To the Graduate Council:

I am submitting herewith a dissertation written by Lori E. Baker entitled "Mitochondrial DNA Haplotype and Sequence Analysis of Historic Choctaw and Menominee Hair Shaft Samples." I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Anthropology.

Lyle W. Konigsberg, Major Professor

We have read this dissertation
and recommend its acceptance:

Karla J. Matteson

Richard L. Jantz

Mariana Ferreira

Accepted for the Council:

Vice Provost and Dean of Graduate Studies

Mitochondrial DNA Haplotype and Sequence Analysis of Historic Choctaw and Menominee Hair Shaft Samples

A Dissertation
Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Lori E. Baker
December 2001

DEDICATION

To Erich,
ILYAMSM.

And to Mom
who always told me:
"You could do anything you put your mind to."

ACKNOWLEDGMENTS

I sincerely appreciate the encouragement and support of my committee. Dr. Karla Matteson with her endless enthusiasm and patience, Dr. Lyle Konigsberg with his challenging, intellectual nature and Dr. Richard Jantz in sharing his unyielding quest for the understanding of New World populations have all greatly influenced and cultivated my intellectual development. I thank Dr. Mariana Ferreira for her strong convictions and her effort to make anthropologists aware of our impact on the individuals we study.

I would also like to express my deepest gratitude to my family, especially my husband, Erich, whose patience, support and help has been limitless. I thank my father, David, who gave me my love for anthropology and my mother, Hazel, who gave me the desire to attain this degree. I would also like to thank my grandmother and grandfather, Mimi and Papa, for all of the years of encouragement and support. I owe a great debt to Erich's parents, Walt and Linda, and my Aunt Kathy and Uncle Alan for their continued support of my aspirations. I would also like to thank, Scott, Danna, Patrick, Kimberly, Angela and Christina.

Finally, I would like to thank Ann, a.k.a. Annie Fanny, for always providing the best distractions from graduate school in the most interesting foreign places. Lastly, I would like to thank my friends, Lee , Danny, Erin, Nikki, Leigh and Samantha not only for their support but for making life much more fun.

ABSTRACT

Franz Boas organized the collection of information for over 15,000 Native Americans for the World's Columbian Exposition in 1893. Of these individuals, 2482 individuals gave hair samples that are currently stored at the American Museum of Natural History in New York. Now, over 100 years and many technical advances later, the information amassed from these Native American individuals is being reanalyzed to provide a genetic dimension to the analysis of this collection.

A new method for the extraction of DNA from aged hair shaft samples was developed and applied in this study. This is the first successful report of research of this kind. Sequence analysis of the mtDNA control region and haplogroup analysis was performed for 19 Mississippi Choctaw and 17 Wisconsin Menominee individuals represented in the Boas hair collection. These samples were then compared with other groups belonging to the same language families for inter-group and intra-group variation.

The main objectives of this research were to successfully extract and analyze mitochondrial DNA from these historic hair shafts and to determine if collections such as the Boas are useful for genetic analysis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1: Review of the Literature

Introduction

In the late 1800s Franz Boas organized the collection of anthropometric data and hair samples from various Native American populations for the World's Columbian Exposition. The purpose of Boas' effort was to preserve as much information as possible about the inhabitants present at the time that Christopher Columbus landed in the New World. Boas was acutely aware of the rapid dwindling of America's diverse and interesting populations and knew that without proper, intensive cataloging many of their unique and independent characteristics would be lost to the homogenization of a growing American culture. Boas recognized both the scientific and the cultural importance of phenotypic and demographic variation of all populations.

The research presented herein analyzes hair samples collected for the Columbian Exposition in an attempt to better understand the genetic diversity of two populations. The groups examined include members of the Mississipian Choctaw and the Wisconsin Menominee (Appendix A). It is the first study to examine population differentiation by use of DNA from historic hair shaft samples, and it is one of the few studies that analyzes both mtDNA haplotype and control region sequence data from non-modern samples.

In summary, the objectives of this research are threefold. First, to develop procedures to successfully extract and analyze both nuclear and mitochondrial DNA from historic hair shaft samples. Secondly, to show that DNA from aged hair shafts can be used to address questions of the genetic diversity in distinct human populations at a fixed point in time. Lastly, this research demonstrates the utility of archived collections, such as the Boas collection, to augment temporal and geographical information in the genetic history of Native American groups.

The Boas Collection

Franz Boas, heralded as the father of American anthropology, was responsible for assembling anthropometric and demographic data from approximately 15,000 Native Americans from over 200 tribal groups. These data were collected primarily for the 1893 World's Columbian Exposition held in Chicago commemorating the 400th anniversary of Columbus's "discovery" of the New World. The Exposition was organized into sections to herald American and European culture, and to document the history of the United States before and after European contact. Boas saw this as an opportunity to preserve a fraction of information about the multitudes of diverse civilizations that were being destroyed by this European contact.

