This is to certify that the
thesis entitled

Predicting Mitochondrial DNA Quality Based on Bone Weathering and Type

presented by

Lisa M. Misner

has been accepted towards fulfillment
of the requirements for the

Master of Science degree in Forensic Science

Major Professor's Signature

8/13/04

Date

PREDICTING MTDNA QUALITY BASED ON BONE WEATHERING AND TYPE

By

Lisa M. Misner

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTERS OF SCIENCE

Forensic Science

2004

ABSTRACT

PREDICTING MTDNA QUALITY BASED ON BONE WEATHERING AND TYPE

By

Lisa M. Misner

Investigation of human skeletal remains often requires DNA analysis by the forensic biologist. This study was designed to determine if the degradative state of a skeleton, bone and/or type of bone can be used to predict the presence of DNA and its qualitative and quantitative condition, reducing the time and expense of analysis. Eighty-nine bone samples originating from 36 skeletons, found at an abandoned cemetery in Pittsburgh, Pennsylvania dating to the mid-1800s were tested. QuantiBlot-based quantification of mtDNA was not successful due to very low DNA yields, thus PCR-based analysis of mtDNA was used. Amplification of 220 bp fragments was most successful, followed by 107 bp, then larger fragments. The weathering condition of whole skeletons and individual bones were not useful in predicting successful DNA amplification, while sex of the individual may play a role in the preservation of DNA, with DNA from male skeletons amplifying more often. The best indicator of the quality of mtDNA obtainable from a bone sample, however, was the type of bone from which DNA was extracted. Amplification occurred most frequently from femurs, followed closely by ribs and distantly by pelvises. DNA sequencing affirmed the skeletal origin of the DNA. The findings indicate that knowledge of bone type is the single most important factor in predicting success of DNA analysis, while other factors, particularly the appearance of bone, are of minimal use.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

An important goal of any forensic investigation involving unidentified human remains is positive identification. Forensic specialists including anthropologists and odontologists may evaluate the remains, estimating the individual's sex, stature, age at death, and ancestry. Any identifiable characteristics such as fingerprints, if flesh is still intact, unique skeletal features and dental arrangement are also noted. However, when antemortem reference records are not available for comparison or remains are fragmented or otherwise in a state in which definitive conclusions cannot be made as to the person's identity, DNA analysis may be required. Indeed, analysis of DNA from human skeletal remains as a means of identification has been used in numerous cases, beginning in 1989 (Pääbo *et al.*, 1989). Advancements in techniques and applications occurred in the 1990's (Boles *et al.*, 1995; Hagelberg *et al.*, 1991; Primorac *et al.*, 1996; reviewed in O'Rourke *et al.*, 2000), and in 1991 the Armed Forces DNA Identification laboratory was established for the identification of the remains of U.S. military personnel (Holland and Parsons, 1999). Ten years later a high-throughput method was developed for the analysis of the thousands of bones connected with the attacks on the World Trade Center (Holland *et al.*, 2003). Bones encountered by the forensic biologist often vary in their degree of degradation, and if more than one bone is available, the scientist may choose which bone to analyze based on its appearance. Although bone degradation and DNA degradation are separate processes, any relationship between the two has not been well defined. Knowledge about such a relationship would help the forensic biologist predict the success of analyses involving skeletal remains. Currently, this proceeds through a series of trial-and-error processes until results are obtained or the samples are abandoned, potentially

wasting laboratory resources and the analyst's time. Knowing the likelihood of obtaining usable data based on the type and condition of the bone would be beneficial to the forensic biologist as it would expedite the analysis procedure and reduce costs.

Bone Degradation

Though the structure and composition of skeletal material allows it to persist long after soft tissue has decomposed, it is still subject to degradative processes. Several studies have addressed the issue of bone degradation and the factors that affect it. At the microscopic level, Shackleford (1996) analyzed the effects of soil condition and geological age of the skeleton on the histology of seven femurs from Native Americans found in Alabama. Five of the femurs were excavated from soil described as a mixture of sand, silt, clay, and shell mound, and were estimated to have been buried during the Mississippian period (500–1000 years ago). Two femurs were estimated to have originated from the Archaic period (3000–5000 years ago); one was removed from a dry, protected bluff, the other was surrounded by a shell mound. Sections of each femur were analyzed using electron microscopy and electron diffraction. All seven samples showed intact mineral constituents similar to that of modern bone, indicating that neither the age of the skeleton nor the composition of the surrounding soil greatly affected the preservation of histological structure.

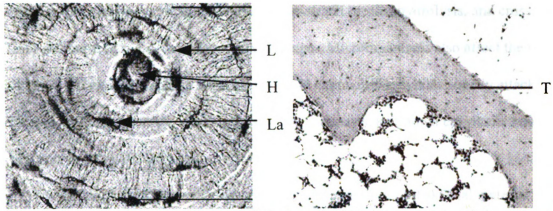
Solomon and Hasse (1967) analyzed the histology of bone from sites in Israel that varied in age, soil composition, and moisture content. A total of six bones, two humeri, three femurs, and one metatarsal were analyzed originating from 3500 B.C. to 200 A.D. The bones were excavated from one of three different soil types: sand/clay from a humid

climate; loess (fine-grained deposit of wind blown dust), lime, and salts from a dry climate; and clay, granite, powder, and salts from a dry climate. Sections of each bone were examined using various stains to compare the histological structure of these samples to fresh bone. Results indicated that histological structure did not vary among bones from the same soil condition regardless of age, further confirming that the time since death of the individual has little effect on bone preservation. However, a difference was noted among bones removed from soils of similar composition but differing humidity. The histological structure of samples from humid climates showed a greater degree of degradation than those of from dry climates, indicating that exposure to moisture had a negative effect on bone preservation.

In an overview of the processes of bone diagenesis, Hedges (2002) summarized the histological and biochemical changes that occur as a bone degrades. One correlation observed was with loss of collagen, the porosity of the bone decreased and the crystallinity increased. Hedges noted that microbial attack (both bacterial and fungal), as determined by histological modification (e.g. re-configuration of the mineral phase), was the major cause of collagen loss, with severely attacked bone having lost at least 80% of its original collagen. In addition, low temperature was found to prevent microbial attack, and extreme hydrology (either very wet or very dry) also inhibited microbes. Analyzing the effects of soil conditions, soils of neutral pH had calcium and phosphate concentrations similar to those found in hydroxyapatite ($Ca_{10}(PO_4)_6(OH)_2$; the major bone mineral), which led to a slower dissolution rate. In contrast, soils of low pH had lower calcium and phosphate concentrations and dissolution occurred faster, as the protons from the soils replaced the calcium ions in the hydroxyapatite of the bone.

At a macroscopic level, Micozzi (1991) summarized the results of several studies which indicated that the survivability of a bone increases with increasing density, and that various factors such as age, size, and bone type affect density. With respect to age, there is a biphasic relationship to bone density, with young and very old individuals having less dense bones. The density of bone is related to its microstructure. There are two main microstructures of bone, compact bone and cancellous (spongy) bone (Figure 1). Compact bone is composed of a system of osteons in which compact concentric layers (lamellae) of hydroxyapatite and collagen are arranged around a central Haversion canal. Between the layers are small spaces (lacunae) where the mature bone cells (osteocytes) reside. Spongy bone consists of trabeculae (spicules) of bone separated by large open spaces, and osteocytes are found within lacunae of the trabeculae. Every bone contains both compact and spongy bone structure though the distribution of each varies depending on bone type. For example, all bones have compact bone along the cortex, however, in long bones, such as the femur, fibula, tibia, humerus, radius, and ulna compact bone continues throughout the thickness from the cortex to the medulla all along the shaft of the bone, while trabecular bone is present beyond the cortex at the articular ends. In contrast, flat bones, such as the rib and pelvis, have spongy trabecular bone beyond the cortex in all portions of the bone.

Figure 1. Microscopic comparison of compact and spongy bone at 100X.



Compact bone is represented on the left. L=lamella, HC=Haversion canal, La=lacuna. Spongy is represented on the right. T=trabecula. (Figures taken from McCutchen, 2003)

Pastron and Clewlow as cited by Micozzi (1991) ranked the survivability of the long bones removed from prehistoric human burials found within caves and rock shelters of a rock wall in New Mexico. They found that the femur was the most durable, followed by the tibia, fibula, humerus, ulna, and radius. Such conclusions indicate that surface area of a bone plays a role in its survivability; and indeed Gill-King (1997) described surface area as an "important factor" in the rate of bone degradation. Bones with a larger surface area to volume ratio have more contact with the surrounding soil, making them more susceptible to ion exchange and the actions of the microorganisms within.

DNA Degradation in Skeletal Remains

As skeletal material undergoes the process of degradation, the organic molecules within the bone, including DNA, also degrade (Gotherstrom *et al.*, 2002). Studies have been carried out to analyze the degradation of the DNA molecule in bone under various conditions (Parsons and Weedn, 1996; Smith *et al.*, 2003). High temperature, high

moisture, and acidic pH accelerate DNA degradation. In addition, chemical modifications occur, such as oxidation, single-strand breaks, hydrolysis, and cross-linking from free radical attack and UV damage. Microorganisms also affect the stability of DNA, producing enzymes such as endo- and exonucleases which directly attack DNA (Rogan and Salvo, 1990). As long as soft tissue persists, additional enzymatic attack can occur from autolytic enzymes.

Although the factors that contribute to the degradation of skeletal material and DNA have been studied individually by numerous researchers, the way in which the two might correlate is not well understood. A limited number of studies have been carried out however, which address the issue (Gotherstrom *et al.*, 2002; Tuross, 1994; others reviewed in Kaestle and Horsburgh, 2002; Parsons and Weedn, 1997; and O'Rourke *et al.*, 2000). In general, researchers have found that the likelihood of obtaining usable DNA from a sample can be predicted, to a degree, from the gross morphology of the sample, with harder, more intact bones providing usable DNA more often than softer, more brittle bones (Kaestle and Horsburgh, 2002; and Parsons and Weedn, 1997 and references within). In addition, a correlation was found between the microscopic preservation of bone samples and DNA recovery, with usable DNA being present in bones with better preserved microstructure (O'Rourke *et al.*, 2000).

Tuross (1994) addressed the issues of time since death and burial conditions by analyzing three collections of bone: a set of three human femurs obtained during surgical procedures; bone samples originating from elephant and wildebeest, which were collected once a year for fifteen years after the animals' deaths; and a variety of human skeletal collections from North and Central America, which had been buried up to 7000

years. DNAs extracted from the samples were quantified by ethidium staining in 4-20% acrylamide gels and compared to modern samples. Results showed the largest difference in DNA content occurred between bones that had been buried immediately (the human skeletons) and those which were exposed to surface conditions (the elephant and wildebeest). Although the unburied samples contained nearly one tenth the amount of DNA of the buried samples, the amount of DNA from the unburied samples remained fairly stable as the years passed, thus suggesting DNA is more stable once the deceased is fully skeletonized. The author proposes this is due to binding of the DNA to hydroxyapatite, which would protect the DNA from hydrolysis. It should be noted that since the DNA extract was quantified via gel analysis, it is unclear whether or not the total amount of DNA extracted was indigenous to the individual, as bacterial and/or fungal DNA may have been coextracted. In addition, comparisons were made among bone samples of three different species, each likely having different microstructures, not to mention the large difference in time since death and in climate, each of which represents a variable that should be considered when evaluating these results. Several studies since Tuross' have shown that environmental conditions play a larger role than time since death in relation to the quality of DNA obtainable from bones (as reviewed by Kaestle and Horsburgh, 2002)

Gotherstrom *et al.* (2002) focused on finding a reliable, fast, and inexpensive way of predicting the preservation of DNA in ancient bones and teeth. Two sets of bones were utilized: one modern, experimentally degraded set of bovid bones and one set of ancient horse bones; both sets were pulverized for analysis. Three measurements were taken from the bones: the crystallinity index for hydroxyapatite as determine by X-ray

diffraction, the presence of mitochondrial DNA (mtDNA) as detected by the polymerase chain reaction (PCR), and the amount of extracted collagen. The results of their study suggested a correlation between crystallinity of the hydroxyapatite and the presence of DNA, with amplifiable DNA decreasing as crystallinity increases, thus supporting Tuross' (1994) suggestion that DNA stability was related to hydroxyapatite binding. A similar relationship between the presence of DNA and the amount of collagen left in the bone was also noted. Based on the research of other scientists, the authors hypothesized that DNA degradation may be related to mineral sorption, and that once hydroxyapatite no longer protects the DNA molecule, it is vulnerable to degradative processes.

Although studies such as these have begun to address the correlation between bone degradation and the amount of DNA degradation, they do not provide the large-scale, statistically testable data necessary to draw significant conclusions. The majority of these studies were performed with small sample sizes, comparing samples from different geographic locations, species, and time since death. Most of the skeletal material used in these studies was of animal origin, and in some cases, weathering conditions were mimicked in a laboratory setting. While these studies were helpful in raising new questions and providing suggestions for a relationship between bone and DNA degradation, a single, large, aged set of skeletal remains is needed to solidify any correlations.

