the combined model. Of the 23 dental morphological variables included in these models, only three maintain their relative degree of importance from the dental variables only model. That is, when the dental variables in the combined model are ordered by importance, only three of them occupy the same position as in the dental variables only model. The majority of the remaining traits maintain the same level of importance relative to one another, but their overall place in the order is different when the cranial variables are included. On average, the magnitude of change in importance in the combined model is greater for dental variables. For example, the four biggest changes in dental variable importance can be seen in shoveling of the lateral mandibular incisor (LI2.SHOV), shoveling of the lateral maxillary incisor (UI2_SHOV), crenulations on the mandibular first molar (LM1.CREN), and lingual cusp number of the mandibular first (third) premolar (LP1.LCUSP). Both shoveling traits increased in importance by seven places, while crenulations decreased by seven places, and lingual cusp number decreased by nine places. On its own, shoveling is a relatively unimportant classificatory variable in the dental models, an effect almost certainly due to lack of Sinodont dentitions in this sample. However, when combined with information from the cranium, shoveling becomes one of the most powerful dental predictors of ancestry. The increased importance of incisor shoveling, coupled with the improved performance in correctly classifying Hispanic individuals, supports the hypothesis that the interaction of cranial and dental traits is key to correct ancestry estimations.

6.5.3 Accuracy

Finally, the results of these models tentatively support the third hypothesis of this research. Using both random forest modeling and naïve Bayes classification, the models based on combined data more accurately classify individuals than any models based only on one type of data. The improvement in accuracy ranges from about 1-13% depending on the models being compared. In all cases, the difference in accuracy between combined data models and dental data models was statistically significant ($p<0.001$) (Table 5.16). The magnitude of these differences is consistently within the range of 10-13%. However, the combined data and cranial data models were never statistically significantly different. Although there is a lack of significant difference between these two types of models, the combined data models consistently classified individuals more accurately than the cranial data models. The difference between the cranial models and combined data models was consistently in the range of 1-4%. The greatest degree of difference was observed in the naïve Bayes classifier, with the combined data model classifying 4.2% more of cases correctly. This improvement suggests a practical difference when data types are combined, at least for the Bayesian classifier, and the comparison between cranial and combined Bayes model is very nearly significant ($z=-1.6275$, $p=0.051869$). This result is not surprising given the superior performance of the Bayes classifier for all data types.

The differences in accuracy observed among models based on different data types are of practical, if not statistical, significance. Of the three data types, the dental morphology based models performed worst, and the sample used in this research lacked the population likely to be best classified using dental morphology, the Asian/Native American population. The cranial data perform well on their own, though not as well as in previous studies (e.g. Hefner and Ousley 2014, Hefner et al. 2014). This too may be in part due to the lack of an Asian/Native American sample. The models that combine the cranial and dental data perform best, and the magnitude of that improvement is expected to increase with a more balanced sample. The analysis of an Asian/Native American sample will likely improve the performance of the dental variables models, and to a lesser extent the cranial variables model. Improvement to both of those models should result in a greater improvement to the combined data model.

Chapter 7:
Conclusions

7.1 Performance of the 'Novel' Traits

The 'novel' traits performed well. Of the four traits explored, only diastemata did not demonstrate any utility in estimating ancestry. Given that the midline diastema has been consistently associated with African ancestry (Irish 1997, Scott and Turner II 1997), the lack of significant differences observed for this trait is surprising. Underperformance of the trait is likely indicative of the need for refinement in the recordation scale. Non-midline diastemata are not strongly associated with ancestry, and the addition of these data may have obscured any pattern inherent in the distribution of the midline diastema.

Dental crowding as an indicator of ancestry was moderately successful. These results offer some support to previous conclusions that there is a genetic component to dental crowding (Normando et al. 2013). The distribution of this trait suggests a pattern related to population and tooth/jaw size discrepancy, not one related to socioeconomic status, as might be expected. However, only crowding in the maxilla demonstrated any significant relationship to ancestry. Mandibular crowding may demonstrate a relationship with ancestry in the analysis of a more balanced and representative sample. Recording this trait on a scale more refined than presence or absence may serve to increase its utility. Adding variations in location and severity of crowding into the scale may serve to further separate different ancestry groups, and improve the classificatory power of this trait.

Palate shape and molar crenulations represent novel scales of recordation for traits already associated with ancestry. Both palate shape and crenulated molars are presented by Rhine (1990) as important ancestry traits, but neither has been explored in as much depth as the more well-established cranial and dental traits. The scale used to record palate shape in this research was previously presented by the author (Maier 2016) and is here tested using a more robust sample. More than other 'novel' traits in this research, the scale presented for palate shape most closely resembles the work presented by Hefner (2009), in that it is an attempt to impose a standard system of recordation on a traditionally subjectively assessed trait. The relatively low intraobserver error and high level of importance in classification for palate shape support the utility of this trait in ancestry estimation proposed by previous studies (e.g. Maier et al. 2015, Rhine 1990).

Like palate shape, molar crenulations have also been suggested as indicative of ancestry (e.g. El-Najjar and McWilliams 1978, Herrick and Walsh-Haney 2010, Kiernberger 1955, Rhine 1990, 1993, Truesdell 2005) without having been thoroughly defined or tested. Although some research has tested the association of crenulations with ancestry (e.g. Truesdell 2005), crenulations were not clearly defined, and group comparisons were somewhat limited. These problems were initially addressed by Pilloud and colleagues (2017), and the scale used in that research is further explored here. Consistent with those results, intraobserver error is low for the recordation of molar crenulations. Additionally, crenulations on at least one molar are highly important in both the dental variables and combined variables models, confirming their utility in ancestry estimates.

Two conclusions can be drawn from the success of the 'novel' traits. First, under-used sources of data, like the dentition, can provide forensic anthropologists with better means of estimating ancestry. Second, many of the traits found on the trait lists (e.g. Rhine 1990) are indicative of ancestry, and can be used successfully in the estimation of ancestry. Two of Rhine's (1990) traits that had been minimally examined, palate shape and molar crenulations, have been proven useful. This conclusion echoes that of Hughes and colleagues (2011), and demonstrates the importance of continued investigation of morphoscopic methods. As 'new' traits are defined and recorded and new methods of analysis continue to expand and improve, the ability of the forensic anthropologist to accurately estimate ancestry must also improve.

7.2 Support of the Hypotheses

Although the models using only cranial morphoscopic traits underperformed those previously published (e.g. Hefner 2007, Hefner et al. 2014, Hefner and Ousley 2014), the results support the continued use of morphoscopic traits in ancestry estimation. Potential reasons for the difference in performance, e.g. differing statistical approaches and sample compositions, have been discussed in the previous chapter. Two important conclusions can be drawn from the results of the cranial trait models. First, as discussed above, palate shape can be an important indicator of ancestry, and more cranial traits should be standardized for use in these types of analyses. The traits defined by Hefner (2009) are not an exhaustive list of the cranial morphoscopic traits that have been associated with ancestry in the course of physical anthropological study. The performance of palate shape in these models demonstrates the utility in exploring more, previously unstandardized, cranial traits.

Second, these results support the value of continuing morphoscopic trait analysis. Morphoscopic methods of ancestry estimation are less popular than metric methods because of the subjectivity inherent in their use. However, if properly standardized and analyzed, cranial morphoscopic traits can provide a wealth of information with respect to ancestry. Recently, there has been a small, but meaningful, push toward reinvigorating the use of cranial morphoscopic traits in ancestry estimation (e.g. Hefner 2009, Hefner et al. 2014, 2015, Maier 2016, Maier et al. 2015). This study furthers that movement. Morphoscopic traits are applicable in some circumstances in which metric methods are not (e.g. Gill 1998, Hefner 2007, Hefner et al. 2012, Rhine 1990). Therefore, ensuring that cranial morphoscopic methods meet the required legal standard provides the anthropologist a greater number of tools at his or her disposal when constructing a biological profile.

One of the most important conclusions is that dental morphology can be used to accurately estimate individual ancestry. Previous authors (e.g. Edgar 2005, 2013, Irish 2015, Scott et al. 2016) have suggested this utility, but none of those studies approached this question in the same way as this research. The relatively high rate of accurate classification supports the further inclusion of dental morphology in forensic estimations of ancestry. Although not as accurate as models based on cranial traits, the dental trait models performed significantly better than chance. These models would likely improve with the analysis of a more representative sample. Additionally, these models support the further exploration of non-standard dental traits in the estimation of ancestry. Molar crenulations appeared among the most important dental variables for estimating ancestry, although they are not currently among the standard dental morphological traits observed. The durability of teeth makes them an excellent ancestry indicator in many circumstances, and dental morphology can be reliably recorded and is largely standardized. The demonstration of their classificatory potential adds support to the use of dental morphology in estimating ancestry.

Finally, the principal hypothesis of this research, that the combination of cranial and dental data will perform better than either type of data alone, is tentatively supported. The models based on both cranial and dental traits significantly outperformed the models based only on dental traits. The combined models classify individuals more accurately than the cranial models, especially in classifying Hispanic individuals, but the magnitude of that improvement is not statistically significant. In classifying European and African individuals, the combined models are about 5% more accurate than the cranial models. However, the combined models increase accurate classification of Hispanic individuals by about 20% when compared to the cranial models. The increased performance of the combined models over the cranial models may be of practical, if not statistical, significance. One reason for the marked improvement in the Hispanic sample may be that the cranium and dentition are providing different information with respect to ancestry. Previous studies (e.g. Hefner et al 2014, Lease and Sciuli 2005) have explored the effects of combining morphological and metric data from a single source (i.e. the cranium or the dentition), and demonstrated increased accuracy in ancestry estimation. The performance of the combined variables models demonstrates that combining data from different sources improves upon the classificatory potential of either source alone. The simultaneous analysis of different datasets can employ combinations of traits and trait interactions that may be ignored in examining a single source, resulting in a more accurate ancestry estimate.

7.3 Future Directions

The use of cranial morphoscopic and dental morphological traits, independently and in combination with each other and other sources of data, in ancestry estimation offers a promising avenue of future research. However, there is room for improvement. The current research could be improved through the analysis of a sample more representative of the four main ancestry designations used by forensic anthropologists. As addressed several times, this sample lacked a significant Asian/Native American component. The Asian/Native American ancestry group represents an important component of the United States' population, and the variation represented by these individuals is not accounted for in the models presented here. The addition of Asian and Native American samples, and more data in general, would serve to improve the accuracy of the models and broaden their applicability. Furthermore, the populations analyzed are exclusively from the United States. It may be informative to collect data on so-called parent populations from continental Africa, Europe, and Asia as a comparative base to the variation expressed in the United States. Such a study may reveal new patterns of variation, or help explain the patterns seen in U.S. populations, especially in populations with complex population histories and documented gene flow, such as the U.S. Hispanic population.

Secondarily, the need for continued reappraisal of existing methods, and the exploration and adoption of new methods has been demonstrated. The list of traits with the potential to be helpful in ancestry estimation has not been exhausted, and new methods continue to be developed. The next step is to refine trait definitions and scales so they are maximally effective in the estimation of ancestry. Fine-tuning the methods available, and implementing new methods, will add to the repertoire of tools available to the forensic anthropologist when making an ancestry estimate.

Finally, ancestry estimation in the forensic context can continue to be improved through the addition of more datasets and the application of different statistical methods. It is only within the last decade that the same type of statistical rigor has been applied to the categorical and ordinal data represented by morphoscopic and morphological traits, as has previously been applied to metric data. Many methods of analysis exist and are appropriate to these types of data, but have largely been inaccessible to anthropologists unfamiliar with statistical software packages. As anthropologists become more versed in statistics and statistical software platforms, the number of 'new' methods applied to skeletal and dental data will only increase.

The subjectivity of ancestry estimation, especially from morphoscopic traits, continues to be a concern for forensic anthropologists. Even as methods are refined and standardized to mitigate subjectivity, the ultimate estimation of individual ancestry remains somewhat subjective. In many cases, the anthropologist must use his or her experience to synthesize an ancestry estimate from a combination of many data types (e.g. craniometrics, cranial morphoscopic traits, dental morphology, and dental metrics). This second source of subjectivity can be eliminated, or at least reduced, by employing methods that consider multiple data sources simultaneously. This approach is already being used successfully in the estimation of other aspects of the biological profile, like age (e.g. Boldsen et al. 2002, Garvin and Passalacqua 2012), and there is little reason a similar approach cannot be applied to the estimation of ancestry. The models presented here represent that approach, and have demonstrated the power of combining data sources in ancestry estimation.

Ancestry in forensic anthropology will always be among the most difficult components of the biological profile to estimate. Improving techniques for DNA extraction and analysis will aid in these estimations, but there will continue to be circumstances in which DNA analysis is not pragmatic. Unknown population histories, changing definitions of ancestry and race, and changing demographics serve to make this estimation more difficult, and more subjective, than other components of the biological profile. In response to this problem a call has been issued to ensure the accuracy, replicability, and reliability of methods in forensic science (*Daubert v. Merrell Dow Pharmaceuticals Inc*. 1993, NAS/NRC 2009). This research has met that call by providing statistically based models for the analysis of cranial morphoscopic and dental morphological data. The models presented here do not conclude the conversation on improving ancestry estimation methods, but rather add to the corpus of methods from which the anthropologist can choose when making an ancestry estimate. Further research in this vein, and continued critical evaluation of existing methods can only serve to make ancestry estimation, and by extension forensic anthropology, more robust.

Appendix A: Frequency Tables

Frequency distribution for Anterior Nasal Spine (ANS).

	European: n=296		**African: n=191**		**Asian/Nat.Am.: n=10**		**Hispanic: n=141**	
ANS	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
1	52	17.57	148	77.49	6	60	53	37.59
2	122	41.22	37	19.37	4	40	57	40.43
3	122	41.22	6	3.14	0	0	31	21.99

Frequency distribution for Inferior Nasal Aperture (INA).

	European: n=330		**African: n=194**		**Asian/Nat.Am.: n=10**		**Hispanic: n=124**	
INA	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
1	4	1.21	76	39.18	3	30	7	5.65
2	16	4.85	67	34.54	1	10	3	2.42
3	84	25.45	42	21.65	5	50	41	33.06
4	109	33.03	7	3.61	1	10	44	35.48
5	117	35.45	2	1.03	0	0	29	23.39

Frequency distribution for Interorbital Breadth (IOB).

	European: n=327		**African: n=194**		**Asian/Nat.Am.: n=10**		**Hispanic: n=153**	
IOB	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
1	143	43.73	9	4.64	1	10	63	41.18
2	160	48.93	101	52.06	7	70	72	47.06
3	24	7.34	84	43.3	2	20	18	11.76

Frequency distribution for Malar Tubercle (MT).

	European: n=327		**African: n=193**		**Asian/Nat.Am.: n=10**		**Hispanic: n=153**	
MT	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	136	41.59	82	42.49	0	0	42	27.45
1	136	41.59	67	34.72	5	50	61	39.87
2	43	13.15	30	15.54	3	30	36	23.53
3	12	3.67	14	7.25	2	20	14	9.15

Frequency distribution for Nasal Aperture Width (NAW).

	European: n=324		**African: n=193**		**Asian/Nat.Am.: n=10**		**Hispanic: n=153**	
NAW	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
1	170	52.47	8	4.15	2	20	38	24.84
2	130	40.12	59	30.57	6	60	98	64.05
3	24	7.41	126	65.28	2	20	17	11.11

Frequency distribution for Nasal Bone Contour (NBC).

	European: n=324		**African: n=161**		**Asian/Nat.Am.: n=10**		**Hispanic: n=147**	
NBC	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	15	4.63	68	42.24	1	10	3	2.04
1	60	18.52	33	20.5	0	0	25	17.01
2	30	9.26	8	4.97	2	20	9	6.12
3	124	38.27	1	0.62	2	20	61	41.5
4	95	29.32	51	31.68	5	50	49	33.33

Frequency distribution for Nasal Overgrowth (NO).

	European: n=314		**African: n=185**		**Asian/Nat.Am.: n=7**		**Hispanic: n=127**	
NO	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	275	87.58	172	92.97	6	85.71	65	51.18
1	39	12.42	13	7.03	1	14.29	62	48.82

Frequency distribution for Post-Bregmatic Depression (PBD).

	European: n=330		**African: n=193**		**Asian/Nat.Am.: n=10**		**Hispanic: n=152**	
PDB	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	318	96.36	161	83.42	8	80	144	94.74
1	12	3.64	32	16.58	2	20	8	5.26

Frequency distribution for Supranasal Suture (SNS).

	European: n=328		**African: n=194**		**Asian/Nat.Am.: n=10**		**Hispanic: n=152**	
SNS	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	167	50.91	145	74.74	4	40	84	55.26
1	46	14.02	4	2.06	0	0	7	4.61
2	115	35.06	45	23.2	6	60	61	40.13

Frequency distribution for Zygomaticomaxillary Suture Shape (ZSS).

	European: n=308		**African: n=180**		**Asian/Nat.Am.: n=8**		**Hispanic: n=148**	
ZSS	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	36	11.69	13	7.22	5	62.5	55	37.16
1	123	39.94	45	25	2	25	44	29.73
2	149	48.38	122	67.78	1	12.5	49	33.11

Frequency distribution for Transverse Palatine Suture Shape (TPS).

	European: n=313		**African: n=175**		**Asian/Nat.Am.: n=10**		**Hispanic: n=146**	
TPS	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	45	14.38	14	8	3	30	63	43.15
1	85	27.16	37	21.14	3	30	56	38.36
2	173	55.27	124	70.86	4	40	25	17.12
3	10	3.19	0	0	0	0	2	1.37

Frequency distribution for Palate Shape (PS).

	European: n=233		**African: n=177**		**Asian/Nat.Am.: n=10**		**Hispanic: n=142**	
PS	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
1	4	1.72	0	0	1	10	28	19.72
2	39	16.74	12	6.78	4	40	60	42.25
3	124	53.22	46	25.99	2	20	36	25.35
4	47	20.17	79	44.63	2	20	16	11.27
5	19	8.15	40	22.6	1	10	2	1.41

Frequency distribution for Winging (WING).

	European: n=137		**African: n=88**		**Asian/Nat.Am.: n=3**		**Hispanic: n=54**	
WING	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
1	10	7.3	6	6.82	1	33.33	18	33.33
2	0	0	0	0	0	0	1	1.85
3	126	91.97	82	93.18	2	66.67	35	64.81
4	1	0.73	0	0	0	0	0	0

Frequency distribution for Maxillary Central Incisor Shoveling (UI1_SHOV).

	European: n=131		**African: n=99**		**Asian/Nat.Am.: n=2**		**Hispanic: n=63**	
UI1.SHOV	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	85	64.89	53	53.54	0	0	12	19.05
1	31	23.66	32	32.32	1	50	14	22.22
2	11	8.4	13	13.13	0	0	16	25.4
3	3	2.29	1	1.01	1	50	13	20.63
4	1	0.76	0	0	0	0	2	3.17
5	0	0	0	0	0	0	6	9.52
6	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0

Frequency distribution for Maxillary Lateral Incisor Shoveling (UI2_SHOV).

	European: n=135		**African: n=127**		**Asian/Nat.Am.: n=4**		**Hispanic: n=69**	
UI2.SHOV	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	74	54.81	49	38.58	0	0	13	18.84
1	39	28.89	45	35.43	1	25	20	28.99
2	17	12.59	28	22.05	1	25	17	24.64
3	4	2.96	4	3.15	1	25	6	8.7
4	0	0	0	0	1	25	10	14.49
5	1	0.74	1	0.79	0	0	2	2.9
6	0	0	0	0	0	0	1	1.45
7	0	0	0	0	0	0	0	0

Frequency distribution for Maxillary Central Incisor Double-Shoveling (UI1.DSHOV).

	European: n=134		**African: n=95**		**Asian/Nat.Am.: n=3**		**Hispanic: n=64**	
UI1. DSHOV	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	133	99.25	95	100	3	100	57	89.06
1	1	0.75	0	0	0	0	4	6.25
2	0	0	0	0	0	0	2	3.13
3	0	0	0	0	0	0	1	1.56
4	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0

Frequency distribution for Maxillary Lateral Incisor Double-Shoveling (UI2.DSHOV).

