Alas, Poor Yorick: A DNA Analysis of Ancestry Using Crania

Claire Hanson
The University Of Montana

Follow this and additional works at: https://scholarworks.umt.edu/etd

Part of the Biological and Physical Anthropology Commons

Let us know how access to this document benefits you.

Recommended Citation
Hanson, Claire, "Alas, Poor Yorick: A DNA Analysis of Ancestry Using Crania" (2020). Graduate Student Theses, Dissertations, & Professional Papers. 11586.
https://scholarworks.umt.edu/etd/11586

This Thesis is brought to you for free and open access by the Graduate School at ScholarWorks at University of Montana. It has been accepted for inclusion in Graduate Student Theses, Dissertations, & Professional Papers by an authorized administrator of ScholarWorks at University of Montana. For more information, please contact scholarworks@mso.umt.edu.
Alas, Poor Yorick: A DNA Analysis of Ancestry In Comparison To Crania
Chairperson: Dr. Meradeth Snow

Abstract
Current methods for estimating ancestry in biological profiles provided to law enforcement and forensic collections rely heavily on the use of metric and non metric methods such as FORDISC 3.1 and “trait lists.” The reliance on these methods has led to inaccurate and ambiguous ancestry estimations, as many metric and non metric methods are subjective and cannot be statistically tested for error rates. As a result, forensic collections, such as the University of Montana Forensic Collection, could potentially house remains that have been inaccurately curated. In order to test the accuracy of curation in the UMFC, DNA samples were extracted from six individual crania. The samples were then sequenced in order to determine mtDNA haplogroups for each individual, which provide insight into the likely geographic region each individual is from. These results were then compared to FORDISC 3.1 and non metric trait list findings in order to analyze the disparity between method results. The greatest disparity resided between FORDISC 3.1 findings and mtDNA haplogroups. This discrepancy suggests that the widely used metric method of FORDISC 3.1 is not a reliable ancestry testing method and should not be the sole method used to determine ancestry for biological profiles. Instead, DNA evidence, such as mtDNA, Y-DNA, or AIMs should be included when assessing ancestry in order to strengthen estimations for law enforcement and forensic collections.
Acknowledgements

First and foremost, I would like to thank my committee chair and advisor Dr. Meradeth Snow for not only her patience and help, but also for training me in lab techniques necessary for this project. I appreciate her guidance and willingness to let me rant in the ancient lab and talk about fantasy novels. Without her, I would not have been able to complete my thesis, perform a bead clean, or survive my Master’s program.

I would also like to thank my committee member, Dr. Kirsten Green-Mink, for teaching me everything I know about osteology, providing multiple opportunities for me to practice osteological lab work, and for giving me constant advice on future opportunities. I would also like to thank Dr. Chris Palmer for his willingness to join my committee late in the game to help a stressed-out grad student. I owe a further debt of gratitude to Dr. Randall Skelton, for providing information on the cases used for this project.

Thank you to my family and friends for listening to me talk about human bones incessantly for two years and for your unending patience and support. I extend thanks to my mother, in particular, had to sit through many a phone call calming down an overwhelmed grad student. Also, thank you to my partner, who never knows what I am talking about, but is always willing to listen to me rant and give me advice.

I express my thanks to ASUM for providing a Research and Creative Scholarship Fund so that my research could continue as well as Dr. David Xing at the University of Montana Genomics core for providing an Illumina Sequencing voucher to cover the costs of DNA sequencing.

Finally, I would like to thank coffee and wine. You have ruined my teeth and intestines, but you got me through quite a few long days and late nights. My sincerest gratitude.
# Table of Contents

Abstract.........................................................................................................................i
Acknowledgments........................................................................................................ii
List of Tables................................................................................................................vi
List of Figures...............................................................................................................vii
1.0 Introduction.............................................................................................................1
   1.1 Forensic Ancestry Assessment..............................................................................2
      1.1a Metric Assessments.......................................................................................3
      1.1b Non Metric Assessments...............................................................................7
      1.1c Postcranial Assessments...............................................................................12
   1.2 Limitations of Morphological Assessments.......................................................13
   1.3 mtDNA Assessments..........................................................................................15
      1.3a Y-Chromosome Lineages..............................................................................17
      1.3b Ancestry Informative Markers......................................................................18
   1.4 NAGPRA.............................................................................................................20
   1.5 Research Significance.........................................................................................20
2.0 Materials and Methods..........................................................................................22
   2.1 Research Sample................................................................................................22
   2.2 FORDISC 3.1....................................................................................................25
   2.3 DNA Extraction.................................................................................................30
      2.3a Drilling............................................................................................................30
      2.3b Extraction, Library prep, and Sequencing.....................................................31
   2.4 Sequencing, Galaxy, and Haplogrep...................................................................35
3.0 Results..................................................................................................................36
  3.1 FORDISC Results...............................................................................................36
  3.2 mtDNA Results..................................................................................................44
4.0 Discussion..............................................................................................................48
  4.1 Metric and Non Metric.......................................................................................48
  4.2 mtDNA Haplogroups.........................................................................................50
  4.3 Further Research...............................................................................................54
5.0 Conclusion..............................................................................................................55
List of Tables

- Table 2a: Measurements taken for all six crania for input into FORDISC 3.1
- Table 2b: Weight of bone dust collected
- Table 2c: Extraction protocol and Qubit results post-extraction
- Table 2d: Library Preparation and Qubit quantification post adapter and indexing
- Table 3a: FORDISC 3.1 results including chi-square and posterior probability
- Table 3b: Key of listed populations in FORDISC 3.1 graphs
- Table 3c: OSSA Scoring Sheet Ancestry Results
- Table 3d: Rhine and Gill (1995) trait list for three distinct ancestral groups
- Table 3e: results of Galaxy and Haplogrep
- Table 3f: Amount of damage to DNA seen on plots and modernity of samples
- Table 4a: Depiction of all methods and respective results for each sample
List of Figures

● Figure 1a: Post-bregmatic depression as seen in OSSA
● Figure 2a: UMFC 80; sample 1
● Figure 2b: UMFC 98; sample 2
● Figure 2c: UMFC 103; sample 3
● Figure 2d: UMFC 155; sample 4
● Figure 2e: UMFC 104; sample 5
● Figure 2f: 141; sample 6
● Figure 3a: Walker (2008) Sex estimation for UMFC 141
● Figure 3b: Example of Standards cranial measurements from sagittal plane and Maximum cranial breadth and bizygomatic diameter
● Figure 3b: UMFC 80 FORDISC 3.1 ancestry graph results
● Figure 3c: UMFC 98 FORDISC 3.1 ancestry graph results
● Figure 3d: UMFC 103 FORDISC 3.1 ancestry graph results
● Figure 3e: UMFC 155 FORDISC 3.1 ancestry graph results
● Figure 3f: UMFC 104 FORDISC 3.1 ancestry graph results
● Figure 3g: UMFC 141 FORDISC 3.1 ancestry graph results
● Figure 3h: mtDNA damage plots for all individuals
1.0 Introduction

The widely used forensic database known as FORDISC 3.1 is home to the cranial and post-cranial measurements of over 3400 individuals, some of which are forensic in nature and were found in the United States and others whose measurements were added due to William White Howells’ research on cranial measurements across 28 historic and archaeological populations worldwide (Ousley and Jantz 2005). This database uses discriminant function analysis in order to determine ancestry of an individual after cranial measurements have been conducted following the standard guidelines (Buikstra and Ubelaker 1994; Langley et al. 2016). FORDISC 3.1 attempts to classify the individual into an ancestral group based on the similarities in cranial measurements between the individual and those individuals already in the database.

Though beneficial in creating a biological profile, FORDISC 3.1 has a notoriously limited scope. Because samples provided by FORDISC 3.1 are from predominantly European (35%) and African individuals (15%), the results provided on remains that do not fall into these categories are often either skewed or they cannot be grouped with a reasonable degree of accuracy (Ousley and Jantz 2005). This is highly problematic for many reasons; in particular, the curation of the remains is dependent on the ancestry of the individual. It begs the question: how many of the remains in forensic or osteological collections have been sorted into an inaccurate biological profile due to FORDISC 3.1’s shortcomings?

Another issue that Forensic Anthropologists face when determining ancestry of remains is the overwhelming probability that the individual was of mixed ancestral descent, though this is admittedly more of a modern problem as mixed ancestral descent is less common in ancestral cases. Genetic admixture in populations is common due to colonialism and increased globalization, so much so that it makes it difficult to discern which metric and nonmetric traits
belong to which population. While most forensic anthropologists remedy this by claiming the remains belonged to a person of “mixed ancestry,” the accuracy of the projected ancestries provided by FORDISC 3.1 are once again called into question.

This thesis project aims to test the validity of FORDISC 3.1 ancestral assessments through the sequencing of the mitochondrial DNA from subjects that have been tested with FORDISC 3.1 to establish ancestry. The results of this testing will be compared to the original metric and nonmetric analyses, as found in Rhine and Gill’s (1990; 1995) trait list and OSSA (Hefner 2009), conducted on the individuals and should display the disparity between the results of ancestry estimation. If FORDISC 3.1 alongside nonmetric methods establishes an ancestry that is not in accordance with the mtDNA results, it will be apparent that the online database is not entirely capable of providing an accurate ancestry estimation of unknown individuals. My goal in conducting this research is to not only shed light on the need for an updated database of forensic samples and address the ethics of the curation process, but also to promote the usage of DNA analysis in the forensic context.

1.1 Forensic Ancestry Assessment

When creating a biological profile, forensic anthropologists must have access to a mostly complete crania and post crania in order to complete analyses of age, sex, ancestry, and stature. Features on the crania and long bones allow forensic anthropologists to determine said biological factors. These features include the size and shape of certain portions of bone, as well as the distance between two points on the crania or long bones. While age and sex differences exhibit distinct features, ancestry is more fluid. Many of these assessments are metric in nature,
conducted with sliding and spreading calipers, but there is a plethora of non-metric assessments that add to the overall decision of each aspect of the biological profile.

There is a wide array of existing literature on ancestral assessments of forensic remains, though this literature has changed exponentially as ideas towards ancestry have shifted within anthropology. Indeed, past anthropologists supported the idea that “race” was a biological concept as a means to justify the racial superiority of European men as well as support nationalist ideologies (Linné and Salvius 1758; Bhopal 2007; Meloni 2017; Dunn et al 2020). Although scientific racism was officially denounced in 1950 by UNESCO in the seminal paper “The Race Question,” antiquated terms and methods were still used well into the 1970’s and 80’s prior to the discovery of mtDNA and Y-DNA as a means to analyze ancestry, which helped separate the terms race and biology (Cann, Stoneking, and Wilson 1987; Meloni 2017). Current anthropologists typically strive to use the term ancestry as opposed to race in order to remove ourselves from scientific racism, and ancestry assessment methods no longer rely on past inherently racist literature. In fact, more recent methods of ancestry assessments are focused on small morphological variation within populations rather than variation between populations and recent ancestral estimation studies acknowledge that ancestry better reflects how and what morphological variations are inherited (Spradley and Jantz 2016; Hefner and Linde 2018; Atkinson and Tallman 2019; Cunha and Ubelaker 2019; Shi et al. 2019; Jin et al. 2020; Dunn et al. 2020).