As described by Gotherstrom *et al.*, 2002 and other authors (Gilbert *et al.*, 2003; Kalmar *et al.*, 2000; Pääbo *et al.*, 1989), mitochondrial DNA (mtDNA) analysis is the primary mode of DNA typing of aged skeletal remains. Mitochondria are cytoplasmic organelles containing a small, maternally-inherited, circular piece of DNA distinct from

the linear chromosomes of the nucleus. Thousands of mitochondria can exist in a single cell, and as such there are thousands of copies of mtDNA in each cell. This is far different then the two copies (maternal and paternal) of a DNA locus that are analyzed in most forensic analyses. Human mtDNA was first sequenced in its entirety in 1981 (Anderson *et al.*) and is approximately 16569 base pairs (bp) in length. The majority of genes in mtDNA encode proteins essential to mitochondrial function, thus these areas are highly conserved among individuals, and are generally not useful for identification. An approximately 1125 bp stretch of non-coding DNA known as the control region contains two regions which mutate frequently and therefore vary among individuals. The sequence variation of these two areas, known as hypervariable region I and II (HVI and HVII), are generally the focus of forensic analysis (reviewed by Holland and Parsons, 1999). Typically, this involves amplification of both hypervariable regions using PCR followed by a sequencing procedure to determine the exact DNA sequence, which is then compared to the Anderson reference sequence. The nucleotide differences between the sample and the reference sequence constitute the mtDNA profile for that sample, which can then be compared back to a known sample or a maternal relative for positive identification.

Description of the Voegtly Samples

Bones for this study were obtained from Dr. Doug Ubelaker of The Smithsonian Institution and sections collected by Jennifer Drier of George Washington University. As described by Ubelaker *et al.* (2003), the skeletal material originated from a burial site, known as the Voegtly Cemetery, in what is now northern Pittsburgh, Pennsylvania. In

1787, an area north of the Alleghany River was designated as Old Alleghany Town. During the early to mid-1800's the area was settled mainly by Swiss-German immigrants. A prominent millwright family, by the name of Voegtly, together with their Swiss neighbors, bought 161 acres along the Alleghany River in Old Alleghany Town in 1822. In 1833, after a donation of land from the Voegtly family, The First Evangelical Church of Alleghany (Voegtly Church) was constructed. The adjacent cemetery was utilized from that point until 1861, when land for a larger cemetery was acquired atop the nearby Troy Hill, just east of the original cemetery. A series of expansions and modifications to the Church occurred over the next century, and in the process, the original cemetery, without permanent markers, was built over.

From 1881 to the early 1900's the church experienced a fluctuation in membership, with continuous decline from the 1950's on. In 1959, it was announced that the Voegtly Church was in the right-of-way for the new East Street Valley Expressway (I-279/I-579). When the last surviving contact with the Voegtly family died in 1972, the parsonage was destroyed, and in November of 1984, the Pennsylvania Department of Transportation acquired the property. In May of 1985 the congregation filed for dissolution and the Voegtly Church was officially disbanded and destroyed the same year. Two years later, in the spring of 1987, the company in charge of construction monitoring discovered five coffins and the associated skeletal remains. A full-scale archaeological excavation was carried out from June to September of 1987 (a site map of the grave locations is shown in Figure 2), in which all remaining skeletal material was removed, recorded (each burial was assigned a consecutive identification number upon discovery), packaged, and taken to The Smithsonian Institution for processing.

Data collection at the Smithsonian began in 1993 following the *Standards for Data Collection from Human Skeleton Remains* (Buikstra and Ubelaker, 1994). At the laboratory the remains were unpacked, cleaned, and estimations of sex, age, and living stature, were recorded when possible. Estimates of sex were made by evaluating cranial and pelvic morphology when possible. When these more reliable features were not available or in poor condition, skeletal robusticity or presence of particular artifacts found with the remains (e.g. earrings, remnants of a moustache) was used. In general, age was estimated by eruption of teeth and dental wear, and stature was estimated by femur length. In addition to these estimates, the degree of weathering was assessed for each individual burial based on a classification system taken from Behrensmeyer (1978). The system is based on a scale from 0 to 5 and briefly includes:

Stage 0: Bone surface shows no signs of cracking or flaking due to weathering.

Stage 1: Bone shows some cracking, usually longitudinal in long bones.

Stage 2: Some cracking and flaking is apparent, especially on the outermost concentric thin layers of the bone.

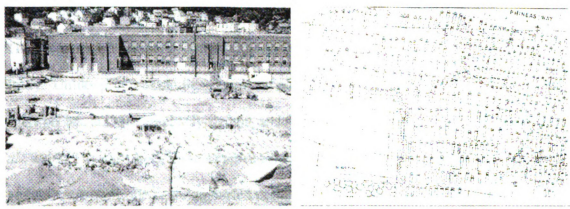
Stage 3: Bone surface has rough patches of weathered compact bone; external concentric layers have been removed, but weathering does not penetrate deeper than 1.0-1.5 mm.

Stage 4: Bone surface is coarse and splinters may exist; weathering reaches into inner cavities.

Stage 5: Bone has large splinters and is easily broken; original bone shape may be undeterminable.

A detailed record of the data collected at the Smithsonian is available in Ubelaker *et al.* (2003).

Figure 2. Excavation Site and Site Map of Grave Locations of the Voegtly Cemetery



(Figures taken from Ubelaker *et al.*, 2003)

According to the church records, which were translated from their original German text, 823 individuals were buried in the original Voegtly Cemetery between 1833 and 1861. Not all of these records contained data on name, sex, or age. In most cases, this lack of information corresponded to records from infants and children. Sixteen of the records with missing information were marked as "not seen", probably indicating a child that was miscarried or stillborn. Of the 799 records in which age was recorded, 311 (38.9%) were less than 1 year old, 192 (24%) were between 1 and 5 years, and 66 (8.3%) were between 5 and 20 years of age. The sex of 774 individuals was recorded, 365 (47.2%) were females and 409 (52.8%) were males.

Of the 724 individuals recovered from the abandoned cemetery, 208 bones from 88 individuals were sampled for analysis. Two small sections (approximately two to four cm^3) were removed from each bone. One set of sections was sent to Brazil for histological examination; the other set was reserved for DNA analysis at Michigan State University.

Research Goals

The goal of this study was to utilize the large set of skeletal remains recovered from the Voegtly Cemetery to determine if a correlation exists between bone type, its level of degradation, and the qualitative and/or quantitative condition of the DNA within. The specimens from the Voegtly Cemetery provide an ideal sample set for a controlled study. Unlike previous studies where comparisons were made between bones of varying age, species, and geographical regions, and sample sizes were small (e.g., Gotherstrom *et al.*, 2002; Tuross, 1994), these samples provide a large sample set in which these variables are controlled. Being buried in the same cemetery, the remains were likely to be exposed to the same soil environment, and through knowledge of church records, all bones had been exposed to this environment for roughly 150 years. Also, the history of the church and its members indicate that the majority of the individuals buried in the cemetery were Swiss-German immigrants, thus making ethnicity a controlled variable as well.

Skeletal remains are the most likely type of sample requiring analysis of the forensic biologist well after an individual's death. Currently, multiple days are needed to assess if usable DNA is obtainable from a particular sample. Without prior knowledge of the likelihood of successful analysis of skeletal material, the scientist may be required to repeat the procedure multiple times on different bones (or portions of) until results are obtained. Being able to make predictions of the amount and quality of DNA from aged bone based on the type of skeletal material (long bone, flat bone, short, or irregular) and its level of weathering, would minimize time and expenses spent on analysis and maximize efficiency.

Materials and Methods

Summary of Samples

A subset of 89 bones from the total set obtained from The Smithsonian Institution was analyzed in this study. Table 1 lists the bones analyzed by the burial number assigned during their recovery. Each bone is described based on information obtained from The Smithsonian Institution, including sex, estimated age, and one of five weathering stages (titled "skeletal weathering stage" herein). As detailed above and by Ubelaker *et al.* (2003), the weathering stages originally assigned to the samples by the anthropologists at The Smithsonian considered the condition of the skeleton as a whole. In addition to this, a new staging system was developed at MSU and assigned to each bone based on visual inspection for the DNA study. This individual bone staging system is separated into four categories:

Stage 1: Bone surface shows minimal flaking. The bone piece is still whole, with no large pieces broken off.

Stage 2: Bone surface shows some flaking, pieces of bone are coming off in sheets. Small pieces of bone are breaking away from the cut piece.

Stage 3: The bone is fragmented into several pieces. At least one large piece is still present.

Stage 4: The bone is extensively fragmented. No large pieces are present.

The bones used in this study spanned five of the six weathering stages assigned by Ubelaker *et al.* (2003): 6 from stage 1, 23 from stage 2, 22 from stage 3, 18 from stage 4, and 20 from stage 5. The three samples from the single individual classified as stage 0 were not tested due to the insufficient sample size and nonexistence of comparison

samples. A total of 2 crania, 28 femurs, 1 fibula, 25 pelves, and 33 ribs were tested, coming from a total of 36 individuals. Of these, 11 were female, 17 were male, and 1 was labeled as "possibly male". This individual had large mastoid processes and supraorbital ridges, but female sex could not be ruled out. With seven samples sex could not be determined.

Preparation of Bones for Subsequent DNA Extraction

Approximately 1 cm^3 of bone was cut from the source section using a Dremel MultiPro tool (model no. 395-76) with cut-off wheel attachment number 409. Cut pieces were collected in a 17 X 100 mm polypropylene tube (FisherScientific) and immersed in 1-2 ml of filter-sterilized wash buffer (1% SDS, 25 mM EDTA, filter sterilized using a 0.2 µm syringe filter (FischerScientific)) and 0.1 mg/ml proteinase K and incubated for one hour at room temperature. Following the incubation, the wash buffer was poured off and each sample was washed with 1ml of sterile dH_2O six consecutive times. Excess water remaining in the tube was removed with a sterilized Pasteur pipette. Samples were dried using compressed air passed through a 0.45 µm vent filter (Millipore) for 15–30 min.

Dried samples were ground to a powder using an IKA A11 Basic Grinder with a tungsten blade (IKA Works, Inc.). The mass of each sample was recorded both before and after grinding, and ground samples were collected in a 2 ml microcentrifuge tube. The grinding mill was cleaned with 10% bleach followed by 70% EtOH after the grinding of each sample and exposed to UV light at 2500 Joules/cm^2 when the next sample was from the same burial number, and thus the same individual. Between

Table 1. Summary of Bones Analyzed
Headings denote burial number, bone type, age, sex (male (M), possibly male (M?), or female (F)), skeletal weathering stage (0–5), and individual bone weathering stage (1–4).

Burial Number	Bone Type	Age	Sex	Skeletal Weathering Stage	Bone Weathering Stage
027	Rib	22	M	5	3
027	Femur	22	M	5	1
030	Femur	40-55	M	5	3
034	Pelvis	35-45	M	4	2
034	Rib	35-45	M	4	4
034	Femur	35-45	M	4	1
047	Pelvis	11	F	4	2
047	Rib	11	F	4	4
047	Rib	11	F	4	1
047	Femur	11	F	4	2
111	Pelvis	30-35	M	2	1
111	Rib	30-35	M	2	3
111	Femur	30-35	M	2	1
114	Femur	10.5	F	5	2
124	Pelvis	28-35	M	3	3
124	Rib	28-35	M	3	4
124	Femur	28-35	M	3	2
126	Femur	25-30	F	3	1
126	Pelvis	25-30	F	3	3
126	Rib	25-30	F	3	3
132	Rib	25-30	M	2	3
132	Femur	25-30	M	2	1
132	Pelvis	25-30	M	2	2
164	Femur	22-26	M	3	1
164	Rib	22-26	M	3	3
164	Pelvis	22-26	M	3	2
167	Rib	15-16	M?	2	2
167	Pelvis	15-16	M?	2	2
167	Femur	15-16	M?	2	1
192	Rib	60-80	M	2	1
192	Pelvis	60-80	M	2	2
192	Femur	60-80	M	2	2
203	Fibula	1.9mos.		1	2
203	Cranium	1.9mos.		1	2
203	Rib	1.9mos.		1	2
256	Pelvis	35-45	M	3	2
256	Rib	35-45	M	3	4
256	Femur	35-45	M	3	3
260	Rib	3		4	2
260	Femur	3		4	2
322	Femur	20-24	M	2	2
322	Rib	20-24	M	2	2
322	Pelvis	20-24	M	2	2
328	Pelvis	40-45	M	2	2
328	Rib	40-45	M	2	2
328	Femur	40-45	M	2	1

Table 1. Continued

331	Rib	15	M	5	3
331	Femur	15	M	5	1
345	Rib	Adult	M	3	4
345	Femur	Adult	M	3	1
345	Rib	Adult	M	3	2
345	Pelvis	Adult	M	3	3
348	Rib	27-35	M	3	4
348	Femur	27-35	M	3	1
348	Pelvis	27-35	M	3	3
349	Femur	5.8		5	3
355	Rib	4		5	4
355	Femur	4		5	3
381	Femur	25-30	M	2	1
381	Rib	25-30	M	2	2
381	Pelvis	25-30	M	2	3
389	Rib			5	3
389	Pelvis			5	3
402	Pelvis	25-40	F	5	3
409B	Femur	30-40	M	3	2
409B	Rib	30-40	M	3	3
409B	Pelvis	30-40	M	3	3
447	Rib	30-40	F	4	3
447	Femur	30-40	F	4	2
448	Femur	25-35	F	5	2
448	Rib	25-35	F	5	4
489	Pelvis	1.5		2	1
489	Rib	1.5		2	1
529A	Rib	18-25	F	4	2
529A	Femur	18-25	F	4	1
529A	Pelvis	18-25	F	4	2
539	Rib	30-40	F	4	4
539	Pelvis	30-40	F	4	2
540	Rib	32-45	F	5	4
540	Pelvis	32-45	F	5	2
545	Femur	25-32	F	1	2
545	Pelvis	25-32	F	1	3
545	Rib	25-32	F	1	4
546	Pelvis	18-21	F	4	2
546	Rib	18-21	F	4	3
686	Cranium	4-5.5		5	2
704	Femur	45-60	M	5	2
704	Pelvis	45-60	M	5	1
704	Rib	45-60	M	5	3

samples of differing burial numbers, the mill was disassembled and with 10% bleach followed by 70% EtOH to remove any trapped bone dust, then reassembled and exposed to UV light at 4167 J/cm^2.