	European: n=142		**African: n=112**		**Asian/Nat.Am.: n=4**		**Hispanic: n=70**	
UI2. DSHOV	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	141	99.3	110	98.21	4	100	58	82.86
1	1	0.7	2	1.79	0	0	8	11.43
2	0	0	0	0	0	0	3	4.29
3	0	0	0	0	0	0	1	1.43
4	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0

Frequency distribution for Maxillary Central Incisor Interruption Groove (UI1.IG).

	European: n=140		**African: n=105**		**Asian/Nat.Am.: n=3**		**Hispanic: n=65**	
UI1.IG	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
O	140	100	105	100	3	100	63	96.92
M	0	0	0	0	0	0	0	0
MD	0	0	0	0	0	0	0	0
Med	0	0	0	0	0	0	1	1.54
D	0	0	0	0	0	0	1	1.54

Frequency distribution for Maxillary Lateral Incisor Interruption Groove (UI2.IG).

	European: n=142		**African: n=129**		**Asian/Nat.Am.: n=4**		**Hispanic: n=73**	
UI2.IG	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
O	129	90.85	117	90.7	3	75	51	69.86
M	1	0.7	3	2.33	0	0	3	4.11
MD	0	0	0	0	0	0	0	0
Med	5	3.52	3	2.33	0	0	5	6.85
D	7	4.93	6	4.65	1	25	14	19.18

Frequency distribution for Maxillary Central Incisor *Tuberculum Dentale* (UI1.TD).

	European: n=138		**African: n=102**		**Asian/Nat.Am.: n=3**		**Hispanic: n=64**	
UI1.TD	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	111	80.43	78	76.47	2	66.67	42	65.63
1	17	12.32	14	13.73	0	0	15	23.44
2	7	5.07	8	7.84	1	33.33	5	7.81

3	2	1.45	2	1.96	0	0	2	3.13
4	1	0.72	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0

Frequency distribution for Maxillary Lateral Incisor *Tuberculum Dentale* (UI2.TD).

	European: n=142		**African: n=129**		**Asian/Nat.Am.: n=4**		**Hispanic: n=71**	
UI2.TD	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	78	54.93	75	58.14	2	50	23	32.39
1	36	25.35	24	18.6	1	25	26	36.62
2	12	8.45	18	13.95	0	0	11	15.49
3	14	9.86	9	6.98	1	25	9	12.68
4	2	1.41	2	1.55	0	0	0	0
5	0	0	0	0	0	0	0	0
6	0	0	1	0.78	0	0	2	2.82

Frequency distribution for Maxillary Canine *Tuberculum Dentale* (UC.TD).

	European: n=180		**African: n=162**		**Asian/Nat.Am.: n=5**		**Hispanic: n=92**	
UC.TD	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	116	64.44	69	42.59	2	40	49	53.26
1	36	20	35	21.6	2	40	23	25
2	15	8.33	27	16.67	1	20	15	16.3
3	9	5	12	7.41	0	0	2	2.17
4	4	2.22	12	7.41	0	0	2	2.17
5	0	0	3	1.85	0	0	1	1.09
6	0	0	4	2.47	0	0	0	0

Frequency distribution for Maxillary Canine Distal Accessory Ridge (UC.DAR).

	European: n=167		**African: n=155**		**Asian/Nat.Am.: n=5**		**Hispanic: n=89**	
UC.DAR	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	132	79.04	53	34.19	2	40	39	43.82
1	24	14.37	39	25.16	1	20	16	17.98
2	7	4.19	29	18.71	2	40	16	17.98
3	4	2.4	12	7.74	0	0	9	10.11
4	0	0	14	9.03	0	0	4	4.49
5	0	0	8	5.16	0	0	5	5.62

Frequency distribution for Maxillary Mesial Canine Ridge/Bushman Canine (BUSH_C).

	European: n=172		African: n=158		Asian/Nat.Am.: n=5		Hispanic: n=91	
BUSH.C	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	164	95.35	112	70.89	5	100	83	91.21
1	4	2.33	24	15.19	0	0	7	7.69
2	4	2.33	14	8.86	0	0	1	1.1
3	0	0	8	5.06	0	0	0	0

Frequency distribution for Maxillary Second (Fourth) Premolar Accessory Cusps (UP2.ACUSP).

	European: n=130		African: n=153		Asian/Nat.Am.: n=7		Hispanic: n=95	
UP2. ACUSP	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	124	95.38	141	92.16	6	85.71	87	91.58
1	6	4.62	12	7.84	1	14.29	8	8.42

Frequency distribution for Maxillary First (Third) Premolar Accessory Cusps (UP1.ACUSP).

	European: n=147		African: n=160		Asian/Nat.Am.: n=7		Hispanic: n=94	
UP1. ACUSP	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	144	97.96	155	96.88	5	71.43	85	90.43
1	3	2.04	5	3.13	2	28.57	9	9.57

Frequency distribution for Maxillary Second (Fourth) Premolar Odontome (UP2.ODONT).

	European: n=112		African: n=149		Asian/Nat.Am.: n=7		Hispanic: n=95	
UP2. ODONT	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	112	100	149	100	7	100	95	100
1	0	0	0	0	0	0	0	0

Frequency distribution for Maxillary First (Third) Premolar Odontome (UP1.ODONT).

	European: n=127		African: n=154		Asian/Nat.Am.: n=7		Hispanic: n=92	
UP1. ODONT	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	127	100	153	99.35	7	100	92	100
1	0	0	1	0.65	0	0	0	0

Frequency distribution for Maxillary Third Molar Hypocone (UM3.HYPO).

	European: n=70		African: n=125		Asian/Nat.Am.: n=6		Hispanic: n=83	
UM3. HYPO	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	17	24.29	12	9.6	0	0	30	36.14
1	18	25.71	28	22.4	3	50	15	18.07
2	18	25.71	32	25.6	0	0	24	28.92
3	10	14.29	36	28.8	3	50	11	13.25
4	6	8.57	12	9.6	0	0	3	3.61
5	1	1.43	5	4	0	0	0	0

Frequency distribution for Maxillary Second Molar Hypocone (UM2.HYPO).

	European: n=153		African: n=153		Asian/Nat.Am.: n=8		Hispanic: n=115	
UM2. HYPO	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	6	3.92	0	0	0	0	1	0.87
1	22	14.38	4	2.61	0	0	9	7.83
2	27	17.65	25	16.34	2	25	30	26.09
3	35	22.88	50	32.68	1	12.5	35	30.43
4	49	32.03	41	26.8	3	37.5	36	31.3
5	14	9.15	33	21.57	2	25	4	3.48

Frequency distribution for Maxillary First Molar Hypocone (UM1.HYPO).

	European: n=145		African: n=144		Asian/Nat.Am.: n=7		Hispanic: n=120	
UM1. HYPO	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	1	0.69	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0
2	1	0.69	0	0	0	0	1	0.83

3	4	2.76	6	4.17	0	0	0	0
4	47	32.41	29	20.14	2	28.57	44	36.67
5	92	63.45	109	75.69	5	71.43	75	62.5

Frequency distribution for Maxillary First Molar Carabelli's Trait (UM1.CARA).

	European: n=148		**African: n=142**		**Asian/Nat.Am.: n=7**		**Hispanic: n=120**	
UM1. CARA	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	67	45.27	59	41.55	4	57.14	58	48.33
1	11	7.43	18	12.68	1	14.29	9	7.5
2	13	8.78	16	11.27	1	14.29	23	19.17
3	22	14.86	20	14.08	1	14.29	13	10.83
4	12	8.11	18	12.68	0	0	5	4.17
5	18	12.16	6	4.23	0	0	11	9.17
6	3	2.03	3	2.11	0	0	1	0.83
7	2	1.35	2	1.41	0	0	0	0

Frequency distribution for Maxillary Second Molar Paramolar Tubercle (UM2.PARA).

	European: n=151		**African: n=156**		**Asian/Nat.Am.: n=8**		**Hispanic: n=119**	
UM2. PARA	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	144	95.36	152	97.44	8	100	117	98.32
1	3	1.99	1	0.64	0	0	1	0.84
2	1	0.66	1	0.64	0	0	0	0
3	1	0.66	0	0	0	0	0	0
4	0	0	1	0.64	0	0	1	0.84
5	2	1.32	0	0	0	0	0	0
6	0	0	1	0.64	0	0	0	0

Frequency distribution for Maxillary First Molar Parastyle (UM1.PARA).

	European: n=145		**African: n=146**		**Asian/Nat.Am.: n=8**		**Hispanic: n=112**	
UM1. PARA	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	144	99.31	144	98.63	8	100	109	97.32
1	1	0.69	2	1.37	0	0	3	2.68
2	0	0	0	0	0	0	0	0

3	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0

Frequency distribution for Maxillary Third Molar Enamel Extensions (UM3.ENEXT).

	European: n=76		**African: n=129**		**Asian/Nat.Am.: n=6**		**Hispanic: n=78**	
UM3. ENEXT	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	72	94.74	124	96.12	1	16.67	53	67.95
1	4	5.26	5	3.88	2	33.33	21	26.92
2	0	0	0	0	3	50	4	5.13
3	0	0	0	0	0	0	0	0

Frequency distribution for Maxillary Second Molar Enamel Extensions (UM2.ENEXT).

	European: n=175		**African: n=159**		**Asian/Nat.Am.: n=8**		**Hispanic: n=124**	
UM2. ENEXT	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	145	82.86	149	93.71	2	25	58	46.77
1	23	13.14	7	4.4	6	75	41	33.06
2	6	3.43	3	1.89	0	0	23	18.55
3	1	0.57	0	0	0	0	2	1.61

Frequency distribution for Maxillary First Molar Enamel Extensions (UM1.ENEXT).

	European: n=166		**African: n=151**		**Asian/Nat.Am.: n=8**		**Hispanic: n=121**	
UM1. ENEXT	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	157	94.58	142	94.04	3	37.5	94	77.69
1	7	4.22	7	4.64	4	50	11	9.09
2	2	1.2	2	1.32	0	0	16	13.22
3	0	0	0	0	1	12.5	0	0

Frequency distribution for Maxillary Third Molar Crenulations (UM3.CREN).

	European: n=65		**African: n=124**		**Asian/Nat.Am.: n=5**		**Hispanic: n=81**	
UM3. CREN	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	25	38.46	16	12.9	1	20	21	25.93
1	33	50.77	38	30.65	2	40	35	43.21
2	7	10.77	70	56.45	2	40	25	30.86

Frequency distribution for Maxillary Second Molar Crenulations (UM2.CREN).

	European: n=138		**African: n=148**		**Asian/Nat.Am.: n=6**		**Hispanic: n=111**	
UM2. CREN	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	119	86.23	53	35.81	3	50	71	63.96
1	19	13.77	64	43.24	2	33.33	37	33.33
2	0	0	31	20.95	1	16.67	3	2.7

Frequency distribution for Maxillary First Molar Crenulations (UM1.CREN).

	European: n=118		**African: n=131**		**Asian/Nat.Am.: n=6**		**Hispanic: n=110**	
UM1. CREN	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	103	87.29	64	48.85	4	66.67	85	77.27
1	13	11.02	47	35.88	1	16.67	22	20
2	2	1.69	20	15.27	1	16.67	3	2.73

Frequency distribution for Maxillary Dental Crowding (CROWD_MAX).

	European: n=142		**African: n=139**		**Asian/Nat.Am.: n=5**		**Hispanic: n=98**	
CROWDING_ MAX	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	67	47.18	102	73.38	3	60	69	70.41
1	75	52.82	37	26.62	2	40	29	29.59

Frequency distribution for Diastema (DIASTEMA).

	European: n=131		**African: n=77**		**Asian/Nat.Am.: n=5**		**Hispanic: n=72**	
DIASTEMA	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	100	76.34	53	68.83	5	100	51	70.83
1	16	12.21	14	18.18	0	0	10	13.89

2	15	11.45	10	12.99	0	0	11	15.28

Frequency distribution for Lower Central Incisor Shoveling (LI1.SHOV).

	European: n=164		**African: n=142**		**Asian/Nat.Am.: n=4**		**Hispanic: n=60**	
LI1. SHOV	**n**	%	**n**	%	**n**	%	**n**	%
0	164	100	141	99.3	3	75	39	65
1	0	0	1	0.7	0	0	14	23.33
2	0	0	0	0	1	25	6	10
3	0	0	0	0	0	0	1	1.67
4	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0

Frequency distribution for Lower Lateral Incisor Shoveling (LI2.SHOV).

	European: n=179		**African: n=161**		**Asian/Nat.Am.: n=4**		**Hispanic: n=72**	
LI2. SHOV	**n**	%	**n**	%	**n**	%	**n**	%
0	179	100	158	98.14	3	75	48	66.67
1	0	0	3	2.11	0	0	19	31.67
2	0	0	0	0	1	25	4	6.67
3	0	0	0	0	0	0	1	1.67
4	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0

Frequency distribution for Lower Canine Distal Accessory Ridge (LC.DAR).

	European: n=189		**African: n=161**		**Asian/Nat.Am.: n=3**		**Hispanic: n=65**	
LC.DAR	**n**	%	**n**	%	**n**	%	**n**	%
0	158	83.6	63	39.13	1	33.33	32	49.23
1	26	13.76	50	31.06	2	66.67	28	43.08
2	4	2.12	29	18.01	0	0	4	6.15
3	1	0.53	9	5.59	0	0	1	1.54
4	0	0	4	2.48	0	0	0	0
5	0	0	6	3.73	0	0	0	0

Frequency distribution for Lower First (Third) Premolar Lingual Cusp Number (LP1.LCUSP).

	European: n=201		**African: n=171**		**Asian/Nat.Am.: n=5**		**Hispanic: n=116**	
LP1. LCUSP	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
A	0	0	0	0	0	0	0	0
0	135	67.16	71	41.52	3	60	50	43.1
1	15	7.46	21	12.28	0	0	15	12.93
2	18	8.96	38	22.22	1	20	18	15.52
3	16	7.96	23	13.45	0	0	16	13.79
4	6	2.99	8	4.68	1	20	6	5.17
5	7	3.48	5	2.92	0	0	7	6.03
6	1	0.5	0	0	0	0	1	0.86
7	0	0	0	0	0	0	0	0
8	3	1.49	5	2.92	0	0	3	2.59
9	0	0	0	0	0	0	0	0

Frequency distribution for Lower Second (Fourth) Premolar Lingual Cusp Number (LP2.LCUSP).

	European: n=167		**African: n=157**		**Asian/Nat.Am.: n=4**		**Hispanic: n=83**	
LP2. LCUSP	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
A	0	0	0	0	0	0	0	0
0	65	38.92	51	32.48	1	25	39	46.99
1	23	13.77	17	10.83	0	0	14	16.87
2	34	20.36	50	31.85	1	25	14	16.87
3	27	16.17	22	14.01	2	50	13	15.66
4	14	8.38	14	8.92	0	0	0	0
5	3	1.8	0	0	0	0	1	1.2
6	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0
8	1	0.6	2	1.27	0	0	1	1.2
9	0	0	1	0.64	0	0	1	1.2

Frequency distribution for Lower First (Third) Premolar Odontome (LP1.ODONT).

	European: n=179		**African: n=166**		**Asian/Nat.Am.: n=5**		**Hispanic: n=85**	
LP1. ODONT	**n**	%	**n**	%	**n**	%	**n**	%
0	179	100	166	100	5	100	85	100
1	0	0	0	0	0	0	0	0

Frequency distribution for Lower Second (Fourth) Premolar Odontome (LP2.ODONT).

	European : n=160		**African: n=158**		**Asian/Nat.Am.: n=4**		**Hispanic: n=79**	
LP2. ODONT	**n**	%	**n**	%	**n**	%	**n**	%
0	160	100	158	100	4	100	78	98.73
1	0	0	0	0	0	0	1	1.27

Frequency distribution for Lower Third Molar Groove Pattern (LM3.GPATT).

	European: n=70		**African: n=122**		**Asian/Nat.Am.: n=5**		**Hispanic: n=62**	
LM3. GPATT	**n**	%	**n**	%	**n**	%	**n**	%
X	54	77.14	104	85.25	5	100	55	88.71
+	2	2.86	1	0.82	0	0	0	0
Y	14	20	17	13.93	0	0	7	11.29

Frequency distribution for Lower Second Molar Groove Pattern (LM2.GPATT).

	European: n=84		**African: n=127**		**Asian/Nat.Am.: n=3**		**Hispanic: n=76**	
LM2. GPATT	**n**	%	**n**	%	**n**	%	**n**	%
X	60	71.43	59	46.46	3	100	52	68.42
+	13	15.48	8	6.3	0	0	16	21.05
Y	11	13.1	60	47.24	0	0	8	10.53

Frequency distribution for Lower First Molar Groove Pattern (LM1.GPATT).

	European: n=60		**African: n=101**		**Asian/Nat.Am.: n=2**		**Hispanic: n=67**	
LM1. GPATT	**n**	%	**n**	%	**n**	%	**n**	%
X	3	5	2	1.98	0	0	6	8.96
+	0	0	0	0	0	0	0	0

Y	57	95	99	98.02	2	100	61	91.04

Frequency distribution for Lower Third Molar Cusp Number (LM3.CUSPNO).

	European: n=72		**African: n=120**		**Asian/Nat.Am.: n=5**		**Hispanic: n=64**	
LM3. CUSPNO	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
4	30	41.67	38	31.67	2	40	23	35.94
5	29	40.28	53	44.17	2	40	32	50
6	13	18.06	29	24.17	1	20	9	14.06

Frequency distribution for Lower Second Molar Cusp Number (LM2.CUSPNO).

	European: n=121		**African: n=132**		**Asian/Nat.Am.: n=3**		**Hispanic: n=79**	
LM2. CUSPNO	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
4	107	88.43	97	73.48	3	100	57	72.15
5	13	10.74	29	21.97	0	0	20	25.32
6	1	0.83	6	4.55	0	0	2	2.53

Frequency distribution for Lower First Molar Cusp Number (LM1.CUSPNO).

	European: n=84		**African: n=100**		**Asian/Nat.Am.: n=3**		**Hispanic: n=67**	
LM1. CUSPNO	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
4	16	19.05	1	1	0	0	1	1.49
5	63	75	93	93	3	100	57	85.07
6	5	5.95	6	6	0	0	9	13.43

Frequency distribution for Lower First Molar Deflecting Wrinkle (LM1.DWRIN).

	European: n=62		**African: n=94**		**Asian/Nat.Am.: n=1**		**Hispanic: n=54**	
LM1. DWRIN	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	53	85.48	72	76.6	1	100	26	48.15
1	4	6.45	13	13.83	0	0	10	18.52
2	5	8.06	6	6.38	0	0	16	29.63
3	0	0	3	3.19	0	0	2	3.7

Frequency distribution for Lower First Molar Anterior Fovea (LM1.ANTFOV).

	European: n=68		**African: n=100**		**Asian/Nat.Am.: n=2**		**Hispanic: n=68**	
LM1. ANTFOV	**n**	**%**	**n**	**%**	**n**	**%**	**n**	%
0	30	44.12	22	22	0	0	34	50
1	23	33.82	41	41	2	100	23	33.82
2	11	16.18	25	25	0	0	11	16.18
3	4	5.88	12	12	0	0	0	0

Frequency distribution for Lower Third Molar Distal Trigonid Crest (LM3.DTC).

	European: n=72		**African: n=120**		**Asian/Nat.Am.: n=5**		**Hispanic: n=61**	
LM3.DTC	**n**	**%**	**n**	**%**	**n**	**%**	**n**	%
0	69	95.83	119	99.17	5	100	60	98.36
1	3	4.17	1	0.83	0	0	1	1.64

Frequency distribution for Lower Second Molar Distal Trigonid Crest (LM2.DTC).

	European: n=114		**African: n=129**		**Asian/Nat.Am.: n=3**		**Hispanic: n=77**	
LM2.DTC	**n**	**%**	**n**	**%**	**n**	**%**	**n**	%
0	113	99.12	129	100	3	100	75	97.4
1	1	0.88	0	0	0	0	2	2.6

Frequency distribution for Lower First Molar Distal Trigonid Crest (LM1.DTC).

	European: n=77		**African: n=96**		**Asian/Nat.Am.: n=1**		**Hispanic: n=62**	
LM1.DTC	**n**	**%**	**n**	**%**	**n**	**%**	**n**	%
0	77	100	94	97.92	1	100	58	93.55
1	0	0	2	2.08	0	0	4	6.45

Frequency distribution for Lower Third Molar Hypoconulid (LM3.CUSP5).