Furthermore, it is extremely rare for only one method to be used to confirm the ancestry of an individual and forensic anthropologists typically utilize multiple methods in order for increased accuracy of ancestry estimation. Therefore, I will focus on the metric and nonmetric methods most commonly used by forensic anthropologists to determine ancestry in addition to
discussing the usefulness of these methods alongside the biases that plague many of these methods. Existing methods must be addressed thoroughly before a comparison can be made with the more modern DNA usage to confirm ancestry.

1.1a Metric Assessments

It is the belief of anthropologists that remains with similar craniometric morphology are more likely to be from the same population than those with differing morphological traits. This assumption has guided much of the study of ancestry in anthropology; there have been countless studies regarding the cranial variation between populations. The most notable of these studies is W.W. Howell’s work with 28 archeological and unknown age populations of varying geographical location, which illuminates the variation of crania size between groups as well as sexual dimorphism within groups (Howells 1973;1989). While studies like Howells’ have shown that populations living in different geographic regions will display differing cranial morphology due to their environment, the research is not definite due to the notably small sample size of a few of the populations (Tsukumo and Ryukyu) and is, admittedly, dated by its terminology and assumptions of ancestral traits.

The metric assessments used when creating a biological profile are employed for both cranial and postcranial remains as well as dental metrics. Ancestral measurements, however, are typically confined to the crania, as there are fewer ancestral markers found on the long bones outside of the obvious platymeria of the femora, which is found in mostly Native American populations (Gill and Rhine 1995). Postcranial measurements, though, can be used when the crania is absent or fragmented/damaged. Metric methods are more reliable than nonmetric methods in forensic casework, as they are inherently less subjective than the non-metric methods.
due to their reliance on distance between landmarks on the bone, which can be substantiated by statistics and therefore removes observer error (Webster and Sheets 2010; Ross, Slice, and Williams 2010; Spradley and Jantz 2016).

Buikstra and Ubelaker compiled *Standards* (1994) which outlines the standardization of data collection in forensic and historical cases, examining 24 cranial measurements and 10 mandibular measurements through the use of sliding calipers, spreading calipers, and an osteometric board. These standards are used only for adult remains, as the juvenile skull and long bones are not completely fused, and prominent features that allow us to assess ancestry and sex are not yet apparent (White and Folkens 2005). If taken correctly, the craniometric method allows for a description of skull size and shape, and potentially presents an individual who can be grouped within a population with similarly formed cranial measurements. Cranial morphology is known to show degrees of genetic divergence and similarities between populations, though the degree to which the morphology is linked to adaptive changes or population history is contested (Hubbe, Hanihara, and Harvati 2009; Smith 2009).

The issue with metric methods lies with the taking of the measurements themselves. Although instructions on where to place the calipers for each measurement are clearly stated, the possibility for error ranges from .05% to >5% (Langley et al. 2018). This error is inherent in measurements such as these, as forensic anthropologists may have conflicting views on where the calipers are to be placed on the skull. The solution to this problem seems to be the utilization of three-dimensional landmark data, known as geometric morphometrics. Geometric morphometrics is particularly useful as it uses 3D-ID and geometric morphometric methods to identify morphological variation. Geometric morphometrics also only relies on the morphology of the trait rather than linear distances and ratios or inter landmark distances to create a
biological profile (Webster and Sheets 2010; Ross, Slice, and Williams 2010; Spradley and Jantz 2016). For example, a past iteration of computer software used to study ancestry, *Morphologika 2*, is used to analyze variation in shape and size of crania from given 3D landmarks, which offer results that are less likely to suffer from interobserver error (O’Higgins and Jones 2006; Smith 2009; Katherine Spradley and Jantz 2016). In one particular instance using *Morphologika 2*, the researchers were able to make a comparison between the distances of variation from the craniometric morphological data and the molecular distances found through sequencing DNA. From the results of this study, it was discovered that the 3D landmarks of the entire crania, rather than just the cranial vault and 2D FORDISC data, are consistent in identifying genetic relationships in humans and, by extension, the supposed ancestry of an individual. Indeed, it seems that geometric morphometrics may be a more accurate way to measure the variation within human crania than long standing caliper-based measurements, since then, more recently developed software by Ousley (2014) known as 3Skull, which allows the user to create their own novel landmarks as well as utilize the collection of 108 bony landmarks, is more commonly accessed than other 3D landmark computer software (Kenyhercz, Klales, and Kenyhercz 2014; Ousley 2014; Dunn et al. 2020). As geometric morphometrics research is furthered, it is clear that the reliance on 3D landmarks to assess ancestry is not only more efficient, but produces far more accurate results with the decrease in interobserver error (<.05%) (Dunn et al. 2020). By studying these 3D landmarks via a digitizer, a more holistic analysis of ancestry may be reached, this in comparison with the potential addition of genetics to the biological profile could potentially ensure a much smaller error rate in ancestral assessments as well as make possible the determination of ancestry outside of the umbrella terms (i.e. Asian, European, African, etc.)
It is no surprise then that the morphology of the long bones is used, alongside cranial morphology, as an indicator of ancestry. As previously mentioned, features on the post-crania that have been dubbed “ancestral markers” are not commonly used to determine ancestry in a forensic context. This is most probably due to the fact that there are very few features on long bones that can accurately predict the population an individual belongs to as well as a lack of standardization, limited scope of population differentiation, and the lack of known error rates as a way to measure reliability (Dunn et al. 2020). Although morphometric data on the postcrania is not often used to assess ancestry, research in the past has suggested that elements of the postcrania such as the os coxae, and femora can indicate a potential ancestry (White and Folkens 2005). One study suggested that measurements taken from the os coxae, including bi-iliac breadth and the transverse diameter and anteroposterior height of the pelvic inlet, could allude to a possible ancestry if the age of the individual was known prior to taking measurements. This research, however, included only European and African individuals and displayed a decrease in accuracy of ancestral classification as compared to previous studies (Işcan and Cotton 1985). Further studies focusing on the intercondylar notch of the distal femur and anterior-posterior diameter of the proximal femur seem to have a higher level of accuracy than other postcranial elements when determining ancestry, although the sample sizes of these studies were relatively small and ancestry of the individuals was already known by researchers conducting said studies (Baker et al 1999; Gill 2001; Shelbourne, Gray, and Benner 2007). While craniometrics appear to be more reliable, Leibenberg et al. (2015) found that the accuracy ratings of determining ancestry from postcranial elements is increased when more than one postcranial element is measured rather than individual bones; this increase is measured at 85%, which is similar to results found from craniometrics (Spradley 2016; Liebenberg, L’Abbé, and Stull 2015). This
however, should not be taken as an indication that metric and nonmetric analyses of the post crania is enough to determine the ancestry of an individual without the crania present.

1.1b Nonmetric Assessments

Typically, in forensic casework, ancestral assessments are conducted via nonmetric means. The study of morphology through nonmetric means is qualitative in nature as nonmetric traits are quasi-continuous and focus on small variations in the morphology of bone (Dunn et al. 2020). The most commonly used methods are detailed in Standards (1994) which describes both primary and supplementary nonmetric cranial traits. The variation found in nonmetric data is reliant on the presence or absence of a few particular traits including metopic suture, parietal foramina, extra bones at pterion, wormian bones, and mylohyoid bridges, which is known as the “trait list” approach (Rhine and Gill 1990; White and Folkens 2005; Dunn et al. 2020). The reliance on the presence or absence of these traits is due to the plethora of evidence suggesting that certain variance in sutures — i.e. wormian bones and metopic sutures — are most often found in two particular populations: Asian and African respectively (White and Folkens 2005). In Rhine and Gill’s (1990;1995) compiled list of useful nonmetric traits, nonmetric traits are separated into European, African, and Native American ancestral groups. These traits, which are confined to cranial nonmetric assessments, rely heavily on the presence or absence of supposed population markers such as inion hooks, post bregmatic depressions, and wormian bones. As expected, the assessments of these traits are highly subjective and do not account for any admixture between these distinct groups as the existence of these markers suggest a higher probability of a certain ancestry, but traits such as wormian bones and a pronounced external occipital protuberance (inion hook) exist in more than one population. They also rely on the
extreme expressions of these traits, which are not representative of an entire population, as most non metric traits such as ossicles, foramina placement, and vault sutures are skeletal anomalies rather than indicators of population origin (Hefner 2009; Atkinson and Tallman 2019; Dunn et al. 2020). If the creation of a biological profile relies on the trait list approach, the analysis is inherently lacking and the determined ancestry can be scrutinized in a court of law.

Further non metric methods include examining variation of the orbital rim shape, which past studies have suggested is indicative of bio-geographic ancestry. Forensic anthropologists have concluded that sloping orbital rims indicate a European ancestry, while rectangular orbital rims indicate African ancestry and round Asian ancestry (Rhine and Gill 1990). As is the case with most non metric assessments, this method is rife with observer bias. It has been found that those analyzing the remains are more likely to “see” a particular eye orbit shape or other ancestral trait based on the gestalt of skull or information provided through law enforcement, even if the trait is not, in fact, present (Hughes et al. 2011; Rubin and DeLeon 2017). In a 2017 study by Rubin and DeLeon it was determined that while the curvilinear relationships of the eye orbits is potentially the most ancestrally informative aspect of orbital rim shape, the exact nature of this relationship could not be determined, further justifying the need to adjust forensic reliance on non metric traits. Perhaps the most informative non metric data comes from teeth and palate shape, due to the fact that tooth size and morphology are likely linked to genetics. In accordance with this, human tooth morphology is also subject to environmental changes, that are either due to individual diet or health, as well as adaptations that may be regional in distribution, which helps to differentiate between living human populations as the majority of skeletal morphology is influenced by the environment (White and Folkens 2005; Rathmann et al. 2017).
The Dahlberg method of trait classification, created in 1956, posits that morphological variables such as a Carabelli’s cusp, shoveling on the incisors, and variation in the hypocone are indicative of bio-geographic ancestry (Scott and Turner 2000). Apart from the Dahlberg method, a common ancestry estimation is done using the palate shape of an individual as described by Rhine and Gill (1990). Those of African ancestry have a widely arched palate, European ancestry shows a moderately arched palate, and both Asian and Native American ancestries present a narrow palate (Rhine and Gill 1990; White and Folkens 2005). When this ancestral estimation was put to the test recently, the research showed that there is not enough variation in palate shape between populations to accurately estimate ancestry. There is, in fact, more variation in palate shape within populations than posited by the Rhine and Gill and Dahlberg methodologies (Clark et al. 2016). That does not mean that assessment of palate shape is not useful; rather, it should not be employed alone. One study showed that analysts were able to correctly assess ancestry via palate shape only 40% of the time (Maier et al. 2015). While there is evidence of secular change in palate shape and depth depending on population, without the aid of digital equipment, this variation is not yet understood (Maier et al. 2015).