DNA Extraction

Four hundred microliters of digestion buffer (20 mM Tris, 100 mM EDTA, 0.1% SDS) and 0.4 mg/ml proteinase K was added to each ground bone sample and incubated overnight at 56°C. After incubation, one volume of phenol was added, each sample was vortexed then centrifuged at 14000 rpm at 4°C for 5 min. The aqueous layer was removed to a clean microcentrifuge tube. If the phenol layer appeared dark brown, a second or third phenol extraction was performed. Following the phenol extraction, one volume of chloroform was added, the sample was vortexed and centrifuged as above, and the aqueous layer was removed to a new microcentrifuge tube. DNA was precipitated with 40 µl of 3 M sodium acetate and 800 µl of 95% EtOH and incubated at -20°C for at least one hour. Samples were centrifuged at 14000 rpm at 4°C for 20 min, after which the supernatant was removed and the pellets were vacuum-dried for approximately 20 min. Each sample was resuspended in TE buffer (10 mM Tris, 1 mM EDTA), using 1 µl/mg of ground bone, and stored at -20°C.

DNA Quantification Using QuantiBlot

To determine the quantity of mtDNA extracted from each sample, the QuantiBlot Human DNA Quanitification Kit (Perkin Elmer) was used with modifications. Three biotinylated probes were created to target human mtDNA (Table 2) (Promega) in place of

the QuantiBlot D17Z1 Probe provided in the kit, which targets nuclear DNA. Optimization of the mtDNA probes led to the use of 30 µl of the individual probe (100 pmol/µl) for each membrane. In addition, the hybridization temperature was brought down from 50°C to 40°C and all other steps were performed at 24°C. One microliter of 10 ng/µl K562 DNA High Molecular Weight (Promega) was used as a positive control. Raccoon (*Procyon lotor*), red fox (*Vulpes vulpes*), badger (*Taxidea taxus*), and salmon (*Oncorhynchus sp.*) DNA were used to test for species specificity. The DNA concentration of these samples were estimated by running them next to 10 ng of cell line DNA on a 1% agarose gel and then diluted to a concentration similar to the cell line source. Thirty microliters of bone extract was loaded onto the membrane. All other steps were followed as described in the QuantiBlot kit protocol; the chemiluminescent method of detection was employed.

Table 2. mtDNA probes used in QuantiBlot procedures

Name	Sequence
DLoop11F	5' CTATCACCCTATTAACCACTC 3'
DLoop31R	5' GAGTGGTTAATAGGGTGATAG 3'
CO1 6281F	5' AACAGTCTACCCTCCCTTAG 3'

F=forward, R=reverse, the numbers correspond to the position of the 5' base of the probe in the complete human mtDNA sequence (Anderson *et al.*, 1981), and DLoop and CO1 (cytochrome *c* oxidase complex 1) refer to the region of the mtDNA to which the probe is targeted.

DNA Amplification of Bone Samples

Preliminary amplification reactions contained 1 U Taq DNA polymerase (Promega), 0.2 mM each of dNTP (Promega), 2 µmol each of mtDNA-specific primers (Genosys) (Table 3), 10X PCR buffer (Promega), 10X $MgCl_2$ (Promega), and 1 µl of sample DNA in a total volume of 20 µl. In early experiments, two concentrations of

sample DNA were used, one undiluted and one 1:20 dilution of the bone extract. As experiments progressed, only the 1:20 dilution was used. Also, as the experiments proceeded, HotMaster Taq (Eppendorf), along with the provided buffer, was used instead of standard Taq. In addition, 5X HotM Enhancer (Eppendorf) was included in later reactions to reduce PCR inhibition. The amplification reaction consisted of denaturation at 94°C for 2 min, followed by 38 cycles of denaturation at 94°C for 30 s, annealing at 56°C for 45 s, and extension at 72°C for 45 s. The last cycle was followed by an additional extension step at 72°C for 5 min. Fragment size and amplification quality were visualized by running 5 µl of the PCR product, or amplicon, on a 2% agarose gel followed by staining with ethidium bromide.

Table 3. mtDNA primers used for amplification and sequencing

Name	Sequence	Region of mtDNA	Size of product
F155	5'TATTTATCGCACCTACGTTC 3'	HVII	329 bp
R484	5'TGAGATTAGTAGTATGGGAG 3'	HVII	
F82	5'ATAGCATTGCGAGACGCTGG3'	HVII	402 bp
F16144	5'TGACCACCTGTAGTACATAA 3'	HVI	107 bp
R16251	5'GGAGTTGCAGTTGATGT 3'	HVI	
F16190	5'CCCCATGCTTACAAGCAAGT 3'	HVI	220 bp
R16410	5'GAGGATGGTGGTCAAGGGGAC 3'	HVI	

F=forward, R=reverse, the numbers refer to the position of the 5' base of the primer in the complete human mtDNA sequence. Headings denote the name of the primer it was assigned when designed, the exact sequence, the region of mtDNA it targets, and the size of the amplicon resulting from the paired forward and reverse primers.

Initially, PCR was performed using F16190 and R16410 for each sample. If a sample did not produce the 220 bp amplicon, F16144 and R16251 were used, to produce a 107 bp amplicon; however, if amplification of the 220 bp amplicon was successful, amplification of progressively larger amplicons (329 bp followed by 402 bp) was attempted. When needed, an additional amplification reaction was performed to increase

the amount of product to be used in sequencing reactions. In these reactions, 1 µl of the original PCR product was used as a template and the number of cycles was reduced to 20.

If a reaction appeared to be inhibited (there was no amplification band and no primer dimer formed), the sample was purified using a Microcon-100 spin column (Millipore) with 300 µl of TE buffer at the recommended speed and time. The samples were washed two more times with TE buffer then the retenate was resuspended to the starting volume using TE.

DNA Sequencing of Bone Samples

When amplification was successful, the remaining 15 µl of PCR product was purified using a Microcon-30 spin column (Millipore) with 300 µl of TE and centrifuged per the manufacturer's instruction. The samples were washed two additional times with TE before retentate was brought back to 15 µl with TE.

Up to 5 µl of template DNA was sequenced using the BeckSeq kit (Beckman-Coulter). Per the manufacturer's recommendations, 50–100 femtograms were used for sequencing reactions; estimations of concentration were made from the yield gel of the original amplification. The primers used for sequencing were the same as those used for amplification, and as a modification to the protocol provided with the kit, the total reaction volume was reduced to 10 µl. The sequencing reaction was followed as outlined in the kit manual. Following sequencing, 4 µl of stop solution (1.5M NaAc, 50 mM EDTA, 2.5 mg/ml glycogen) and 30 µl of cold 95% EtOH was added to each reaction, which was then centrifuged at 14000rpm for 15 min. The supernatant was removed and the sample was washed twice with 200 µl of cold 70% EtOH with a 2 min, 14000rpm

spin between washes. After the final wash, the supernatant was removed and the sample was vacuum-dried for approximately 10 min, then resuspended in 40 µl Sample Loading Solution (SLS) (Beckman-Coulter). Samples were sequenced on a CEQ 8000 Genetic Sequence Analyzer (Beckman-Coulter), using the LFR-1-60 program (capillary temperature 50°C, denature 120 s at 90°C, inject 15 s at 2.0kV, and separate 60 min at 4.2kV). Resulting sequences were aligned using BioEdit Sequence Alignment Editor (Hall, 2004) and compared to the Anderson reference sequence. Any polymorphisms within a sequence were entered for comparison to the FBI mtDNA database to determine the frequency of the specific base change(s) (Monson *et al.*, 2002).

Statistical Methods

Statistical analysis was performed using Microsoft Excel 2002. Single factor Analysis of Variance (ANOVA) was chosen to examine the effect of weathering stage, sex, age, and bone type on the amplification results. ANOVA is used when the independent variables (in this case, age, sex, weathering condition, and bone type) have discrete categories and the dependent variable (amplification success) is continuous or binary (yes/no). This method of analysis involves comparing the difference among the means of two or more sample sets, while considering the variance around each mean. The analysis tests the hypothesis that the means of the groups are equal. For the analysis, the age of the individual was categorized into 5 groups: infant (less than 1 year old), child (1-13 years old), adolescent (14-17 years old), young adult (18-24 years old), adult (25-49 years old), and older adult (50+). Results were considered significant at $p<.05$.

Results

Sampling

Samples were chosen based on the number and bone type available for a particular burial number. As a result, there was a bias in sampling towards individuals with more than one bone available for analysis. Those individuals with at least three bone types, preferably femur, rib, and pelvis, were selected first, followed by those with only two of these bones. One individual (203) had three bone types available (fibula, cranium, and rib), two of which did not correspond directly to the set categories. The fibula was grouped into the femur category for statistical analysis due to its similar structure as a compact bone. The cranium remained in a category of its own. Five individuals from skeletal weathering stage five (30, 114, 349, 402, 686) were analyzed despite there being only one bone type for each individual. These samples were included to bring the sample size at that stage to a number closer to that of the other stages.

Observations during Bone Preparation

In general all of the bones were brown in color on the outer surface. Sections of bone originating from femurs appeared white beyond the cortical layer. In contrast, all pieces originating from pelves and ribs were brown throughout. During the cutting process, sections of femur came away in solid pieces, with the outer layer flaking off at times. The porous nature of the pelvic bone led to the production of several smaller pieces separating from the main sample when contact was made with the Dremel tool. For the most part, sections cut from ribs separated in one piece from the source sample, with the outermost layer flaking off at times.

During the washing procedure, small amounts of dirt were removed in the first wash step, with all of the rinse steps with water remaining clear. A difference in the amount of dirt removed was noticed only for the rib and pelvis samples from individual 704. With these samples, the wash solution turned opaquely brown and did not completely clear until the second or third water rinse. This individual had been rated as a stage five with the skeletal weathering stage criteria. Using the individual bone weathering stage criteria and rating each bone separately, the rib was a stage three, and the pelvis was a stage one. Despite the noted difference in the amount of dirt removed from these two samples, amplification of DNA was not affected, as both samples produced an amplicon.

During the milling process, the femurs were the hardest bone type to crush. In most cases, the sections were broken down into smaller pieces ranging from 1 mm to powder. Both ribs and pelves were more friable than the femurs and thus ground more easily. The size of rib and pelvis pieces after grinding ranged from powdered to 1–2 mm slivers of bone. The cortical layer of rib and pelvis sections tended to break off from the main part of the bone in sheets. If this outer layer was present at the grinding stage, it was not broken down into powder, but retained its fibrous structure and was broken into slivers of bone approximately 0.5 mm in diameter and 2 mm in length. The mass of the sample before and after grinding was recorded. On average there was a 25% loss of bone during the grinding procedure and transfer to tubes for extraction.

DNA Extraction

Many samples needed to be extracted two or three times with phenol before the phenol layer no longer appeared dark brown. In almost all instances, the samples which

required multiple phenol extractions came from either rib or pelvis. The color of the DNA pellet after extraction was noted for each sample, which varied in translucency and color; out of solution DNA should be colorless if it is clean. Translucent samples ranged from colorless to light yellow or light brown, while opaque samples ranged in color from white to light yellow or light brown. On occasion some samples would have a darker brown spot in an area of the pellet, presumably residual dirt. In general, DNAs from femurs were lighter in color than those from pelves or ribs, with only one femur sample having a dark brown spot in an otherwise clear pellet. The darker, brown pellets resulted exclusively from extractions of pelves and ribs. Pellet color did not seem to affect amplification in a predictable manner though, as rib extracts amplified significantly more often than pelvis extracts (details presented below).