	European: n=73		**African: n=119**		**Asian/Nat.Am.: n=5**		**Hispanic: n=64**	
LM3. CUSP 5	**n**	**%**	**n**	**%**	**n**	**%**	**n**	%
0	31	42.47	35	29.41	2	40	25	39.06
1	0	0	3	2.52	0	0		0
2	3	4.11	11	9.24	0	0	12	18.75
3	12	16.44	26	21.85	0	0	14	21.88

4	17	23.29	21	17.65	2	40	6	9.38
5	10	13.7	23	19.33	1	20	7	10.94

Frequency distribution for Lower Second Molar Hypoconulid (LM2.CUSP5).

	European: n=120		**African: n=132**		**Asian/Nat.Am.: n=3**		**Hispanic: n=80**	
LM2. CUSP5	**n**	**%**	**n**	**%**	**n**	**%**	**n**	%
0	106	88.33	94	71.21	3	100	55	68.75
1	1	0.83	1	0.76	0	0	1	1.25
2	3	2.5	6	4.55	0	0	5	6.25
3	4	3.33	9	6.82	0	0	13	16.25
4	4	3.33	19	14.39	0	0	6	7.5
5	2	1.67	3	2.27	0	0	0	0

Frequency distribution for Lower First Molar Hypoconulid (LM1.CUSP5).

	European: n=83		**African: n=100**		**Asian/Nat.Am.: n=3**		**Hispanic: n=68**	
LM1. CUSP5	**n**	**%**	**n**	**%**	**n**	**%**	**n**	%
0	15	18.07	1	1	0	0	1	1.47
1	2	2.41	1	1	0	0	0	0
2	6	7.23	3	3	0	0	2	2.94
3	16	19.28	10	10	1	33.33	10	14.71
4	25	30.12	33	33	1	33.33	31	45.59
5	19	22.89	52	52	1	33.33	24	35.29

Frequency distribution for Lower Third Molar Cusp 6 (LM3.CUSP6).

	European: n=72		**African: n=120**		**Asian/Nat.Am.: n=5**		**Hispanic: n=64**	
LM3. CUSP6	**n**	**%**	**n**	**%**	**n**	**%**	**n**	%
0	59	81.94	91	75.83	4	80	54	84.38
1	2	2.78	4	3.33	0	0	2	3.13
2	6	8.33	18	15	0	0	2	3.13
3	3	4.17	5	4.17	0	0	2	3.13
4	2	2.78	2	1.67	1	20	3	4.69
5	0	0	0	0	0	0	1	1.56

Frequency distribution for Lower Second Molar Cusp 6 (LM2.CUSP6).

	European: n=120		**African: n=132**		**Asian/Nat.Am.: n=3**		**Hispanic: n=78**	
LM2. CUSP6	**n**	%	**n**	%	**n**	%	**n**	%
0	118	98.33	126	95.45	3	100	76	97.44
1	1	0.83	1	0.76	0	0	0	0
2	1	0.83	3	2.27	0	0	1	1.28
3	0	0	2	1.52	0	0	1	1.28
4	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0

Frequency distribution for Lower First Molar Cusp 6 (LM1.CUSP6).

	European: n=83		**African: n=99**		**Asian/Nat.Am.: n=3**		**Hispanic: n=68**	
LM1. CUSP6	**n**	%	**n**	%	**n**	%	**n**	%
0	77	92.77	93	93.94	3	100	59	86.76
1	3	3.61	1	1.01	0	0	3	4.41
2	1	1.2	3	3.03	0	0	4	5.88
3	2	2.41	1	1.01	0	0	1	1.47
4	0	0	1	1.01	0	0	1	1.47
5	0	0	0	0	0	0	0	0

Frequency distribution for Lower Third Molar Cusp 7 (LM3.CUSP7).

	European: n=72		**African: n=121**		**Asian/Nat.Am.: n=5**		**Hispanic: n=64**	
LM3. CUSP7	**n**	%	**n**	%	**n**	%	**n**	%
0	72	100	110	90.91	5	100	64	100
1	0	0	1	0.83	0	0	0	0
1A	0	0	0	0	0	0	0	0
2	0	0	4	3.31	0	0	0	0
3	0	0	3	2.48	0	0	0	0
4	0	0	3	2.48	0	0	0	0

Frequency distribution for Lower Second Molar Cusp 7 (LM2.CUSP7).

	European: n=120		**African: n=132**		**Asian/Nat.Am.: n=3**		**Hispanic: n=79**	
LM2. CUSP7	**n**	**%**	**n**	**%**	**n**	**%**	**n**	%
0	120	100	129	97.73	3	100	79	100
1	0	0	2	1.52	0	0	0	0
1A	0	0	0	0	0	0	0	0
2	0	0	1	0.76	0	0	0	0
3	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0

Frequency distribution for Lower First Molar Cusp 7 (LM1.CUSP7).

	European: n=85		**African: n=100**		**Asian/Nat.Am.: n=3**		**Hispanic: n=69**	
LM1. CUSP7	**n**	**%**	**n**	**%**	**n**	**%**	**n**	%
0	83	97.65	85	85	3	100	67	97.1
1	0	0	2	2	0	0	0	0
1A	0	0	0	0	0	0	0	0
2	0	0	9	9	0	0	0	0
3	1	1.18	4	4	0	0	1	1.45
4	1	1.18	0	0	0	0	1	1.45

Frequency distribution for Lower Third Molar Protostylid (LM3.PROTO).

	European: n=87		**African: n=133**		**Asian/Nat.Am.: n=5**		**Hispanic: n=61**	
LM3. PROTO	**n**	**%**	**n**	**%**	**n**	**%**	**n**	%
0	81	93.1	126	94.74	4	80	48	78.69
1	3	3.45	4	3.01	0	0	8	13.11
2	0	0	2	1.5	1	20	0	0
3	1	1.15	0	0	0	0	2	3.28
4	1	1.15	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0
6	1	1.15	1	0.75	0	0	1	1.64
7	0	0	0	0	0	0	2	3.28

Frequency distribution for Lower Second Molar Protostylid (LM2.PROTO).

	European: n=154		**African: n=143**		**Asian/Nat.Am.: n=3**		**Hispanic: n=86**	
LM2. PROTO	**n**	**%**	**n**	**%**	**n**	**%**	**n**	%
0	141	91.56	121	84.62	2	66.67	68	79.07
1	10	6.49	16	11.19	1	33.33	16	18.6
2	0	0	6	4.2	0	0	1	1.16
3	2	1.3	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	1	1.16
6	1	0.65	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0

Frequency distribution for Lower First Molar Protostylid (LM1.PROTO).

	European: n=122		**African: n=117**		**Asian/Nat.Am.: n=4**		**Hispanic: n=53**	
LM1. PROTO	**n**	**%**	**n**	**%**	**n**	**%**	**n**	%
0	100	81.97	75	64.1	2	50	47	88.68
1	22	18.03	35	29.91	2	50	2	3.77
2	0	0	7	5.98	0	0	3	5.66
3	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	1	1.89
6	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0

Frequency distribution for Lower Third Molar Enamel Extensions (LM3.ENEXT).

	European: n=87		**African: n=127**		**Asian/Nat.Am.: n=5**		**Hispanic: n=52**	
LM3. ENEXT	**n**	**%**	**n**	**%**	**n**	**%**	**n**	%
0	68	78.16	118	92.91	1	20	35	67.31
1	16	18.39	9	7.09	3	60	12	23.08
2	3	3.45	0	0	1	20	5	9.62
3	0	0	0	0	0	0	0	0

Frequency distribution for Lower Second Molar Enamel Extensions (LM2.ENEXT).

	European: n=165		**African: n=144**		**Asian/Nat.Am.: n=3**		**Hispanic: n=89**	
LM2. ENEXT	**n**	**%**	**n**	**%**	**n**	**%**	**n**	%
0	128	77.58	119	82.64	1	33.33	34	38.2
1	26	15.76	20	13.89	1	33.33	37	41.57
2	10	6.06	5	3.47	1	33.33	17	19.1
3	1	0.61	0	0	0	0	1	1.12

Frequency distribution for Lower First Molar Enamel Extensions (LM1.ENEXT).

	European: n=143		**African: n=120**		**Asian/Nat.Am.: n=5**		**Hispanic: n=83**	
LM1. ENEXT	**n**	**%**	**n**	**%**	**n**	**%**	**n**	%
0	133	93.01	114	95	4	80	50	60.24
1	10	6.99	6	5	0	0	19	22.89
2	0	0	0	0	1	20	12	14.46
3	0	0	0	0	0	0	2	2.41

Frequency distribution for Lower Third Molar Crenulations (LM3.CREN).

	European: n=69		**African: n=120**		**Asian/Nat.Am.: n=5**		**Hispanic: n=63**	
LM3. CREN	**n**	**%**	**n**	**%**	**n**	**%**	**n**	%
0	25	36.23	11	9.17	1	20	12	19.05
1	30	43.48	27	22.5	1	20	23	36.51
2	14	20.29	82	68.33	3	60	28	44.44

Frequency distribution for Lower Second Molar Crenulations (LM2.CREN).

	European: n=108		**African: n=125**		**Asian/Nat.Am.: n=3**		**Hispanic: n=76**	
LM2. CREN	**n**	**%**	**n**	**%**	**n**	**%**	**n**	%
0	87	80.56	25	20	1	33.33	51	67.11
1	15	13.89	39	31.2	1	33.33	22	28.95
2	6	5.56	61	48.8	1	33.33	3	3.95

Frequency distribution for Lower First Molar Crenulations (LM1.CREN).

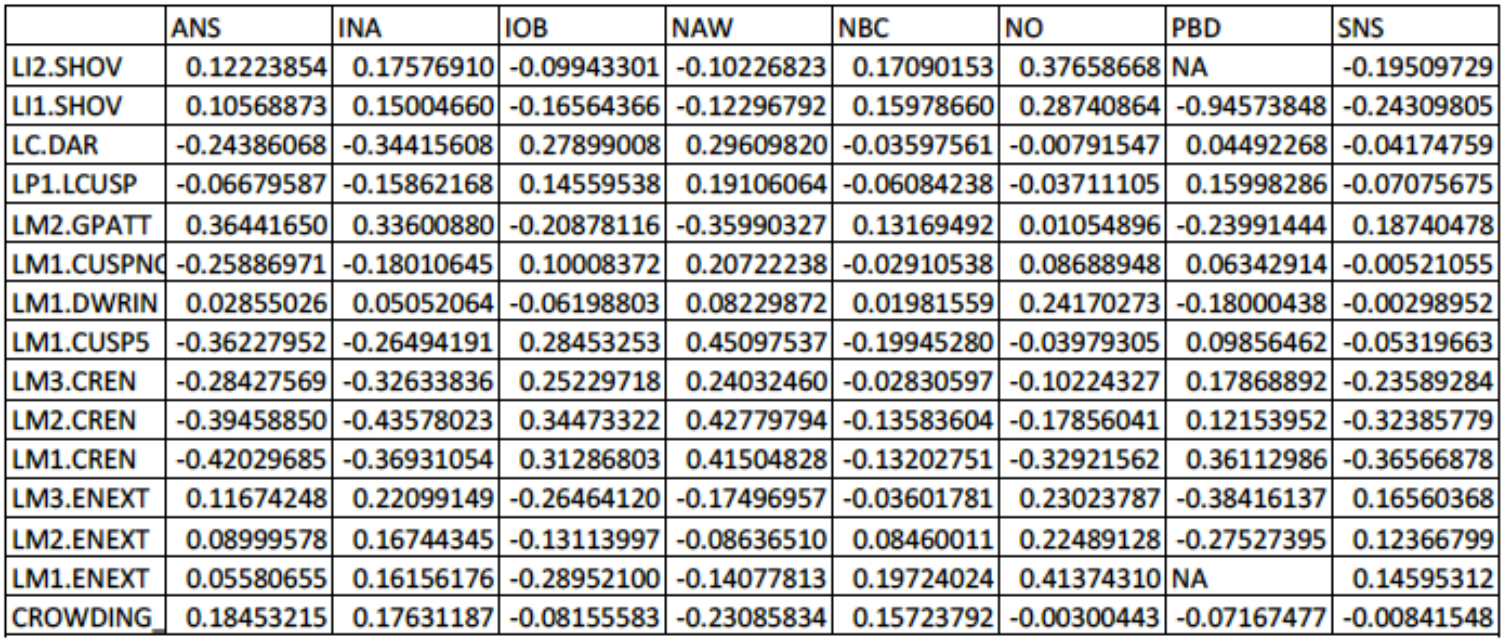
	European: n=76		**African: n=97**		**Asian/Nat.Am.: n=1**		**Hispanic: n=60**	
LM1. CREN	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	58	76.32	23	23.71	0	0	40	66.67
1	14	18.42	44	45.36	0	0	17	28.33
2	4	5.26	30	30.93	1	100	3	5

Frequency distribution for Mandibular Dental Crowding (CROWDING_MAND).

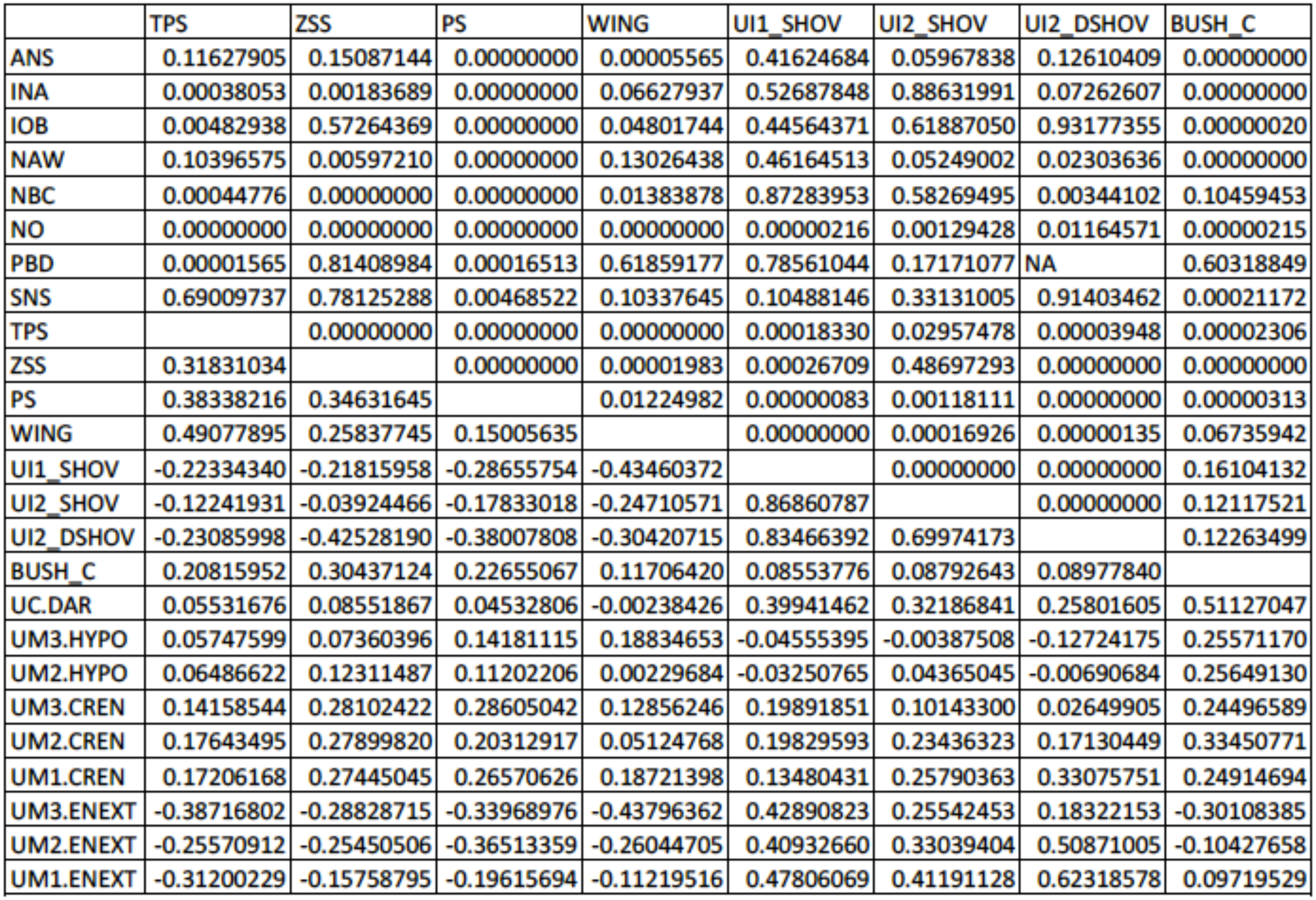
	European: n=178		**African: n=149**		**Asian/Nat.Am. : n=5**		**Hispanic: n=81**	
CROWDING_ MAND	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
0	38	21.35	50	33.56	0	0	19	23.46
1	140	78.65	99	66.44	5	100	62	76.54

Appendix B: Correlation Matrix with p-values in the Upper Triangle

	ANS	INA	IOB	NAW	NBC	NO	PBD	SNS
ANS		0.00000000	0.00000000	0.00000000	0.00000000	0.90600646	0.00000000	0.00000064
INA	0.72227369		0.00000000	0.00000000	0.00000000	0.25015735	0.00000000	0.00000869
IOB	-0.34084072	-0.46651528		0.00000000	0.00000000	0.07884305	0.00000000	0.00334784
NAW	-0.61669515	-0.65330017	0.59754661		0.00000000	0.68250888	0.00000000	0.00001282
NBC	0.28184707	0.31201135	-0.30200645	-0.35722826		0.00757535	0.00149274	0.00000248
NO	-0.00486732	0.04584630	-0.06990462	0.01634099	0.10613407		0.00000000	0.05774105
PBD	-0.31095958	-0.30815452	0.25625316	0.25385441	-0.12265213	-0.23128384		0.27467635
SNS	0.19616117	0.16938125	-0.11219393	-0.16669739	0.18053643	0.07546674	-0.04194870	
TPS	-0.06424026	-0.13970981	0.11127052	0.06448106	-0.13944307	-0.37151180	0.16995695	0.01577843
ZSS	-0.05876419	-0.12272952	-0.02227162	0.10856684	-0.24046716	-0.26096256	0.00930596	-0.01097191
PS	-0.28394469	-0.34283649	0.24953536	0.29392761	-0.25990173	-0.26747696	0.15910172	-0.11966064
WING	0.24726198	0.10991258	0.11913081	-0.09163381	-0.14937136	-0.37791760	0.02998729	-0.09824482
UI1_SHOV	-0.04894880	-0.03717720	-0.04519485	0.04385430	0.00955623	0.28318261	0.01604153	-0.09575252
UI2_SHOV	-0.10572269	-0.00787566	-0.02760963	0.10782985	0.03068411	0.18135584	0.07541485	-0.05380748
UI2_DSHOV	0.08692410	0.09956073	0.00479717	-0.12743580	0.16358157	0.14452188	NA	0.00603930
BUSH_C	-0.41291102	-0.41704000	0.25115689	0.36988010	-0.07987893	-0.23522799	0.02543636	-0.18003881
UC.DAR	-0.28720623	-0.31810757	0.17661472	0.28594753	-0.01558410	0.08567039	0.01510988	0.09897451
UM3.HYPO	-0.08696003	-0.16306459	0.21104146	0.13412192	-0.12323838	-0.35730814	0.18456691	-0.03523223
UM2.HYPO	-0.13933872	-0.15486261	0.16364577	0.13508170	-0.15718094	-0.14024777	0.21659341	-0.00497484
UM3.CREN	-0.29027547	-0.28758730	0.07565129	0.31125717	-0.10276585	-0.09575812	0.09822630	-0.32719247
UM2.CREN	-0.46276231	-0.32509178	0.20179739	0.39941719	-0.10265100	-0.09672509	0.24900245	-0.21472305
UM1.CREN	-0.37690735	-0.27799333	0.19179783	0.24638890	-0.08326765	-0.15780071	0.23774310	-0.29380566
UM3.ENEXT	0.09720200	0.10052942	-0.16039321	-0.17710508	0.07045443	0.42417154	-0.16160892	-0.02781285
UM2.ENEXT	0.17453145	0.21512846	-0.23337732	-0.16794311	0.17454712	0.25508460	-0.36131494	0.03155772
UM1.ENEXT	0.06575335	0.12175228	-0.17551991	-0.10782352	0.08805810	0.02935240	-0.33261231	0.06147361