In regard to cranial non metric traits and methods, extreme expression of morphoscopic traits, traits that are reliant on shape and form, are not valid in determining ancestry because they do not fall in line with typical ancestral assumptions of trait morphology (Hefner 2009). Extreme morphoscopic traits are, however, often cited in forensic casework. There is also evidence to suggest that our lack of knowledge on morphoscopic trait variation within a population leads professionals to assume there is admixture when, in fact, there is none (Hefner 2009). In order to account for the potential failings of morphoscopic trait variation, the Optimized Summed Scored Attributes method (OSSA) was created, which estimates ancestry for Black and White
individuals from identified morphoscopic traits. OSSA uses six morphoscopic traits in which each trait is dichotomized in order to illuminate between group differences. In OSSA, each trait can be ranked 1-3 or 0 in the case of absent traits (most commonly post-bregmatic depression).

![Figure 1a: Post-bregmatic depression as seen in OSSA](image)

This system suggests that the reliance on macromorphoscopic traits reduces interobserver error so long as the practitioners have adequate experience scoring traits of a crania. It has been suggested that OSSA, in conjunction with Hefner’s (2009) data collection procedures, allow professionals to assess ancestry of European and African remains with near complete certainty (Klales and Kenyhercz 2015; Hefner and Ousley 2014; Atkinson and Tallman 2019). There are, however, several weaknesses that plague the OSSA method. As outlined in Hefner and Ousley’s (2014) research, the OSSA method gives equal weight to each score, relies on the importance of a post-bregmatic depression, requires that the crania being scored is fully intact, does not calculate posterior probabilities, and is limited to African or European identification. Therefore, this method is most useful when the population is guaranteed to be either of European or African descent and does not include individuals who may be of mixed ancestry. While the use of OSSA
in comparison with FORDISC 3.1 results could potentially offer accurate results, particularly if
the individual was of European or African descent, it is difficult to assume the ancestry of the
individual from solely OSSA and FORDISC 3.1. In fact, neither can be used individually with
complete confidence due to the small sample size of possible ancestry identifications.

The issue remains that most non metric evaluations are conducted on European and
African American individuals with little research on Asian Americans and Native Americans.
Recent studies, however, propose that Asian and Asian-derived individuals express differences
in their cranial morphology, meaning that it is possible for forensic anthropologists to potentially
differentiate Asian ancestries rather than relying on the vague umbrella term, “Asian” (Atkinson
and Tallman 2019). With more nuanced non metric methods, there is potential that researchers
could accurately differentiate between Asian and Asian derived populations such as Native
Americans (Asian-derived) and populations such as the Japanese, Thai, and Filipino as seen in
Atkinson and Tallman’s (2019) study. While promising and incredibly beneficial to forensic
studies, the identification of ethnic groups within a population using only non metric traits
requires further study and much larger sample sizes before it can be wholly implemented. With
further refinement, the non metric traits found between ethnic groups within a widespread
population has the potential to change the way forensic anthropologists’ approach and create a
biological profile

1.1c Postcranial Assessments

As it pertains to postcranial ancestral assessments, the metric and non metric methods
used are often forgotten in forensic casework, though not without reason. Finnegan (1978)
provided an initial description of thirty nonmetric traits that could be useful in determining ancestry using the Terry Collection located at the Smithsonian. The traits studied, however, were highly susceptible to environmental change, as the research mostly relied on the shape, size, and location of articular surfaces (White and Folens 2005). Since then, further research has been conducted regarding the utility of postcranial traits in estimating ancestry, though there is minimal evidence of ancestry-specific postcrania traits outside of a bifurcated spinous process and the presence of a septal aperture (Trotter 1934; Asvat 2012; Spiros 2019). Similarly to the aforementioned research, a recent study compiling 11 postcranial traits showed the possibility of certain postcranial traits being linked to ancestry and found that only four traits displayed substantial variation between groups: C3 and C4 spinous process bifurcation, the septal aperture, the third trochanter, and the anterior and middle calcaneal facets (Spiros 2019). Although postcranial non metric methods have potential to be useful to forensics, these traits should not be used in isolation, and require more time and research before they should be implemented. For the purposes of this research, only cranial metric and non metric assessments will be completed due in part to the lack of available post cranial elements and also the inherent unreliability in both metric and non metric postcrania ancestry assessments.

1.2 Limitations of Morphological Assessments

It cannot be denied that the study of morphological traits has been beneficial to understanding the relation of shape and size change to bio-geographic ancestry. In fact, more current research on geometric morphometrics provides a promising outlook on the future of ancestral traits research. It has been noted, however, that morphological change on the bone does
not guarantee a correct determination of ancestry, indeed the reliance on cranial and postcranial elements to assess ancestry is concerning and problematic. This is in part due to the lack of acknowledging that many variations of morphological traits, such as: mandible shape, cranial size, long bone length and shape, and the os coxae, are the results of secular change rather than an effect of genetics (Cole 2003; Larnkjaer et al. 2007; Martin and Danforth 2009; Driscoll 2010). Secular traits are non-genetic changes and are therefore not due to evolution, but instead correspond with changes in living conditions that affect an entire population (Albanese 2010), which makes it more difficult, but not impossible to assess what traits can be considered “ancestral” and what traits are simply the consequence of secular change. Another issue with relying on morphological traits is the role that environmental stressors play on human cranial and postcranial morphology. Clearly, some variations in the human form are the product of environmental stress, such as bone health and density as well as stature (Ackhert-Bicknell and Karasik 2013). Malnutrition during youth can affect the growth of not only the long bones, but also the crania and dentition, causing potential malformations that could affect ancestral metric and non metric assessments (White and Folkens 2005). A final issue to consider is the ever-popular trait list approach that many anthropologists employ when creating a biological profile. Trait lists, as seen in Rhine (1990) and Rhine and Gill’s (1995) research, are lists of possible ancestral traits to choose from. After accounting for the presence and form of each trait listed, the ancestry with the most traits checked off is the ancestry the individual is most likely from. In the case of Rhine and Gill (1990; 1995) the determinations are between European, African, and Asian populations. The trait list does not provide a robust analysis of ancestry and does not rely on statistics to measure error rate (Dunn et al. 2020). Traits are also often assigned post hoc and have a high potential of observer-specific bias, further diminishing the value of the commonly
used list (Hefner 2009). Further ancestral research on the importance of morphological variation should not incorporate outdated, deficient methods, but instead should rely on mixed method models as well as statistical methods and computer software to reduce observer error.

1.3 mtDNA Assessments

Ancestry estimation has been more efficient since the discovery of mitochondrial DNA (mtDNA) and Y-Chromosome DNA as informants on lineage. In 1987, Cann et al proposed the Mitochondrial Eve hypothesis or the Recent African Origin hypothesis, which suggested all mtDNA in contemporary humans could be traced back to one African ancestor living some 200,000 years ago. Because mtDNA is maternally inherited, has a high mutation rate, and does not recombine like nuclear DNA, it is highly effective in population studies (Stoneking 1994). The accumulation of genetic variance through maternal lineages form phylogenetically related haplotypes known as mitochondrial haplogroups (Wallace, Brown, and Lott 1999; Wallace 2015; Chinnery and Gomez-Duran 2018). These haplogroups or lineages that began with “Mitochondrial Eve” started with a split into four lineages: L0, L1, L2, and L3. From these original lineages, it has been discovered that humans have since branched into a myriad of sub-haplogroups; globally it is proposed that the number of mtDNA haplogroups is 4,000. It should be noted, however, that the way these mtDNA haplotypes are grouped is not entirely consistent. Therefore, there could potentially be more or less haplotypes depending on the way they have been grouped. These groups can be found in the comprehensive phylogenetic tree on the online database Phylotree (van Oven and Kayser 2009; Brotherton et al. 2013; Chinnery and Gomez-Duran 2018).
In terms of identifying region of ancestry, when a haplogroup exists at a high frequency in one geographic location, it is possible that the haplogroup originated there, though this conclusion is not always proven to be correct. In fact, as populations continue to spread and migrate and as new mutations arise within said populations, it becomes increasingly difficult to pinpoint the exact origin of a mutation/haplogroup (Chiaroni et al 2009). High frequencies of specific haplogroups in certain geographic areas is beneficial for ancestral assessments, as it shows a reasonably accurate portrayal of an individual’s lineage. In fact, recent studies on mtDNA lineages have focused on discerning self-ascribed ancestry compared to mitochondrial haplogroups sequenced from an individual. Lao et al (2010) found that by studying matrilineal markers in several populations, they could not only discern rates of admixture, but also detect whether said admixture was sex-biased, thus enabling anthropologists to reconstruct past lifeways. Matrilineal lineage studies are also beneficial in differentiating lineages of highly admixed groups, such as Mestizos in South and Central America (Martínez-Cortés et al. 2013). This sequencing of mtDNA aids in identifying past practices and migration events, as haplogroups associated with different matrilineal lineages vary in frequency depending on geography and past population structure (Martínez-Cortés et al. 2013).

While mtDNA sequencing grants a multitude of research possibilities regarding lineages and linked traits, mtDNA sequencing is not entirely infallible. In fact, there are several limitations to mtDNA research. For instance, mtDNA haplogroups are often widespread and can be found in relatively high frequencies on more than one continent. This, as well as the fact that mtDNA is inherited as a unit and therefore must be treated as a single marker, can make assigning ancestry based only on mtDNA problematic. Research on mitochondrial haplogroups can be time-consuming and is prone to error; there is also always the chance that false
associations will be made due to sample mix-up or contamination. These errors, while concerning, have been lessened by the recommendation that mtDNA sequences be double checked for phylogenetic consistency, as well as insistence on heightened lab protocols to avoid possible contamination (Bandelt et al 2009; Kloss-Brandstätter et al. 2011). Furthermore, time constraints created by identifying mtDNA haplogroups are potentially being rectified as databases and software such as Phylotree and HaploGrep seek to streamline the process of identifying mtDNA haplogroups from part of — if not the entire — mitochondrial genome (Kloss-Brandstätter et al. 2011).

### 1.3a Y-Chromosome Lineages

In combination with mtDNA, Y-chromosomal DNA, which passes through paternal lines, offers further insight into human ancestry. Y-chromosomal DNA (Y-DNA), however, exhibits differences when compared to mtDNA. Unlike the maternally inherited DNA, Y-DNA recombines during the process of meiosis (apart from the proposed Male Specific Region or MSY). This patrilineal inherited polymorphism also displays fewer mutations than those found within mtDNA, though mutations do accumulate in the MSY and paint a picture of patrilineal genealogy (Oven et al. 2013; Bradbury 2017). Despite these differences, Y-DNA is undoubtedly beneficial in migration and genealogical studies, as the MSY region of the Y-chromosome is particularly sensitive to drift and founder effects (Underhill et al. 2001).