DNA Quantification Using QuantiBlot

DNA yields from seven samples were analyzed using QuantiBlot. Optimization of the system with the mtDNA probes designed for this experiment was completed using 1 µl and 10 µl of human cell line DNA at a concentration of 10 ng/µl. Testing probe concentration, it was found that faint detection of the DNA was possible with the mtDNA probes DLoop11F, DLoop31R, and CO1 6281F, beginning at a 10 pmol/µl concentration and increasing with the 100 pmol/µl concentration using the manufacturer's suggested hybridization and first wash step temperature of 50°C. When the hybridization temperature was dropped to 45°C and the first wash step was dropped to 24°C, detection of DNA improved greatly, hence even lower hybridization temperatures were tested. The SDS in the hybridization solution began to precipitate at a 30°C hybridization

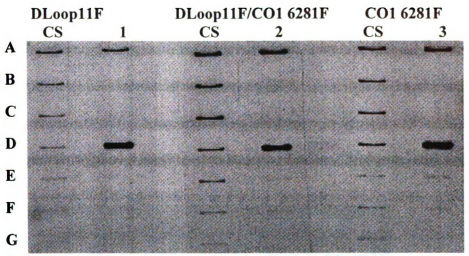
temperature, although hybridization did occur. Increasing the hybridization temperature to 40°C, prevented the SDS from precipitating and hybridization was improved compared to the 45°C experiment. The human positive control DNA hybridized to all three probes at all hybridization temperatures.

DNAs from raccoon, red fox, badger, and salmon were used to test the specificity of each of the mtDNA probes; DLoop11F and DLoop31R are designed to be human-specific, while CO1 6281F is designed to be mammal-specific. Preliminary results indicated that at 40°C, raccoon, red fox, and salmon DNA hybridized to the DLoop11F and DLoop31R probes, while all these species plus badger hybridized to the CO1 6281F probe. These DNAs were at a much higher concentration than the human control DNA however, and diluting them to 10 ng/μl decreased the amount of hybridization, while raising this temperature to 45°C prevented hybridization completely when DLoop11F and CO 6281F were used, but not DLoop31R. Therefore, DLoop11F and CO1 6281F were used in the quantification procedures for preliminary testing on seven bone samples; DLoop31R was not used due to its non-specific binding. In addition, the 40°C hybridization temperature was used to improve the sensitivity of the probes.

Six of the samples tested did not generate an interpretable signal under these conditions, as the strength of the signal was not above the lowest standard, and thus was considered to be below the lower limit for the sensitivity of the experiment (Figure 3). Only DNA extracted from the femur of burial number 47 produced a positive result: a dark band, stronger in intensity than the most concentrated standard. Given this rather surprising result, it was felt that this result may have been artifactual (see Discussion), and in combination with the negative results obtained for the 6 other samples, as well as

the large amount of each sample required for each analysis, quantification via QuantiBlot was not continued.

Figure 3. Quantification of mtDNA Using QuantiBlot



MtDNA quantification using the DLoop11F and CO1 6281F probes separately and together (DLoop11F/CO1 6281F). CS=Concentration Standard, A–G represents serial 1:2 dilutions of the human cell line DNA beginning with 10 ng. Lane A in columns 1–3 represents the positive control human cell line DNA (1 µl loaded). Lanes D–F in lanes 1–3 contain 30 µl of DNA from bone samples 47 femur, 345 femur, and 489 rib. Results from 47 femur represent usable data with a band greater in intensity than the most concentrated standard. Results from 345 femur and 489 rib were not above the lowest standard, and thus were considered to be below the lower limit for the sensitivity of the experiment.

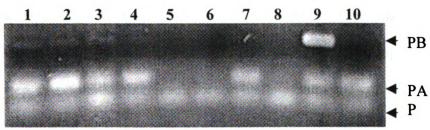
DNA Amplification

MtDNA fragments were successfully amplified from 55 of the 89 samples. Typical PCR results can be seen in Figure 4. A result was positive when a band for the targeted amplicon was observed (positive band, PB) and negative when no band was observed, but primer activity (PA) was present. Primer activity (including a process known as primer-dimer formation) stems from the two PCR primers interacting to form a small PCR product. The product should be seen even when when no bone DNA is added.

If primer activity does not occur and only the primers themselves are seen (P), something is inhibiting the PCR reaction.

On occasion the reagent blanks from the bone extraction procedures produced positive results after PCR. Each time this occurred, the mtDNA from the reagent blank was sequenced along with the bone samples that were processed following the last negative reagent blank. In no instances did mtDNA from a reagent blank match any of the sequences of the bone samples or laboratory personnel involved in sample processing.

Figure 4. Typical Amplification Results: Positive, Negative, and Inhibited



Lanes 1–8 represent amplification results from bone samples using the 220 bp primer set. Lanes 1–4 and 7 are positive (target band (PB) present). Lanes 5, 6, and 8 represent inhibited results (no primer activity (PA) observed). Lane 9 is the positive control. Lane 10 is the negative control.

Originally, 31 bones were amplified using standard *Taq* polymerase, of which only one produced a band using the 220 bp primer set. In an effort to improve amplification success, standard *Taq* was replaced with HotMaster *Taq*. An additional 6 samples of this original group of 31 produced a 220 bp amplicon. During these preliminary experiments it was noted that a 1:20 dilution of the extracted DNA improved PCR results; in some cases amplification occurred at a 1:20 dilution, but not with undiluted DNA. Further, if amplification did occur with the undiluted DNA, amplification also occurred at the 1:20 dilution, but with increased band intensity. Therefore only the 1:20 DNA dilution was used for the remaining experiments.

In addition to the change in *Taq*, a PCR enhancer was added that might help reduce PCR inhibition. As a result, six samples lost PCR inhibition, three of which successfully amplified. Samples that continued to show PCR inhibition were cleaned via a Microcon-100 spin column. Eight of the twelve remaining inhibited samples successfully amplified following this procedure. Results are shown in Table 4. In the end, four samples out of the total 89 remained inhibited after all steps to remove PCR inhibition were performed, three were pelves (328, 345, and 402) and one was a rib (539). Of interest, only one femur (409B) inhibited PCR before any improvement steps were taken, which was overcome with the addition of the enhancer. One sample, 489 rib, inhibited PCR after the addition of the enhancer, but the sample was lost in the Microcon step, therefore the final result for this sample is not known; however the available information for this sample was included in the overall PCR results.

Table 4. Improvement in Inhibition

Burial Number	Bone Type	Inhibited	Amplification	Skeletal Stage	Individual Bone Stage
047	Rib	No	Yes*	4	4
124	Rib	No	Yes*	3	4
126	Pelvis	No	No	3	3
126	Rib	No	No	3	3
132	Pelvis	No	Yes*	2	2
192	Pelvis	No	Yes	2	2
322	Pelvis	No	Yes	2	2
328	Pelvis	Yes	No*	2	2
328	Rib	No	Yes*	2	2
345	Pelvis	Yes	No*	3	3
402	Pelvis	Yes	No*	5	3
409B	Rib	No	Yes*	3	3
409B	Femur	No	Yes	3	2
489	Rib	Yes	No**	2	1
539	Rib	Yes	No*	4	4
545	Pelvis	No	No	1	3
545	Rib	No	Yes*	1	4
704	Rib	No	Yes*	5	3
704	Pelvis	No	Yes*	5	1

List of the 18 samples which initially showed PCR inhibition arranged by burial number and bone type. Inhibited column indicates whether PCR reactions with these samples remained inhibited following dilution of the DNA extract. * indicates samples that were processed with a Microcon-100 spin column. ** indicates the sample for which final results are not known as the sample was lost during the Microcon step.

The final PCR results, arranged by burial number and bone type, are shown in Table 5. If amplification was successful, it was noted by the maximum fragment length obtainable. Samples that were subject to Microcon-100 processing are denoted by a star. Thirty-four samples produced the 220 bp amplicon, all of which were additionally analyzed using the 329 bp set of primers. Five samples (203 cranium, 164 femur and rib, and 345 rib 1 and 2) produced the 329 bp amplicon. PCR of those five samples with the 402 bp primers was carried out, but negative results were obtained from all samples. Of the 50 samples that did not produce a 220 bp band, but did not inhibit PCR, 21 produced a 107 bp amplicon.

Table 5. Overall PCR Results
Headings denote burial number, bone type, and the maximum obtainable amplicon size. Those samples in which an amplicon was not produced and there was no primer activity are labeled as "inhibited". Samples which did not produce an amplicon but showed primer activity were considered negative results and all boxes were left blank.
* indicates samples that were processed with a Microcon-100 spin column.

Table 5.

Burial Number	General bone	107 bp	220 bp	329 bp	Inhibited
027	Femur		X		
027	Rib		X		
030	Femur		X		
034	Femur	X*			
034	Pelvis				
034	Rib		X		
047	Femur				
047	Pelvis		X		
047	Rib		X*		
047	Rib				
111	Femur	X			
111	Rib	X			
111	Pelvis	X			
114	Femur		X		
124	Femur	X			
124	Pelvis	X			
124	Rib		X*		
126	Femur				
126	Pelvis				
126	Rib				
132	Femur	X			
132	Pelvis		X*		
132	Rib	X			
164	Femur			X	
164	Pelvis				
164	Rib			X	
167	Femur		X		
167	Pelvis				
167	Rib				
192	Femur		X		
192	Pelvis		X		
192	Rib				
203	Cranium			X	
203	Fibula		X		
203	Rib				
256	Femur				
256	Pelvis				
256	Rib				
260	Femur	X			
260	Rib	X			
322	Femur	X			
322	Pelvis		X		
322	Rib	X			
328	Femur	X			
328	Pelvis				X*
328	Rib		X*		

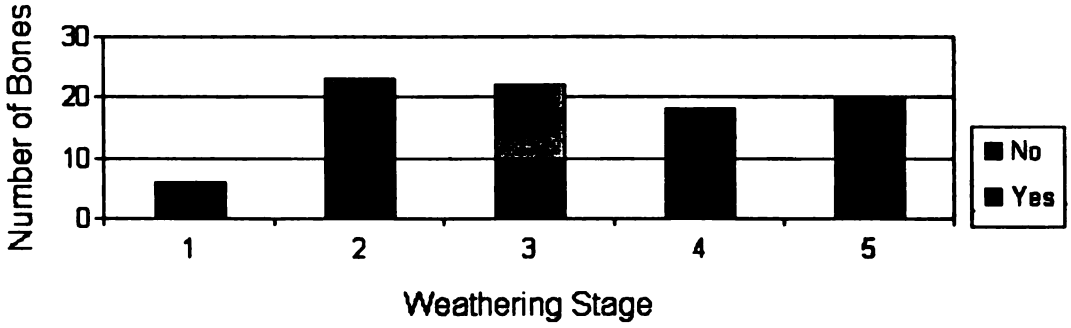
Table 5. Continued.

331	Femur	X			
331	Rib	X			
345	Femur				
345	Pelvis				X*
345	Rib			X	
345	Rib			X	
348	Femur				
348	Pelvis				
348	Rib	X			
349	Femur		X		
355	Femur				
355	Rib				
381	Femur	X			
381	Pelvis				
381	Rib	X			
389	Pelvis				
389	Rib				
402	Pelvis				X*
409B	Femur	X			
409B	Pelvis				
409B	Rib		X*		
447	Femur		X		
447	Rib		X		
448	Femur		X		
448	Rib		X		
489	Pelvis				
489	Rib				X
529A	Femur		X		
529A	Pelvis				
529A	Rib		X		
539	Pelvis				
539	Rib				X*
540	Pelvis		X		
540	Rib				
545	Femur	X			
545	Pelvis				
545	Rib		X*		
546	Pelvis		X		
546	Rib				
686	Cranium		X		
704	Femur		X		
704	Pelvis	X*			
704	Rib		X*		
Total		21	29	5	5

Skeletal Weathering Stage and Amplification Success

DNA amplification of a sample was considered positive regardless of the amplicon size produced. Amplification results broken down by skeletal weathering stage, which was based on the condition of the whole skeleton, are displayed in Figure 5 and show that there is no obvious relationship between weathering stage and amplification success. Bones from stage five, the most weathered, amplified the most often at 70% (14 out of 20), followed by stage two at 69.57% (16 out of 23), stage one at 66.67% (4 out of 6), and stage four at 61.11% (11 out of 18). Samples from stage three amplified the least often, at 45.45% (10 out of 22). The difference in amplification among all weathering stage categories is not significant (p=0.460).

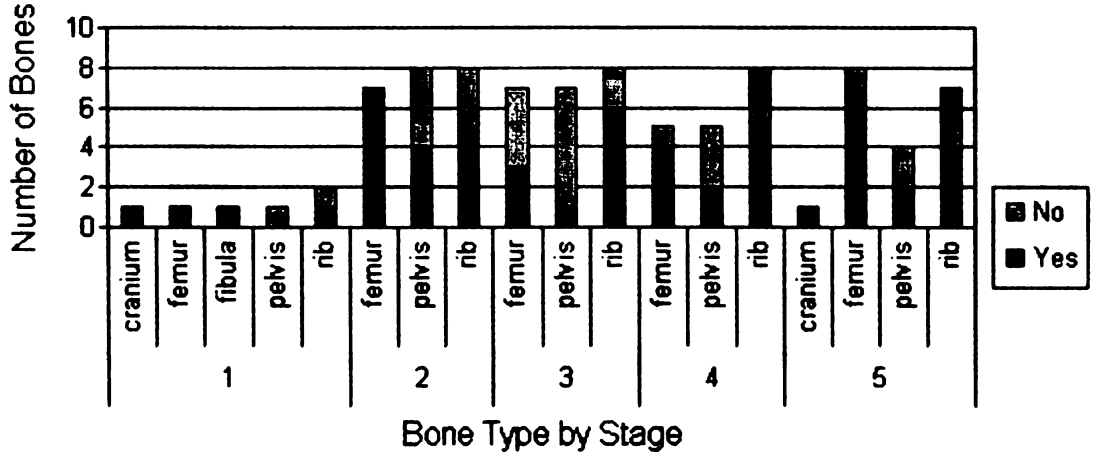
Figure 5. Amplification Results in Comparison to Skeletal Weathering Stage



Graph depicting amplification success (yes) and failure (no) arranged by the five skeletal weathering stages.