Frederic Putnam, the organizer of the anthropology exhibit at the Columbian Exposition, hired Boas in 1891 to serve as his chief assistant (Darnell, 1970). The reasons behind the large and aggressive cataloging effort is best described by Putnam in a letter to the Chicago Tribune on May 31, 1890. Putnam wrote that the Exposition "will be the first bringing together on a grand scale of representatives of the peoples who where living on the continent when it was discovered by Columbus."

Boas organized and trained 50 anthropologists to collect data at the major Amerindian sites located in the US and Canada. The observers collected data using a standardized data form that included 12 anthropometric measurements of the body, head and face. The forms also included information such as name, age, sex, birthplace, the location where the sample was taken, tribal affiliation, parent's tribal affiliations and the individuals relation to others in the study.

In several articles, Boas attempted to analyze a portion of the data collected (Boas, 1895; 1899; 1905); however, without the aid of computers, the vastness of the information was difficult to manage. In the following seven decades there was little work done with the data (Sullivan, 1920; Hall and MacNair, 1972). Fortunately, in the 1980s, Dr. Richard Jantz led an effort to computerize the data into a searchable Paradox database. The resulting databases are available from Dr. Jantz at the University of Tennessee, Knoxville. Since the computerization of

the data, the Boas database has served as a reservoir for many scientific and academic efforts. (Hunt et al., 1988; Prince, 1989; Falsetti, 1989; Falsetti and Jantz, 1990; Stivers, 1990; Jantz et al., 1992; Konigsberg and Ousley, 1995; Szathmary, 1995; Jantz, 1995; Hall and Hall, 1995; Ousley, 1995; Comuzzie, 1995).

In addition to the anthropometric and demographic information, hair samples from roughly 2482 individuals representing over 70 different groups were also collected and are now stored at the American Museum of Natural History in New York. The hair samples were rediscovered in 1995 and have been matched with the corresponding individuals in the database. The original data files do not make any mention of hair being sampled and, therefore, for years the hair samples have gone virtually unnoticed.

Each hair sample consists of a few strands of cut hair sealed in a paper envelope with identifying information hand written on the outside. When the collection was inventoried, it was noticed that many envelopes have "World's Columbian Exposition" as part of a return address in the upper left-hand corner, leaving little doubt as to the sample's origin. Additionally, the observer's name, individual's name, an identifying number and various other information on the envelope corresponds to that of the original data files.

While there are over 15,000 individuals that participated in the anthropometric study, hair samples from less than 2,500 individuals were found. It is uncertain why there are hair samples from only a subset of individuals. There are several possibilities. First, for the Choctaw (Mississippi and Oklahoma) in particular, there are roughly 522 individuals who participated in the study but only 150 hair samples. Information for the 522 Choctaws studied was recorded by four different observers; however, hair samples exist only for individuals studied by one of the observers (WAH). It is possible that only some observers collected hair samples along with the anthropometric data. Out of the 50 observers who collected data, only 24 were represented as collectors of the hair stored at the museum.

Secondly, there is the possibility that some individuals made the decision not to participate in the donation of hair samples. The original data files included a portion of files on which "declined" had been written for those who were not interested in participating in the anthropometric part of the study. Since there was not a section on the original files concerning hair sampling, it is impossible to know if individuals may have declined that particular part of the study. Of the roughly 1,600 Mississippi Choctaws only 329 participated in the anthropometric study and, of these, only 152 donated hair samples. These numbers may indicate that those individuals who participated in the anthropometric

analysis were not necessarily willing to have a part of their hair cut and collected. It is always a concern when dealing with historic collections such as these, whether individuals had the power to decline participation. Of particular concern in Native American collections, is whether the participation of younger individuals residing at boarding schools had the power to decline inclusion in the data process.

Yet another possibility for explaining the disparity of data files to hair samples may be that there are additional hair samples that have not yet been found. The original data sheets were stored at several institutions and this might also be the case for the hair samples. However, it is most likely a combination of each of the above-mentioned reasons that has contributed to this discrepancy.

The Boas collection has been a valuable source for population information because the anthropometric data set included a record of the parents' tribal heritage and the blood quantum of that heritage. Now, with the discovery of the hair samples, molecular information likewise becomes available.