The haplogroups found through patrilineal studies indicate that, similar to matrilineal lineages, there is a Y-DNA haplogroup that points to a MRCA, frequently referred to as “Y-chromosome Adam,” some 250 kya (Mendez et al. 2013). Also, the discovery of a small cluster
of haplogroups found in Eurasia and Oceania at 47-52 kya, are consistent with the out of Africa bottleneck, allowing for an accurate depiction of out of Africa migration events (Karmin et al. 2015). Therefore, Y-chromosome DNA polymorphisms elucidate global dispersals as well as Y-chromosome lineage differentiation. This data can then be used in forensic and genealogical contexts, as haplogroups can infer bio-geographic ancestry. Patrilineal lineage has been explored in abundance, from Caucasian genetic diversity to the emergence of current Y-chromosome haplogroups in modern Japan and how they reflect migration and cultural shifts (Nasidze et al. 2004; Hammer et al. 2006; Oven et al. 2013). This is best exemplified through surname studies, as surnames are typically passed from fathers to sons alongside Y-DNA polymorphisms. Past research in Spain and the British Isles discovered that surname frequency corresponds greatly with Y-chromosome haplotypes (Martinez-Cadenas et al. 2016). Y-chromosome studies also offer information on the possible amount of admixture within a population, particularly those populations with high rates of migration and past exogamous behavior. This is evidenced in Indian populations in particular, as multiple ethnic groups within India have members that belong to more than one haplogroup (Mahal and Matsoukas 2018). This information, in conjunction with evidence obtained through mtDNA sequencing, presents not only information on ancient migration patterns and cultural shifts, it also displays a highly accurate portrait of bio-geographic ancestry of an individual that is exceptionally valuable in forensic studies.

1.3b Ancestry Informative Markers

Ancestry Informative Markers (AIMs) are single nucleotide polymorphisms (SNPs) at loci within the nuclear genome and appear in differing frequencies between populations in
different geographic regions (National Human Research Institute n.d.), hence AIMs are valuable in estimating biogeographical ancestry and admixture (Halder et al. 2008). When compared to mtDNA and Y-DNA ancestral assessments, AIMs are more beneficial when studying individual genomic admixture due to the fact that they can trace through maternal and paternal and are less affected by the environment (particularly for neutral loci). In the early phases of this research, Halder (2008) established a panel of 176 AIMs with a high accuracy rate of distinguishing admixture proportions by using ancestry-phenotype associations. However, as technology has improved and gotten cheaper, current ancestry identifying services, such as 23andMe, work with 7400-50,000 AIMs per chromosome (“Ancestry Composition - 23andMe” n.d.). AIMs research has also been furthered by reliable ancestry informative panels regarding Asian ancestries that have similar external visual characteristics and are also closely genetically related. The accuracy of the AIMs discriminating between Chinese, Han, Japanese, and Korean individuals of said research reaches 90-99%, indicating that AIMs panels are not only valuable for identifying continental populations, but also major regional groups (Shi et al. 2019). In a recent study by Jin et al (2019), research aimed to discover the biogeographic ancestors of admixed individuals of three continental populations. AIMs proved again to be useful in this context as the AIMs panels were able to accurately assess biogeographic origins of major regional populations, particularly the Uychur group of Eastern Asia. Like mtDNA and Y-DNA, AIMs also have the potential to be useful in a forensic context as AIMs are best suited to assess proportions of admixture. In fact, when considering AIMs, the inclusion of data found from AIMs is possibly more beneficial than the inclusion of mtDNA. As discussed, mtDNA has a limited scope when compared to AIMs because the latter uses more than one marker, which is ostensibly treated as independent from the other SNPs around it, to track ancestry. Because AIMs studies are capable of simultaneously
studying thousands of SNPs at different loci across multiple populations, assignation of ancestry is likely more accurate than an assessment done from a single marker. Further research should work to include AIMs in order to increase accuracy of ancestral determinations, as well as identify admixture events in the recent or distant past.

1.4 NAGPRA

In 1990, Congress fully recognized that Native American and Native Hawaiian remains, funerary objects, sacred objects, and objects of cultural patrimony were to be repatriated to the tribes from which they belong (NPS n.d.). In accordance with NAGPRA, any known Native American human remains must be inventoried prior to consulting lineal descendants or Indian tribes and moving towards potential repatriation. With the ever-growing knowledge of mtDNA haplogroups, geneticists have been, for some time, able to identify maternal haplogroups that are potentially linked to Native American ancestry. The identification of maternal haplogroups, while beneficial, do not necessarily point to the exact bio-geographic ancestry of an individual. In this case, some Native American associated haplogroups are currently also found in variable frequencies in Asia, this is not wholly surprising as it is well known that groups of people crossed into modern day Americas via the Bering Strait from Asia in multiple migration events (Llamas, Harkins, and Fehren-Schmitz 2017). Because of this, identifying Native American remains must be a comparison between DNA findings and where the remains were found or received from in order to most ethically address curation of the remains.

1.5 Research Significance
By incorporating mitogenomes sequenced from bone samples, ancestry will be assessed through the corresponding haplogroups identified in the analysis. These mtDNA haplogroups provide information on where this DNA is most commonly found. This offers more evidence on the ancestry of the individual, rather than relying on morphological cranial traits. With this knowledge, the cases used in my research will be more accurately curated within the University of Montana Forensic Collection. If evidence is found of the individual being of a different ancestry than the FORDISC 3.1 assessments gave, then steps will be taken in order for the remains to be re-curated within the collection under correct information. Due to the known discrepancies in FORDISC 3.1 assessments, it is likely that the estimations provided by the database will not be in accordance with other estimations such as mtDNA and non metric methods. If the FORDISC 3.1 assessments of the UMFC individuals are vastly different from the other results, it is possible that many of the individuals housed within the UMFC are not correctly curated. The inclusion of mtDNA analysis, though not all powerful, could only benefit the study of ancestry of the UMFC individuals as the mtDNA data could help indicate the level of discrepancy in FORDISC 3.1 ancestral estimations as well as estimations made exclusively from cranial and post cranial morphology.

Furthermore, the MiSeq data will potentially contribute toward a larger ancestry identification database. The goal of this database will be to utilize machine learning in order to better estimate ancestry of skeletal remains, with genetic data to be part of the equation. The haplogroup data that will potentially be available through the MiSeq will spur further research into the disparity between FORDISC 3.1 and DNA assessments and be a vital component in establishing the ancestry of individuals. Hopefully, with this information, this research will
continue on a broader scale and will be able to work towards creating a more accurate forensic
database that will aid forensic anthropologists, and in turn, law enforcement, with estimating
ancestry in order to aid in forensic cases.

2.0 Materials and Methods

2.1 Research Sample

The University of Montana houses a Forensic Collection (UMFC) of cranial, postcranial,
and fragmented human remains. The ancestry, age, and sex of these individuals has been
assessed by students in osteology courses; the remains are often handled by both undergraduate
and graduate students. Of these remains, six individuals were chosen from UMFC for the
purposes of this research. These crania were decided upon for several reasons, including
bleaching of the remains, completeness of the crania, amount of handling, and treatment of the
remains with paint or clay; all of which could affect if DNA could be extracted from the bone.
Of the six individuals chosen, three (UMFC 103, 104, and 141) are of supposed Asian ancestry
— Chinese in particular — as that is where the remains originated and were housed prior to their
time in the UMFC. The remaining three crania (UMFC 155, UMFC 98, UMFC 80) are cited as
“peoples of India,” as they were originally used as anatomical specimens and the majority of
anatomical specimens in the UMFC are of Indian descent. The knowledge of where these
individuals originated was obtained both through past case reports in the UMFC as well as
correspondence with the head of the UMFC who supplied the order reports for these individuals.
Photographs of each crania were taken in order to illuminate the differences in shape and size of each crania.

Figure 2a: UMFC 80; sample 1

Figure 2b: UMFC 98; sample 2
Figure 2c: UMFC 103; Sample 3

Figure 2d: UMFC 155; Sample 4
2.2 FORDISC 3.1
After the crania were chosen, metric and non-metric methods were used in order to assign each individual a possible ancestry. The metric methods incorporated were solely the cranial and mandibular measurements found in *Standards* (Buikstra and Ubelaker 1994). These measurements were taken using the tools suggested in *Standards*: a sliding and spreading caliper. The mandibular measurements were taken for all but UMFC 80 and UMFC 98, as no mandible was present with these crania. The lack of mandibular measurements should not affect the ancestral assessment of FORDISC 3.1, as little evidence has been found to suggest that mandibular traits are highly indicative of ancestry, but were taken in order to stay in accordance with the *Standards* measurement methods. Metric methods of ancestral estimation are all taken to the nearest millimeter and are taken on the left side of the crania, though the right side of the crania may be substituted if need be (Buikstra and Ubelaker 1994). All measurements taken were then input into FORDISC 3.1 using only the Forensic Database (FDB).

The sex of the individuals was determined before utilizing FORDISC 3.1 and was determined through the size and robusticity of cranial and mandibular features, in order to lend a more accurate interpretation of the results (Walker 2008; Buikstra and Ubelaker 1994). FORDISC 3.1 determined the sex of the individual during ancestral assessments. Because of this, both sexes and all available ancestries in FORDISC 3.1 were used before final identification.
FORDISC 3.1 relies on discriminant functions analysis in order to sort measurements into an ancestral group. A discriminant functions analysis falls under the statistical methods for the optimal separation of groups and classification of unknowns using measurements. Discriminant functions analyses involve reference groups with known identification and then uses the known identities as a reference to classify the unknowns (Ousley and Jantz 2005; 2016). The database itself houses 13 population groups and over 2400 individuals; most were born after 1930 and are grouped as White, African American, or Hispanic. Furthermore, FORDISC 3.1 provides posterior probabilities: data which evaluate the probability that the unknown individual comes from each reference group. The database also provides typicality probability — particularly chi square — which estimates how likely it is that an unknown individual belongs to each group. It should be noted that if typicality probability is low (below 0.01), then all posterior
probabilities should be ignored as the typicality probabilities evaluate the overall “fit” of an individual into a classification (Ousley and Jantz 2005; 2016).

FORDISC 3.1 determined likely sex of the individual as well as possible ancestry. As soon as sex was determined, all measurements of ancestries regarding the opposite sex of the one determined were eliminated from the process.

Figure 3a: Example of Standards cranial measurements from sagittal plane and Maximum cranial breadth and bizygomatic diameter.