Figure 6 shows the same data divided into bone type. Amplification success is similar across the five weathering stages, with and pelvis and rib samples amplifying the least often at stage one (0 out of 1 and 1 out of 2, respectively), while femur samples amplifying the least often at stage three (42.86%, 3 out of 7). There is no significant difference in the pelvis, rib, or femur data among the five stages (p =0.604, p=0.956, and p=0.076, respectively).

Figure 6. Amplification Results in Comparison to Skeletal Weathering Stage and Bone Type

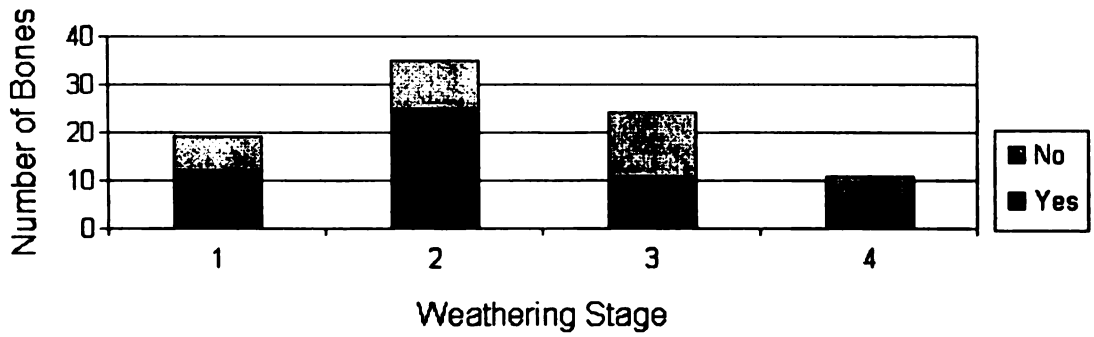


Graph depicting amplification success (yes) and failure (no) arranged by bone type within each of the five skeletal weathering stages.

Individual Bone Weathering Stage and Amplification Success

Figure 7 shows amplification results based on the individual bone weathering criteria in which each bone was staged individually. Amplification success across the stages (without considering bone type) was similar, with the highest at stage two (71.43%, 25 out of 35) followed by stage four (63.63%, 7 out of 11), stage one (63.16%, 12 out of 19), and stage three (45.83%, 11 out of 24). Statistical analysis comparing all four weathering stages shows the variance in amplification success across the stages is not significantly different ($p=0.269$).

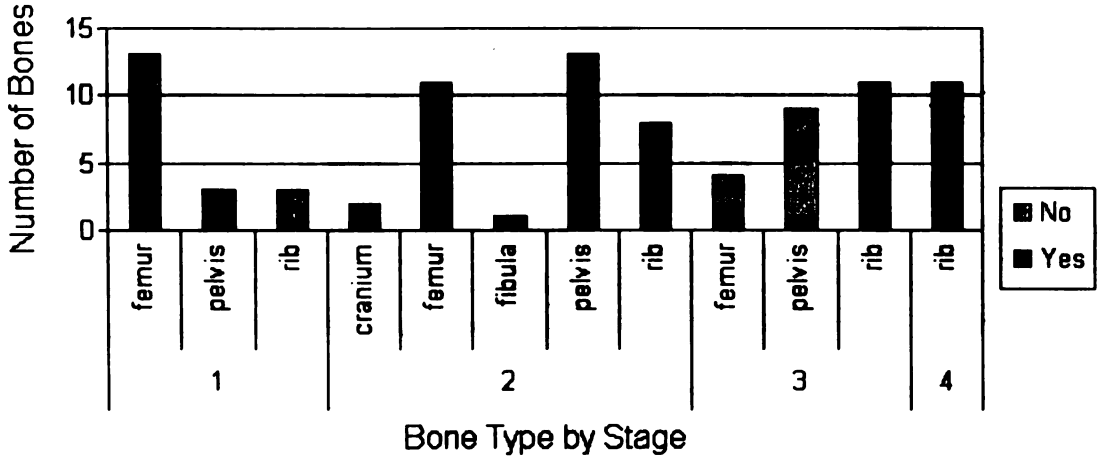
Figure 7. Amplification Results in Comparison to Individual Bone Weathering Stage



Graph depicting amplification success (yes) and failure (no) in relation to the four individual bone weathering stages.

Amplification success of the different bone types varies across the four weathering stages (Figure 8), however the differences are not significant. For example, stage two femurs have the highest amplification success at 90.91% (10 out of 11), followed by those at stage one at 76.92% (10 out of 13), and stage three at 50% (2 out of 4) (p=0.213). The pelvis samples show a decrease in amplification success from stage one to stage three (as with femurs, none of the pelvis pieces analyzed in this study were categorized as stage four) (66.67%, 2 out of 3; 46.15%, 6 out of 13; and 11.11%, 1 out of 9 for stages one, two, and three, respectively) (p=0.130). Rib samples have the highest amplification success at stage two (75%, 6 out of 8), followed by stage three (72.73%, 8 out of 11), stage four (63.64%, 7 out of 11), and stage one (0%, 0 out of 3) (p=0.111).

Figure 8. Amplification Results in Comparison to Individual Bone Weathering Stage and Bone Type

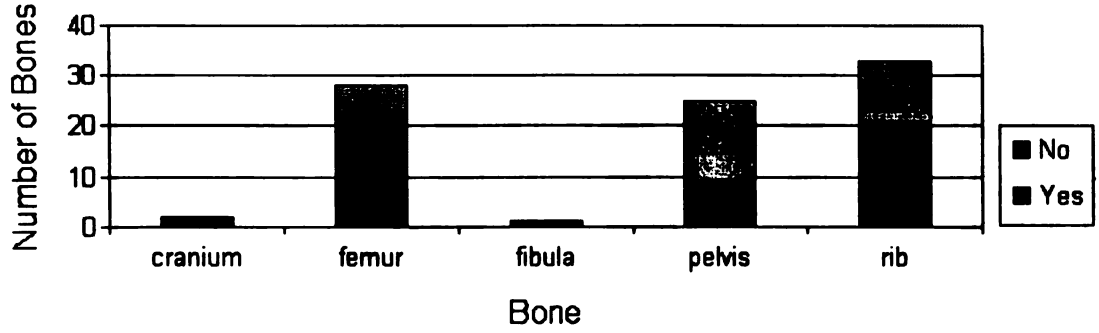


Graph depicting amplification success (yes) and failure (no) arranged by bone type within each of the four individual bone weathering stages.

Bone Type and Amplification Success

The strongest influence on amplification success appears to originate from bone type. As shown in Figure 9, amplification occurred in the 2 crania samples (100%), 79.31% (23 out of 29) of the femur/fibula samples, 36% (9 out of 25) of the pelvis samples, and 63.64% (21 out of 33) of the rib samples. Across bone types there is strong statistical evidence ($p=0.006$) suggesting a difference in amplification success, with the greatest difference between femurs and pelves ($p=0.0009$), followed by ribs and pelves ($p=0.037$). There is no significant difference between ribs and femurs ($p=0.181$). Crania were not included in the comparisons due to the low sample number.

Figure 9. Amplification Results in Comparison to Bone Type

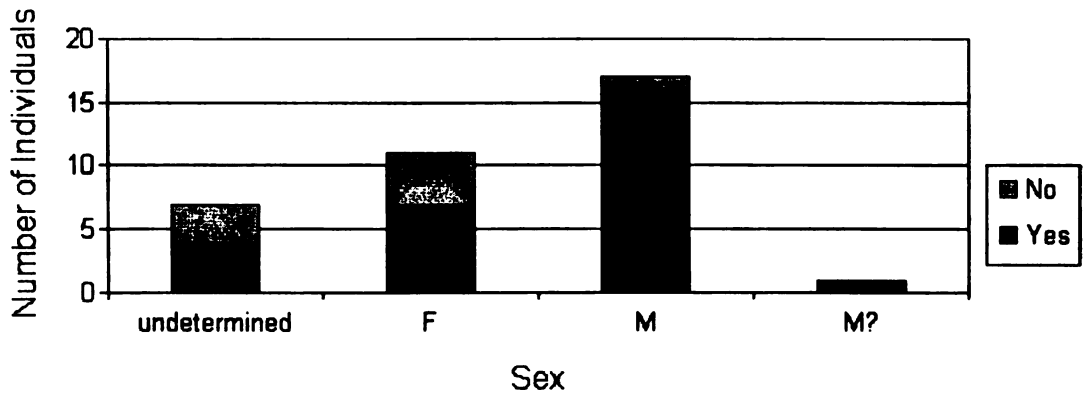


Graph depicting amplification success (yes) and failure (no) arranged by bone type.

Sex Estimates and Amplification Success

Amplification results for each individual are shown in Figure 10 by the four sex categories (female, male, "possible male", and undetermined). Amplification occurred in 57.14% (4 out of 7) of the undetermined individuals, 63.64% (7 out of 11) of the females, and 94.12% (16 out of 17) of the males. The one individual classified as "possibly male" (M?) also amplified. There is no significant difference in amplification success among the four categories (p=0.120). However, removing the undetermined individuals from the analysis reveals a significant difference in amplification success between males and females (p=0.041), with male samples amplifying more often. Adding the "possible male" to the male category drops the p-value to 0.034.

Figure 10. Amplification Results in Comparison to Sex of the Individual



Graph depicting amplification success (yes) and failure (no) in relation to the anthropologically estimated sex of the individual. Four categories were used to describe the individuals sampled in this study: undetermined (sex of the individual could not be determined), female (F), male (M), and possible male (M?).

Amplification success for each bone was compared with respect to sex and skeletal weathering stage. Six bones were classified as stage one, three originating from undetermined individuals and three from females, with each category having an amplification success of 66.67% (2 out of 3); no male skeletons were classified as stage one, eliminating a comparison between sexes. At stage two, 25% (1 out of 4) of bones from the undetermined category amplified, 83.33% (15 out of 18) of bones from male skeletons amplified, and 33.33% (1 out of 3) bones from the "possible male" amplified; no female skeletons were designated as stage two, nullifying a statistical comparison between sexes. Bones from male skeletons at stage three amplified 52.63% (10 out of 19) of the time, whereas bones from female skeletons at stage three amplified 0% (out of 3) ($p=0.097$). All the bones from undetermined individuals at stage four amplified, 53.85% (7 of the 13) of female bones amplified, and 66.67% 2 out of 3 male bones amplified ($p=0.710$). Finally, at stage five, amplification occurred in 40% (2 out of 5) of

bones from undetermined individuals, 66.67% (4 out of 6) of female bones, and 100% (out of 8) of male bones, providing no evidence for a significant difference in amplification success between the two sexes (p=0.088). Thus, no significant difference between the sexes existed at stage three and stage five, while comparisons could not be made at stage one and two.

Similar comparisons were made for the individual bone weathering categories. However, unlike the skeletal weathering categories, these included bones from male and female individuals in all stages. Thirty-three percent (1 out of 3) of the female bones from stage one amplified and 78.57% (11 out of 14, one of which was from the "possible male") of male bones amplified (p=0.134). Stage two bones from females amplified 72.72% of the time (8 out of 11) and from males 66.67% of the time (12 out of 18, two of which were from the "possible male") (p=0.743). Female bones at stage three amplified 16.67% of the time (1 out of 6), while male bones amplified 64.29% of the time (9 out of 14) (p=0.054). At the most weathered stage, 60% (3 out of 5) of the female and 80% (4 out of 5) of the male bones amplified (p=0.545). Thus, at all four stages of the bone staging system, there is no significant difference in amplification success between male and female bones.