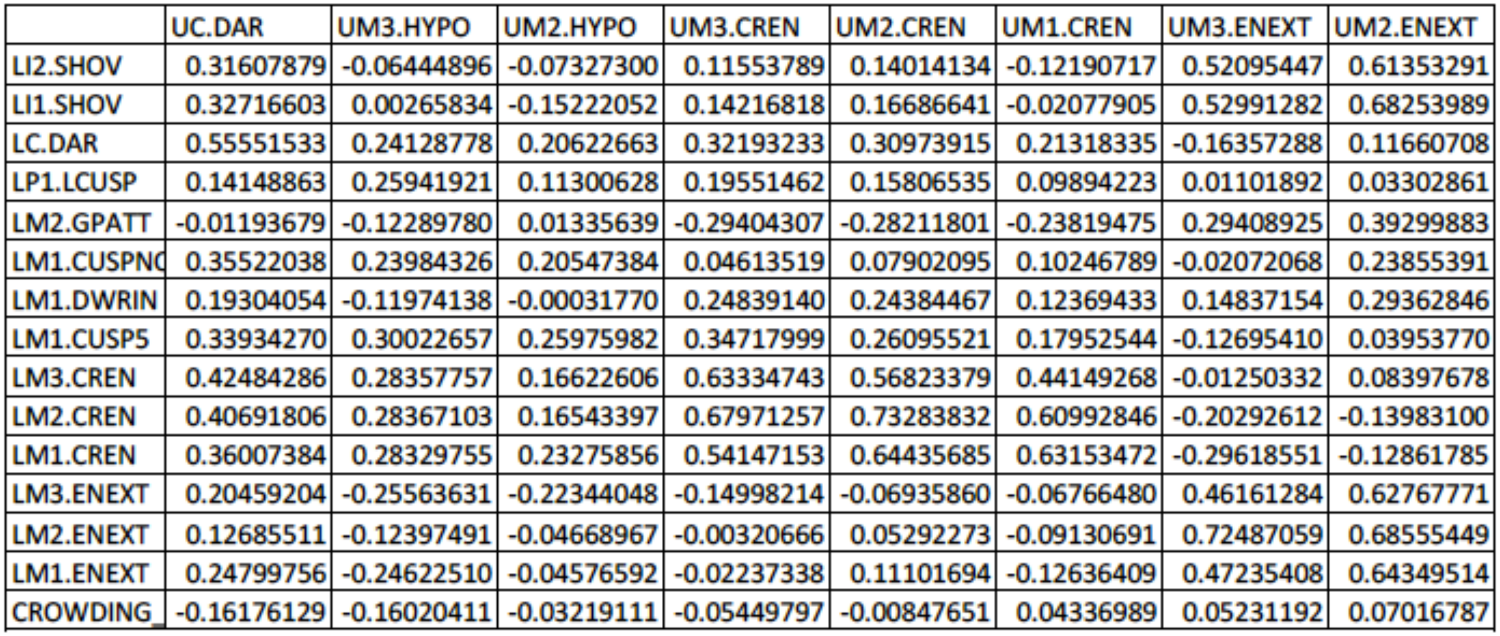
	ANS	INA	IOB	NAW	NBC	NO	PBD	SNS
LI2.SHOV	0.12223854	0.17576910	-0.09943301	-0.10226823	0.17090153	0.37658668	NA	-0.19509729
LI1.SHOV	0.10568873	0.15004660	-0.16564366	-0.12296792	0.15978660	0.28740864	-0.94573848	-0.24309805
LC.DAR	-0.24386068	-0.34415608	0.27899008	0.29609820	-0.03597561	-0.00791547	0.04492268	-0.04174759
LP1.LCUSP	-0.06679587	-0.15862168	0.14559538	0.19106064	-0.06084238	-0.03711105	0.15998286	-0.07075675
LM2.GPATT	0.36441650	0.33600880	-0.20878116	-0.35990327	0.13169492	0.01054896	-0.23991444	0.18740478
LM1.CUSPNC	-0.25886971	-0.18010645	0.10008372	0.20722238	-0.02910538	0.08688948	0.06342914	-0.00521055
LM1.DWRIN	0.02855026	0.05052064	-0.06198803	0.08229872	0.01981559	0.24170273	-0.18000438	-0.00298952
LM1.CUSP5	-0.36227952	-0.26494191	0.28453253	0.45097537	-0.19945280	-0.03979305	0.09856462	-0.05319663
LM3.CREN	-0.28427569	-0.32633836	0.25229718	0.24032460	-0.02830597	-0.10224327	0.17868892	-0.23589284
LM2.CREN	-0.39458850	-0.43578023	0.34473322	0.42779794	-0.13583604	-0.17856041	0.12153952	-0.32385779
LM1.CREN	-0.42029685	-0.36931054	0.31286803	0.41504828	-0.13202751	-0.32921562	0.36112986	-0.36566878
LM3.ENEXT	0.11674248	0.22099149	-0.26464120	-0.17496957	-0.03601781	0.23023787	-0.38416137	0.16560368
LM2.ENEXT	0.08999578	0.16744345	-0.13113997	-0.08636510	0.08460011	0.22489128	-0.27527395	0.12366799
LM1.ENEXT	0.05580655	0.16156176	-0.28952100	-0.14077813	0.19724024	0.41374310	NA	0.14595312
CROWDING	0.18453215	0.17631187	-0.08155583	-0.23085834	0.15723792	-0.00300443	-0.07167477	-0.00841548



	TPS	ZSS	PS	WING	UI1_SHOV	UI2_SHOV	UI2_DSHOV	BUSH_C
ANS	0.11627905	0.15087144	0.00000000	0.00005565	0.41624684	0.05967838	0.12610409	0.00000000
INA	0.00038053	0.00183689	0.00000000	0.06627937	0.52687848	0.88631991	0.07262607	0.00000000
IOB	0.00482938	0.57264369	0.00000000	0.04801744	0.44564371	0.61887050	0.93177355	0.00000020
NAW	0.10396575	0.00597210	0.00000000	0.13026438	0.46164513	0.05249002	0.02303636	0.00000000
NBC	0.00044776	0.00000000	0.00000000	0.01383878	0.87283953	0.58269495	0.00344102	0.10459453
NO	0.00000000	0.00000000	0.00000000	0.00000000	0.00000216	0.00129428	0.01164571	0.00000215
PBD	0.00001565	0.81408984	0.00016513	0.61859177	0.78561044	0.17171077	NA	0.60318849
SNS	0.69009737	0.78125288	0.00468522	0.10337645	0.10488146	0.33131005	0.91403462	0.00021172
TPS		0.00000000	0.00000000	0.00000000	0.00018330	0.02957478	0.00003948	0.00002306
ZSS	0.31831034		0.00000000	0.00001983	0.00026709	0.48697293	0.00000000	0.00000000
PS	0.38338216	0.34631645		0.01224982	0.00000083	0.00118111	0.00000000	0.00000313
WING	0.49077895	0.25837745	0.15005635		0.00000000	0.00016926	0.00000135	0.06735942
UI1_SHOV	-0.22334340	-0.21815958	-0.28655754	-0.43460372		0.00000000	0.00000000	0.16104132
UI2_SHOV	-0.12241931	-0.03924466	-0.17833018	-0.24710571	0.86860787		0.00000000	0.12117521
UI2_DSHOV	-0.23085998	-0.42528190	-0.38007808	-0.30420715	0.83466392	0.69974173		0.12263499
BUSH_C	0.20815952	0.30437124	0.22655067	0.11706420	0.08553776	0.08792643	0.08977840	
UC.DAR	0.05531676	0.08551867	0.04532806	-0.00238426	0.39941462	0.32186841	0.25801605	0.51127047
UM3.HYPO	0.05747599	0.07360396	0.14181115	0.18834653	-0.04555395	-0.00387508	-0.12724175	0.25571170
UM2.HYPO	0.06486622	0.12311487	0.11202206	0.00229684	-0.03250765	0.04365045	-0.00690684	0.25649130
UM3.CREN	0.14158544	0.28102422	0.28605042	0.12856246	0.19891851	0.10143300	0.02649905	0.24496589
UM2.CREN	0.17643495	0.27899820	0.20312917	0.05124768	0.19829593	0.23436323	0.17130449	0.33450771
UM1.CREN	0.17206168	0.27445045	0.26570626	0.18721398	0.13480431	0.25790363	0.33075751	0.24914694
UM3.ENEXT	-0.38716802	-0.28828715	-0.33968976	-0.43796362	0.42890823	0.25542453	0.18322153	-0.30108385
UM2.ENEXT	-0.25570912	-0.25450506	-0.36513359	-0.26044705	0.40932660	0.33039404	0.50871005	-0.10427658
UM1.ENEXT	-0.31200229	-0.15758795	-0.19615694	-0.11219516	0.47806069	0.41191128	0.62318578	0.09719529

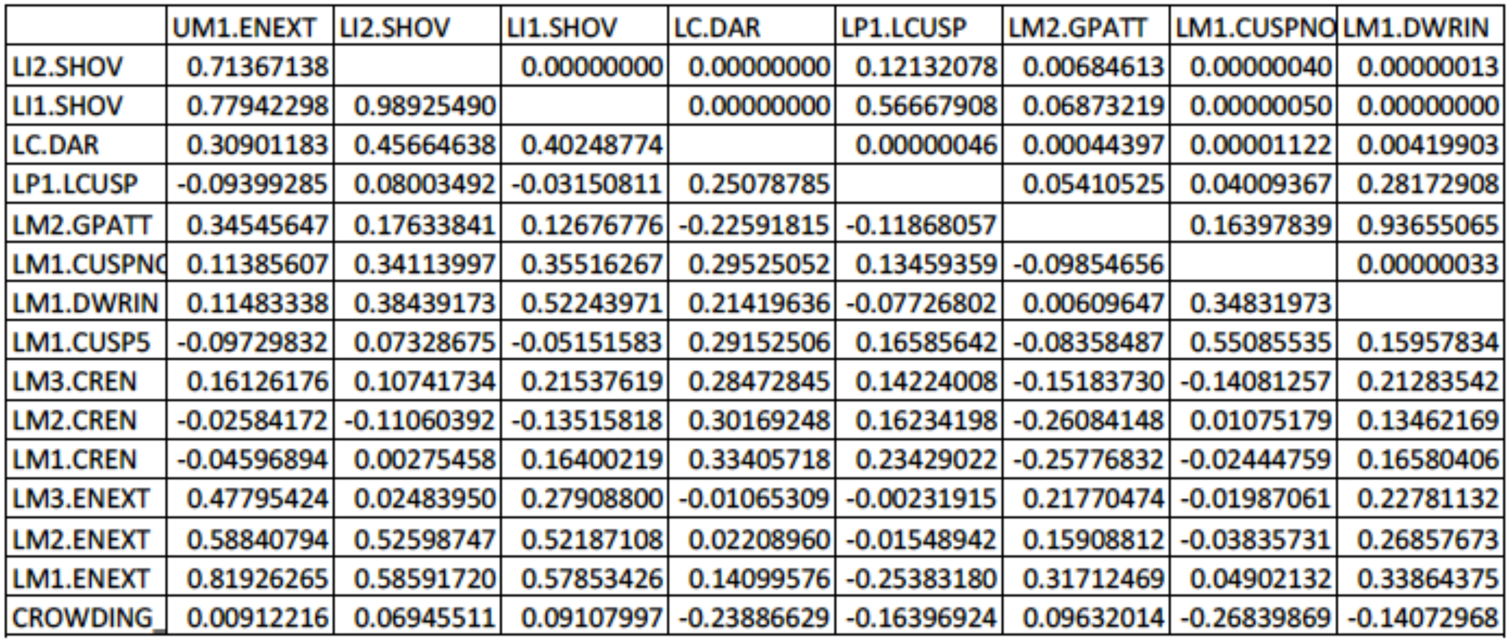
	TPS	ZSS	PS	WING	UI1_SHOV	UI2_SHOV	UI2_DSHOV	BUSH_C
LI2.SHOV	-0.51450164	-0.46810304	-0.36798241	-0.33766288	0.77344532	0.69886046	0.69330651	-0.13534633
LI1.SHOV	-0.56591114	-0.45987517	-0.37070232	-0.42866639	0.86506956	0.76477853	0.71202238	-0.34513357
LC.DAR	0.02312957	-0.05153456	0.04820671	-0.13766363	0.33327309	0.27855846	0.37483423	0.44358879
LP1.LCUSP	0.15379872	0.07910128	0.01178820	0.00363061	-0.02074027	-0.03109110	-0.10270482	0.25752596
LM2.GPATT	-0.18196104	-0.09275716	-0.27292055	-0.08848880	0.14145280	0.11710916	0.12589247	-0.05938772
LM1.CUSPNC	0.01521278	-0.01551038	-0.14895270	-0.09922156	0.28799085	0.33319625	0.32734243	0.20239179
LM1.DWRIN	-0.16355951	-0.12082383	-0.15432338	-0.14882277	0.37684487	0.24149226	0.37248606	0.01277300
LM1.CUSP5	0.14702000	0.17858779	0.00414992	0.02968602	0.17704853	0.11422288	-0.09913912	0.32829336
LM3.CREN	0.22006383	0.09553330	0.20294530	0.08421810	0.28406479	0.13696910	0.11271407	0.27869648
LM2.CREN	0.27024891	0.22632808	0.39481712	0.02782193	0.11511658	0.09177179	-0.10890279	0.43518045
LM1.CREN	0.30206805	0.20720322	0.33152035	0.16107462	0.22954441	0.22206257	-0.03498177	0.36331324
LM3.ENEXT	-0.05549956	-0.17407288	-0.15530938	-0.11641672	0.33448749	0.12300922	0.46813284	-0.16309931
LM2.ENEXT	-0.21125836	-0.15832602	-0.20121087	-0.20710577	0.50060633	0.31864204	0.53251692	-0.06380855
LM1.ENEXT	-0.36237156	-0.26295208	-0.19479360	-0.06270547	0.48118797	0.41321404	0.45596065	0.05897609
CROWDING_	-0.10773628	-0.02838447	0.01397394	0.00303306	-0.01628440	-0.03780720	0.02702827	-0.15838476

	UC.DAR	UM3.HYPO	UM2.HYPO	UM3.CREN	UM2.CREN	UM1.CREN	UM3.ENEXT	UM2.ENEXT
ANS	0.00000001	0.15036988	0.00518687	0.00000134	0.00000000	0.00000000	0.10583283	0.00025864
INA	0.00000000	0.00588083	0.00134454	0.00000123	0.00000000	0.00000008	0.08802342	0.00000299
IOB	0.00033744	0.00035900	0.00072872	0.21275874	0.00005125	0.00025681	0.00656309	0.00000042
NAW	0.00000000	0.02454818	0.00550087	0.00000016	0.00000000	0.00000245	0.00269544	0.00031092
NBC	0.75482397	0.04113572	0.00133450	0.09378348	0.04302918	0.11785341	0.23910355	0.00019519
NO	0.09195660	0.00000000	0.00540821	0.13027961	0.06308258	0.00378441	0.00000000	0.00000009
PBD	0.76033950	0.00192695	0.00000714	0.10664927	0.00000052	0.00000542	0.00607066	0.00000000
SNS	0.04545455	0.55572394	0.91874457	0.00000003	0.00001594	0.00000001	0.63891666	0.49957140
TPS	0.27213864	0.34679546	0.19211003	0.02214093	0.00052266	0.00127173	0.00000000	0.00000005
ZSS	0.09253582	0.22457959	0.01386003	0.00000337	0.00000003	0.00000025	0.00000106	0.00000008
PS	0.36289744	0.01717725	0.02120057	0.00000154	0.00004559	0.00000031	0.00000000	0.00000000
WING	0.97081279	0.02280693	0.97310241	0.12731083	0.46444469	0.01209253	0.00000002	0.00004418
UI1_SHOV	0.00000000	0.57228735	0.62078414	0.01308977	0.00281181	0.05456259	0.00000002	0.00000000
UI2_SHOV	0.00000001	0.95825004	0.47178155	0.17304188	0.00013268	0.00006324	0.00038930	0.00000001
UI2_DSHOV	0.00000742	0.09054246	0.91072984	0.72700973	0.00641003	0.00000046	0.01381929	0.00000000
BUSH_C	0.00000000	0.00007333	0.00000124	0.00017530	0.00000000	0.00001445	0.00000203	0.04501858
UC.DAR		0.11517598	0.01074751	0.00000000	0.00000000	0.00000000	0.36320215	0.00032538
UM3.HYPO	0.10348380		0.00000000	0.00053330	0.06547776	0.01540890	0.02120483	0.00232971
UM2.HYPO	0.13797905	0.49978987		0.51340095	0.01350638	0.01500117	0.38353921	0.02221243
UM3.CREN	0.37929446	0.20864732	0.04127530		0.00000000	0.00000000	0.76034583	0.89282750
UM2.CREN	0.42792858	0.11668370	0.12388500	0.76282980		0.00000000	0.57734027	0.24161939
UM1.CREN	0.33665002	0.16316664	0.13400452	0.58377544	0.78671532		0.00022609	0.92923978
UM3.ENEXT	0.05945240	-0.14150899	-0.05436869	-0.01915616	-0.03577448	-0.25013531		0.00000000
UM2.ENEXT	0.18787246	-0.18387771	-0.11064628	0.00834726	0.05853380	0.00483370	0.75523545	
UM1.ENEXT	0.26401363	-0.02926174	-0.04790051	0.05512462	0.09184822	0.01118692	0.46961357	0.69941748