<table>
<thead>
<tr>
<th>Measurement</th>
<th>UMFC 80</th>
<th>UMFC 98</th>
<th>UMFC 103</th>
<th>UMFC 155</th>
<th>UMFC 104</th>
<th>UMFC 141</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max Cranial L.</td>
<td>164</td>
<td>170</td>
<td>159</td>
<td>170</td>
<td>183</td>
<td>160</td>
</tr>
<tr>
<td>Max Breadth</td>
<td>128</td>
<td>134</td>
<td>131</td>
<td>148</td>
<td>130</td>
<td>120</td>
</tr>
<tr>
<td>Measure</td>
<td>Value 1</td>
<td>Value 2</td>
<td>Value 3</td>
<td>Value 4</td>
<td>Value 5</td>
<td>Value 6</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>Bizygomatic Diameter</td>
<td>114</td>
<td>N/A</td>
<td>104</td>
<td>115</td>
<td>122</td>
<td>115</td>
</tr>
<tr>
<td>Basion-Bregma Height</td>
<td>131</td>
<td>125</td>
<td>128</td>
<td>123</td>
<td>133</td>
<td>116</td>
</tr>
<tr>
<td>Cranial Base L</td>
<td>120</td>
<td>125</td>
<td>95</td>
<td>92</td>
<td>101</td>
<td>97</td>
</tr>
<tr>
<td>Basion-Prosthion Length</td>
<td>91</td>
<td>84</td>
<td>84</td>
<td>87</td>
<td>96</td>
<td>103</td>
</tr>
<tr>
<td>Maxi-Alveolar Breadth</td>
<td>63</td>
<td>56</td>
<td>48</td>
<td>64</td>
<td>61</td>
<td>61</td>
</tr>
<tr>
<td>Maxilla-Alveolar Length</td>
<td>42</td>
<td>41</td>
<td>45</td>
<td>40</td>
<td>49</td>
<td>53</td>
</tr>
<tr>
<td>Biauricular Breadth</td>
<td>N/A</td>
<td>112</td>
<td>112</td>
<td>125</td>
<td>120</td>
<td>115</td>
</tr>
<tr>
<td>Upper Facial Height</td>
<td>63</td>
<td>59</td>
<td>60</td>
<td>73</td>
<td>70</td>
<td>58</td>
</tr>
<tr>
<td>Minimum Frontal Breadth</td>
<td>86</td>
<td>92</td>
<td>89</td>
<td>90</td>
<td>85</td>
<td>88</td>
</tr>
<tr>
<td>Upper Facial Breadth</td>
<td>97</td>
<td>94</td>
<td>100</td>
<td>98</td>
<td>97</td>
<td>101</td>
</tr>
<tr>
<td>Nasal Height</td>
<td>50</td>
<td>43</td>
<td>50</td>
<td>57</td>
<td>55</td>
<td>50</td>
</tr>
<tr>
<td>Nasal Breadth</td>
<td>24</td>
<td>22</td>
<td>26</td>
<td>24</td>
<td>23</td>
<td>27</td>
</tr>
<tr>
<td>Orbital Breadth</td>
<td>40</td>
<td>38</td>
<td>32</td>
<td>33</td>
<td>35</td>
<td>34</td>
</tr>
<tr>
<td>Orbital Height</td>
<td>33</td>
<td>35</td>
<td>32</td>
<td>34</td>
<td>36</td>
<td>36</td>
</tr>
<tr>
<td>Biorbital Breadth</td>
<td>90</td>
<td>87</td>
<td>87</td>
<td>89</td>
<td>91</td>
<td>95</td>
</tr>
<tr>
<td>Interorbital Breadth</td>
<td>22</td>
<td>18</td>
<td>22</td>
<td>22</td>
<td>24</td>
<td>25</td>
</tr>
<tr>
<td>Frontal Chord</td>
<td>108</td>
<td>114</td>
<td>113</td>
<td>113</td>
<td>121</td>
<td>115</td>
</tr>
<tr>
<td>Parietal Chord</td>
<td>103</td>
<td>109</td>
<td>95</td>
<td>94</td>
<td>108</td>
<td>104</td>
</tr>
<tr>
<td>Occipital Chord</td>
<td>83</td>
<td>90</td>
<td>82</td>
<td>80</td>
<td>87</td>
<td>83</td>
</tr>
<tr>
<td>Foramen Magnum L</td>
<td>36</td>
<td>41</td>
<td>28</td>
<td>38</td>
<td>38</td>
<td>36</td>
</tr>
<tr>
<td>Foramen Magnum Breadth</td>
<td>27</td>
<td>29</td>
<td>28</td>
<td>33</td>
<td>32</td>
<td>33</td>
</tr>
<tr>
<td>Mastoid Length</td>
<td>35</td>
<td>29</td>
<td>30</td>
<td>37</td>
<td>40</td>
<td>38</td>
</tr>
</tbody>
</table>
Group means that were flagged by FORDISC 3.1 as being too high or too low — at least two standard deviations from the groups — were re-measured before processing the results. UMFC 80 basion to bregma height was labeled as too low while the mastoid length was labeled by the database as too high. UMFC 98 displayed a low maxilla-alveolar length while UMFC 103 displayed a low bizygomatic breadth. UMFC 155 maxilla-alveolar length was also flagged as being too low and UMFC 104 orbital breadth was flagged as too low to match the rest of the measurements. Each of these measurements were re-taken in order to ensure no errors were made during measurements. However, the flagged measurements of UMFC 103 and UMFC 155 remain unchanged as re-measuring the distance between the landmarks did not change from the previous measurement. FORDISC 3.1 flagging these measurements means that UMFC 103 and UMFC 155 do not fit the expected measurements when compared to the individuals represented in the database.

Following these measurements, the non metric methods of OSSA (Hefner and Ousley 2014) in conjunction with Rhine and Gill (1995) were utilized in order to provide a more complete interpretation of ancestry, as well as to juxtapose the results of metric and non metric ancestry assessments before DNA results were given. Sex was first determined via the Walker (2008) method, in which sex is estimated based off visual assessments and ordinal scores (1-5) of the glabella, mastoid process, superorbital margin, mental eminence, and nuchal crest before being inputted into a discriminant function analysis in order to more accurately predict a possible ancestry using only morphoscopic traits. These non metric methods were chosen in part for reliability, and in part because they are widely used by anthropologists when determining ancestry. OSSA, however, only differentiates between “White” and “Black,” while the Rhine and Gill (1995) method, though more dated, differentiates between European, African, and Asian
descent. Due to the FORDISC 3.1 results determining each of the individuals to be either of European or African ancestry, OSSA was deemed to be a relevant non metric method for the purposes of this research.

2.3 DNA Extraction

2.3a Drilling

The petrous portion of the six crania were drilled into using a Dremel and a HP8 oral surgery drill bit. Due to the portion of the temporal bone being more dense and, therefore, better preserved, the petrous portion is the most suitable for DNA extraction—in fact, the petrous bone can potentially yield 100x more DNA than other less dense bones in the body (Pinhasi et al. 2015). Although drilling into the petrous portion is technically a destructive method, due to its placement in the inner ear, the effects of drilling are not easily visible on the crania. Prior to drilling, each crania was wiped down with DNA Away (Thermo Fisher) in order to rid the crania of any modern contamination as each of the six individuals are frequently handled by students. The drilling station was cleaned with bleach and lined with waxed butcher paper before usage. Each crania was assigned a number 1-6 for future extraction and library prep. The petrous portion was drilled into until an adequate weight of dust was met—in this case approximately 0.02 g—and collected in a weight boat. The bone dust was then transferred into 2ml low-bind plastic tubes.

Each collection of dust required all used items to be UV’ed prior to and between usage, including the weight boats for dust collection, dremel and drill bits, and 2ml tubes used to collect
DNA. For collection, appropriate aDNA lab precautions were used which include wearing nitrile gloves, masks, and suits to protect the lab from outside contaminants.

Table 2b: Weight of bone dust collected in grams

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Number</th>
<th>Weight of Dust in Grams</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>1</td>
<td>0.035g</td>
</tr>
<tr>
<td>98</td>
<td>2</td>
<td>0.02g</td>
</tr>
<tr>
<td>103</td>
<td>3</td>
<td>0.023g</td>
</tr>
<tr>
<td>155</td>
<td>4</td>
<td>0.022g</td>
</tr>
<tr>
<td>104</td>
<td>5</td>
<td>0.037g</td>
</tr>
<tr>
<td>141</td>
<td>6</td>
<td>0.08g</td>
</tr>
</tbody>
</table>

2.3b Extraction, Library Prep, and Sequencing

The following DNA extraction method follows the protocol outlined by Dabney et. al (2011). DNA extraction for the samples began by adding one milliliter of EDTA to each sample of bone dust contained in 2ml tubes. An extract control was created for the purpose of identifying any possible contamination introduced through the extraction and library preparation processes. For the purposes of DNA extraction, EDTA aids in the breakdown of the samples. The samples were then vortexed in order to ensure that the EDTA was incorporated into the sample before storing the samples in soaking samples drawer within the aDNA lab. This storage method kept the samples out of the UV light of the lab that are utilized daily to eliminate contamination.

Following this storage method, proteinase K was added in order to lyse the cells. The sample tubes were then covered in a thin layer of parafilm in order to prevent any leakage during
incubation. The samples were vortexed and taped to the incubation rack, which incubated the samples at 55 degrees Celsius, rotating slowly at 4 rpm overnight (~16 hours).

The DNA extraction process must follow directly after the addition of proteinase K to the samples, as leaving it in the sample for too long will begin to damage the sample. To finalize the extraction, 13ml of PB buffer, containing guanidinium thiocyanate, was added to a 15ml tube. Approximately 1ml of supernatant was then added to the buffer prior to the sample being put into a Qiagen spin column and spun through the silica extraction filter. This was followed by two rinses with PE Qiagen buffer (an ethanol wash). A dry spin was then conducted in order to remove PE buffer from the spin column in order not to interfere with future analysis. Water was then added to the samples and left to sit until spun through into a final collection tube (repeated twice with 50ul of ddH20). Following extraction, the DNA in the samples was quantified through the use of qubit, with the addition of 195ul of QB buffer and 5ul of DNA to the sample, which was then placed in the qubit fluorometer. The concentration results are in Table 2c.

The library prep protocol followed in Dr. Meredith Snow’s aDNA lab is a hybrid protocol with Meyer-Kircher (2010) and KAPA (Roche). This protocol begins with blunt-end repair and requires first that a master mix be prepared for 50ul reactions. The volumes of the components given in this protocol were first multiplied by the number of samples, or 7.5, in order to accurately calculate the total needed for the reactions. The samples were then incubated at 25 degrees Celsius in the thermal cycler prior to a bead clean in order to isolate and clean the DNA. This bead clean was composed of 2x SeraPure beads added to the reaction. After an incubation period of 5 minutes, the beads are then placed on the magnetized rack, which separates the beads from the rest of liquid so that they may be rinsed by a 70% ethanol wash. After completing two 70% ethanol washes, 15ul of EB was then added to each sample. The
samples were then magnetized a final time in order to obtain a solution with purified DNA. The solution was then transferred into new PCR tubes. After the bead clean, the DNA was resuspended in 18ul of EB buffer.