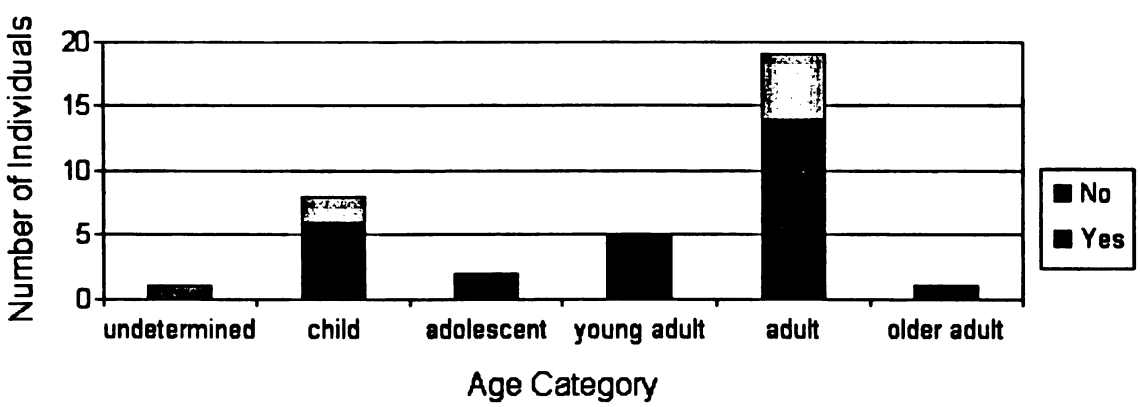
Amplification results in comparison to bone type and sex of the individual were also analyzed. In all, 7 femurs, 8 pelves, and 10 ribs were sampled from females; 18 femurs, 15 pelves, and 18 ribs were sampled from males. The male numbers include three bones, one of each type, that were collected from the "possible male". Neither of the two crania nor the single fibula was given a sex classification. There is no significant

difference in amplification success between the sexes when comparing femurs, pelves, or ribs (p=0.372, p=0.912, and p=0.065, respectively).

Age and Amplification Success

The effect of estimated age of the individual on amplification success is displayed in Figure 11. There is no significant difference in the amplification success across the five age categories (p=0.748). The amplification results of each age category are as follows: 6 out of 8 (75%) children, 2 out of 2 (100%) adolescents, 5 out of 5 (100%) young adults, 14 out of 19 (73.68%) adults, and 1 out of 1 (100%) older adult amplified. The age of one individual (from burial number 389) was not estimated at the time of recovery and therefore is not included in the statistical analysis.

Figure 11. Amplification Results in Comparison to Estimated Age



Graph depicts amplification success (yes) and failure (no) of each bone sampled in relation to the five age categories. The age of one individual was not estimated.

Amplification success in relation to age and bone type was also analyzed. Successful amplification from femurs occurred 66.67% (4 out of 6) of the time in children, 100% in adolescents, young adults, and older adults (out of 2, 4, and 1 sample, respectively), and 75% (12 out of 16) of the time in adults. With ribs, 33.33% (2 out 6)

of samples from children amplified, as did 50% (1 out of 2) from adolescents, 80% (4 out of 5) from young adults, 77.78% (14 out of 18) from adults, and 0% (0 out of 1) from older adults. Pelvis samples amplified 50% of the time from children (1 out of 2) and young adults (2 out of 4), 0% (out of 1) from adolescents, 31.25% (5 out of 16) from adults, and 100% (out of 1) from older adults. A significant difference is not observed among the age categories regardless of bone type ($p=0.674$, 0.172, and 0.626 for femur, rib, and pelvis, respectively).

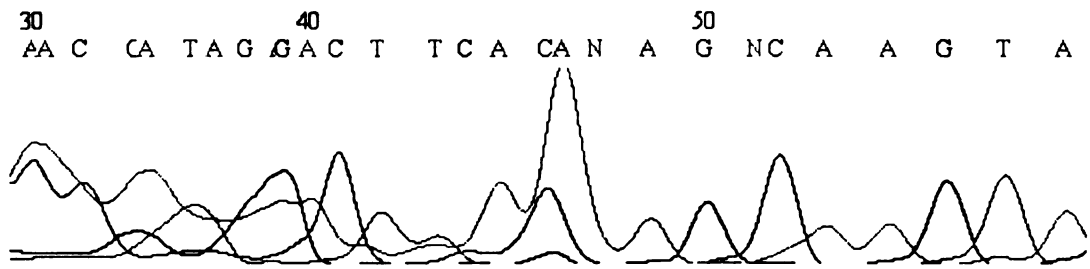
The effect of age on amplification success shows no significant difference among the age categories at any of the five stages of the skeletal weathering system ($p=1$, 0.057, 0.451, 0.948, and 0.624, respectively). Similarly, when the same analysis is performed based on the bone weathering stages, no significant difference among age classes is found at weathering stages two through four ($p=0.150$, 0.465, and 0.695, respectively); however at weathering stage one there is a significant difference in amplification success among the age categories ($p=0.026$). With this stage, bones from adolescent to adult amplified the best (100%, 100%, and 70%, respectively), while bones from the child and older adult categories did not amplify at all.

DNA Sequencing

To confirm that PCR product generated in these experiments actually originated from the bone being tested, positive samples were processed for sequencing. Sequences were produced for any individual that produced PCR products from more than one bone so that sequences comparisons within an individual could be made. Sequencing was performed on 31 PCR products, ten 107 bp amplicons and twenty-one 220 bp amplicons,

using the same primers as amplification. The 5' end of both the forward and reverse reactions produced poor quality data which, in some cases, meant that overlapping sequences between the two reactions was not possible (Figure 12). A full sequence of the targeted region was obtained from 15 (48.39%) samples, none of which originated from a 107 bp fragment.

Figure 12. Electropherogram Depicting Typical 5' End Data



Portion of electropherogram showing poor quality data at the 5' end. Peaks correspond to each of the four bases: adenine (A), guanine (G), thymine (T), and cytosine (C). N represents a situation in which one of the four bases could not be called. Bases are number by tens along the top. Analyzable data begin at base 50.

Forward and reverse sequences were aligned with the Anderson reference sequence and compared to determine if the DNA extracted from the different bones originated from the same source. Fifteen of the 21 (71.43%) 220 bp samples produced analyzable sequence from both the forward and reverse primers. Four (19.05%) 220 bp samples (27 femur, 192 pelvis, 448 femur, and 448 rib), produced analyzable sequence from only one DNA strand. In each case, the readable strand originated from the forward primer, with the exception of 448 rib. Two (9.52%) 220 bp samples (345 rib and 409B rib) did not produce an analyzable sequence from either primer. Eight of the ten (80%) 107 bp samples did not produce analyzable sequences. Two (20%) 107 bp samples (111 femur and 260 femur) gave readable sequence with the forward primer only.

Comparable sequences, those in which there was at least a partial sequence from more than one bone of an individual, were obtained from eight individuals (Table 6). Sequences within an individual were consistent in all of these. Three individuals (27, 164, and 203) differed from the Anderson reference sequence by one base across the analyzable section, one individual (447) differed by two bases, and four samples (192, 448, 529A, and 704) did not differ from the reference sequence. All sequence polymorphisms were unique to the mtDNA database.

Table 6. Sequence Polymorphisms In Individuals Generating PCR products form More Than One Bone

Burial Number and Bone Type	Sequence Interval	Sequence Polymorphism
27 femur F rib F/R	16121-16232 16272-16376	 16340C
164 femur F/R rib F/R	16173-16216 16247-16385	 16267+T
192 femur F/R pelvis F	16167-16390	None
203 fibula F/R cranium F/R	16167-16388	16270+T
447 femur F/R rib F/R	16169-16260 16299-16378	 16248T, 16340C
448 femur F rib R	16177-16382	None
529A femur F/R rib F/R	16169-16387	None
704 femur F/R rib F/R	16186-16378	None

The sequence interval available for comparison between at least two bones of an individual, with polymorphisms from the human mtDNA reference sequence noted. F=forward, R=reverse sequence.

Discussion

The goal of this study was to determine if there is a relationship between the outward appearance of skeletal material (considering bone type and level of weathering) and the quality of the DNA within. Being able to predict the amount of obtainable DNA and its level of degradation (i.e., an appropriate target sized amplicon) would assist the forensic scientist in providing a more efficient and productive analysis of the material.

DNA Pellet Color

One step of DNA preparation after extraction is to bring the DNA out of solution and centrifuge it down to a pellet. Out of solution, DNA is colorless or white if it is clean. Discolored pellets may be dirty and inhibit PCR. The DNA pellets of the Voegtly samples were sometimes discolored; however, there was no apparent trend between the degree of discoloration and amplification success. Although DNA pellets from femur extractions were often lighter and DNAs from femurs amplified more often, pellets from pelvis and rib extracts showed equal amounts of brown discoloration, yet DNAs from ribs amplified significantly more often than that from pelves. Therefore, pellet color does not seem to be a good indicator of predicting amplification success.

DNA Quantification

In a typical forensic lab setting, nuclear DNA yield for a particular sample is estimated before amplification so that extracts can be diluted for optimal results; the ABI QuantiBlot system is a method commonly employed for the quantification of nuclear DNA. However, no such method exists for mtDNA. Modifications were made to the

QuantiBlot procedure to allow for the quantification of mtDNA, including the design of three mtDNA probes and reduction of the hybridization and wash temperatures. Although the modified procedure did not prove useful for DNA quantification of the Voegtly samples, the procedure was optimized and shown to specifically detect human mtDNA using the DLoop11F probe and mammalian mtDNA with the CO1 6281 probe. Oddly, the DLoop31R probe did not bind specifically to human mtDNA, but also bound to raccoon, salmon, red fox, and badger DNA, even at elevated hybridization temperatures. This is surprising as DLoop31R is the exact complement of DLoop11F; the reason why one would bind specifically but not the other is unknown. The single interpretable signal from a bone extract may have been artifactual, as DNA from that extract (47 femur) was not amplifiable. If the result was not artifactual, and was indeed of human origin, then the DNA from that extract should amplify given the intensity of the band observed (Figure 3). However, amplification was not successful, therefore the source of this of the DNA could not be confirmed and could be the result of contamination.

In addition, the 10 ng/μl concentration of the human cell line DNA given by the manufacturer describes the total DNA concentration, both mitochondrial and nuclear. Without knowing the exact mtDNA concentration of standards used, only comparative results can be made (e.g. sample 1 is less concentrated than sample 2, but more than sample 3). It is quite possible that the detection limits of this method of quantification are too high for the amount of mtDNA contained in the Voegtly samples, indicating that a more sensitive test would be required. A separate study is being conducted to determine

if an alternate method for quantification, real-time PCR, which amplifies the DNA and thus has a lower detection limit, could prove useful for analyzing these samples.

DNA Amplification

Since quantification via QuantiBlot was unsuccessful, amplification procedures were performed without knowing the DNA yield of the extracts, therefore, a more empirical approach was needed to optimize the amount of DNA extract added to each reaction. In early experiments, side-by-side PCR reactions were run, one using the DNA extract undiluted and the other at a 1:20 dilution. During these experiments, it was noticed that when amplification was successful with the undiluted DNA, amplification also occurred with the 1:20 dilution, while the band intensity generally increased with the diluted sample. On occasion PCR was successful only with the diluted DNA and not the undiluted, and in some cases, an undiluted sample was inhibited, but inhibition was overcome by dilution. Therefore, as experiments progressed, only a 1:20 DNA dilution was used. The improvement in amplification at this concentration likely resulted from the dilution of components within the DNA extract that either inhibited PCR or prevented amplification of the target region. For example, with degraded samples, the many small pieces of DNA could interfere with PCR because they contain the target sequence for the primers and are preferentially amplified. Small pieces of DNA contained in the extract may also act as random primers and result in non-specific amplification. In either case, dilution of the extracts would lower the concentration of these components to a level at which they do not interfere with the designed amplification.

In addition to the 1:20 DNA dilution, a PCR enhancer was incorporated into the reaction in an effort to reduce PCR inhibition that remained after the dilution of the DNA extract. The ingredients of the enhancer are proprietary, so one can only speculate on what exactly is occurring with its addition to the reaction. One possible reason for inhibition is the presence of compounds that bind to proteins. As a protein, *Taq* polymerase would be a target for such inhibitors. Another protein, bovine serum albumin (BSA) is often added to reduce PCR inhibition. BSA acts as a competitor for protein-binding inhibitors; thus if enough BSA is added, *Taq* polymerase will be free to perform its polymerase activity and PCR will no longer be inhibited. BSA and/or other similar components may be present in the PCR enhancer. Although addition of the enhancer did not improve all inhibited reactions, it improved some, and for that reason it was added as a routine component to the PCR reaction.

The final addition made to the PCR reaction to optimize amplification results was hot-start *Taq* polymerase. Standard *Taq* polymerase has an optimal activation temperature of 72°C; however, some activity occurs at lower temperatures. Thus, if primers, including the aforementioned small DNAs acting as primers, bind non-specifically to the template DNA at lower temperatures, this early activity may lead to spurious amplification of the template. These undesirable products have the potential to monopolize the activity of the polymerase, leading to reduced amplification of the targeted region. Initiating PCR at a higher temperature prevents such non-specific primer binding and amplification. The hot-start *Taq* polymerase used in this study is coupled with a temperature-dependent inhibitor which is released at 94°C, making the polymerase inactive at temperatures below 60°C, thus reducing non-specific amplification due to low

temperature activity. When standard *Taq* polymerase was used in initial experiments, only 1 of 31 samples produced a positive PCR product. Using hot-start *Taq* polymerase improved the amplification results of those 31 samples six fold. Thus, the use of hot-start *Taq*, together with a 1:20 DNA dilution, and a PCR enhancer greatly improved amplification success.