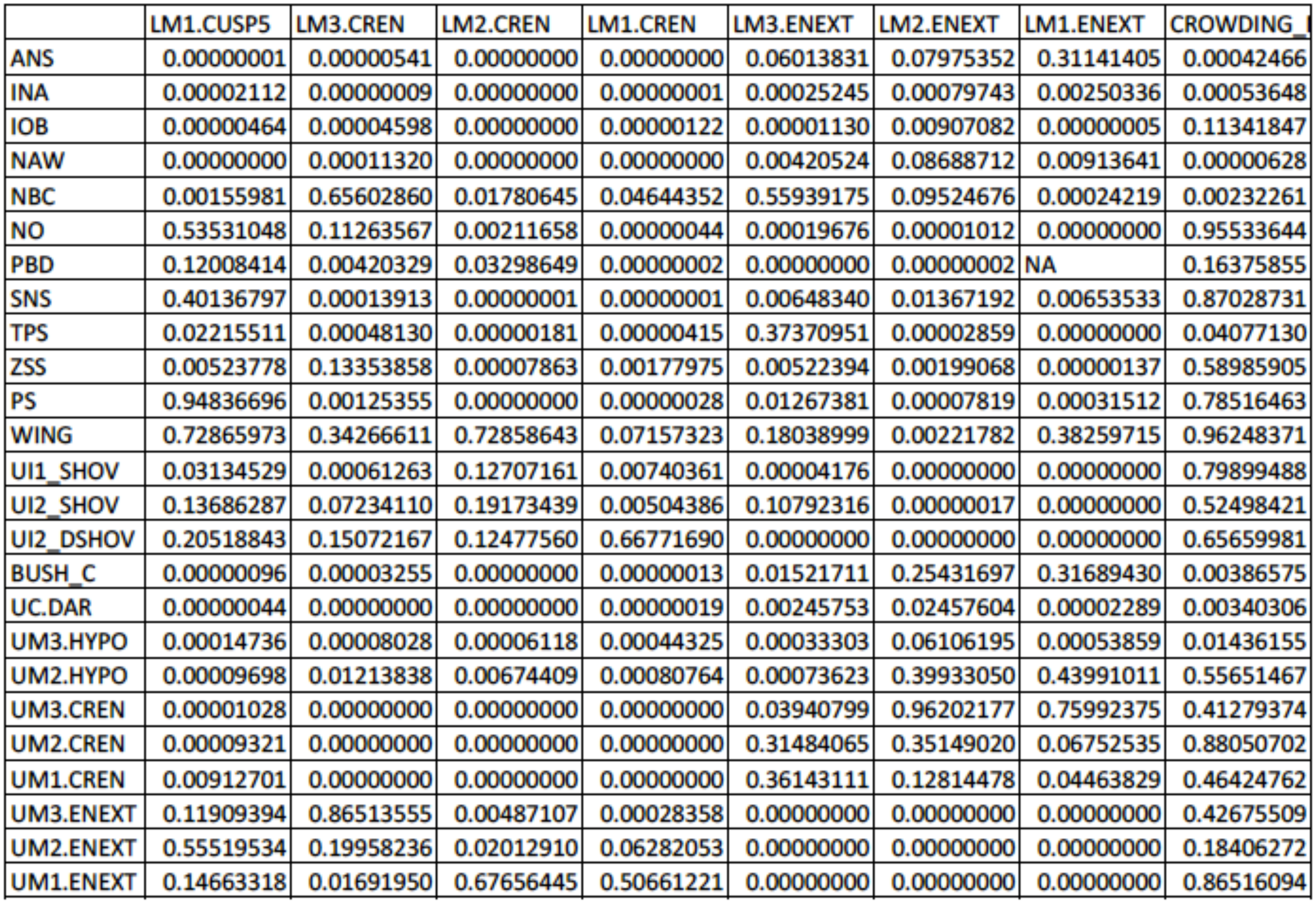


	UC.DAR	UM3.HYPO	UM2.HYPO	UM3.CREN	UM2.CREN	UM1.CREN	UM3.ENEXT	UM2.ENEXT
LI2.SHOV	0.31607879	-0.06444896	-0.07327300	0.11553789	0.14014134	-0.12190717	0.52095447	0.61353291
LI1.SHOV	0.32716603	0.00265834	-0.15222052	0.14216818	0.16686641	-0.02077905	0.52991282	0.68253989
LC.DAR	0.55551533	0.24128778	0.20622663	0.32193233	0.30973915	0.21318335	-0.16357288	0.11660708
LP1.LCUSP	0.14148863	0.25941921	0.11300628	0.19551462	0.15806535	0.09894223	0.01101892	0.03302861
LM2.GPATT	-0.01193679	-0.12289780	0.01335639	-0.29404307	-0.28211801	-0.23819475	0.29408925	0.39299883
LM1.CUSPNC	0.35522038	0.23984326	0.20547384	0.04613519	0.07902095	0.10246789	-0.02072068	0.23855391
LM1.DWRIN	0.19304054	-0.11974138	-0.00031770	0.24839140	0.24384467	0.12369433	0.14837154	0.29362846
LM1.CUSP5	0.33934270	0.30022657	0.25975982	0.34717999	0.26095521	0.17952544	-0.12695410	0.03953770
LM3.CREN	0.42484286	0.28357757	0.16622606	0.63334743	0.56823379	0.44149268	-0.01250332	0.08397678
LM2.CREN	0.40691806	0.28367103	0.16543397	0.67971257	0.73283832	0.60992846	-0.20292612	-0.13983100
LM1.CREN	0.36007384	0.28329755	0.23275856	0.54147153	0.64435685	0.63153472	-0.29618551	-0.12861785
LM3.ENEXT	0.20459204	-0.25563631	-0.22344048	-0.14998214	-0.06935860	-0.06766480	0.46161284	0.62767771
LM2.ENEXT	0.12685511	-0.12397491	-0.04668967	-0.00320666	0.05292273	-0.09130691	0.72487059	0.68555449
LM1.ENEXT	0.24799756	-0.24622510	-0.04576592	-0.02237338	0.11101694	-0.12636409	0.47235408	0.64349514
CROWDING	-0.16176129	-0.16020411	-0.03219111	-0.05449797	-0.00847651	0.04336989	0.05231192	0.07016787

	UM1.ENEXT	LI2.SHOV	LI1.SHOV	LC.DAR	LP1.LCUSP	LM2.GPATT	LM1.CUSPNO	LM1.DWRIN
ANS	0.18125383	0.01532252	0.04850995	0.00000058	0.16481005	0.00000000	0.00004754	0.68744315
INA	0.01032064	0.00032617	0.00391476	0.00000000	0.00066595	0.00000000	0.00420146	0.46864152
IOB	0.00021920	0.04445907	0.00154044	0.00000000	0.00186912	0.00036964	0.11372188	0.37374929
NAW	0.02435228	0.03942746	0.01960085	0.00000000	0.00004428	0.00000000	0.00100460	0.23958799
NBC	0.06715561	0.00054333	0.00235987	0.46107333	0.19765350	0.02701155	0.64762716	0.77740346
NO	0.55389996	0.00000000	0.00000004	0.87243092	0.43850266	0.86200465	0.17521203	0.00052935
PBD	0.00000000	NA	0.00000000	0.35330036	0.00060573	0.00004012	0.31785140	0.00927626
SNS	0.19808355	0.00006848	0.00000261	0.38894585	0.13137671	0.00139886	0.93453610	0.96581651
TPS	0.00000000	0.00000000	0.00000000	0.64011125	0.00130927	0.00219694	0.81386446	0.02065739
ZSS	0.00126115	0.00000000	0.00000000	0.29906362	0.09982190	0.12421494	0.80990338	0.08673957
PS	0.00003424	0.00000000	0.00000000	0.33865545	0.80942888	0.00000404	0.01941907	0.02753463
WING	0.09172127	0.00000011	0.00000000	0.03415857	0.95538029	0.29157101	0.24519020	0.11243000
UI1_SHOV	0.00000000	0.00000000	0.00000000	0.00000007	0.74219310	0.07531647	0.00038618	0.00001260
UI2_SHOV	0.00000000	0.00000000	0.00000000	0.00000144	0.59356027	0.11044345	0.00000847	0.00366533
UI2_DSHOV	0.00000000	0.00000000	0.00000000	0.00000000	0.08348446	0.09037466	0.00001773	0.00000584
BUSH_C	0.06776333	0.01343835	0.00000000	0.00000000	0.00000073	0.36682551	0.00300527	0.86524546
UC.DAR	0.00000056	0.00000001	0.00000001	0.00000000	0.00758756	0.85742395	0.00000011	0.00962520
UM3.HYPO	0.64253707	0.35039974	0.97111785	0.00035665	0.00005704	0.09291321	0.00273562	0.16030730
UM2.HYPO	0.34920531	0.19749787	0.01075307	0.00018983	0.03483063	0.83290181	0.00219088	0.99657557
UM3.CREN	0.38835033	0.09654498	0.05355645	0.00000189	0.00296628	0.00004855	0.57120173	0.00319425
UM2.CREN	0.08223171	0.01619174	0.00617810	0.00000003	0.00393891	0.00000614	0.24420813	0.00079666
UM1.CREN	0.83155684	0.04658409	0.74724215	0.00037997	0.08982181	0.00033219	0.13889034	0.09901152
UM3.ENEXT	0.00000000	0.00000000	0.00000000	0.01515150	0.86459190	0.00004196	0.80063043	0.08590430
UM2.ENEXT	0.00000000	0.00000000	0.00000000	0.02963950	0.52707404	0.00000000	0.00030544	0.00004119
UM1.ENEXT		0.00000000	0.00000000	0.00000001	0.07908085	0.00000003	0.08912332	0.11659891



	UM1.ENEXT	LI2.SHOV	LI1.SHOV	LC.DAR	LP1.LCUSP	LM2.GPATT	LM1.CUSPNO	LM1.DWRIN
LI2.SHOV	0.71367138		0.00000000	0.00000000	0.12132078	0.00684613	0.00000040	0.00000013
LI1.SHOV	0.77942298	0.98925490		0.00000000	0.56667908	0.06873219	0.00000050	0.00000000
LC.DAR	0.30901183	0.45664638	0.40248774		0.00000046	0.00044397	0.00001122	0.00419903
LP1.LCUSP	-0.09399285	0.08003492	-0.03150811	0.25078785		0.05410525	0.04009367	0.28172908
LM2.GPATT	0.34545647	0.17633841	0.12676776	-0.22591815	-0.11868057		0.16397839	0.93655065
LM1.CUSPNO	0.11385607	0.34113997	0.35516267	0.29525052	0.13459359	-0.09854656		0.00000033
LM1.DWRIN	0.11483338	0.38439173	0.52243971	0.21419636	-0.07726802	0.00609647	0.34831973	
LM1.CUSP5	-0.09729832	0.07328675	-0.05151583	0.29152506	0.16585642	-0.08358487	0.55085535	0.15957834
LM3.CREN	0.16126176	0.10741734	0.21537619	0.28472845	0.14224008	-0.15183730	-0.14081257	0.21283542
LM2.CREN	-0.02584172	-0.11060392	-0.13515818	0.30169248	0.16234198	-0.26084148	0.01075179	0.13462169
LM1.CREN	-0.04596894	0.00275458	0.16400219	0.33405718	0.23429022	-0.25776832	-0.02444759	0.16580406
LM3.ENEXT	0.47795424	0.02483950	0.27908800	-0.01065309	-0.00231915	0.21770474	-0.01987061	0.22781132
LM2.ENEXT	0.58840794	0.52598747	0.52187108	0.02208960	-0.01548942	0.15908812	-0.03835731	0.26857673
LM1.ENEXT	0.81926265	0.58591720	0.57853426	0.14099576	-0.25383180	0.31712469	0.04902132	0.33864375
CROWDING_	0.00912216	0.06945511	0.09107997	-0.23886629	-0.16396924	0.09632014	-0.26839869	-0.14072968



	LM1.CUSP5	LM3.CREN	LM2.CREN	LM1.CREN	LM3.ENEXT	LM2.ENEXT	LM1.ENEXT	CROWDING_
ANS	0.00000001	0.00000541	0.00000000	0.00000000	0.06013831	0.07975352	0.31141405	0.00042466
INA	0.00002112	0.00000009	0.00000000	0.00000001	0.00025245	0.00079743	0.00250336	0.00053648
IOB	0.00000464	0.00004598	0.00000000	0.00000122	0.00001130	0.00907082	0.00000005	0.11341847
NAW	0.00000000	0.00011320	0.00000000	0.00000000	0.00420524	0.08688712	0.00913641	0.00000628
NBC	0.00155981	0.65602860	0.01780645	0.04644352	0.55939175	0.09524676	0.00024219	0.00232261
NO	0.53531048	0.11263567	0.00211658	0.00000044	0.00019676	0.00001012	0.00000000	0.95533644
PBD	0.12008414	0.00420329	0.03298649	0.00000002	0.00000000	0.00000002	NA	0.16375855
SNS	0.40136797	0.00013913	0.00000001	0.00000001	0.00648340	0.01367192	0.00653533	0.87028731
TPS	0.02215511	0.00048130	0.00000181	0.00000415	0.37370951	0.00002859	0.00000000	0.04077130
ZSS	0.00523778	0.13353858	0.00007863	0.00177975	0.00522394	0.00199068	0.00000137	0.58985905
PS	0.94836696	0.00125355	0.00000000	0.00000028	0.01267381	0.00007819	0.00031512	0.78516463
WING	0.72865973	0.34266611	0.72858643	0.07157323	0.18038999	0.00221782	0.38259715	0.96248371
UI1_SHOV	0.03134529	0.00061263	0.12707161	0.00740361	0.00004176	0.00000000	0.00000000	0.79899488
UI2_SHOV	0.13686287	0.07234110	0.19173439	0.00504386	0.10792316	0.00000017	0.00000000	0.52498421
UI2_DSHOV	0.20518843	0.15072167	0.12477560	0.66771690	0.00000000	0.00000000	0.00000000	0.65659981
BUSH_C	0.00000096	0.00003255	0.00000000	0.00000013	0.01521711	0.25431697	0.31689430	0.00386575
UC.DAR	0.00000044	0.00000000	0.00000000	0.00000019	0.00245753	0.02457604	0.00002289	0.00340306
UM3.HYPO	0.00014736	0.00008028	0.00006118	0.00044325	0.00033303	0.06106195	0.00053859	0.01436155
UM2.HYPO	0.00009698	0.01213838	0.00674409	0.00080764	0.00073623	0.39933050	0.43991011	0.55651467
UM3.CREN	0.00001028	0.00000000	0.00000000	0.00000000	0.03940799	0.96202177	0.75992375	0.41279374
UM2.CREN	0.00009321	0.00000000	0.00000000	0.00000000	0.31484065	0.35149020	0.06752535	0.88050702
UM1.CREN	0.00912701	0.00000000	0.00000000	0.00000000	0.36143111	0.12814478	0.04463829	0.46424762
UM3.ENEXT	0.11909394	0.86513555	0.00487107	0.00028358	0.00000000	0.00000000	0.00000000	0.42675509
UM2.ENEXT	0.55519534	0.19958236	0.02012910	0.06282053	0.00000000	0.00000000	0.00000000	0.18406272
UM1.ENEXT	0.14663318	0.01691950	0.67656445	0.50661221	0.00000000	0.00000000	0.00000000	0.86516094

	LM1.CUSP5	LM3.CREN	LM2.CREN	LM1.CREN	LM3.ENEXT	LM2.ENEXT	LM1.ENEXT	CROWDING_
LI2.SHOV	0.29046175	0.11981223	0.08091593	0.96959284	0.71407633	0.00000000	0.00000000	0.23352248
LI1.SHOV	0.48026297	0.00284257	0.04236366	0.02963162	0.00007456	0.00000000	0.00000000	0.13404820
LC.DAR	0.00001463	0.00002808	0.00000087	0.00000161	0.87402083	0.69020190	0.01608897	0.00002415
LP1.LCUSP	0.01122360	0.02757474	0.00629083	0.00051695	0.97113147	0.77118041	0.00000562	0.00373597
LM2.GPATT	0.23930740	0.03363181	0.00001015	0.00031813	0.00291252	0.00663192	0.00000113	0.15640369
LM1.CUSPNC	0.00000000	0.07389079	0.87516129	0.71348739	0.81312808	0.55848963	0.43663649	0.00014840
LM1.DWRIN	0.02228301	0.00990247	0.06770461	0.01836179	0.00970387	0.00014684	0.00000060	0.07053103
LM1.CUSP5		0.00000183	0.00256683	0.00639054	0.29638856	0.76893853	0.85085908	0.00000335
LM3.CREN	0.36364554		0.00000000	0.00000000	0.85917912	0.28531819	0.07238584	0.06344791
LM2.CREN	0.20418741	0.81341590		0.00000000	0.05111613	0.36769182	0.00544570	0.00100575
LM1.CREN	0.17972167	0.57283576	0.80163239		0.00104196	0.21850579	0.59556859	0.03723863
LM3.ENEXT	-0.08730680	-0.01200214	-0.14210014	-0.27521101		0.00000000	0.00000000	0.01259484
LM2.ENEXT	0.01926397	0.07137725	-0.05117003	-0.08348276	0.71746614		0.00000000	0.52725823
LM1.ENEXT	0.01185564	0.13064803	-0.17707050	-0.03487462	0.50945023	0.68253760		0.00016253
CROWDING_	-0.32582572	-0.13386830	-0.21187851	-0.15370467	0.17747907	-0.03670230	0.22795468	

Appendix C: Conditional Probabilities for Each Trait

Anterior Nasal Spine			
Ancestry	**1**	**2**	**3**
Black	0.77372263	0.20437956	0.02189781
Hispanic	0.375	0.39423077	0.23076923
White	0.16666667	0.41428571	0.41904762

Inferior Nasal Aperture					
Ancestry	**1**	**2**	**3**	**4**	**5**
Black	0.36170213	0.3475177	0.23404255	0.04964539	0.00709220
Hispanic	0.07142857	0.16964286	0.29464286	0.28571429	0.17857143
White	0.01702128	0.06382979	0.23829787	0.34042553	0.34042553

Interorbital Breadth			
Ancestry	**1**	**2**	**3**
Black	0.0647482	0.51798561	0.41726619
Hispanic	0.37272727	0.48181818	0.14545455
White	0.43722944	0.47186147	0.09090909

Nasal Aperture Width			
Ancestry	**1**	**2**	**3**
Black	0.04347826	0.31884058	0.63768116
Hispanic	0.22727273	0.66363636	0.10909091
White	0.53070175	0.37719298	0.09210526

Nasal Bone Contour					
Ancestry	**0**	**1**	**2**	**3**	**4**
Black	0.34285714	0.17857143	0.03571429	0.17857143	0.26428571
Hispanic	0.03636364	0.15454545	0.05454545	0.38181818	0.37272727
White	0.05652174	0.17826087	0.10434783	0.36086957	0.3

Nasal Overgrowth

Ancestry	0	1
Black	0.92424242	0.07575758
Hispanic	0.51648352	0.48351648
White	0.87837838	0.12162162

Post-bregmatic Depression

Ancestry	0	1
Black	0.8467153	0.1532847
Hispanic	0.9266055	0.0733945
White	0.9527897	0.0472103

Supranasal Suture

Ancestry	0	1	2
Black	0.71223022	0.02877698	0.25899281
Hispanic	0.51818182	0.05454545	0.42727273
White	0.51515152	0.14718615	0.33766234

Transverse Palatine Suture Shape

Ancestry	0	1	2	3
Black	0.078740157	0.181102362	0.732283465	0.007874016
Hispanic	0.420560748	0.401869159	0.14953271	0.028037383
White	0.117647059	0.248868778	0.597285068	0.036199095

Zygomaticomaxillary Suture Shape

Ancestry	0	1	2	3
Black	0.07751938	0.170542636	0.744186047	0.007751938
Hispanic	0.348623853	0.348623853	0.293577982	0.009174312
White	0.097222222	0.393518519	0.50462963	0.00462963

Palate Shape

Ancestry	1	2	3	4	5
Black	0.00763359	0.07633588	0.26717557	0.46564886	0.18320611
Hispanic	0.19230769	0.41346154	0.24038462	0.125	0.02884615
White	0.01190476	0.1845238	0.51190476	0.19642857	0.09523810

Incisor Winging

Ancestry	1	2	3	4
Black	0.07575758	0.01515152	0.89393939	0.01515152
Hispanic	0.33333333	0.04761905	0.5952381	0.02380952
White	0.06930693	0.00990099	0.91089109	0.00990099

Incisor Shoveling (Maxillary Lateral Incisor)

Ancestry	0	1	2	3	4	5	6	7
Black	0.3711340	0.3402062	0.1958763	0.0412371	0.0103093	0.0206186	0.0103093	0.0103093
Hispanic	0.1818182	0.2545455	0.2545455	0.0727273	0.1454546	0.0363636	0.0363636	0.0181818
White	0.53	0.29	0.11	0.02	0.01	0.02	0.01	0.01

Double Shoveling (Maxillary Lateral Incisor)

Ancestry	0	1	2	3	4	5	6
Black	0.91666667	0.02380952	0.01190476	0.01190476	0.01190476	0.01190476	0.01190476
Hispanic	0.74074074	0.12962963	0.05555556	0.01851852	0.01851852	0.01851852	0.01851852
White	0.93457944	0.01869159	0.00934579	0.00934579	0.00934579	0.00934579	0.00934579

Mesial Canine Ridge (Bushman Canine)

Ancestry	0	1	2	3
Black	0.697478992	0.12605042	0.109243697	0.067226891
Hispanic	0.841269841	0.111111111	0.031746032	0.015873016
White	0.943089431	0.016260163	0.032520325	0.008130081

Distal Accessory Ridge (Maxillary Canine)						
Ancestry	**0**	**1**	**2**	**3**	**4**	**5**
Black	0.33898305	0.24576271	0.22033898	0.05932203	0.09322034	0.04237288
Hispanic	0.44444444	0.17460317	0.17460317	0.11111111	0.04761905	0.04761905
White	0.752	0.128	0.064	0.04	0.008	0.008

Hypocone (Maxillary 3rd Molar)						
Ancestry	**0**	**1**	**2**	**3**	**4**	**5**
Black	0.07368421	0.24210526	0.26315789	0.27368421	0.09473684	0.05263158
Hispanic	0.28125	0.203125	0.328125	0.140625	0.03125	0.015625
White	0.26315789	0.24561404	0.24561404	0.12280702	0.0877193	0.03508772

Hypocone (Maxillary 2nd Molar)						
Ancestry	**0**	**1**	**2**	**3**	**4**	**5**
Black	0.00869565	0.02608696	0.14782609	0.33913044	0.26086957	0.21739130
Hispanic	0.02352941	0.07058824	0.23529412	0.29411765	0.32941177	0.04705882
White	0.04464288	0.15178571	0.17857143	0.1875	0.33928571	0.09821429

Crenulations (Maxillary 3rd Molar)			
Ancestry	**0**	**1**	**2**
Black	0.1098901	0.3076923	0.5824176
Hispanic	0.3103448	0.3965517	0.2931034
White	0.36	0.48	0.16