The library prep continues with adapter ligation and adapter fill-in. During the adapter ligation processes, a master mix was created of 22.5ul molecular grade water, 150ul quick ligase buffer, and 1.5ul of Meyer-Kircher annealed adapters. After mixing, 24ul of the master mix was aliquoted into six PCR tubes before adding 15ul of the blunt-end repaired DNA and 1ul of Quick Ligase. The PCR tubes were then incubated at 25 degrees Celsius in the thermal cycler, followed by a bead clean. The adapter fill-in process for this protocol consisted of adding 33ul of master mix, created from 37.5ul of ThermoPol Buffer, 4.5ul of dNTPs, and 205.5ul of molecular grade water, to each PCR tube. 15ul of the purified DNA from the adapter ligation process was then added following 2ul of Bst DNA polymerase to each sample. The samples were incubated at 37 degrees Celsius for 30 minutes then 80 degrees Celsius for 20 minutes in order to inactivate the Bst.

Following this, indexing PCR took place in which 42ul of the master mix was added to each sample before the addition of 5ul of P5 and P7 adapters to each sample. The samples were then purified through another bead clean and eluted with 20ul of EB. Moving into the protocol outlined by KAPA, the sample library was then amplified via LM-PCR. A Pre-capture LM-PCR master mix of 30ul was added to each PCR tube prior to running the samples through the thermocycler according to the program given for Pre-capture LM-PCR. A further bead clean is conducted on the samples in order to create a purified, amplified library. The samples were then run through the Qubit procedure in order to analyze the amount of DNA in the sample post-indexing. These amounts were recorded in Table 2d. The samples were pooled by adding equal
amounts of the amplified libraries and were pooled together, the final pooling volumes are labeled in Table 2d. The pooled samples were then cleaned with yet another bead wash and added to hybridization probes that captured the mitogenome from the other DNA present in the sample. The samples were then run through a series of bead wash buffers. The Capture Beads used for this protocol were washed 2x with 1x Bead Wash Buffer in order for the Capture Beads to be able to bind to the captured DNA added via the hybridization sample. After incubation, the Capture Beads and bead bound DNA were then given washes with 1x Stringent Wash Buffer, 1x Wash Buffer, Wash Buffer II, and Wash Buffer III. The bead bound DNA was then amplified using LM-PCR using the Post-Capture LM-PCR program outlined in KAPA. A further purification of the sample was conducted via a bead clean before a final Qubit of the samples could be conducted.

**Table 2c: Extraction protocol and Qubit results post-extraction**

<table>
<thead>
<tr>
<th>Sample</th>
<th>Sample ID</th>
<th>Notes</th>
<th>EDTA</th>
<th>ProtK</th>
<th>Extraction</th>
<th>Qubit ng/ul</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>80</td>
<td></td>
<td>1/16/20</td>
<td>1/23/20</td>
<td>1/24/2020</td>
<td>0.9</td>
</tr>
<tr>
<td>2</td>
<td>98</td>
<td></td>
<td>1/16/20</td>
<td>1/23/20</td>
<td>1/24/2020</td>
<td>0.48</td>
</tr>
<tr>
<td>3</td>
<td>103</td>
<td></td>
<td>1/16/20</td>
<td>1/23/20</td>
<td>1/24/2020</td>
<td>0.16</td>
</tr>
<tr>
<td>4</td>
<td>155</td>
<td></td>
<td>1/16/20</td>
<td>1/23/20</td>
<td>1/24/2020</td>
<td>0.059</td>
</tr>
<tr>
<td>5</td>
<td>104</td>
<td></td>
<td>1/16/20</td>
<td>1/23/20</td>
<td>1/24/2020</td>
<td>0.026</td>
</tr>
<tr>
<td>6</td>
<td>141</td>
<td></td>
<td>1/16/20</td>
<td>1/23/20</td>
<td>1/24/2020</td>
<td>0.12</td>
</tr>
<tr>
<td>7</td>
<td>Extract Control</td>
<td></td>
<td>1/16/20</td>
<td>1/23/20</td>
<td>1/24/2020</td>
<td>too low</td>
</tr>
</tbody>
</table>

**Table 2d: Library Preparation and Qubit quantification post adapter and indexing**

<table>
<thead>
<tr>
<th>Sample</th>
<th>Sample ID</th>
<th>Extraction Qubit ng/ul</th>
<th>P5 adapter</th>
<th>P7 adapter</th>
<th>Post-Indexing Qubit ng/ul</th>
<th>Pooling ul</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>80</td>
<td>0.9</td>
<td>P5-01</td>
<td>P7-01</td>
<td>53</td>
<td>9.433962264</td>
</tr>
<tr>
<td>2</td>
<td>98</td>
<td>0.48</td>
<td>P5-02</td>
<td>P7-02</td>
<td>36.8</td>
<td>13.58695652</td>
</tr>
<tr>
<td>3</td>
<td>103</td>
<td>0.16</td>
<td>P5-03</td>
<td>P7-03</td>
<td>32.7</td>
<td>15.29051988</td>
</tr>
</tbody>
</table>
2.4 Sequencing, Galaxy, and Haplogrep

In order to sequence the samples, next generation sequencing services at the Genomics Core at the University of Montana and Illumina Sequencing were utilized. The samples were submitted to the Genomics Core following the appropriate library prep protocol specified for the MiSeq platform. The Genomics Core offers MiSeq data from Illumina Sequencing. For the purposes of this research, the MiSeq 300 Nano v2 was used, which offers read lengths of 300 bp.

After the samples were sequenced, the data was uploaded to Galaxy — a web-based platform for bioinformatics — using the pipeline created for Dr. Meradeth Snow’s aDNA lab. The pipeline itself follows a ten-step process, though for the purposes of this research, steps such as mapDamage and PMDtools were not necessary. After inputting the raw data, a FastQC is run prior to merging and adapter trimming. The next step of the pipeline trims adapters and merges forward and reverse reads before the next step which maps reads to the reference genome. The overall read depth, sequence coverage, and amount of reads is then generated via Qualimap. The pipeline then creates genotype calls before generating a consensus file and estimating overall contamination in the samples, which finishes out the analysis before Haplogrep (Colon 2016; Blohm 2019).
As previously discussed, the database Haplogrep streamlines the process of identifying mtDNA haplogroups. By submitting the finished Galaxy analysis into Haplogrep, the database then classifies haplogroups itself based on pre-calculated phylogenetic weights that correspond to the occurrence per position in Phylotree as well as reflecting the mutational stability of a variant (Kloss-Brandstätter et al. 2011; Weissensteiner et al. 2016). The results from Haplogrep were then recorded and incorporated into a summary spreadsheet with all relevant Galaxy data.

3.0 Results

3.1 FORDISC Results

The first individuals measurements input into FORDISC 3.1 were those of UMFC 80 and they show the possibility of European descent (Table 3a). Following this, UMFC 103, UMFC 155, and UMFC 98 all displayed European typical cranial traits, according to the database. In comparison, the cranial and mandibular traits of UMFC 104 and UMFC 141 were labeled as being from African ancestry. The sex of the individual was also determined through FORDISC 3.1, as the robusticity and shape of particular traits are variable depending on sex. However, FORDISC 3.1 often has difficulty determining the sex of an individual based off the measurements input into the database, in fact one study found that FORDISC 3.1 sex determination reliability was only 60-71% (Urbanova et al. 2014). This is in part due to the fact that cranial measurements for certain populations are smaller regardless of sex and that FORDISC 3.1 contains more European reference measurements and therefore can only accurately determine the sex of European samples (Ramsthaler, Kreutz, and Verhoff 2007). Sex
as determined by FORDISC 3.1 is also made less accurate as cranial traits are not the best indicators of sex of an individual. If an os coxae is present, sex is best estimated via the Phenice methodology (1969) which can then be compared to the cranial morphology as described by Walker (2008).

Alongside the typicality and posterior probability, FORDISC 3.1 created graphs to visualize where the measurements of these individuals fall when compared with the measurements of certain ancestries in the database. The lower the posterior probability was for certain ancestries, the more the graphs were manually trimmed down in order to offer more clarity to the graph. For example, Figure 3e displays a graph with only three ancestries labeled. While all ancestries were included in the initial assessment, those with a low posterior probability were removed from the assessment, leaving only the ancestries with the highest posterior probabilities to be shown in the graph. The black “X” is used to mark where the individual’s measurements sorted them in comparison to the measurements of 13 possible ancestries. The ancestry, chi-square, and posterior probability were then recorded in Table 3a.

Table 3a: FORDISC 3.1 results including chi-square and posterior probability

<table>
<thead>
<tr>
<th>Sample</th>
<th>FORDDIC</th>
<th>Chi Square</th>
<th>Posterior Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>UMFC 80</td>
<td>European male</td>
<td>0.00</td>
<td>.998</td>
</tr>
<tr>
<td>UMFC 103</td>
<td>European female</td>
<td>0.00</td>
<td>.946</td>
</tr>
<tr>
<td>UMFC 104</td>
<td>African female</td>
<td>0.00</td>
<td>.647</td>
</tr>
<tr>
<td>UMFC 141</td>
<td>African male</td>
<td>0.00</td>
<td>.979</td>
</tr>
</tbody>
</table>
The following figures depict the placement of the individual’s cranial measurements in comparison with other ancestries and those of the same sex. Due to FORDISC 3.1 operating system, the database assigns an ancestry to an individual no matter how dissimilar the case is to existing measurements within the database. Therefore, while the graphs created by the database offer a clear look at the difference between measurements, in order to confirm an ancestry for the individual measurements, one should focus their attention on the typicality and posterior probabilities, as well as the Mahalanobis distance. Each graph displays circles which are labeled with a certain ancestral group and color. Within each circle, FORDISC 3.1 displays shapes with a corresponding color to their labeled ancestry in order to show the number of individuals in the database within a particular ancestry.