Genetic analysis of this collection is significant because of its extensive documentation of demographic and kinship information. This detailed collection removes the need to estimate admixture rates due to the known biological heritage of individuals and their parents. The antiquity of the sample, itself, implies that the admixture rates will be

somewhat less than that found in modern groups. In addition, these samples are important because they provide a genetic evaluation of a population at a fixed point in time. In the examination of specimens from archaeological sites, dates have to be estimated by a form of dating that requires the attachment of a confidence interval. Also, the analyses of specimens from archaeological sites are restricted by the availability of specimens that have been excavated for that particular time period. This makes it difficult to compare specimens of the same age to one another, leaving many temporal gaps in the understanding of past and present Native Americans. Because all of the 2482 individuals donated their hair in a one-year period, the 73 different groups in the collection can be compared in relation to each other at the same point in history. Perhaps the most limiting factors of the Boas hair collection are the sample sizes and the small amount of DNA contained in the hair itself. It is apparent that the multitude of information about each individual, such as place of birth, age, sex, tribal affiliations, parents heritage, blood quantum and anthropometric data combined with DNA analysis provides the potential for a very well-rounded look at a relatively large number of individuals from these historic populations.

Ethical Considerations

While the Boas hair collection offers a unique opportunity to glance into the genetic past of over 2,400 individuals, it also introduces several issues for ethical consideration. These issues are important and are not overlooked. Precautions were taken to insure that the primary ethical obligations to the people represented in this study were fulfilled according to the American Anthropological Associations code of ethics. While one would like to assume that individuals gave their hair samples with the idea that they would be stored for future study, informed consent was not really a consideration in the 1890s. Advances in technology could not foresee the implication of DNA analysis. Therefore, because genetic studies were unforeseen at the time of data collection, this study strives to provide as much protection as possible to each individual that was a participant. The present study does not divulge the names of individuals included to protect the privacy of these individuals. In addition, the identity of those individuals included in the analysis is not disclosed in order to protect the living descendents from any possible adverse effects of the generated information. While a person's identity comes from many areas such as community ties and associations, the effect of how a person feels about that identity may be altered by the knowledge of their genetic make-up. By not disclosing the identity of the individuals genetically analyzed, information that could potentially

influence the genetic self-knowledge of living descendents is hopefully avoided (Condit, 1999; Finkler, 2000).

MtDNA data in the present study has no known biomedical implications and therefore is non-diagnostic in any way. The information from the present study merely shows a small bit of information of the maternal genetic heritage of each individual. The information is then compiled with that of other individuals from the same group to determine the amount of maternal diversity for the group. This information is then compared to the diversity observed in other language-related groups to search for similarities and differences. The genetic history presented herein does not aim to replace the oral history of these groups, it is just another aspect describing part of the population's past. Each individual gave a significant gift to the future in their participation and this study strives to honor each individual and their descendents. The present study aims to generate additional information about the individuals who were part of Boas' preservation efforts.

Cultural Background

Choctaw

There are several Choctaw migration legends that explain their arrival in Mississippi. In one account, the Choctaws and Chickasaws lived in a country in the Far West, under the leadership of two brothers,

Chahta and Chikasa (Halbert, 1899). The population became so large that the land was no longer providing sufficient food. When the prophets declared that there were fertile lands far to the east, the entire population decided to make the journey to this bountiful land. In order to find enough supplies along the way, the people were divided into two groups each separated by a day's travel. A prophet led the tribe, carrying a pole and planting it each night in front of the camp. The direction that it leaned the next morning pointed the way for the next day's journey. After many months, they arrived at Nanih Waiya, where the pole stood erect. That same day the group led by Chikasa camped across the creek from Nanih Waiya. But that night a great storm occurred, and it rained for several days, making the waters impassible. A few days later, messengers were sent to Chikasa to tell him of the good news: they had found their destination. Unfortunately, Chikasa had already gone; he had continued on to the Tombigbee River, where he and his followers halted and consequently became a separate nation. Although there are many versions of the migration legend, each involve a journey from the West and a settling at Nanih Waiya, which today is a large mound located in Mississippi.

The Choctaw are a part of the Muskogean language family, which is in the Macro-Algonquian phylum, and their social organization is similar to that of the other Mississippi tribes (Bettersworth, 1959). There

were two classes of people, an upper and a lower, each possessing elaborate subdivisions. The tribe was separated into two moieties, the Imoklasha and the Iholahata, each of which consisted of several clans (Gibson, 1973). The clans were exogamous and in matters of inheritance and descent, the Choctaws were matrilineal.

Hernando De Soto led the first European expedition that contacted the Choctaw in 1540. However, it was two and a half centuries later before there were any extreme changes in their social institutions due to European contact (McKee and Schlenker, 1980).

The Choctaws resided primarily in the central and southern part of Mississippi and a portion of southwestern Alabama prior to forced relocation to the Indian Territory (later Oklahoma) in the 1830s. The Choctaws, together with the Chickasaws, Creeks, Cherokees, and Seminoles were known as the "Five Civilized Tribes." They are considered closely related to the Chickasaws and more distantly related to the Creeks and the Seminoles.

Prior to European contact the Choctaws were skilled agriculturists living on very fertile land. Although, they owned less land than surrounding groups, they had a higher rate of production and often sold their surplus to their neighbors. The Choctaws easily accepted the idea of increasing their trading partners to include the Europeans. They

welcomed the new domestic plants and animals, as well as European tools and weapons that the foreigners had to offer.