Crenulations (Maxillary 1st Molar)			
Ancestry	**0**	**1**	**2**
Black	0.4742268	0.3814433	0.1443299
Hispanic	0.75324675	0.19480519	0.05194805
White	0.85227273	0.125	0.02272727

Enamel Extensions (Maxillary 2nd Molar)

Ancestry	0	1	2	3
Black	0.922413793	0.043103448	0.025862069	0.00862069
Hispanic	0.494382022	0.348314607	0.146067416	0.011235955
White	0.817460317	0.142857143	0.031746032	0.007936508

Enamel Extensions (Maxillary 1st Molar)

Ancestry	0	1	2	3
Black	0.92920354	0.03539823	0.026548673	0.008849558
Hispanic	0.76744186	0.104651163	0.11627907	0.011627907
White	0.94214876	0.033057851	0.016528926	0.008264463

Incisor Shoveling (Mandibular Lateral Incisor)

Ancestry	0	1	2	3	4	5	6	7
Black	0.9145299	0.034188	0.0085470	0.0085470	0.0085470	0.0085470	0.0085470	0.00854701
Hispanic	0.6071429	0.2678571	0.0357143	0.0178571	0.0178571	0.0178571	0.0178571	0.0178571
White	0.9469697	0.0075758	0.0075758	0.0075758	0.0075758	0.0075758	0.0075758	0.0075758

Distal Accessory Ridge (Mandibular Canine)

Ancestry	0	1	2	3	4	5
Black	0.39316239	0.29914530	0.16239316	0.05128205	0.03418803	0.05982906
Hispanic	0.38333333	0.35	0.16666667	0.05	0.03333333	0.01666667
White	0.80714286	0.14285714	0.02142857	0.01428571	0.00714286	0.00714286

Lingual Cusp Number (Mandibular 1st Premolar)

Ancestry	A	0	1	2	3	4	5	6	7	8	9
Black	0.007575758	0.401515152	0.128787879	0.204545455	0.136363636	0.045454545	0.03030303	0.007575758	0.007575758	0.022727273	0.007575758
Hispanic	0.015625	0.53125	0.171875	0.09375	0.03125	0.0625	0.015625	0.015625	0.015625	0.03125	0.015625
White	0.006578947	0.631578947	0.092105263	0.078947368	0.078947368	0.032894737	0.039473684	0.013157895	0.006578947	0.013157895	0.006578947

Groove Pattern (Mandibular 2nd Molar)

Ancestry	Y	+	X
Black	0.49462366	0.07526882	0.43010753
Hispanic	0.11111111	0.24074074	0.64814815
White	0.140625	0.171875	0.6875

Cusp Number (Mandibular 1st Molar)

Ancestry	4	5	6
Black	0.02631579	0.89473684	0.07894737
Hispanic	0.03846154	0.86538462	0.09615385
White	0.14516129	0.77419355	0.08064516

Deflecting Wrinkle (Mandibular 1st Molar)

Ancestry	0	1	2	3
Black	0.77777778	0.09722222	0.08333333	0.04166667
Hispanic	0.47619048	0.19047619	0.30952381	0.02380952
White	0.77083333	0.10416667	0.10416667	0.02083333

Hypoconulid (Mandibular 1st Molar)

Ancestry	0	1	2	3	4	5
Black	0.02531646	0.02531646	0.05063291	0.08860759	0.32911392	0.48101266
Hispanic	0.03571429	0.01785714	0.01785714	0.17857143	0.46428571	0.28571429
White	0.125	0.03125	0.09375	0.234375	0.265625	0.25

Crenulations (Mandibular 3rd Molar)

Ancestry	0	1	2
Black	0.1136364	0.2386364	0.6477273
Hispanic	0.1836735	0.4285714	0.3877551
White	0.2941176	0.4705882	0.2352941

Crenulations (Mandibular 1st Molar)			
Ancestry	**0**	**1**	**2**
Black	0.28	0.42666667	0.29333333
Hispanic	0.65217391	0.26086957	0.08695652
White	0.71929825	0.21052632	0.07017544

Enamel Extensions (Mandibular 3rd Molar)				
Ancestry	**0**	**1**	**2**	**3**
Black	0.92473118	0.05376344	0.01075269	0.01075269
Hispanic	0.65853659	0.19512195	0.12195122	0.02439024
White	0.7761194	0.17910448	0.02985075	0.01492537

Enamel Extensions (Mandibular 2nd Molar)				
Ancestry	**0**	**1**	**2**	**3**
Black	0.822429907	0.121495327	0.046728972	0.009345794
Hispanic	0.384615385	0.446153846	0.153846154	0.015384615
White	0.758333333	0.175	0.058333333	0.008333333

Maxillary Dental Crowding		
Ancestry	**0**	**1**
Black	0.68	0.32
Hispanic	0.6865672	0.3134328
White	0.4951456	0.5048544

Works Cited

Adams D. 2015. Odontometric differentiation between Southwest Hispanics, Native Americans, and European Americans. Master's Thesis, Boston University, Boston, MA.

Algee-Hewitt BFB. 2011. If and how many races? The application of mixture modeling to worldwide human craniometric variation. PhD Dissertation. Department of Anthropology. University of Tennessee, Knoxville.

Allen JA. 1877. The influence of physical conditions in the genesis of species. Radical Rev 1:108-140.

American Association of Physical Anthropologists [AAPA]. 1996. AAPA statement on the biological aspects of race. Am J Phys Anthropol 101:569-570

Angel JL. 1944. A racial analysis of the ancient Greeks: an essay on the use of morphological types. Am J Phys Anthropol 2:329-376.

Angel JL. 1946. Skeletal change in ancient Greece. Am J Phys Anthropol 4:69-98.

Angel JL, and JO Kelly. 1990. Inversion of the posterior edge of the jaw ramus: New race trait. In: Gill G, and Rhine S, editors. Skeletal Attribution of Race: Methods for Forensic Anthropology. Albuquerque, New Mexico: Maxwell Museum of Anthropology. pp. 33–39.

Armelagos G, and van Gerven DP. 2003. A century of skeletal biology and paleopathology: Contrasts, contradictions, and conflicts. Am Anthropol 105:51-62.

Armelagos GJ, and Goodman AH. 1998. Race, racism, and anthropology. In: Goodman AH, and Leatherman TL, editors. Building a new biocultural synthesis. Ann Arbor, MI: University of Michigan Press. pp. 359-377.

Auerbach BM. 2014. The William W. Howells craniometric data set. < http://web.utk.edu/~auerbach/HOWL.htm>. Accessed 2016 October 17.

Baker SJ, Gill GW, and Keiffer DA. 1990. Race and sex determination from the intercondylar notch of the distal femur. In: Gill G, and Rhine S, editors. Skeletal Attribution of Race: Methods for Forensic Anthropology. Albuquerque, New Mexico: Maxwell Museum of Anthropology. pp. 91–96.

Bamshad M, Wooding S, Salisbury BA, and Stephens JC. 2004. Deconstructing the relationship between genetics and race. Nat Rev Genet 5:598-609.

Bass WM. 2005. Human osteology: A laboratory and field manual, 5th ed. Columbia: Missouri Archaeological Society.

Bates M. 1952. Physical anthropology: Origin and evolution of man. Cold Spring Harbor Symposia on quantitative biology. Am Anthropol 54:87-88.

Beals KL, Smith CL, and Dodd SM. 1983. Climate and the evolution of brachycephalization. Am J Phys Anthropol 62:425-437.

Beals KL, Smith CL, and Dodd SM. 1984. Brain size, cranial morphology, climate and time machines. Curr Anthropol 25:301-330.

Bendyshe T, and Blumenbach JF. 1865. The Anthropological Treatises of Johann Friedrich Blumenbach. Bendyshe, T, editor. London: Anthropological Society.

Berger LR, Hawks J, de Ruiter DJ, Churchill SE, Schmid P, Delezene LK, Kivell TL, Garvin HM, Williams SA, DeSilva JM, Skinner MM, Musiba CM, Cameron N, Holliday TW, Harcourt-Smith W, Ackermann RR, Bastir M, Bogin B, Bolter D, Brophy J, Cofran ZD, Condon KA, Deane AS, Dembo M, Drapeau M, Elliot MC, Feuerriegel EM, Garcia-Martinez D, Green DJ, Gurtov A, Irish JD, Kruger A, Laird MF, Marchi D, Meyer MR, Nalla S, Negash EW, Orr CM, Radovcic D, Schroeder L, Scott JE, Throckmorton Z, Tocheri MW, VanSickle C, Walker CS, Wei P, Zipfel B. 2015. *Homo naledi*, a new species of the genus *Homo* from the Dinaledi Chamber, South Africa. eLife 4: e09560.

Bergmann K. 1847. Über die Verhältnisse der wärmeökonomie der Thiere zu ihrer Grösse. Göttinger Studien 3: 595-708.

Bertoni B, Budowle B, Sans M, Barton SA, and Chakraborty R. 2003. Admixture in Hispanics: distribution of ancestral population contributions in the continental United States. Hum Biol 75:1-11.

Birdsell JB. 1963. Review: The origin of human races. Q Rev Biol 38:178-185.

Birkby WH. 1966. An evaluation of race and sex identification from cranial measurements. Am J Phys Anthropol 24:21-27.

Birkby W, Fenton T, and Anderson B. 2008. Identifying Southwest Hispanics using nonmetric traits and the cultural profile. J Forensic Sci 53: 29-33.

Blakey M. 1987. Skull doctors: intrinsic social and political bias in the history of American physical anthropology; with special reference to the work of Ales Hrdliçka. Crit Anthropol 7:7-35.

Blumenbach JF. 1775. De generis humani varietate native liber. Gottingae, Germany:

Bergman Publishers.

Boas F. 1912. Changes in the bodily form of descendants of immigrants. Am Anthropol 14:530-562.

Boas F. 1918. Notes on the anthropology of Sweden. Am J Phys Anthropol 1:415-426.

Boldsen JL, Milner GR, Konigsberg LW, and Wood JW. 2002. Transition analysis: a new method for estimating age from skeletons. In: Hoppa RD, and Vaupel JW, editors. Paleodemography: age distributions from skeletal samples. Cambridge: Cambridge University Press. p 73-106.

Bonilla C, Gutierrez G, Parra EJ, Kline C, and Shriver MD. 2005. Admixture analysis of a rural population of the state of Guerrero, Mexico. Am J Phys Anthropol 128:861-869.

Boyd WC. 1940. Critique of methods of classifying mankind. Am J Phys Anthropol 27:333-364.

Boyd WC. 1950. Genetics and the races of man: an introduction to modern physical anthropology. Boston: Little Brown.

Boyd WC. 1963. Genetics and the human race. Science 140:1057-1064.

Brace CL. 1981. Tales of the phylogenetic woods: the evolution and significance of evolutionary trees. Am J Phys Anthropol 56:411-429.

Brace CL. 1982. The roots of the race concept in American physical anthropology. In: Spencer F, editor. A history of American physical anthropology: 1930-1980. New York: Academic Press. pp 11-29.

Brace CL. 1995. Region does not mean "Race"--reality versus convention in forensic anthropology. J Forensic Sci 40:171-175.

Brace CL. 2005. “Race” is a four-letter word: the genesis of the concept. Oxford: Oxford University Press.

Brattain M. 2007. Race, racism, and antiracism: UNESCO and the politics of presenting science to the postwar public. Am Hist Rev 112:1386-1413.

Breiman L. 2001. Random forests. Mach Learn 45(1):5–32.

Breiman L, and Cutler A. 2004. Random forests. Berkeley, http://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm

Breiman L, Friedman J, Stone CJ, and Olshen RA. 1984. Classification and regression trees. Boca Raton, FL: CRC Press.

Broca M. 1862. The projections of the head and a new procedure in cephalometry. Bull Soc Anthropol Paris 3:514.544.

Brues AM. 1977. People and Races. New York: MacMillan Publishing Co. Inc.

Brues AM. 1990. The once and future diagnosis of race. In: Gill G, and Rhine S, editors. Skeletal Attribution of Race: Methods for Forensic Anthropology. Albuquerque, New Mexico: Maxwell Museum of Anthropology. pp. 1-7.

Brues AM. 1992. Forensic diagnosis of race – general race vs. specific populations. Soc Sci Med 34:125-128.

Brues AM. 1993. The objective view of race. NAPA Bull 13:74-78.

Buikstra JE. 2006. A historical introduction. In: Buikstra JE, and Beck LA, editors. Bioarchaeology: The contextual analysis of human remains. Burlington, MA: Academic Press. pp 7-25.

Buikstra JE, and Beck LA. 2006. Bioarchaeology: the contextual analysis of human remains. Burlington, MA: Academic Press.

Burns KR. 2007. Forensic anthropology training manual, 2nd edition. Upper Saddle River, NJ: Pearson-Prentice Hall.

Butler PM. 1939. Studies of the mammalian dentition and of the differentiation of the post-canine dentition. P Zool Soc Lond 109:1-36.

Byers SN, Churchill SE, and Curran B. 1997. Identification of Euro-Americans, Afro-Americans, and Amerindians from palatal dimensions. J Forensic Sci 42:3-9.

Byers SN. 2011. Introduction to forensic anthropology, 4th edition. New York: Pearson.

Camper P. 1791. Over het natuurlijk verschil der wezenstrekken in menschen van onderscheiden landaart en ouderdom; over het schoon in antyke beelden en gesneedene stenen. Utrecht, Netherlands: Wild and Altheer.

Carbonell VM. 1965. The teeth of the Neanderthal child from Gibraltar: a re-evaluation. Am J Phys Anthropol 23:41-49.

Carter K, and Worthington S. 2015. Morphologic and demographic predictors of third molar agenesis: A systematic review and meta-analysis. J Dent Res 94: 886-894.

Cartmill M. 1998. The status of the race concept in physical anthropology. Am Anthropol 100:651-660.

Caspari R. 2009. 1918: Three perspective on race and human variation. Am J Phys Anthropol 139:5-15.

Cavalli-Sforza LL, Menozzi P, Piazza A. 1994. The History and Geography of Human Genes. Princeton, NJ: Princeton University Press.

Cavalli-Sforza LL. 1997. Genes, peoples, and languages. Proc Natl Acad Sci 94:7719-24.

Cerda-Flores RM, Budowle B, Jin L, Barton SA, Deka R, and Chakraborty R. 2002. Maximum likelihood estimates of admixture in northeastern Mexico using 13 short tandem repeat loci. Am J Hum Biol 14:429-439.

Christensen A. 2004. The impact of Daubert: Implications for testimony and research in forensic anthropology (and the use of frontal sinuses in personal identification). J Forensic Sci 49:1–4.

Christensen AM, Leslie WD, and Baim S. 2014. Ancestral differences in femoral neck axis length: Possible implications for forensic anthropological analyses. Forensic Sci Int 236:e1-e4.

Cohen J .1960. A coefficient of agreement for nominal scales Edu Psychol Meas 20: 37–46.

Cohen J. 1968. Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. Psychol Bull 70:213-220.

Comas J. 1960. Manual of physical anthropology. Springfield, IL: Charles C. Thomas.

Committee on Identifying the Needs of the Forensic Sciences Community, National Research Council. 2009. Strengthening forensic science in the United States: A path forward. Doc No:228091. Washington D.C.: U.S. Department of Justice.

Cook DC. 2006. The old physical anthropology and the New World: A look at the accomplishments of an antiquated paradigm. In: Buikstra JE, and Beck LA, editors. Bioarchaeology: The contextual analysis of human remains. Burlington, MA: Academic Press. pp. 27-71.

Coon CS. 1962. The origin of races. New York: Alfred A. Knopf.

Corruccini RS, Townsend GC, Richards LC, and Brown T. 1990. Genetic and environmental determinants of dental occlusal variation in twins of different nationalities. Hum Biol 62:353-367.

Craig EA. 1995. Intercondylar shelf angle: A new method to determine race from the distal femur. J Forensic Sci 40:777-782.

Dahlberg AA. 1945. The changing dentition of man. J AM Dent Assoc 32:676-690.

Dahlberg AA. 1951. The dentition of the American Indian. In: Laughlin WS, editor. Papers on the physical anthropology of the American Indian. New York: Viking Fund. pp. 138-176.

Dahlberg AA. 1956. Materials for the establishment of standards for classifications of tooth characters, attributes and techniques in morphological studies of the dentition. Chicago, IL: Zooler Laboratory of Dental Anthropology, University of Chicago.

Darwin C. 1859. On the origin of species. London: John Murray Press.

Daubenton L. 1764. Mémoire sur les différences de la situation du grand trou occipital dans l'homme et dans les animaux (Memoir on the Different Positions of the Occipital Foramen in Man and Animals). In: Académie royale des sciences, editor. Histoire de l'Académie royale des sciences. Oxford: Oxford University Press.

Daubert v. Merrell Dow Pharmaceuticals Inc. 1993. US Supreme Court 509.U.S.579,113S.Ct.2786, 125L. Ed.2d 469.

Deng H, Runger G, and Tuv E. 2011. Bias of importance measures for multi-valued attributes and solutions. Proceedings of the 21st International Conference on Artificial Neural Networks.

DiBennardo R, and Taylor JV. 1983. Multiple discriminant function analysis of sex and race in the postcranial skeleton. Am J Phys Anthropol 61:305-314.

Dirkmaat DC, Cabo LL, Ousley SD, and Symes SA. 2008. New perspectives in forensic anthropology. Am J Phys Anthropol 137:33-52.

Dobra A, and Gehrke J. 2001. Bias correction in classification tree construction. In: Brodley CE, and Danyluk AP, editors. Proceedings of the Eighteenth International Conference on Machine Learning. Williamstown, MA: Morgan Kaufmann. p. 90-97.

Dobzhansky T. 1963. Two views of Coon's Origin of Races with comments by Coon and replies: Possibility that Homo sapiens evolved independently 5 times is vanishingly small. Curr Anthropol 4:360, 364-366.

Domingos P, and Pazzani M. 1997. On the optimality of the simple Bayesian classifier under zero-one loss. Mach Learn 29:103-130.

Drennan RD. 2009. Statistics for archaeologists: A common sense approach, 2nd edition. New York: Springer.

Dumont ER. 1999. Crenulated molar enamel in primates: a functional model. Presented to the 68th Meeting of the American Association of Physical Anthropologists. Columbus, OH, April 26-May 1.

Dunn LC, and Dobzhansky T. 1952. Heredity, race, and society. New York: The New American Library.

Edgar H. 2005. Prediction of race using characteristics of dental morphology. J Forensic Sci 50:269-273.

Edgar H. 2007. Microevolution of African American dental morphology. Am J Phys Anthropol 132:535-544.

Edgar H. 2009. Testing the utility of dental morphological traits commonly used in the forensic identification of ancestry. In: Koppe T, Meyer G, Alt K, editors. Comparative dental morphology. Frontiers of oral biology. Basel, Switzerland: Karger. p. 49-54

Edgar H. 2013. Estimation of ancestry using dental morphological characteristics. J Forensic Sci 58: S3-S8.

Edgar H. 2015. Dental morphological estimation of ancestry in forensic contexts. In: Berg GE, and Ta'ala SC, editors. Biological affinity in forensic identification of human skeletal remains: Beyond black and white. Boca Raton, FL: CRC Press. pp. 191-208.

Edgar H, and Hunley KL. 2009. Race reconciled?: how biological anthropologists view human variation. Am J Phys Anthropol 139:1-4.

Edwards AW. 2003. Human genetic diversity: Lewontin's fallacy. Bioessays 25:798-

801.

El-Najjar MY, and McWilliams KR. 1978. Forensic anthropology: the structure, morphology, and variation of human bone and dentition. Illinois: Charles C. Thomas.

Eller E. 1990. Population substructure and isolation by distance in three continental regions. Am J Phys Anthropol 108:147-159.

Elliot M, and Collard M. 2009. FORDISC and the determination of ancestry from cranial measurements. Biol Lett 5:849-852.

Evans JD. 1996. Straightforward statistics for the behavioral sciences. Pacific Grove, CA: Brooks/Cole.

Falconer DS. 1965. The inheritance of liability to certain diseases, estimated from distances among relatives. Ann Hum Genet 29:51-76.

Falconer DS. 1967. The inheritance of liability to disease with variable age of onset, with particular reference to diabetes mellitus. Ann Hum Genet 31:1-20.

Farman AG. 2007. Tooth eruption and dental impaction. In: Farman AG, editor. Panoramic Radiology: Seminars on Maxillofacial Imaging and Interpretation. New York: Springer. pp 73-82.