Table 3b: Key of listed populations in FORDISC 3.1 graphs

<table>
<thead>
<tr>
<th>African</th>
<th>European</th>
<th>Chinese</th>
<th>Guatemalan</th>
<th>Hispanic</th>
<th>Japanese</th>
<th>American Indians</th>
<th>Vietnamese</th>
</tr>
</thead>
<tbody>
<tr>
<td>BM/BF</td>
<td>WM/WF</td>
<td>CM</td>
<td>GTM</td>
<td>HM/HF</td>
<td>JM/JF</td>
<td>AM/AF</td>
<td>VM</td>
</tr>
</tbody>
</table>
Figure 3b: UMFC 80 FORDISC 3.1 ancestry graph results. GTM (light green), VM (pink), JM (blue), AM (blue triangle), CRM (black), BM (green), WM (red), HM (purple)

Figure 3c: UMFC 98 FORDISC 3.1 ancestry graph results. JM (purple), WM (blue), AF (pink), BM (blue triangles), WF (red), JF (light green), AM (pink), HM (green), HF (black)
Figure 3d: UMFC 103 FORDISC 3.1 ancestry graph results. JF (blue), WF (red), HF (green), BF (purple), AF (light green)

Figure 3e: UMFC 155 FORDISC 3.1 ancestry graph results. VM (green), WF (red), WM (purple), HM (blue)
Figure 3f: UMFC 104 FORDISC 3.1 ancestry graph results. JF (green), WF (blue), BF (red)

Figure 3g: UMFC 141 FORDISC 3.1 ancestry graph results. GTM (green), WM (blue), CM (purple), BM (red)
Following the results of FORDISC 3.1, the results of the non metric methods were recorded. The OSSA scoring sheet uses a varying ranking system of 1-3, 1-4, 1-5, and 0-1 of cranial non metric traits including: anterior nasal spine, nasal aperture width, inferior nasal aperture (1-5), nasal bone contour (1-4), interorbital breadth, and post-bregmatic depression (0-1). After scoring each of the crania using this sheet, the results are as follows:

Table 3c: OSSA Scoring Sheet Ancestry Results

<table>
<thead>
<tr>
<th>Sample</th>
<th>OSSA Ancestry</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>European</td>
</tr>
<tr>
<td>98</td>
<td>European</td>
</tr>
<tr>
<td>103</td>
<td>European</td>
</tr>
<tr>
<td>155</td>
<td>European</td>
</tr>
<tr>
<td>141</td>
<td>African</td>
</tr>
<tr>
<td>104</td>
<td>European</td>
</tr>
</tbody>
</table>

The Rhine and Gill (1995) scoring method was also utilized in identifying the presence or absence of ancestry specific traits. Much like the OSSA method, the results of the Rhine and Gill scoring method suggest European ancestry for most of the individuals apart from UMFC 141. Though each cranium exhibited a myriad of traits, most if not all of crania did not exhibit solely European, Asian, or African cranial traits as described by Rhine and Gill (1995).

Table 3d: Rhine and Gill (1995) trait list for three distinct ancestral groups
For example, the palate shape for the individuals of supposed Chinese ancestry (UMFC 103, 104, 141), did not exhibit strictly round palate shapes nor could their cranial sutures be described as serpentine. These individuals, however, did exhibit oval orbital shapes, medium nasal apertures, and wormian bones. The presence of the Rhine and Gill (1995) traits on these crania is not enough to confirm or deny a particular ancestry.

### 3.2 mtDNA results

The MiSeq results gained through Galaxy and Haplogrep are displayed in Table 3c. The results found through these databases have been cut down to the most relevant information including percent endogenous (mapped unique total reads); percent of mapped reads kept (after eliminating duplicates; the average length (how long fragments are in terms of basepairs) and

<table>
<thead>
<tr>
<th>Feature</th>
<th>European Descent</th>
<th>Asian Descent</th>
<th>African Descent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orbit shape</td>
<td>lateral slope</td>
<td>oval</td>
<td>rectangle/square</td>
</tr>
<tr>
<td>Nasal root</td>
<td>depressed</td>
<td>tented</td>
<td>shallow</td>
</tr>
<tr>
<td>Nasal bridge</td>
<td>high</td>
<td>low</td>
<td>low</td>
</tr>
<tr>
<td>Nasal spine</td>
<td>prominent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasal sill</td>
<td>present</td>
<td>variable</td>
<td>guttered</td>
</tr>
<tr>
<td>Nasal aperture</td>
<td>narrow</td>
<td>medium</td>
<td>broad</td>
</tr>
<tr>
<td>Metopic suture</td>
<td>rare</td>
<td>often</td>
<td>rare</td>
</tr>
<tr>
<td>Suture pattern</td>
<td></td>
<td>serpentine</td>
<td></td>
</tr>
<tr>
<td>Wormian bones</td>
<td>rare</td>
<td>common</td>
<td>rare</td>
</tr>
<tr>
<td>Prognathism</td>
<td>orthognathic</td>
<td></td>
<td>prognathic</td>
</tr>
<tr>
<td>Zygomatics</td>
<td>retreating</td>
<td>projecting</td>
<td></td>
</tr>
<tr>
<td>EAM shape</td>
<td>round</td>
<td>oval</td>
<td>round</td>
</tr>
<tr>
<td>Dental</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bite</td>
<td>overbite</td>
<td>direct</td>
<td></td>
</tr>
<tr>
<td>Palate</td>
<td>parabolic</td>
<td>elliptical/round</td>
<td>rectangular/square</td>
</tr>
<tr>
<td>Ramus</td>
<td>eversion</td>
<td>inversion</td>
<td></td>
</tr>
<tr>
<td>Indicies</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cranial index</td>
<td>dolicephalic</td>
<td>brachycephalic</td>
<td>mesocephalic</td>
</tr>
<tr>
<td></td>
<td>&lt;74.99</td>
<td>&gt;80</td>
<td>75-79.99</td>
</tr>
<tr>
<td>Nasal index</td>
<td>leptorrhine</td>
<td>mesorrhine</td>
<td>platyrhine</td>
</tr>
<tr>
<td></td>
<td>&lt;47.9</td>
<td>48-52.9</td>
<td>&gt;53</td>
</tr>
</tbody>
</table>
depth of the reads (how many times each basepair has been sequenced by separate reads); the percent of reference sequence covered (both >1x and >2x; those basepairs that are used in haplogroup assignment and read depth); variant sites with just 1x coverage (sites that vary from the reference that are only sequenced once); variant sites with more than 1x coverage; and finally the corresponding haplogroups. Alongside these results, damage plots were also provided to indicate not only the level of damage in the DNA samples, but also the possible “age” of samples, i.e. the modernity of the DNA. According to said damage plots, there are indications of possible ancient, or at least damaged, DNA within the samples. The damage plots for the samples as seen in Figure 3h are known as fragment misincorporation plots. The final plot in each damage plot figure exhibits a red and blue line which should indicate C-->T substitutions and G-->A substitutions. If the sample is modern both lines should lay relatively flat on the graph. A curve of the line (a “smile”) indicates damage to the DNA, which occurs most frequently towards the end of the strands, and is typically found when said DNA is old or damaged (Overballe-Petersen, Orlando, and Willerslev 2012; Briggs et al. 2007).

Table 3e: results of Galaxy and Haplogrep

<table>
<thead>
<tr>
<th>Sample</th>
<th>% endogenous</th>
<th>% Mapped</th>
<th>Mapped read avg length</th>
<th>Mean Read Depth</th>
<th>% ref seq &gt;1x</th>
<th>% ref seq &gt;2x</th>
<th>Variant sites 1x coverage</th>
<th>Variant sites &gt;1x coverage</th>
<th>Haplogroups</th>
</tr>
</thead>
<tbody>
<tr>
<td>103</td>
<td>19.88%</td>
<td>68.28%</td>
<td>101.035</td>
<td>69.986X</td>
<td>100%</td>
<td>100%</td>
<td>0</td>
<td>21</td>
<td>D4</td>
</tr>
<tr>
<td>104</td>
<td>3.42%</td>
<td>95.20%</td>
<td>82.86</td>
<td>97.5945X</td>
<td>100%</td>
<td>99.9%</td>
<td>0</td>
<td>32</td>
<td>U2a1b</td>
</tr>
<tr>
<td>141</td>
<td>16.86%</td>
<td>77.06%</td>
<td>132.55</td>
<td>71.9442X</td>
<td>100%</td>
<td>99.9%</td>
<td>0</td>
<td>12</td>
<td>H</td>
</tr>
<tr>
<td>155</td>
<td>6.79%</td>
<td>92.39%</td>
<td>117.10</td>
<td>221.422</td>
<td>100%</td>
<td>100%</td>
<td>0</td>
<td>14</td>
<td>H1u1</td>
</tr>
<tr>
<td></td>
<td>8X</td>
<td>%</td>
<td>45.9957</td>
<td>%</td>
<td>99.9</td>
<td>%</td>
<td>0</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>---------</td>
<td>-------</td>
<td>------</td>
<td>-------</td>
<td>---</td>
<td>----</td>
<td>-------</td>
</tr>
<tr>
<td>80</td>
<td>68.96%</td>
<td>54.17%</td>
<td>84.82</td>
<td>100</td>
<td>99.9</td>
<td>8%</td>
<td>0</td>
<td>27</td>
<td>R8α1α1</td>
</tr>
<tr>
<td>98</td>
<td>13.07%</td>
<td>85.98%</td>
<td>123.70</td>
<td>100</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>46</td>
<td>C1δ1α1</td>
</tr>
<tr>
<td>Control Extraction</td>
<td>3.14%</td>
<td>70.13%</td>
<td>125.429</td>
<td>0.053X</td>
<td>5.3%</td>
<td>0%</td>
<td>2</td>
<td>0</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Figure 3h: mtDNA damage plots for all individuals
Table 3f: Amount of damage to DNA seen on plots and modernity of samples

<table>
<thead>
<tr>
<th>UMFC 80</th>
<th>UMFC 98</th>
<th>UMFC 103</th>
<th>UMFC 104</th>
<th>UMFC 141</th>
<th>UMFC 155</th>
</tr>
</thead>
<tbody>
<tr>
<td>Significant damage,</td>
<td>Significant damage,</td>
<td>Some damage,</td>
<td>Little to no damage,</td>
<td>Little to no damage,</td>
<td>Little to no damage,</td>
</tr>
<tr>
<td>potentially older</td>
<td>potentially older</td>
<td>potentially older</td>
<td>modern</td>
<td>modern</td>
<td>modern</td>
</tr>
</tbody>
</table>

4.0 Discussion

4.1 Metric and Non Metric

The metric and non metric methods employed offer an interesting juxtaposition between the results of each method. Although the ancestry of each individual was supposedly known prior to utilizing these metric and non metric methods, these ancestry assessments completed do not reflect the initial assigned ancestries of Chinese and Indian. This is due to several factors, but the most likely reason is that FORDISC 3.1 and OSSA are notably limited in the populations.
they are able to assess. While FORDISC 3.1 can potentially differentiate between 13 populations groups, the proportion of White, Black, and Hispanic Americans is much higher compared to other populations. Because of this, unknown individuals are more likely to be sorted in one of these three main groups as they provide a wider array of measurements to the database. As previously discussed, OSSA and the Rhine and Gill (1995) method are severely limited in their available population assessments, which makes the discovery of a potential ancestry of the remains both more difficult and considerably less accurate. There was, however, considerable overlap in the results between these methods, with the only disparity being UMFC 104, which was rated as African by FORDISC 3.1 and European by OSSA. This disparity is most likely due to the fact that UMFC 104 does not display a post-bregmatic depression, a typical African trait according to the OSSA scoring sheet, although there is conflicting evidence about the importance of a post-bregmatic depression and assessing ancestry (Hefner and Ousley 2014; Klales and Kenyhercz 2015; Coelho et al. 2017).