During the eighteenth century, the Choctaws were town dwellers with each town controlled by a council of clan elders. The towns were joined together into a nation with a general council and three principal chiefs, each representing a geographic division of the nation. By mid-1700 there was a rapid increase in mixed-blood families, since the French and British had been living for several generation among the Mississippi tribes. In the last half of the century French and British settlers were mating with the Mississippi Indians giving rise to mixed-blood families (Gibson, 1973). The changes, which are apparent from the increase of mixed bloods among the Choctaws, did not reach full velocity until the time of relocation. Gradually the mixed bloods, who identified more with their European fathers than their Choctaw mothers, rose to become a type of aristocracy in the tribe. Although they remained a part of the Choctaw community they practiced a way of life that was more European than Choctaw culturally.

During the 1750s European traders and settlers brought African slaves to the Choctaws (Gibson, 1973). Slaveholding was practiced and was somewhat dominated by the 'mixed-blood aristocracy' who dominated in Choctaw society. (Edwards, 1932; Debo, 1934; Hudson, 1934; Morrison, 1987).

As the European population of settlers increased around the lands of the Choctaw, the American government increased its pressure to buy the fertile land from the Choctaw. Ultimately, this pressure resulted in the Treaty of Dancing Rabbit Creek, which negotiated the removal of the Choctaw from Mississippi. The 14th article of the treaty asserted that if Choctaws became citizens of the United States they could remain in Mississippi. In order to remain in Mississippi, each person was required to file claim for their land within six months after ratification of the treaty, and had to live on the land for no less than five years.

Many Choctaws remained in Mississippi, but thousands were removed to the Indian Territory (Oklahoma). It is estimated that 12,500 Choctaw migrated along with 248 slaves of these 2,500 died, and about 5,000 Choctaws remained in Mississippi (Debo, 1934; McKee and Schlenker, 1980). The Choctaw population in Mississippi was reduced to one-forth its original size. Throughout the nineteenth century, some Mississippi Choctaws continued to migrate to Oklahoma.

These historical movements lead to the question: Do the Choctaw exhibit a lower amount of genetic diversity because of this known recent bottleneck event and will it be more evident in the Mississippi Choctaw than in the Oklahoma Choctaw? Unfortunately, there are no Oklahoma Choctaw represented in the hair collection for comparison; although,

there are roughly 329 represented in the anthropometric and demographic data.

Menominee

The Menominee creation myth not only traces their ancestry, but also serves as base for their social organization, family precedence, and civil government. The Great Mystery, who made the earth, permitted the Underground White Bear with a copper tail to emerge from the earth at Mini'kani near the mouth of the Menominee River and to assume human form. The White Bear's need for a brother prompted the Great Mystery to summon the spirits, known as Thunderers, from above the earth to join the bear man. Eagle descended and became a man. Other Thunderers joined the pair and Nama'kumkiu, the Beaver Woman, also became a generic brother to the group, as did the Wolf. The entire family lived together along the waters of the Menominee, where wild rice grew and sturgeon abounded (Hoffman, 1896; Keesing, 1939; Ouranda, 1979).

The Menominee are part of the Algonquian language family as are many other groups in the northeastern United States. The Menominee lived mostly as sedentary village dwellers with their social relationships organized around a number of patrilineal, exogamous clans apparently linked in two moieties, Bear and Thunderbird (Hosmer, 1999). The largest of the Menominee villages stood at the mouth of the Menominee

River where it empties into Green Bay, near the present-day city of Green Bay, Wisconsin. This lush area provided an abundance of "wild rice," fish and wild game, so mobility was not necessary. The Menominee vied with the Chippewa, Ottawa, Potawatomi and the Winnebago for this much-desired territory around the Green Bay area.

The first European contact with the Menominee was by the French Sieur Jean Nicollet who arrived in Green Bay in 1634. Thirty years after this contact, other Frenchmen entered the region and developed strong ties and relationships with the Menominee. The sedentary lifestyle of the Menominee was altered as traders moved into the area and desired the pelts of fur-bearing animals. The Menominee became scattered into mobile bands and began to hunt for animals bearing fur to trade with the Europeans.

The relationship between the Menominee and the French settlers was generally amiable and ties grew stronger with the increase in marriages of Frenchmen to Menominee women (Hoffman, 1896). The French supplied arms, ammunition and clothing to the Menominee in great abundance which slowly modified their original lifestyles. In 1763, the French withdrew from the area and were replaced by the English and the supply of European goods changed. The English did not enjoy as friendly a relationship with the Menominee as the French had; the

Menominee nevertheless sided with the English in the war of 1812 (Hoffman, 1896).