Federal Rules of Evidence (1975); 2014.

Feldesman M. 2002. Classification trees as an alternative to linear discriminant analysis. Am J Phys Anthropol 119:257–75.

Fox J. 2016. polycor: polychoric and polyserial correlations. R package version 0.7-9. https://CRAN.R-project.org/package=polycor.

Freund RJ, Wilson WJ, Mohr DL. 2010. Statistical methods, 3rd edition. New York: Academic Press.

Fried D, Spradley MK, Jantz RL, Ousley SD. 2005. The truth is out there: how NOT to use Fordisc. Am J Phys Anthropol S40:103.

Friedman N, Geiger D, Goldszmidt M. 1997. Bayesian network classifiers. Mach Learn 29:131-163.

Frye v. United States. 1923. D.C. Circuit Court 293.F.1013,1014.

Gamer M, Lemon J, and Singh IFP. 2012. irr: various coefficients of interrater reliability and agreement. R package version 0.84. https://CRAN.R-project.org/package=irr.

Garn S. 1965. Human races. Springfield, IL: C.C. Thomas.

Garn SM, and Giles E. 1995. Earnest Albert Hooton, November 20, 1887-May 3, 1954. Bio Memoirs 68:167-179.

Garvin HM, and Passalacqua NV. 2012. Current practices by forensic anthropologists in adult skeletal age estimation. J Forensic Sci 57:427-433.

George R. 2015. Non-metric and metric dental assessment of ancestry in contemporary Mexican individuals: an aid to identifying undocumented border crossers. Master's Thesis. California State University: Chico, Chico, CA.

Gilbert R, And Gill GW. 1990. A metric technique for identifying American Indian femora. In: Gill G, and Rhine S, editors. Skeletal Attribution of Race: Methods for Forensic Anthropology. Albuquerque, New Mexico: Maxwell Museum of Anthropology. pp. 97-99.

Giles E, and Elliot O. 1962. Racial identification from cranial measurements. J Forensic Sci 7:147-57.

Gill GW. 1971. The prehistoric inhabitants of northern coastal Nayarit: Skeletal analysis and description of burials. Ph.D. dissertation, University of Kansas.

Gill GW. 1986. Craniofacial criteria in forensic race identification. In: Reichs K, editor. Forensic osteology: advances in the identification of human remains. Springfield, IL: Charles C. Thomas. p 143-159.

Gill GW. 1990. Introduction. In: Gill GW, and Rhine S, editors. Skeletal attribution of race. Albuquerque, NM: Maxwell Museum of Anthropology. pp. vii-xii.

Gill GW. 1995. Challenge on the frontier: Discerning American Indians from Whites osteologically. J Forensic Sci 40:783-788.

Gill GW. 1998. Craniofacial criteria in forensic race identification. In: Reichs KJ, editor. Forensic Osteology: Advances in the Identification of Human Remains. Springfield, IL: Charles C. Thomas. p. 143-59.

Gill GW. 2001. Racial variation in the proximal and distal femur: Heritability and forensic utility. J Forensic Sci 46:791-799.

Gill GW, and Rhine S. 1990. Skeletal attribution of race. Albuquerque, NM: Maxwell Museum of Anthropology.

Gill GW, Hughes SS, Bennett SM, and Gilbert BM. 1988. Racial identification from the midfacial skeleton with special reference to American Indians and whites. J Forensic Sci 33:92-99.

Gotelli NJ, and Ellison AM. 2004. A primer of ecological statistics. Massachusetts: Sinauer.

Gould SJ. 1994. The geometer of race. Discover 15:65-69.

Gould SJ. 1996. The mismeasure of man. New York: W.W. Norton and Company.

Gravlee CC, Bernard HR, and Leonard WR. 2003. Heredity, environment, and cranial form: a re-analysis of Boas's immigrant data. Am Anthropol 105:125-38.

Grine F. 1981. Occlusal morphology of the mandibular permanent molars of the South African Negro and the Kalahari San (Bushman). Ann S Afr Mus 86:157-215.

Grivas CR, and Komar DA. 2008. Kumho, Daubert, and the nature of scientific inquiry: Implications for forensic anthropology. J Forensic Sci 53: 771-776.

Halie-Selassie Y. 2001. Late Miocene hominids from the Middle Awash, Ethiopia. Nature 412:178-181.

Hanihara K. 1967. Racial characteristics in the dentition. J Dental Res 46: 923-926.

Hanihara T. 1993. Population prehistory of East Asia and the Pacific as viewed from craniofacial morphology: the basic populations in East Asia, VII. Am J Phys Anthropol 91:173-187.

Hanihara T. 1996. Comparison of craniofacial features of major human groups. Am J Phys Anthropol 99:389-412.

Hanihara T, and Ishida H. 2005. Metric dental variation of major human populations. Am J Phys Anthropol 128:287-298.

Hapfelmeier A, Hothorn T, Ulm K, and Strobl C. 2014. A new variable importance measure for random forests with missing data. Stat Comput 24:21-34.

Harris EF, and Foster CL. 2015. Size matters: Discrimination between American blacks and whites, males and females, using tooth crown dimensions. In: Berg GE, and Ta'ala SC, editors. Biological affinity in forensic identification of human skeletal remains: Beyond black and white. Boca Raton, FL: CRC Press. pp. 209-238.

Hattab FN, and Abu Alhaija ESJ. 1999. Radiographic evaluation of mandibular third molar eruption. Oral Surg Oral Med O 88: 285-291.

Hefner JT. 2003. Assessing nonmetric cranial traits currently used in the forensic determination of ancestry. Master's Thesis, The University of Florida, Gainesville, FL.

Hefner JT. 2007. The statistical determination of ancestry using cranial nonmetric traits. PhD Dissertation. Gainesville, FL: University of Florida.

Hefner JT. 2009. Cranial nonmetric variation and estimating ancestry. J Forensic Sci 54:985-95.

Hefner JT. 2017. Introducing the macromorphoscopic databank (MaMD): a data collection and analytical tool for the analysis of macromorphoscopic trait data. Proceedings of the 69th Annual Scientific Meeting of the American Academy of Forensic Sciences. New Orleans, LA: February 13-18.

Hefner J, and Ousley SD. 2006. The statistical determination of ancestry using cranial non-metric traits. Proceedings of the 58th Annual Scientific Meeting of the American Academy of Forensic Sciences. Seattle, WA.

Hefner JT, and Ousley SD. 2014. Statistical classification methods for estimating ancestry using morphoscopic traits. J Forensic Sci 59:883-890.

Hefner JT, Ousley SD, Dirkmaat DC. 2012. Morphoscopic traits and the assessment of ancestry. In: Dirkmaat, DE, editor. A companion to forensic anthropology. Malden, MA: John Wiley & Sons Inc. pp. 287-310.

Hefner JT, Spradley MK, and Anderson B. 2014. Ancestry Assessment Using Random Forest Modeling. J Forensic Sci 59:583-589.

Hefner JT, Pilloud MA, Black CJ, and Anderson BE. 2015. Morphoscopic trait expression in "Hispanic" populations. J Forensic Sci 60:1135-1139.

Heim KN. 2013. Secular change in the rate of dental impaction. Master's Thesis. Baton Rouge, LA: Louisiana State University.

Hellman M. 1928. Racial characters in human dentition. Proc Am Philos Soc 67:157-174.

Herrick CE, and Walsh-Haney HA. 2010. Molar crenulation as an attribute of ancestry in forensic cases: Identification and accuracy. Proceedings of the 62nd Annual Meeting of the American Academy of Forensic Sciences. Seattle, WA, February 22-27.

Hilaire EG. 1795. Mémoire sur une nouvelle division des mammifères, et les principes qui doivent servir de base dans cette sorte de travail, lu à la Société d'Histoire naturelle, le premier floréal de l'an troisième. In: Magasin Encyclopédique, volume VI. Paris: de l'Imprimerie Biliographique.

Hinkes MJ. 1990. Shovel–shaped incisors in human identification. In: Gill G, and Rhine S, editors. Skeletal Attribution of Race: Methods for Forensic Anthropology. Albuquerque, New Mexico: Maxwell Museum of Anthropology. pp. 21–26.

Hinkes MJ. 1993. Race, ethnicity, and forensic anthropology: realities of racial determination in a forensic setting. NAPA Bull 13:48-53.

Hinkle DE, Wiersman W, and Jurs SG. 2003. Applied statistics for the behavioral sciences, 5th ed. Boston: Houghton Mifflin.

Holliday TW, and Falsetti AB. 1999. A new method for discriminating African–American from European–American skeletons using postcranial osteometrics reflective of body shape. J Forensic Sci 44: 926–930.

Hooton EA. 1918. On certain Eskimoid characters in Icelandic skulls. AM J Phys Anthropol 1:53-76.

Hooton EA. 1926. Lecture notes of EA Hooton. Peabody Museum archives. Cambridge, MA.

Hooton EA. 1930. Indians of Pecos Pueblo: a study of their skeletal remains. New Haven, CT: Yale University Press.

Hooton EA. 1931a. Lecture notes of EA Hooton. Unpublished manuscripts of the Peabody Museum archives of E. A. Hooton. Cambridge, MA.

Hooton EA. 1931b. Up from the ape. New York: MacMillan Company.

Hooton EA. 1936. Plain statements about race. Science 83:511-513.

Hooton EA. 1946. Lecture notes of EA Hooton. Unpublished manuscripts of the Peabody Museum archives of E. A. Hooton. Cambridge, MA.

Hothorn T, Buehlmann P, Dudoit S, Molinaro A, and Van der Laan M. 2006a. Survival ensembles. Biostatistics 7:355-373.

Hothorn T, Hornik K, and Zeileis A. 2006b. Unbiased recursive partitioning: a conditional inference framework. J Comput Graph Stat 15:651-674.

Howells WW. 1970. Multivariate analysis for the identification of race from the crania. In: Stewart TD, editor. Personal Identification in Mass Disasters. Washington, D. C: Smithsonian Institution Press. pp. 111–112.

Howells WW. 1973. Cranial variation in man: A study by multivariate analysis of patterns of difference among recent human populations. Papers of the Peabody Museum of Archaeology and Ethnology, Harvard University, vol. 67. Cambridge, MA: The Peabody Museum.

Howells WW. 1989. Skull shapes and the map: craniometric analyses in the dispersion of modern Homo. Papers of the Peabody Museum of Archaeology and Ethnology, Harvard University, vol. 79. Cambridge, MA: The Peabody Museum.

Howells WW. 1995. Who's who in skulls: Ethnic identification of crania from measurements. Papers of the Peabody Museum of Archaeology and Ethnology, Harvard University, vol. 82. Cambridge, MA: The Peabody Museum.

Hrdlička A. 1920. Shovel-shaped teeth. Am J Phys Anthropol 3:429-465.

Hrdlička A. 1921. Further studies of tooth morphology. Am J Phys Anthropol 4:141-176.

Hubbe M, and Neves WA. 2007. On the misclassification of human crania: are there any implications for assumptions about human variation?. Curr Anthropol 48:285-288.

Hughes CE, Juarez CA, Hughes TL, Galloway A, Fowler G, and Chacon S. 2011. A simulation for exploring the effects of the "trait list" method's subjectivity on consistency and accuracy of ancestry estimations. J Forensic Sci 55:1094-1106.

Hughes T, Thomas C, Richards L, and Townsend G. 2001. A study of occlusal variation in the primary dentition of Australian twins and singletons. Arch Oral Biol 46:857-864.

Hulse FS. 1963. Book Review: The origin of races. Am Anthropol 65:685-687.

Hunt DR. 2016. Terry Collection. NMNH Anthropology: Physical Anthropology: Terry collection [Internet]. Available from: http://anthropology.si.edu/cm/terry.htm

Huxley J. 1942. Evolution: the modern synthesis. London: George Alien and Unwin Ltd.

Huxley TH. 1863. Evidence as to man's place in nature. London: Williams and Norgate.

Irish J. 1993. Biological affinities of the Late Pleistocene through modern African aboriginal populations: the dental evidence. Ph.D. dissertation. Arizona State University, Tempe, AZ.

Irish J. 1997. Characteristic high- and low- frequency dental traits in Sub-Saharan African populations. Am J Phys Anthropol 102:455-67.

Irish J. 2013. Afridonty: the 'Sub-Saharan African dental complex' revisited. In: Scott Gr, and Irish J, editors. Anthropological Perspectives on Tooth Morphology: Genetics, Evolution, Variation. Cambridge: Cambridge University Press. pp 296-318.

Irish JD. 2015. Dental nonmetric variation around the world: Using key traits in populations to estimate ancestry in individuals. In: Berg GE, and Ta'ala SC, editors. Biological affinity in forensic identification of human skeletal remains: Beyond black and white. Boca Raton, FL: CRC Press. pp. 165-190.

İşcan MY, and Steyn M. 1999. Craniometric determination of population affinity in South Africans. Int J Leg Med 112:91-97.

Jackson JP. 2001. In ways unacademical: the reception of Carleton S. Coon's "The Origin of Races." J Hist Biol 34:247-285.

Jacob T. 1967. Racial identification of the Bronze Age human dentitions from Bali, Indonesia. J Dent Res 46:903-910.

Jantz RL, and Jantz LM. 2000. Secular change in craniofacial morphology. Am J Hum Biol 12:327-38.

Jensen DD, and Cohen PR. 2000. Multiple comparisons in induction algorithms. Mach Learn 38:309-338.

Jonke E, Prossinger H, Bookstein FL, Schaefer K, Bernhard M, and Freudenthaler JW. 2007. Secular trends in the facial skull from the 19th century to the present, analyzed with geometric morphometrics. Am J Orthod Dentofac 132:63-70.

Jorde LB, and Wooding SP. 2004. Genetic variation, classification and 'race'. Nature 36: S28-S33.

Kass GV. 1980. An exploratory technique for investigating large quantities of categorical data. J R Stat Soc 29:119-127.

Kawala B, Antoszewska J, and Necka A. 2007. Genetics or environment? A twin-method study of malocclusions. World J Orthod 8:406-410.

Keene HJ. 1963. Distribution of diastemas in the dentition of man. Am J Phys Anthropol 21:437-441.

Kennedy KAR. 1995. But professor, why teach race identification if races don't exist? J Forensic Sci 40:797-800.

Kiernberger A. 1955. Morphologische Untersuchungen am Gebiss und den Zähnen der Buschmann-Hottentotten-Gruppe. Vienna: R.M. Rohrer.

Kim H, and Loh W. 2001. Classification trees with unbiased multiway splits. J Am Stat Assoc 96:589-604.

Kimminau KS. 1993. Dental variation and racial estimation: problems and practical forensic applications. NAPA Bull 13:68-73.

King MC, and Motulsky AG. 2002. Mapping human history. Science 298:2342-2343.

Kirveskari P, and Alvesalo L. 1982. Dental morphology in Turner's syndrome (45, X) females. In: Kurtén B, editor. Teeth: form, function and evolution. New York: Columbia University Press. p. 298-303.

Klales AR, and Kenyhercz MW. 2015. Morphological assessment of ancestry using cranial macromorphoscopics. J Forensic Sci 60:13-20.

Klepinger LL. 2006. Fundamentals of forensic anthropology. Hoboken: John Wiley & Sons Inc.

Klimentidis YC, Miller GF, and Shriver MD. 2009. Genetic admixture, self-reported ethnicity, self-estimated admixture, and skin pigmentation among Hispanics and native Americans. Am J Phys Anthropol 138:375-383.

Komar DA, and Grivas C. 2008. Manufactured populations: what do contemporary reference skeletal collections represent? A comparative study using the Maxwell Museum documented collection. Am J Phys Anthropol 137:224-233.

Konigsberg LW, Algee-Hewitt BFB, and Steadman DW. 2009. Estimation and evidence in forensic anthropology: Sex and race. Am J Phys Anthropol 139:77-90.

Kosiba S. 2000. Assessing the efficacy and pragmatism of "race" designation in human skeletal identification: a test of FORDISC 2.0 program. Am J Phys Anthropol S30:200.

Krogman WM. 1939. A guide to the identification of human skeletal remains. FBI Law Enforcement Bulletin 3:3–31.

Krogman WM. 1962. The Human Skeleton in Forensic Medicine. Springfield, IL: Charles C. Thomas.

Krogman WM, and İşcan MY. 1986. The Human Skeleton in Forensic Medicine, 2nd edition. Springfield, IL: Charles C. Thomas.

Krzanowski, WJ. 2000. Principles of multivariate analysis: A user's perspective. London: Oxford University Press.

Landis JR, and Koch GG. 1977. The measurement of observer agreement for categorical data. Biometrics 33:159-174.

Langley P, Iba W, and Thompson K. 1992. An analysis of Bayesian classifiers. Aaai 90:223-228.

Lasker GW. 1950. Genetic analysis of racial traits of the teeth. Cold Spring Harbor Symp Quant Biol 15:191-203.

Lasker GW, and Lee MMC. 1957. Racial traits in human teeth. J Forensic Sci 2:401-419.

Lavelle CLB. 1970. The distribution of diastemas in different human population samples. Eur J Oral Sci 78:530-534.

Lavelle CLB. 1973. Variation in the secular changes in the teeth and dental arches. The Angle Orthodontist 43:412-421.

Lavelle CLB, and Foster TD. 1969. Crowding and spacing of the teeth in an adult British population. Dent Pract Dent Rec 19:239-242.

Leakey LSB, Tobias PV, Leakey MD, and Butler PM. 1965. Olduvai Gorge. Cambridge: Cambridge University Press.

Lease LR, and Sciulli PW. 2005. Brief communication: Discrimination between European-American and African-American children based on deciduous dental metrics and morphology. Am J Phys Anthropol 126:56-60.

Lewontin RC. 1972. The apportionment of human diversity. Evol Biol 6:381-398.

Lewontin RC. 1995. Human diversity. New York: W.H. Freeman and Company.

Liaw A, and Wiener M. 2002. Classification and regression by Random Forest. Rnews 2:18–22.

Linnaeus C. 1759. Systema Natura. Lugduni Batavorum: Theodorum Haak.

Lisker R, Perez-Briceno R, Granados J, Babinsky V, de Rubens J, Armendares S, and Buentello L. 1986. Gene frequencies and admixture estimates in a Mexico City population. Am J Phys Anthropol 71:203-207.

Lisker R, Ramirez E, and Babinsky V. 1996. Genetic structure of autochthonous populations of Mesoamerica: Mexico. Hum Biol 68:395-404.

Lisker R, Ramirez E, Briceno RP, Granados J, and Babinsky V. 1990. Gene frequencies and admixture estimates in four Mexican urban centers. Hum Biol 62:791-801.

Livingstone FB, and Dobzhansky T. 1962. On the non-existence of human races. Curr Anthropol 3:279-281.

Louisiana (State). Legislature. Assembly. House Bill 1140, Act 227. 2006. Louisiana repository for unidentified and missing persons information program.

Loving v. Virginia. 1967. US Supreme Court 388.U.S. 1, 87S. Ct.1817, 18L. Ed.2d 1010.

Maier CA. 2016. A novel method for recording palate shape in the estimation of ancestry. Proceedings of the 68th Annual Meeting of the American Academy of Forensic Sciences. Las Vegas, NV, February 22-27.

Maier CA, Zhang K, Manhein MH, Li X. 2015. Palate shape and depth: A shape-matching and machine learning method for estimating ancestry from human skeletal remains. J Forensic Sci 60: 1129-1134.

Manly BFJ. 1985. The statistics of natural selection on animal populations. New York: Chapman and Hall.

Marks J. 1995. Human biodiversity: genes, race, and history. New Brunswick, NJ: Transaction Publishers.

Marks J. 2010. The two 20th-century crises of racial anthropology. In: Little MA, and Kennedy KAR, editors. Histories of American physical anthropology in the twentieth century. Lanham, MD: Lexington Books. pp. 187-206.

Martini FH, Timmons MJ, and Tallitsch RB. 2009. Human anatomy, 6th edition. London: Benjamin Cummings.

Mayhall JT, Saunders SR, and Belier PL. 1982. The dental morphology of North American whites: a reappraisal. In: Kurtén B, editor. Teeth: Form, Function, and Evolution. New York: Columbia University. pp 245-258.

Mayr E. 1962. Review: Origin of human races. Science 138:420-422.

Mercer DA. 2013. Nonparametric discriminant analysis in forensic ancestry estimation: an assessment of utilized and alternative statistical methods. Knoxville, TN: University of Tennessee, Knoxville.