The glaring issues with these results are the posterior and typicality probability. Because FORDISC 3.1 sorts the measurements no matter if they accurately fit with other populations, the database alerts the user when the measurements are too dissimilar to be accurately sorted. In this case, every individual received this feedback after the measurements were uploaded to the database. In the case of the reported chi-square probabilities, each individual received 0.00. In FORDISC 3.1, if a chi-square is less than >0.01, then the posterior probabilities should not be trusted (Ousley and Jantz 2005; Langley et al. 2016). Although the posterior probabilities reported are relatively high, in this case, they should be ignored for the sake of accuracy.

The issues within the FORDISC 3.1 assessments are easily visible in Figures 3a-3f, as the measurements of the remains are often overlapping or completely away from other populations.
measurements. UMFC 98 in particular is clearly apart from all other measurements within the database, while individuals such as UMFC 80 and UMFC 104 display multiple overlapping ancestries. This confusion in results is most likely due to genetic admixture in an individual as well as the low total members in a population within the FORDISC 3.1 database. The difficulty the database has with determining sex could also lead to incorrect ancestral assessments. If a forensic anthropologist were to rely solely on FORDISC 3.1 assessments or solely non metric assessments, the accuracy would surely be dismal. However, the combination of both metric and non metric methods is more promising, though not wholly accurate. As expected, the disparity in results is likely the result of past genetic admixture, as well as common ancestry in the distant past, which results in a variation of cranial traits that cannot be pinned down as one particular ancestry. This, in turn, makes it difficult for forensic anthropologists to rely only on metric and non metric methods as law enforcement often push for a distinct ancestry to be listed on a biological profile.

The difficulty here is that much, if not most, of the population of America is the result of waves upon waves of relatively recent, separate migration events into the continent which increase the chances of genetic admixture, and there is no accounting for admixture within metric and non metric methods. Until such a time that this admixture is accounted for in these methods or a better understanding of the metric and non metric differences between ethnic groups and whole populations is gained, complete accuracy of ancestry determination is not feasible. With the inclusion of mtDNA haplogroups, a clearer picture of ancestry should be made.

4.2 mtDNA Haplogroups
The mtDNA results gained from the sequencing of the bone samples offer some interesting insight into not only the possible lineage of these individuals, but also their age since death. In regard to the modernity of these cases, every individual was curated into the UMFC as “modern,” or within the last 50 years. However, the damage plots of UMFC 80, UMFC 98, and UMFC 103 (Figure 3g) suggest that these cases may be older rather than modern. While interesting, this has yet to be confirmed, as it is also possible that damage occurred to the DNA due to damaging properties such as bleach or UV exposure. According to what is known by the University of Montana, UMFC 103 comes from a distributor of skeletal remains for the purposes of learning in China, while UMFC 98 is an anatomical specimen donated to the UMFC. As for UMFC 80, which also served as an anatomical specimen at one point, the crania was illegally possessed and was eventually added to the UMFC after being confiscated by local police. If the remains are, in fact, ancient, then further research into their history prior to their acquisition by the University of Montana must be conducted in order for confirmation.

Furthermore, the haplogroups as given by Haplogrep identify a wide variation of possible maternal lineages for these individuals. In the case of UMFC 80, which is classified as Indian according to the UMFC, the associated haplogroup is listed as R8a1a1. In accordance with the UMFC, this individual's mtDNA haplogroup appears to stem from East India, as the subclade R8a appears most often in Indian populations (Larruga et al. 2017). Although haplogroup R is widespread as carriers of mtDNA haplogroup R within the clade U, which arose from haplogroup R, spread throughout the entire Eurasian continent some 50kya (Larruga et al. 2017). While promising that the ancestry listed for UMFC 80 and the haplogroup assigned match, the disparity between the ancestry found via metric and non metric assessments is concerning.
UMFC 104, which is curated in the UMFC as having an Asian ancestry, the haplogroup assigned is U2a1b. This haplogroup, a clade of mtDNA haplogroup R, is most typically found in Europe, with 90% of the population belonging to macro-haplogroups HV, U, and JT. Alongside this, haplogroup U makes up nearly 20% of Caucasian populations which includes modern populations in Europe and Western Asia (Chinnery and Gomez-Duran 2018). However, the subclade U2 is found in South Asia, Western Asia, North Africa, and Europe. This indicates that it is entirely possible that UMFC is of Indian or Chinese descent, but it also alludes to possible European ancestry which would meet the assessments of both metric and non metric methods. In a case such as this, it is important to rely on where the UMFC received the remains from until further analysis can be conducted in order to form a more definite biological profile. Therefore, it is most likely that UMFC 104 was of Indian descent as haplogroup U2 is less commonly found in China, the probability of the individual being from South or Western Asia is higher.

UMFC 141, listed as Chinese, and UMFC 155, listed as Indian, have both been shown to be from the mtDNA haplogroup H and H1u1, most commonly found in areas of Europe (Brotherton et al. 2013). However, haplogroup H exhibits a wide spread throughout the European continent, perhaps the largest of distributions, making it particularly difficult to discern a discrete ancestry or ethnicity of these individuals, though haplogroup H reaches a frequency of 45% within Europe (Chinnery and Gomez-Duran 2018). There is potentially more synchronization between the metric and non metric results with the haplogroup results here, as the metric and non metric methods listed these individuals as of European and African descent respectively.

Finally, both UMFC 98 and UMFC 103 have been found to be from the haplogroups C1d1a1 and D4 respectively. There is a possible issue with these results, as both mtDNA haplogroups are often of Native American descent. Haplogroup D4, while appearing in Native
American populations, is also incredibly common in areas of Eastern Asia, such as China and Korea (Derenko et al. 2010). In this case, UMFC 103 was received from China and has been labeled within the UMFC as being of Asian descent. So, it is unlikely that UMFC 103 is of Native American descent.

UMFC 98, in comparison, is of the mtDNA haplogroup C, which is common in both Northeast Asia and within the Americas. The subclade here, C1d, is often found in Native American populations and is considered a founding American haplogroup and continues to be mostly specific to the Americas (Figueiro, Hidalgo, and Sans 2011; Kumar et al. 2011) it is not entirely absent from Asian lineages; in fact, haplogroup C is derivative of a subclade of the Asian specific M clade (Derenko et al. 2010; Nikitin, Newton, and Potekhina 2012) making it entirely possible for an individual with an mtDNA haplogroup C to be of Asian ancestry. In a case such as this, it is important that the acquisition of the individual be researched more thoroughly to ensure that the remains have been properly curated and that if officially labeled as Native American, that the University of Montana is in compliance with NAGPRA. Due to the mtDNA group of C1d1, it is likely that UMFC 98 is of Native American descent. Further investigation will be conducted through the University of Montana to confirm where the remains were acquired from, when the remains were acquired, and potential tribal affiliation.

While the results of the mtDNA haplogroups are quite different from the listed, metric, and non metric methods, the haplogroups do not necessarily confirm the ancestry of these remains, as mtDNA results are not concrete depictions of ancestry. This is due to the nature of using mtDNA haplogroups to determine ancestry. Although the proportion of mtDNA haplogroups in a region enhances the possibility that an individual is of that biogeographical ancestry, it is not always the case. When looking at an individual’s haplotype, such as UMFC
141 and UMFC 155, both of which are from a lineage that is found in high rates in Europe, it is possible that their biogeographical ancestry is not in fact European, but potentially Asian or African, as multiple continents see reasonably high rates of haplogroup H. Even with this limitation, the mtDNA results, when in direct comparison with FORDISC 3.1 highlight the essential need for change within the FORDISC 3.1 database if it is to be used as the most accurate ancestry determination by forensic anthropologists. Of the six individuals, FORDISC 3.1 results were in line with the mtDNA haplogroups twice, or 33.3% of the time.

In order to truly understand a person’s ancestral lineage, the Y-DNA haplogroups and AIMs could offer unparalleled insight when combined with mtDNA, as mtDNA is limited in its own right. Such additions would be beneficial in the cases of UMFC 98 and UMFC 103 to analyze further the potential Native American ancestry of these individuals.

<table>
<thead>
<tr>
<th>Method</th>
<th>UMFC 80</th>
<th>UMFC 98</th>
<th>UMFC 103</th>
<th>UMFC 104</th>
<th>UMFC 141</th>
<th>UMFC 155</th>
</tr>
</thead>
<tbody>
<tr>
<td>UMFC label</td>
<td>Indian descent</td>
<td>Indian descent</td>
<td>Chinese descent</td>
<td>Chinese descent</td>
<td>Chinese descent</td>
<td>Indian descent</td>
</tr>
<tr>
<td>FORDISC 3.1</td>
<td>European</td>
<td>European</td>
<td>European</td>
<td>African</td>
<td>African</td>
<td>European</td>
</tr>
<tr>
<td>Non Metric</td>
<td>European</td>
<td>European</td>
<td>European</td>
<td>European</td>
<td>African</td>
<td>European</td>
</tr>
<tr>
<td>mtDNA</td>
<td>R8a1a1</td>
<td>C1d1a1</td>
<td>D4</td>
<td>U2a1b</td>
<td>H</td>
<td>H1u1</td>
</tr>
</tbody>
</table>

### 4.3 Further research
Furthering this research has the potential to be highly beneficial to forensic anthropologists when creating a biological profile. However, a larger sample of crania must be analyzed using this research method in order for the results to reflect the deficiency of FORDISC 3.1 assessments. Further research into this subject should also include Y-DNA haplogroups, as a nuclear SNPs panels, as well as further study into the acquisition of the remains into the UMFC, which could help confirm both ancestry and modernity.

**5.0 Conclusion**

FORDISC 3.1 and other non metric means of ancestry assessment, while valid and essential to the creation of biological profiles, are admittedly lacking due to their inherent subjectivity and the potential for human error, as well as the fact that there are not unique enough features to reliably distinguish differing ancestries using these methods. To reconcile this, the inclusion of mtDNA research would be inherently valuable and would decrease the chance that the remains are being inaccurately and unethically curated. In order to reach that point, however, more research must be conducted on the disparity between the metric and non metric results with the results of sequencing to confirm the need for DNA evidence. As the shortcomings of the metric and non metric methods are widely known, a new or updated database should be created that utilizes DNA evidence to provide a more accurate depiction of ancestry.
Literature Cited


Blohm, Tre. 2019. “Snow aDNA Lab Galaxy Bioinformatics Tutorial.”


Dunn, Rhian R., Micayla C. Spiros, Kelly R. Kamnikar, Amber M. Plemons, and Joseph T.


Hubbe, Mark, Tsunehiko Hanihara, and Katerina Harvati. 2009. “Climate Signatures in the


https://doi.org/10.5962/bhl.title.542.


Shi, Cheng-Min, Qi Liu, Shilei Zhao, and Hua Chen. 2019. “Ancestry Informative SNP Panels for Discriminating the Major East Asian Populations: Han Chinese, Japanese and


Webster, Mark, and H. David Sheets. 2010. “A Practical Introduction to Landmark-Based Geometric Morphometrics.” *Quantitative Methods in Paleobiology* 16: 163-188. DOI: https://doi.org/10.1017/S1089332600001868