The use of the Microcon-100 spin column was incorporated as part of DNA preparation when a sample inhibited the PCR reaction. The pore size of the filters within the Microcon-100 columns is small enough that double-stranded DNA molecules 125 bp (300 bp single-stranded) and larger are retained while smaller molecules flow through and are removed. During this process, smaller pieces of DNA, which, as mentioned above, may compete with the targeted DNA for primer binding or act as random primers during PCR, are removed. In addition, other, smaller water soluble inhibitors that are not removed during the organic preparation would be eliminated. Use of the spin column removed the inhibitory factors in 2/3 of the remaining inhibited reactions (Table 4). With this success rate, Microcon-100 spin columns do provide an easy and quick way of removing PCR inhibition, and thus should be used as a routine step when a sample inhibits PCR.

DNA Quality

The quality of mtDNA in each sample was assessed by determining the largest amplifiable fragment of DNA within the sample. Twenty-nine of the 89 samples produced a 220 bp fragment, 21 produced a 107 bp amplicon, 5 samples produced the 329 bp fragment, while no samples amplified at the 402 bp size. The 29 samples which

did not produce a 220 bp amplicon, but were not inhibited, were not tested for amplification of fragments smaller than 107 bp. The degraded state of the DNA likely reflects the age and condition of the samples and is in line with the typical fragment length isolated from ancient DNA (<300–500 bp) (O'Rourke *et al.*, 2000). It is interesting to note that three of the five bones that produced a 329 bp fragment were ribs, while general amplification (regardless of size) was most successful with femurs (discussed below). This could indicate that the DNA within ribs and femurs degrades on different timelines due to the difference in bone structure. Flat bones such as ribs are composed largely of spongy bone, with more accessible DNA. Also, there may be more cells by weight in the spongy portion of a flat bone than in the compact portion of a long bone. Both of these characteristics would lead to the extraction of higher quality DNA from flat bones, although this relationship might only exist while the bones are well-preserved. The larger distribution of spongy bone within flat bones also makes them more susceptible to degradation and as time passes and weathering occurs, the structures that protect the DNA will be destroyed. Bones that are composed largely of compact bone, like femurs, degrade slower, and while there may be fewer cells within a section, or DNA is less accessible, the microstructure stays intact longer, allowing for amplification of DNA from long bones after it is no longer successful from flat bones. Similar results have been found regarding quantity of DNA. Parsons and Weedn (1997) found that fresh, spongy bones, such as ribs, have the potential to yield 10- to 20-fold more DNA than compact bone; however such yields are less reliable with increased time since death.

DNA Sequencing

When amplification product was available for at least two bones from any one individual, amplicons were sequenced to ensure that the DNA extracted from the different bones originated from the same source. The limited amplicon size obtained from these aged samples prevented the successful analysis of all 31 samples that were processed for sequencing. Smaller amplicons proved to be the most problematic. This was due to the poor quality of data at the 5' end of every reaction, which can be attributed to the large amount of unincorporated ddNTPs that pass through the capillary at the beginning of the sequencing procedure. Deletion of this poor quality data before sequence analysis rarely left enough quality data to be successfully analyzed from the 107 bp reactions; only two samples produced readable sequence, and only in the forward direction. Therefore, sequencing of small amplicons using the cleanup methods recommended by Beckman-Coulter is not advised. Better methods are needed for the cleanup step to remove unincorporated ddNTPs. Sequence analysis of the 220 bp fragments did provide sufficient data to show that the PCR results were from the bones and not contamination (Table 6). In the eight individuals analyzed, the sequences being compared were consistent within the individual, indicating the DNA extracted from different bones originated from the same skeleton. Four individuals showed these consistencies with no nucleotide differences (polymorphisms) from the Anderson reference sequence. The remaining four had at least one polymorphism, (one individual had two) within the sequenced interval. In no instance did the DNA isolated from bone samples match the lab personnel involved in the extraction and analysis procedures and all sequence polymorphisms were unique to the FBI mtDNA database.

Weathering Condition and Amplifiable DNA

One of the most important goals of this study was to determine if the visual appearance of a skeleton or bone could be used to estimate the likelihood of obtaining a PCR product. Originally, the anthropologists at The Smithsonian Institution systematically graded the condition of each skeleton based on the guidelines of Behrensmeyer (1978). To address the question of whether this overall classification was predictive of DNA quality, bones were chosen with the intent of equally sampling five of the six weathering stages (small sample size prevented the analysis of stage 0 skeletons). It is interesting that as bones were being processed for DNA analysis, there did not seem to be a correlation between how a skeleton was graded and the condition of a specific bone sample. On the other hand, it was noticeable that the most distinguishing differences occurred among bone types. Femurs tended to be in better condition than ribs and pelves; they were rarely discolored beyond the outermost layer, and in no instance was there a femur in several small pieces. In contrast, ribs and pelves were brown throughout. In addition, the outermost layer of most pelves was missing in portions of the samples or not present at all, leaving the spongy interior exposed and susceptible to breakage. Similar flaking occurred in ribs, but to a lesser degree; however, ribs displayed the greatest amount of breakage and in some cases, only tiny pieces of bone were present. Some of this breakage may have resulted from the fragile nature of ribs, which would affect how they break down under natural conditions or how well they stand up to handling and transport. Comparing individual bone types among skeletal weathering stages, it was not unusual to have a femur that came from a more degraded skeleton (e.g.

stage four) that appeared to be in the same or better condition as one which came from a less degraded skeleton (e.g. stage two). In comparison, the state of ribs and pelves seemed to coincide more often with the skeletal ratings, although on occasion a rib or pelvis was found that was in better or worse condition than the others of the same stage. The difference between femurs and ribs or pelves in discoloration and condition could be attributed to the structure of each bone type. Femurs and other long bones are composed of compact bone from the cortex to the medullary cavity along the entire diaphysis (shaft) of the bone and spongy bone beneath the cortex at the articular ends. In contrast, ribs and pelves have spongy bone beneath the cortex along all portions of the bone, with compact bone present only in the cortex. The compact nature of the femur diaphysis prevents discoloring compounds from the soil to penetrate much beyond the cortical layer, while the spongy nature of the interior of ribs and pelves would not prevent such action once the cortical layer is penetrated or removed. In addition, denser, compact areas of bones are more resistant to degradative processes than less dense, spongy areas, leaving femurs in better condition than the rest of the skeleton.

ANOVA was used to examine the difference in amplification success among the different categories of skeletal weathering, individual bone weathering, bone type, sex, and age. No difference in amplification success was detected among the five skeletal weathering stages. Interestingly, although stages one, two, four, and five were similar in success (approximately 70%), bones from stage five skeletons amplified the most often, while bones from stage three skeletons amplified the least often (45.45%) (Figure 5). One way to account for this is chance; if no relationship exists between bone appearance and DNA degradation, a random distribution of the data would allow for one category to

be the highest and another the lowest. A second possibility is that the DNA from more degraded skeletons, such as those at stage five, is more accessible for extraction owing to the fact that these bones are already broken down. Alternatively, all bones within a skeleton do not degrade in the same manner with better preserved bone existing in an otherwise highly degraded skeleton. The better preserved bones may contain more amplifiable DNA, and the presence of these bones in the stage five skeletons increases the amplification success of this category above that which would be found if only poor bones were selected. If stage five skeletons provide the best source of DNA due to the structure of a specific bone type rather that the condition of the skeleton as a whole, this should be considered by the DNA analyst. Whatever the reason for the distribution of amplification success across the stages of skeletal weathering, it is clear that the condition of a skeleton is not indicative of the quality of DNA within the bones.

To incorporate the potential influence of specific bone types on amplification success, each skeletal weathering stage was further divided by bone type. No significant difference was detected for femurs, ribs, or pelves (Figure 6), further indicating that staging based on skeletal weathering does not aid in the prediction of the quality of DNA that can be extracted from the bone. Though skeletal rating is informative to the anthropologist, it may not be a useful classification system for the DNA analyst.

A new four stage weathering classification system was created during the course of this study and applied to each bone. Reclassification of the individual bones confirmed that the degradation of a skeleton was not consistent throughout the skeleton (Table 4). For example, in burial number 34, the skeleton was determined to be at stage four under the original weathering system; however, when each bone was classified on its

own, the pelvis was classified as stage two, the rib as stage four, and the femur as stage one. In general, all three bones within a skeleton were reclassified into different stages. On occasion, a pelvis and rib from a single skeleton were reclassified as the same stage; only rarely were all three bones reclassified as the same stage. If the level of degradation is predictive of the quality of DNA within the bone, a bone rating system would reveal a correlation that was masked when the whole skeleton rating system was used. However, when the amplification success data were rearranged according to the bone staging system, a significant difference among the stages was still not detectable (Figure 7). It should be noted that as a consequence of reclassification, the sample sizes across the four weathering stages was not equal. While there were close to 20 bones available for analysis from each of the skeletal rating stages and all three bone types were represented at each of the stages, stage one and stage three of the bone rating system had sample sizes close to 20, while stage two had over 30 and stage four had only 10. The decrease in sample size at stage four is due to the fact that no femurs or pelves were placed into this category. Because of this unequal sampling, differences among the stages may be hard to uncover, particularly among the most weathered bones. First, with a smaller sample size, the bones at stage four may not be representative of all stage four bones, and indeed, pelves and femurs are not represented. In addition, comparing the data from stage four to those from stages with two or three times the number of samples would affect the significance of any differences among the stages; the more equal the sample sizes are among stages, the more significant a difference will be if it exists. Regardless, the data available from this study do not appear to follow a trend and indicate that individual bone weathering condition is not predictive of DNA quality.

The issue of unequal sample sizes among the stages of the bone staging system continued when samples were further categorized by bone type (Figure 8); supporting the idea that skeletal staging does not match bone staging. The difference in femur amplification success among the three stages was not significant, although the small sample size of stage three femurs, compared to the other two stages, may influence this finding. Data from stage one and stage two femurs, which have larger sample sizes, appeared to be very close in amplification success, suggesting that weathering condition of femurs is not indicative of DNA quality. A larger sample size of stage three femurs in addition to a representative sample of stage four femurs for comparison would solidify this result.

Pelves were the only bone type for which amplification success appeared to follow a trend across the bone weathering stages with a decrease from stage one to stage three (no pelves being classified as stage four); however the difference among the three stages was not significant. Again, these data may be skewed by the uneven distribution of the pelves across the weathering stages. In this instance, stage one had the fewest number of pelves. A larger sample size from stage one may confirm that there is no difference in DNA amplification success among the different stages of pelves or may reveal that stage one pelves contain the best quality DNA. In addition, having a set of stage four femurs would help strengthen either conclusion, depending on the results.

Data from ribs did not appear to follow a downward trend; the lowest amplification success occurred at stage one and the difference among the four stages was not significant. This suggests that the degree of weathering of ribs also does not influence the quality of the DNA within. Interestingly, considering only stages two

through four, amplification success decreased with increasing degradation; stage two and stage three had nearly the same success while success at stage four dropped by 10%. The sample size of stage one ribs was small and could influence the results found at this stage. Alternatively, the small sample size of ribs at stage one may not have affected the results, in which case, amplification success truly was worse with the most well-preserved bones, perhaps due to the DNA being trapped in the better preserved material. DNA may be released more easily from the more degraded ribs, with their more fragile microstructure. A larger sample size of stage one ribs would confirm such a result or solidify that amplification success was the same among ribs of different stages.

A commonality among femurs, pelves, and ribs when categorized by bone staging was unequal sample sizes among the stages and small sample size at one or more stages, as opposed to the equal sampling using skeletal staging. However, the reason for the difference between the two staging methods varied depending on bone type. For example, most femurs staged by the bone rating system were categorized as either stage one or two, with only four rated as stage three and none as stage four. This indicates that poorly preserved skeletons may still have well preserved long bones. In contrast, more pelves were categorized as stage two or three by bone staging, with very few were categorized as stage one. Thus, it appears that pelves of moderate degradation were part of better preserved skeletons based on the skeletal rating system. The one problem with this view is that in no instance was a pelvis placed in stage four of the new system, while pelves were found in association with stage five skeletons. This may be due in part to the selection process of skeletal material; certainly only whole pieces of pelvis would be collected from the burials, and skeletons without this desired bone type were less likely to

be sampled. Finally, only three ribs were rated as stage one of the bone rating system. There are at least three possible explanations for this; the first is that ribs were generally in a more degraded state compared to the rest of the skeleton due to their thinness and overall fragile nature, and reclassifying them using the bone staging system revealed that. Second, it is possible that ribs, due to their small size and spongy microstructure beyond the cortex, were broken during transportation; thus ribs that appeared to be in better condition at the excavation site or at The Smithsonian arrived at MSU in worse condition. If this is true, then the number of stage four ribs would have increased as a result of initially well-preserved ribs breaking during transportation. The microstructure of these initially well-preserved ribs would still be intact and could increase the amplification success of stage four ribs. The third possibility, as mentioned above, is that the small sample size of ribs at stage one did not affect the results and amplification success was the lowest with the most well-preserved bones due to the microstructure of the bone. It would be interesting to base a study of DNA quality solely on the appearance of each individual bone, equaling sample size among categories, to fully determine whether the differences or lack thereof among bones of various weathering stages are real or not. From the data generated here, it would be predicted that individual bone differences would not influence DNA results.