In the treaty of 1854, the Menominee were pressured into giving up their lands and were relocated to a reservation on the Wolf and the Oconto rivers in north central Wisconsin (Keesing, 1987; Ourada, 1979). The treaty provided for the erection of a sawmill on the Menominee reservation, thus introducing the Menominee to timber as a valuable resource. Although the Menominee were now located on a reservation, contact remained open between the Menominee and the whites due to the lumber trade.

By the mid-nineteenth century, one-half to two-thirds of the 1,700 Menominees were mixed bloods (Hosmer, 1999) and by 1885 officials calculated that 675 of the 1,308 (or 52 percent) Menominee were of mixed blood. The approximate population totals for the time of the Columbian Exposition can best be understood in a report by the Commissioner of Indian Affairs for 1892, the "whole number of Menomini reported on the reservation is 1,335, with 343 children of school age." This number does not count the approximately 300 persons of Menominee lineage living in scattered enclaves to the east of the reservation in places like Oconto and Menomonee. The 300 undercount raises the population estimate to 1,635 in line with estimates taken during the summer of 1893 (Hoffman, 1896). Although there are 273 of

the 1,635 Menominee individuals represented in the anthropometric database, hair samples for only 51 of these individuals were found. Genetic analysis of 17 of these individuals is included in the present study.

Technical Aspects

Hair

The advent of polymerase chain reaction makes possible the molecular analysis of minute quantities of intact or partially degraded DNA. One benefit of this technology is the examination of biological samples that are limiting in their concentration of DNA. The hair shafts in the Boas collection are just such samples.

Hair shafts are highly keratinised structures produced by a single hair follicle. The hair bulb, representing the base of the growing hair shaft, consists of actively dividing epithelial cells. Because the hair shaft contains significantly less DNA than is found in the hair bulb, molecular techniques in the past restricted hair studies to the DNA extracted from the hair bulb (Barbujani et al., 1996; Vigilant et al., 1990).

Recently, new methods were reported for the extraction of DNA from modern hair shafts. In 1995, Wilson and colleagues report the use of an organic extraction method for isolation of mtDNA from hair shaft samples. In another 1995 article, Wilson et al. provided the validation of

this methodology for use in forensic casework. In 1998, Allen and co-workers documented that mtDNA from shed hairs can be positively matched with the mtDNA taken from criminal suspects. Pfeiffer et al. (1999) examined the sequencing success rate of shed hairs from different areas of the body, and Huhne et al. showed the reliability of mtDNA from hair shafts by examining intra-individual differences. Lastly, a study by Baker and colleagues (2001), using the same silica-guanidine thiocyanate extraction procedure as employed in the present study, showed that mtDNA extracted from teeth and hair shafts taken from forensic casework provide the same mtDNA sequence in every instance. The same sequence patterns were observed for each tissue type tested including samples that had undergone a variety of degrative processes. These studies demonstrate that the mitochondrial DNA from human hair shafts is adequate for genetic analysis.

More and more frequently, hair is being recovered from archaeological sites. Humans shed hundreds to thousands of hairs per day and some animal species shed even more. After a period of prolonged occupation, vast amounts of hair can be scattered throughout a site. Therefore, attention is being paid to the recovery of hair as a biological artifact. Most often the provenience of human hair is a problem because it is difficult to know if a hair is indigenous to the site, or if it has been deposited by one of the excavators or other individuals

that may have passed through the area. Consequently, there is an increasing movement among archaeologists to develop site excavation methods that minimize possible human hair contamination. It is a difficult proposition and extreme care must be taken to make sure that a hair is true to the site. Because of the inevitable problems of human contamination of sites, previous studies of archaeological hair focused on non-human samples. Large amounts of animal hair can be found during a careful excavation. Therefore, hair has been used mostly to determine what species have occupied a site in the past (Kalbe, 1988).

Shed hair as a source for genetic testing has become increasingly important with the enactment of the Native American Graves and Repatriation Act (NAGPR). NAGPRA addresses the improper collection, use, display and study of human remains from American Indian groups. The law restricts the use of human burial remains without proper permissions, and it has traditionally been difficult to obtain permissions for studies that prove destructive to the sample. These include methods employed by radiocarbon dating and genetics. Because humans shed hundreds of hairs each day, most individuals pay little attention to this lost biological entity; consequently, shed hair that is not associated with burials does not have the same protection provided by NAGPRA. Therefore, some researchers are exploring alternative sources for genetic material such as shed hair. Although some groups like the Coquille of

Oregon, the Northwestern Band of Shoshone, the Aleut Corporation and the Chaluka Corporation (O'Rourke, 2000) to name a few have taken an interest in genetic testing for a greater understanding of their past, many Native American groups remain opposed. Permissions for genetic study of Native American groups have become increasingly difficult to obtain due to past relationships of geneticists with various groups. Data has not always been shared freely or research has been conducted that has no benefit to the groups participating.