Meyer D, Dimitriadou E, Hornik K, Weigessel A, and Leisch F. 2015. e1071: misc. functions of the department of statistics, probability theory group (formerly: e1071), TU Wien. R package version 1.6-7. https://CRAN.R-project.org/package=e1071.

Mizoguchi Y. 1985. Shoveling: a statistical analysis of its morphology. Tokyo: University of Tokyo Press.

Montagu A. 1963. Two views of Coon's Origin of Races with comments by Coon and replies: What is remarkable about varieties of man is likeness, not differences. Curr Anthropol 4:361-364.

Moorees CFA, and Reed RB. 1954. Biometrics of crowding and spacing of teeth in the mandible. Am J Phys Anthropol 12:77-88.

Morton SG. 1839. Crania Americana; or a comparative view of the skulls of various aboriginal nations of North and South America: to which is prefixed an essay on the varieties of the human species. Philadelphia: J. Dobson.

Morton SG. 1844. Crania Aegyptica: observations on Egyptian ethnography. Philadelphia: J. Pennington.

Müller-Wille S. 2007. Race et appurtenance ethnique: la diversité humaine et l'UNESCO déclarations sur la race (1950 et 1951). In: 60 ans d'histoire de l'UNESCO. Actes du colloque international, Paris, 16-18 Novembre 2005. Paris: UNESCO. pp 211-220.

Niblett T. 1987. Constructing decision trees in noisy domains. Proceedings of the Second European Working Session on Learning. Bled, Yugoslavia: Sigma. p 67-78.

Nichol CR, and Turner II CG. 1986. Intra- and interobserver concordance in classifying dental morphology. Am J Phys Anthropol 69:299-315.

Nicholson E, and Harvati K. 2006. Quantitative analysis of human mandibular shape using three-dimensional geometric morphometrics. Am J Phys Anthropol 131:368-383.

Normando D, Almeida MAO, Quintão CCA. 2013. Dental crowding: the role of genetics and tooth wear. Angle Orthod 83:10-15.

Novembre J, Johnson T, Bryc K, Kutalik Z, Boyko A, Auton A, Indap A, King KS, Bergmann S, Nelson MR, Stephens M, and Bustamante CD. 2008. Genes mirror geography within Europe. Nature 456:98-101.

Olivier G. 1969. Practical anthropology. Springfield, IL: CC Thomas.

Oschinsky L. 1963. A critique of "The Origin of Races" by C.S. Coon. Anthropologica 5:111-116.

Ousley SD, and Jantz RL. 2006. FORDISC 3.0: personal computer forensic discriminant functions. The University of Tennessee Knoxville.

Ousley SD, and Jantz RL. 2012. Fordisc 3 and statistical methods for estimating sex and ancestry. In: Dirkmaat, DE, editor. A companion to forensic anthropology. Malden, MA: John Wiley & Sons Inc. pp. 311-329.

Ousley SD, Jantz RL, Freid D. 2007. Exploring human craniometric variation: statistical, mensural, biological, and historical considerations. Am J Phys Anthropol Suppl 44:182.

Ousley SD, Jantz RL, and Fried D. 2009. Understanding race and human variation: why forensic anthropologists are good at identifying race. Am J Phys Anthropol 139:68-76.

Parra EJ, Marcini A, Akey J, Martinson J, Batzer MA, Cooper R, Allison DB, Deka R, Ferrell RE, and Shriver MD. 1998. Estimating African American admixture proportions by use of population-specific alleles. Am J Hum Genet 63:1839-1851.

Parra EJ, Teixeira Ribeiro JC, Caeiro JLB, Riveiro A. 1995. Genetic structure of the populations of Cabo Verde (West Africa): Evidence of substantial European admixture. Am J Phys Anthropol 91:381-389.

Pearson K. 1900. Mathematical contributions to the theory of evolution. VII. On the correlation of characters not quantitatively measurable. Philos T R Soc S-A 195:1-405.

Pedersen PO. 1949. The east Greenland Eskimo dentition, numerical variations and anatomy: a contribution to comparative ethnic odontography. CA Reitzel 142(3).

Pilloud MA, Hefner JT, Hanihara T, Hayashi A. 2014. The use of tooth crown measurements in the assessment of ancestry. J Forensic Sci 59:1493-1501.

Pilloud MA, Maier C, Scott GR. 2017a. Molar crenulation trait definition and population variation. Proceedings of the 69th Annual Scientific Meeting of the American Academy of Forensic Sciences. New Orleans, LA: February 13-18.

Pilloud MA, Scott GR, Navega DS, de Oliveira Coelho JPV, Cunha E, and Irish JD. 2017b. A new application for estimating ancestry based on dental morphology. Presented to the 69th Annual Scientific Meeting of the the American Academy of Forensic Sciences. New Orleans, LA, February 13-18.

Pink C, Maier C, Pilloud MA, Hefner JT, 2016. Cranial nonmetric and morphoscopic data sets. In: Pilloud MA, and Hefner JT, editors. Biological distance analysis: forensic and bioarchaeological perspectives. San Diego, CA: Academic Press. pp. 91-108.

Quinlan JR. 1987. Simplifying decision trees. Int J Man Mach Stud 27:221-234.

R Core Team. 2016. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org.

Ramsthaler F, Kreutz K, Verhoff MA. 2007. Accuracy of Metric Sex Analysis of Skeletal Remains Using FORDISC based on a Recent Skull Collection. Int J Leg Med 121:477-82.

Rangel-Villalobos H, Munoz-Valle JF, Gonzalez-Martin A, Gorostiza A, Magana MT, Paez-Riberos LA. 2008. Genetic admixture, relatedness, and structure patterns among Mexican populations revealed by the Y-chromosome. Am J Phys Anthropol 135:448-461.

Relethford JH. 1994. Craniometric variation among modern human populations. Am J Phys Anthropol 95:53-62

Relethford JH. 2004. Global patterns of isolation by distance based on genetic and morphological data. Hum Biol 76:499-513.

Relethford JH. 2009. Race and global patterns of phenotypic variation. Am J Phys Anthropol 139:16-22.

Relethford JH. 2010. Race and the conflicts within the profession of physical anthropology during the 1950s and 1960s. In: Little MA, and Kennedy KAR, editors. Histories of American physical anthropology in the twentieth century. New York: Lexington Books. Pp. 207-219.

Relethford JH, and Harpending HC. 1994. Craniometric variation, genetic theory, and modern human origins. Am J Phys Anthropol 95:249-270.

Retzius A. 1843. Om formen af nordboernes cranier. Stockholm: P.A. Norstedt.

Rhine S. 1990. Non-metric skull racing. In: Gill G, and Rhine S, editors. Skeletal Attribution of Race: Methods for Forensic Anthropology. Albuquerque, New Mexico: Maxwell Museum of Anthropology. pp. 9-20.

Rhine S. 1993. Skeletal criteria for racial attribution. Natl Assoc Prac Anthropol Bull 13:64-67.

Rieger A, Hothorn T, and Strobl C. 2010. Random forests with missing values in the covariates. URL http://epub. ub. uni-muenchen. de/11481.

Ripley WZ. 1899. The races of Europe: a sociological study. New York: D. Appleton and Company.

Rish I, Hellerstein J, and Thathachar J. 2001. An analysis of data characteristics that affect naïve Bayes performance.

Ritchie-Scott A. 1918. The correlation coefficient of a polychoric table. Biometrika 12:93-133.

Roseman CC, and Weaver TD. 2004. Multivariate apportionment of global human craniometric diversity. Am J Phys Anthropol 126:257-263.

Rosenberg NA, Pritchard JK, Weber JL, Cann HM, Kidd KK, Zhivotovsky LA, and Feldman MW. 2002. Genetic structure of human populations. Science 298:2381-2385.

Ross AH, Slice DE, Ubelaker DH, and Falsetti AB. 2004. Population affinities of 19th century Cuban crania: Implications for identification criteria in south Florida Cuban Americans. J Forensic Sci 49: 1-6.

Rubi-Castellanos R, Martinez-Cortes G, Munoz-Valle JF, Gonzalez-Martin A, Cerda-Flores RM, Anaya-Palafox M, and Rangel-Villalobos H. 2009. Pre-Hispanic Mesoamerican demography approximates the present-day ancestry of Mestizos throughout the territory of Mexico. Am J Phys Anthropol 139:284-294.

Saker M, Ogle OE, and Dym H. 2008. Basic and complex exodontia and surgical management of impacted teeth. In: Fonseca R, Marciani R, and Turvey T, editors. Oral and Maxillofacial Surgery vol. 1, 2nd edition. Philadelphia: Elsevier. pp 185-211.

Sauer NJ. 1992. Forensic anthropology and the concept of race: If races don't exist, why are forensic anthropologists so good at identifying them? Soc Sci Med 34:107-111.

Sauer NJ. 1993. Applied anthropology and the concept of race: A legacy of Linnaeus. NAPA Bull 13:79-84.

Scheuer L, and Black S. 2004. The juvenile skeleton. San Diego: Elsevier Academic Press.

Schmidt CW. 2008. Forensic dental anthropology: issues and guidelines. In: Irish JD, and Nelson GC, editors. Technique and application in dental anthropology. Cambridge: Cambridge University Press. p 266-292.

Scott GR. 1973. Dental morphology: a genetic study of American white families and variation in living Southwest Indians. PhD dissertation. Tempe, AZ: Arizona State University.

Scott GR. 1977. Classification, sex dimorphism, association and population variation of the canine distal accessory ridge. Hum Biol 49:453-469.

Scott GR. 1980. Population variation of Carabelli's trait. Hum Biol 52:63-78.

Scott GR, and Irish JD. 2017. Human tooth crown and root morphology. New York: Cambridge University Press.

Scott GR, and Turner II CG. 1988. Dental anthropology. Ann Rev Anthropol 17:99-126.

Scott GR, and Turner II CG. 1997. The anthropology of modern human teeth: dental morphology and its variation in recent human populations. Cambridge: Cambridge University Press.

Scott GR, Anta A, Schomberg R, and de la Rúa C. 2013. Basque dental morphology and the 'Eurodont' dental pattern. In: Scott GR, and Irish JD, editors. Anthropological Perspectives on Tooth Morphology: Genetics, Evolution, Variation Cambridge: Cambridge University Press. p. 296-318.

Scott GR, Navega D, Coelho J, Cunha E, Irish JD, 2016. rASUDAS: A New Method for Estimating Ancestry from Tooth Crown and Root Morphology. Proceedings of the American Association of Physical Anthropologists 85th Annual Meeting. Atlanta, GA, April 12-16.

Shih YS. 2004. A note on split selection bias in classification trees. Comp Stat Data An 45:457-466.

Shortell T. 2001. An Introduction to Data Analysis and Presentation: Correlations. Sociology 712 [Internet]. Available from: http://www.shortell.org/book/chap18.html.

Smay D and Armelagos GJ. 2000. Galileo wept: A critical assessment of the use of race in forensic anthropology. Trans Anthropol 9:19-29.

Smith BH, Garn SM, Hunter WS. 1986. Secular trends in face size. The Angle Orthodontist 56:196-204.

Snow CC, Hartman S, Giles E, and Young FA. 1979. Sex and race determination of crania by calipers and computer: A test of the Giles and Elliot discriminant functions in 52 forensic science cases. J Forensic Sci 24: 448–460.

Sofaer JA, MacLean CJ, and Bailit HL. 1972a. Heredity and morphological variation in early and late developing human teeth of the same morphological class. Arch Oral Biol 17:811-816.

Sofaer JA, Niswander JD, MacLean CJ, and Workman PL. 1972b. Population studies on Southwestern Indian tribes V. tooth morphology as an indicator of biological distance. Am J Phys Anthropol 37: 357-366.

Sparks CS, and Jantz RL. 2002. A reassessment of human cranial plasticity: Boas revisited. PNAS 99:14636-14639.

Spencer F. 1981. The rise of academic physical anthropology in the United States (1880-1980): a historical overview. Am J Phys Anthropol 56:353-364.

Spradley MK, and Weisensee K. 2013. Why do forensic anthropologists estimate ancestry, and why is it so controversial?. In: Tersigni-Tarrant MTA, and Shirley NR, editors. Forensic anthropology: An introduction. pp. 231-243.

Spradley MK, Jantz RL, Robinson A, and Peccerelli F. 2008. Demographic change and forensic identification: problems in metric identification of Hispanic skeletons. J Forensic Sci 53:21-8.

Stewart TD. 1961. UNESCO statements on race. Science 133:1632-1634.

Stewart TD. 1962 Anterior femoral curvature: Its utility for race identification. Hum Biol 34:49-62.

Stini WA. 2010. Sherwood L. Washburn and 'the new physical anthropology'. In: Little MA, and Kennedy KAR, editors. Histories of American physical anthropology in the twentieth century. Lanham, MD: Lexington Books. pp 173-185.

Strobl C, Boulesteix A, Kneib T, Augustin T, and Zeileis A. 2008. Conditional variable importance for random forests. BMC Bioinformatics 9:307-318.

Strobl C, Boulesteix A, Zeileis A, and Hothorn T. 2007. Bias in random forest variable importance measures: illustrations, sources and a solution. BMC Bioinformatics 8:25-46.

Strobl C, Malley J, and Tutz G. 2009. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. Psych Meth 14:323-348.

Swindler DR, and Ward S. 1988. Evolutionary and morphological significance of the deflecting wrinkle in the lower molars of the hominoidea. Am J Phys Anthropol 75: 405-411.

Ta'ala SC. 2015. A brief history of the race concept in physical anthropology. In: Berg GE, and Ta'ala SC, editors. Biological affinity in forensic identification of human skeletal remains: Beyond black and white. Boca Raton, FL: CRC Press. pp. 1-16.

Templeton AR. 1998. Human races: a genetic and evolutionary perspective. Am Anthropol 100:632-650.

The Scientific Working Group for Forensic Anthropology (SWGANTH). 2013. SWGANTH Home Page.< http://swganth.org/index.html>. Accessed 2015 October 7.

Thomas DH. 2001. A short history of scientific racism in America. In: Skull wars: Kennewick Man, archaeology, and the battle for Native American Identity. pp 37-43.

Townsend G, and Brook A. 2008. Genetic, epigenetic, and environmental influences on dental development. Ortho Tribune 3:4-6.

Townsend G, Hughes T, Luciano M, Bockmann M, and Brook A. 2009. Genetic and environmental influences on human dental variations: a critical evaluation of studies involving twins. Arch Oral Biol 54:S45-S51.

Townsend G, Yamada H, and Smith P. 1990. Expression of the entoconulid (sixth cusp) on mandibular molar teeth of an Australian aboriginal population. Am J Phys Anthropol 82:267-274.

Trigger BG. 1965. Coon's theory on "The Origin of Races." Anthropologica 7:179-187.

Truesdell ND. 2005. Secular change in the skull between American Blacks and Whites. Master's Thesis. Baton Rouge, LA: Louisiana State University.

Turner II CG. 1967. Dental genetics and microevolution in prehistoric and living Koniag Eskimo. J Dent Res 46:911-917.

Turner II CG. 1971. Three-rooted mandibular first permanent molars and the question of American Indian origins. Am J Phys Anthropol 34:229-242.

Turner II CG. 1983. Sinodonty and Sundadonty: a dental anthropological view of Mongoloid microevolution, origin, and dispersal into the Pacific Basin, Siberia, and the Americas. In: Vasilievsky RS, editor. Late Pleistocene and Early Holocene Cultural Connections of Asia and America. Novosibirsk, Siberia: USSR Academy of Sciences, Siberian Branch. pp 72-76.

Turner II CG. 1985. The dental search for Native American origins. In: Kirk R, and Szathmary E, editors. Out of Asia: Peopling the Americas and the Pacific. Canberra, Australia: The Journal of Pacific History, Inc. pp 51-78.

Turner II CG. 1987. Late Pleistocene and Holocene population history of East Asia based on dental variation. Am J Phys Anthropol 73:305-321.

Turner II CG. 1990. Major features of Sundadonty and Sinodonty, including suggestions about East Asian microevolution, population history, and Late Pleistocene relationships with Australian aboriginals. Am J Phys Anthropol 82:295-318.

Turner II CG, and Scott GR. 1977. Dentition of Easter Islanders. In: Dahlberg AA, and Graber TM, editors. Orofacial growth and development. The Hague: Mouton. p 229-249.

Turner II CG, Nichol CR, and Scott GR. 1991. Scoring Procedures for Key Morphological Traits of the Permanent Dentition: The Arizona State University Dental Anthropology System. In: Kelley MA, and Larsen CS, editors. Advances in Dental Anthropology. New York: John Wiley and Sons, Inc.

U.S. Customs and Border Patrol. 2012. United States Border Patrol Sector Profile – Fiscal Year 2012. U.S. Border Patrol.

UNESCO. 1951. Statement on the nature of race and race differences, by physical anthropologists and geneticists—June 1951. In: The race concept: results of an inquiry. Westport, CT: Greenwood Press. pp 11-16.

United States Census Bureau. 2010. United States Census 2010. <http://www.census.gov/2010census/>. Accessed 9 October 2015.

Visweswaran K. 1998. Race and the culture of anthropology. Am Anthropol 100:70-83.

Vogel ER, van Woerden JT, Lucas PW, Utami Atmoko SS, van Schaik CP, and Dominy NJ. 2008. Functional ecology and evolution of hominoid molar enamel thickness: Pan troglodytes schweinfurthii and Pongo pygmaeus wurmbii. J Hum Evol 55:60-74.

von Cramon-Taubadel N. 2011. Global human mandibular variation reflects differences in agricultural and hunter-gatherer subsistence strategies. Proc Natl Acad Sci 108:19546-19551.

von Koenigswald GH. 1982. Distribution and evolution of the orangutan, Pongo pygmaeus (Hoppius). In: de Boer LEM, editor. The Orangutan: its biology and conservation. Boston, MA: Dr. W. Junk Publishers. p 1-15.

Warren MW, Parr NM, Skorpinski KE, and Zambrano CJ. 2011. Bare bones: A survey of forensic anthropology. Dubuque, IA: Kendall Hunt Publishing.

Washburn SL. 1951. The new physical anthropology, Trans NY Acad Sci 13:298-304.

Weidenreich F. 1937. The dentition of *Sinanthropus pekinensis*: a comparative odontography of the hominids. Geological Survey of China.

Weidenreich F. 1946. Apes, giants, and man. Chicago: University of Chicago Press.

Welkowitz J, Ewen RB, Cohen J. 1976. Introductory statistics for the behavioral sciences. New York: Academic Press.

Wescott DJ. 2005. Population variation in femur subtrochanteric shape. J Forensic Sci 50:286-293.

Wescott DJ, and Jantz RL. 2005. Assessing craniofacial secular change in American Blacks and Whites using geometric morphometry. In: Slice DE, editor. Modern morphometrics in physical anthropology. New York: Kluwer Academic/Plenum Publishers. p 231-245.

Wescott DJ, and Moore-Jansen PH. 2001. Metric variation in the human occipital bone: Forensic anthropological applications. J Forensic Sci 46:1159-1163.

White AP, and Liu WZ. 1994. Technical note: bias in information-based measures in decision tree induction. Mach Learn 15:321-329.

Willermet C, Daniels J, and Edgar H. 2016. Seeing RED: a new statistical solution to an old categorical data problem. Poster presented to the American Association of Physical Anthropologists 85th Annual Meeting. Atlanta, GA, April 12-16.

Williams FL, Belcher RL, and Armelagos GJ. 2005. Forensic misclassification of Ancient Nubian crania: implications for assumptions about human variation. Curr Anthropol 46:340-346.

Wilson RJ, Algee-Hewitt B, and Jantz LM. 2007. Demographic trends within the Forensic Anthropology Center's body donation program. Proceedings of the 67th Annual Meeting of the American Association of Physical Anthropologists. Philadelphia, PA, March 28-31.

Witkowski JA. 2017. 1950: Origin and Evolution of Man, Vol XV. Cold Spring Harbor Symposia on Quantitative Biology [cited 2017 February 20]. Available from: http://symposium.cshlp.org/site/misc/topic15.xhtml.

Wolpoff MH, and Caspari R. 1997. Race and human evolution: A fatal attraction. New York: Simon and Schuster.

Wright S. 1921. Systems of mating. III. Assortative mating based on somatic resemblance. Genetics 6:144-161.