Bone Type and Amplifiable DNA

The effect of bone type was analyzed on its own, showing a strong difference in amplification success among pelves, femurs, and ribs. The strongest difference occurred between femurs and pelves, followed by the difference between ribs and pelves, while

there was no statistical difference between ribs and femurs. This may be explained by the structural differences between the different bone types. The structure of long bones gives them less surface area per volume from which degradation can begin. In addition the microstructure of compact bone may be conducive to preserving DNA more so than that of spongy bone. With the osteocytes of compact bone embedded in spaces between concentric layers of hydroxyapatite and collagen, the cells, and thus the DNA within, may be well-protected. While the osteocytes of spongy bone are embedded in spaces within the trabeculae, this level of protection is not further reinforced by multiple layers of bone matrix, but instead is surrounded by open spaces; thus the less-protected osteocytes of spongy bone may be more susceptible to degradation. With this in mind, DNA from flat bones, such as ribs and pelves, will be better protected only in the cortex, where the structure is compact. If this cortical layer is not present, the sample may not provide useable DNA. In contrast, the compact microstructure extends beyond the cortical layer in the diaphyses of long bones, such as femurs and fibulae, providing a larger area of protected DNA. The difference in amplification success between ribs and pelves may be attributed to the rib's smaller, linear structure with a smaller surface area to volume ratio. The pelvis is irregularly shaped with a large surface area, making the cortical layer more susceptible to degradation. Gotherstrom *et al.* (2002) observed that DNA preservation was connected to both hydroxyapatite and collagen preservation. These two components were not studied at the microscopic level in this study; however, at a macroscopic level, femurs were better preserved than either ribs or pelves. A separate study has begun in Brazil to analyze the microscopic qualities of the Voegtly samples.

Sex and Amplifiable DNA

The difference between males and females was significant, with samples from males amplifying more often. There were insufficient data to compare male and female samples among each of the stages of the skeletal staging system; many stages contained no bones from one of the sexes. At each stage of the bone rating system, male bones amplified more often than female bones; however, this difference was strongest at stage three. The difference was not specific to any particular bone type as there was no difference between the sexes when comparing bone type. Of the 36 individuals analyzed, 11 were female, 17 were male, 7 were not determined, and the "possible male" was also sampled. Though the sampling between the sexes was fairly even, the traits available for sexing the Voegtly samples were generally not ideal. The degraded state of the skeletons lead to the estimation of sex from less sexually dimorphic qualities of the skeleton such as the robustness of the skeleton, and not on the more reliable features of the pelvis or the skull. On occasion, female sex was assigned if an earring was present within the burial, and male sex was determined if remnants of a moustache were present. With this in mind, it is possible that the skeletons were sexed incorrectly when judgment was based solely on robustness and /or burial artifacts. Relying on robustness could be particularly problematic for judging the sex of these individuals due to their ethnicity. The people buried in the Voegtly Cemetery were known to be of Swiss-German heritage. In a population that is more robust than average, there may not be enough sexual dimorphism to rely on this trait for sex determination. In this case, more robust females would be misclassified as male (as would more gracile males in a less robust population), changing the distribution of amplification success. Again, the anthropological estimates of sex

must be confirmed to determine if these data are correct. A separate study has begun to analyze nuclear DNA of the Voegtly samples to determine if the original sex determinations were correct or not. Preliminary results indicate that on occasion, they were not. If the anthropological estimates were generally correct, then male skeletons amplified more often most likely because they were larger and more robust, and therefore less susceptible to degradation.

Age and Amplifiable DNA

Age alone did not appear to affect amplification results in general. All aspects of the data (weathering stages and bone type) in comparison to age were considered. Small sample sizes may have been a factor in the analysis of these data however, as bones were distributed across five age categories. Eight children (16 bones), 2 adolescents (5 bones), 5 young adults (13 bones), 19 adults (50 bones), and one older adult (3 bones) were analyzed. Further dividing theses data based on weathering condition and bone type generated categories containing only one or two samples, creating a problem of small sample size. A significant difference was detected between amplification success and age at stage one of the bone staging system, with bones from children and older adults amplifying less often than those from adolescents, young adults, and adults. This may be an artifact of the small sample size of the child and older adult categories (3 bones and 1 bone, respectively), or perhaps indicate a difference in the quality of preservation between extreme and intermediate age categories. The latter idea supports Micozzi's (1991) description of a biphasic relationship between age and bone density, and thus survivability. Beyond random chance, this may be explained by the structure of the

bones at those respective ages. For example, at very young ages, calcification has not been completed and most of the skeleton is still constructed of cartilage. The structure of the bones is more susceptible to degradation and may not provide a stable environment in which DNA is protected. Such degradation would explain why of the over 300 individuals recorded in the Voegtly Church burial records of being less than one year old, only 175 were recovered and of these, seven were available for DNA analysis. Similarly, old-aged individuals were not well represented in this sample set largely due to the fact that most individuals buried in the cemetery did not live long enough to be placed in the older age category. The average life expectancy of individuals once adulthood was reached was estimated to be in the early to mid-30s, and of the 230 individuals who lived beyond the age of 20, only 80 individuals passed the age of 50 (Ubelaker *et al.*, 2003). At older ages the bones are well into the degenerative process, making them more prone to degradation and possibly breakdown of DNA Interestingly, no significant difference among the age categories was observed at any other weathering stage, and thus again the biphasic finding for stage one bones could be artifactual. On the other hand, this biphasic relationship might only be discernable in well-preserved bones, and not be detectable once bones have progressed to a higher level of degradation, meaning that there is a threshold of degradation at which there is a distinction among age categories. Finally, dividing the samples among the various stages and ages results in very small sample sizes for many of the categories, thus spurious results would not be unusual.

Conclusions

The results of this study indicate that data on age of an individual at death, weathering condition of the whole skeleton, and weathering condition of an individual bone are not useful for predicting the quality of DNA contained within the bones. Sex of the individual may play a role in the preservation of DNA, with DNA from male skeletons amplifying more often; however, the anthropological estimate of sex for the skeletons used in this study would ideally be confirmed genetically before this conclusion was reached. Bone type does appear to be useful in predicting the quality of DNA within the bone. Reclassifying the bones by the condition of each led to small sample sizes in several categories. Larger sample sizes based on individual bones would be useful to determine if the difference in amplification success with increased bone degradation is statistically significant; however, data from the stages with larger samples sizes, such as stage one and two femurs and stage two and three pelves and ribs, suggest that bone appearance is not indicative of DNA quality. Using bone type alone as a means of predicting DNA quality, femurs were found to be the best bone from which to obtain amplifiable DNA, and other long bones (fibulae, tibias, humeri, radii, and ulnas) may show the similar results. Ribs were found to be the next best bone type, followed by pelves. The difference in amplification success of both of femurs and ribs in comparison to pelves was found to be strongly significant. This was true of bones of all age categories and both sexes. Thus, with the knowledge of bone type, the forensic biologist should be able to predict the quality of the DNA within the sample. This would allow the scientist to choose the best available bone if more than one exists. In addition, guidelines could be provided to the individuals collecting the bone samples, so those most likely to

contain usable DNA are not overlooked and are sent to the forensic biologist. All of these factors have the potential to increase the efficiency of DNA analysis of bone.

References

Anderson, S., A.T. Bankier, B.G Barrell, M.H.L. deBruijn, A.R. Coulson, J. Drouin, I.C. Eperon, D.P. Nierlich, B.A. Rose, F. Sanger, R.H. Schreier, A.J.H. Smith, R. Staden, I.G. Young. 1981. Sequence and organization of the human mitochondrial genome. *Nature*. 290:457-465.

Behrensmeyer, A.K. 1978. Taphonomic and Ecologic Information from Bone Weathering. *Paleobiology*. 4(2):150-162.

Buikstra J.E. and D.H. Ubelaker, editors, 1994. Standards for Data Collection from Human Skeletal Remains, Proceedings of a Seminar at The Field Museum of Natural History. Arkansas Archeological Survey Research Series No. 44. Fayetteville: Arkansas Archeological Survey.

Boles, T., C. Snow, and E. Stover. 1995. Forensic DNA testing of skeletal remains from mass graves: A pilot study in Guatemala. *Journal of Forensic Sciences*. 40:349-355.

Gilbert, M.T.P., E. Willerslev, A.J. Hansen, I. Barnes, L. Rudbeck, N. Lynnerup, and A. Cooper. 2003. Distribution patterns of postmortem damage in human mitochondrial DNA. *American Journal of Human Genetics*. 72:32-47.

Gotherstrom, A., M.J. Collins, A. Angerbjorn, and K. Liden. 2002. Bone preservation and DNA amplification. *Archaeometry*. 44(3):395-404.

Hagelberg, E., I.C. Gray, A.F. Jeffreys. 1991. Identification of the skeletal remains of a murder victim by DNA analysis. *Nature*. 352:427-429.

Hall, T. 2004 July 4. BioEdit sequence alignment editor for Windows 95/98/NT/2K/XP. Homepage. <http://www.mbio.ncsu.edu/BioEdit/bioedit.html> Accessed 2004 July 13.

Hedges, R.E.M. 2002. Bone diagenesis: an overview of process. *Archaeometry*. 44(3):319-328.

Hochmeister, M.N., B. Budowle, U.V. Borer, O. Rudin, M. Bohnert, R. Dirnhofer. 1995. Confirmation of the identity of human skeletal remains using multiplex PCR amplification and typing kits. *Journal of Forensic Sciences*. 40:701-705.

Holland, M.M. and T.J. Parsons. 1999. Mitochondrial DNA sequence analysis—validation and use for forensic casework. *Forensic Science Review*. 11:22-49.

Holland, M.M., C.A. Cave, C.A. Holland, and T.W.Bille. 2003. Development of a quality, high throughput DNA analysis procedure for skeletal samples to assist

with the identification of victims from the World Trade Center attacks. *Croation Medical Journal*. 44(3):264-272.

Kaestle, F.A., and K.A. Horsburgh. 2002. Ancient DNA in anthropology: methods, applications and ethics. Yearbook of Physical Anthropology. 45:92-130.

Kalmar, T., C.Z. Bachrati, A. Marcsik, and I. Rasko. 2000. A simple and efficient method for PCR amplificable DNA extraction from ancient bones. *Nucleic Acid Research*. 28(12):e67.

McCutchen, L. 2003 Fall. Biology 2401: Human anatomy and physiology. Homepage. <http://blc1.kilgore.cc.tx.us/kcap1/practical_2_photos.htm> Accessed 2004 July 28.

Micozzi, M.S. 1991. Postmortem change in human and animal remains: A systematic approach. Springfield: Charles C. Thomas.

Monson, K.L., K.W.P. Miller, M.R. Wilson, J.A. DiZinno, B. Budowle. 2002. The mtDNA population database: an integrated software database resource for forensic comparison. *Forensic Science Communications*. 4(2). <http://www.fbi.gov/hq/lab/fsc/backissu/april2002/miller1.htm> Accessed 2004 August 12.

O'Rourke D.H., G. Hayes, S.W. Carlyle. 2000. Ancient DNA Studies in Physical Anthropology. *Annual Review of Anthropology*. 29:217-242.

Parsons, T.J. and V.W. Weedn. 1997. Preservation and recovery of DNA in postmortem specimens and trace samples. In Forensic taphonomy: the postmortem fate of human remains, ed. W.D. Haglund and M.H Sorg. Boca Raton:CRC.

Pääbo, S., R.G. Higuchi, A.C. Wilson. 1989. Ancient DNA and the polymerase chain reaction: The emerging field of molecular archaeology. *Journal of Biological Chemistry*. 264:9709-9712.

Primorac, D., S. Andelinovic, M. Definis-Gojanovic, I. Drmic, B. Rezic, M.M. Baden, M.A. Kennedy, M.S. Schanfield, S.B. Skakel, H.C. Lee HC. 1996. Identification of war victims from mass graves in Croatia, Bosnia, and Herzegovina by the use of standard forensic methods and DNA testing. *Journal of Forensic Sciences*. 41:891-894.

Rogan, P.K. and J.J. Salvo. 1990. Study of Nucleic acids isolated from ancient remains. *Yearbook of Physical Anthropology*. 33:195-214.

Shackleford, J.M. 1966. The ultrastructure of Mississippian and Archaic Indian bones from various soil and drainage conditions. American Journal of Phys. Anthro. 25:291-98.

Solomon C.D. and N. Hasse. 1967. Histological and histochemical observations of decalcified sections of ancient bones from excavations in Israel. *Israel Journal of Medical Science*. 3:747-54.

Smith, C.I. *et al.* 2003. The thermal history of human fossils and the likelihood of successful DNA amplification. *Journal of Human Evolution*. 45:203-217.

Stout S.D. and S.L. Teitelbaum. 1976. Histological analysis of undecalcified thin sections of archaeological bone. *American Journal of Physical Anthropology*. 44:263-69.

Tuross N. 1994. The biochemistry of ancient DNA in bone. *Experentia*. 50:530-535.

Ubelaker, D.H., E.B. Jones, Editors. Diane Beynon Landers, Associate Editor for Archaeology. 2003. Human remains from Voegtly Cemetery, Pittsburgh, Pennsylvania. Smithsonian Contributions to Anthropology Series, Number 46, Smithsonian Institution Press, Washington, D. C.