The Boas collection is important because it allows Native Americans a chance examine the genetics of their ancestors without the destruction of burial remains. Concurrently, researchers studying Native American populations and genetics can also extract significant information that may benefit future generations of Native Americans.

MtDNA

Previous analyses of mtDNA from hair shafts have reported the use of organic solvents, usually phenol/chloroform, in the extraction methodology (Wilson et al., 1995a; Uchihi et al., 1992; Wilson et al., 1995b). However, analysis of DNA from ancient and historical anthropological samples has employed a silica-based method for DNA extraction (Hoss and Paabo, 1993). Therefore, for this study, a method of DNA extraction was developed specifically for use with historic and

ancient hair shaft samples using the proven methodology of silica-based protocol.

Human mitochondrial DNA (mtDNA) is a maternally transmitted circular genome approximately 16,569 bp in length (Hutchinson et al., 1974) (Figure 1-1). The most polymorphic region of the human mtDNA genome is concentrated in two hypervariable segments within the non-coding, D-loop region (Greenberg et al., 1983). The mutation rate in this area is 5-10 times higher than that of nuclear DNA (Brown et al., 1982; Cann et al., 1987). The high mutation rate of mtDNA in this area combined with the maternal mode of transmission allows comparison of mtDNA sequences between maternally related individuals, and is routinely employed for identification purposes (Allen et al., 1998; Ginther et al., 1992; Higuchi et al., 1988; Holland et al., 1993; Piercy et al., 1993; Stoneking et al., 1991; Sweet et al., 1995; Wilson et al., 1993; Wilson et al., 1995). In addition, mtDNA is present in hundreds to thousands of copies per cell and therefore may be better suited than nuclear DNA for ascertaining genetic information in cases where DNA amounts may be limited and/or degraded. Furthermore, mtDNA does not undergo recombination (Giles et al., 1980) allowing an estimate as to how many maternal lineages there are in each population tested.

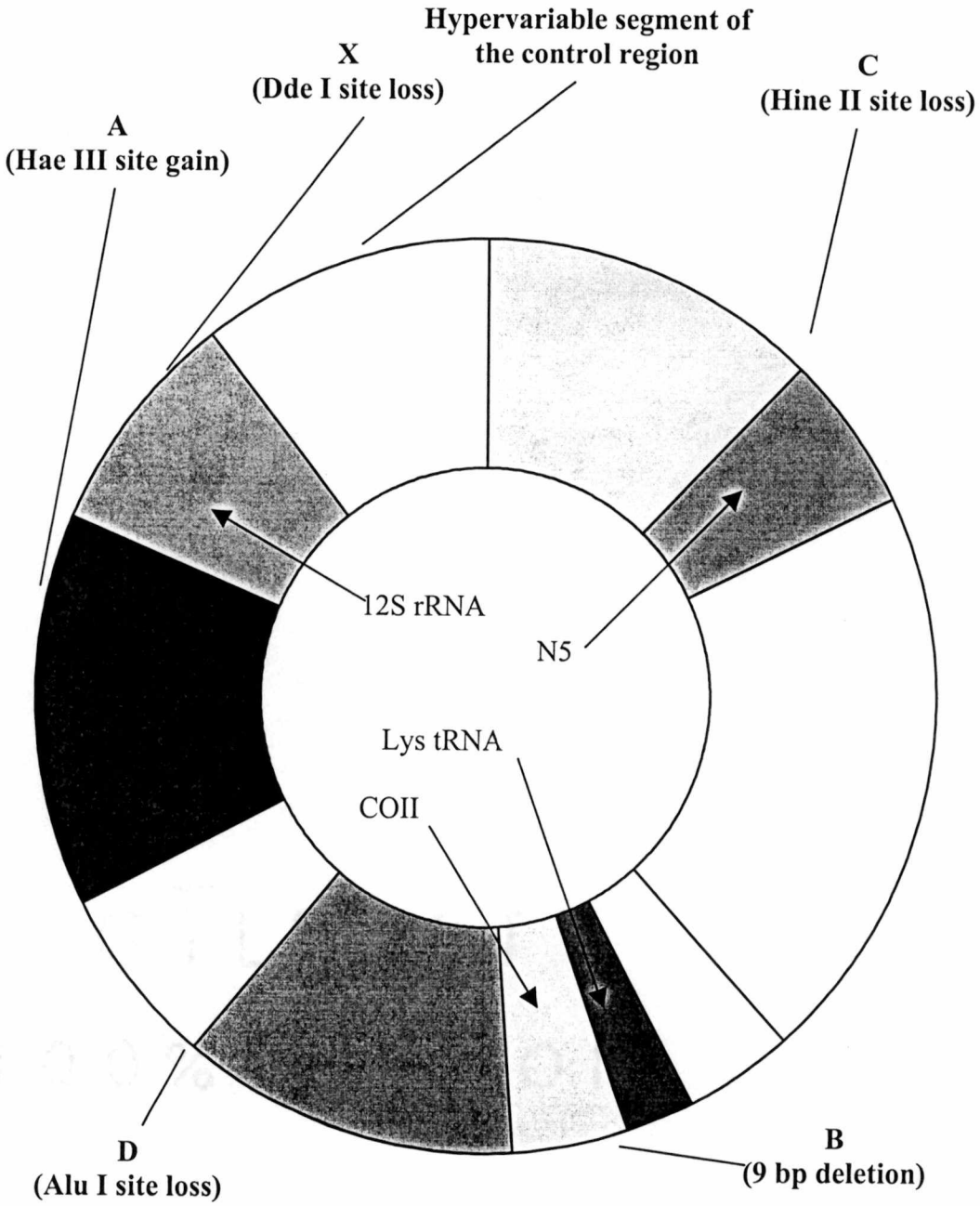

Figure 1-1: The Mitochondrial DNA Genome. Mitochondrial DNA is a circular molecule. The top portion represents the Control Region that contains the two hypervariable segments. Four restriction site locations and one deletion location that are used to classify individuals into haplogroups are depicted.

MtDNA research dealing with Native Americans focuses on both regions of the mtDNA genome: the coding and the non-coding regions. The coding region supplies several areas of genetic mutation that, when analyzed by restriction digestion, help to classify worldwide population types. For the Americas, thus far, there are 5 haplogroups that have been defined and confirmed using this technology: A, B, C, D, and X (Wallace et al., 1985; Schurr et al., 1990; Torroni et al., 1992; Horai et al., 1993; Scozzari et al., 1997; Brown et al., 1998; Stone and Stoneking, 1998; Kaestle and Smith, 2001). Each of these 5 haplogroups is defined by a particular set of polymorphisms (Table 1-1). For example, haplogroup A is defined by a Hae III restriction site gain at position 663. The change can also be determined by direct sequencing of the region.

Table 1-1: RFLP, Deletion and Control Region Polymorphisms

Haplo-Group	Polymorphic Restriction and Deletion Sites	Polymorphic CR Nucleotides
A	+HaeIII 663	16111T, 16223T, 16290T, 16319G, 16362C
B	COII/tRNA^{LYS}9-bp Deletion	16189C, 16217C
C	+DdeI 10394, +AluI 10397, -HincII 13259, +AluI 13262	16223T, 16298C, (16325C), 16327T
D	-AluI 5167, +Dde 10394, +AluI 10397	16223T, (16325C), 16362C
X	-DdeI 1715, +AccI 14465	16189C, (16213G), 16223T, 16278T, 200G

Populations can further be classified by an analysis of the control region, specifically hypervariable region I (HVR I) and hypervariable region II (HVR II). Sequence analysis provides the nucleotide code that can help to further delineate within and between the haplogroups. The control region is under reduced functional constraints and therefore it can have a high rate of change without causing debilitating or adverse effects.

Specific Aims

The main goal of this research was to develop a technique for the extraction and analysis of ancient DNA (aDNA) from human hair shafts. In doing this, the Boas collection as a source of genetic data is evaluated. It is the first study to examine population differentiation by use of DNA from historic hair shaft samples. The archived hair shaft samples are proven to be useful in providing genetic information about past people.

The objectives of this research are threefold. First, to develop procedures to successfully extract and analyze mitochondrial DNA from historic hair shaft samples. Secondly, to show that DNA from aged hair shafts can be used to address questions of the genetic diversity in distinct human populations at a fixed point in time. Lastly, this research demonstrates the utility of archived collections, such as the Boas

collection, to supplement temporal and geographical data in the genetic history of Native American groups.

Chapter 2: Materials and Methods

Overview of Method

DNA was extracted from hair shaft samples by a modification (Baker et al., 2001; Baker in press) of the silica/guanidine thiocyanate (GuSCN) method described in Boom et al. (1990). Höss and Pääbo (1993) previously reported another modification of the Boom et al. method for use on ancient samples.

A particular benefit of the silica technique is that the strong DNA binding capacity of silica allows the recovery of small amounts of DNA. This characteristic of silica extraction allows for the extraction of highly degraded fragmentary DNA such as that from hair shaft samples. Unfortunately, as a result of its specificity, the method is also very sensitive to contamination by contemporary DNA. The method therefore, necessitates extra care in all extraction procedures.

DNA Extraction Buffer Preparation

Extraction buffer (L6 lysis buffer) [10 M GuSCN, 0.1 M Tris-HCL (pH 6.4), 0.2 M EDTA (pH 8.0), 1.3% Triton X-100] was prepared according to Boom et al. (1990).

Eight grams of guanidine thiocyanate (GuSCN) is dissolved into 6.67 mL of 0.1 M Tris hydrochloride (pH 6.4) in a 50 mL screw top tube. To

facilitate the dissolution of the GuSCN into the Tris, water is heated in a glass beaker in the microwave to 60-65 C and then the 50 mL tube is placed in the glass beaker until the GuSCN dissolves. Shaking the tube periodically is helpful. To this add 1.47 mL of 0.2 M EDTA (pH 8.0) and invert several times to mix. Finally, add 1.67 mL of Triton X-100 (Sigma Chemical Co., St. Louis, MO) and invert several times to mix. Contaminating DNA is removed from the buffer by the addition of 1.5 grams of silica (Sigma Chemical Co., St. Louis, MO), mixing thoroughly followed by centrifugation at 2,000 X g for 15 minutes to pellet the silica. The extraction buffer was then distributed into single use aliquots (~200 µL).

Laboratory preparation of silica described by Boom et al. (1990) is time consuming, difficult and contamination is easily introduced. Therefore, in order to simplify the extraction procedure, GlassMilk®, from a GeneClean II kit (Bio 101, La Jolla, CA) is substituted for the prepared silica.

A series of blank extractions (all of the reagents are used but no sample DNA was added) are performed to test for contamination. If the subsequent PCR results show signs of DNA, the extraction buffer is once again silica purified and re-tested. If the contamination is not eliminated all solutions and buffers are remade.

Sample Preparation

In a pre-cleaned, pre-UV treated laminar flow hood, six to eight centimeters of a single strand of hair from each individual is chosen. The hair samples are cut hair shafts without the hair root. The hair was washed with 100% ethanol and rinsed with water that had undergone reverse osmosis (here after referred to as RO water) to eliminate any external contamination that might be on the hair. The sample was then cut into 1 cm portions using stainless steel surgical scissors and forceps and placed into a pretreated 0.2 mL glass tissue homogenizer (Kontes Glass, Vineland, NJ). After all portions are in the tissue homogenizer, 100 µL of extraction buffer is added and the hair is ground until no distinct portions remain. The homogenate is then transferred to a UV-irradiated microcentrifuge tube. Another 100 µL of extraction buffer is added to the homogenizer and used to rinse the remainder of any product off of the sides. The wash is subsequently pooled with the sample. Samples are placed in an incubator at 60°C with slight agitation overnight (10-24 hours).

The DNA is then isolated using a GeneClean II kit (Bio 101, Vista, CA). Three volumes of sodium iodide and 5 µl of GlassMilk® are added to the sample and incubated at 57°C for 15 min. with slight agitation. Following incubation, the sample is centrifuged for 5 min. at 12,000 X g and the supernatant discarded. The pellet is washed by re-suspending it

in 500 μL of BIO 101 New Wash® solution (from the GENECLEAN II kit) and then centrifuged for 5 minutes at 12,000 X g and the supernatant discarded. This step is repeated once. Any remaining New Wash® is removed by centrifugation for an additional 3 min. The sample is eluted by the addition of 30 μL of 10 mM Tris/1 mM EDTA buffer (pH 7.6) (TE) and incubation at 56°C for 10 min. The sample is then centrifuged for 5 min. and the supernatant moved to a UV irradiated microcentrifuge tube that is stored at -20°C until amplification.

Contamination Precautions

Stringent precautions were employed to eliminate the possibility of contamination. This is of particular concern because the small amount of DNA in the hair shaft and the age of the samples is easily masked by the preferential amplification of modern DNA.

Disposable lab coats, masks, gloves and surgical caps are worn while doing extractions and PCR reactions. All equipment (e.g. hood, microcentrifuge, scale, pipettes, trays, surgical steel scissors, forceps, weigh boats, pens, spatulas, beakers, graduated cylinders, pipette tip boxes) is cleaned with 10% bleach, rinsed with 100% ethanol and then UV irradiated for 24 hours prior to use. In addition, all microcentrifuge tubes are UV irradiated for approximately 12-24 hours before use. Sterile, pre-loaded, filtered pipette tips are always used.

Buffers and solutions are made with molecular biology grade (MBG) water in a laminar flow hood using sterilized equipment and aliquoted into single use amounts. Upon receiving the GENECLEAN II kit, NaI is aliquoted into single-use centrifuge tubes and stored at room temperature. The New Wash concentrate is mixed according to the manufacture's instructions, aliquoted into single-use containers, and stored at –20 C until use. The extraction buffer is prepared as described herein, aliquoted and stored in the dark for no more than three weeks. All of the preceding is done in a laminar flow hood with pre-treated equipment. Glass grinders are soaked in bleach, rinsed with 100% EtOH and then RO water, and UV irradiated for 24 hours prior to each use. A mock extraction is done before any sample material is attempted to check for contamination.

All amplifications included blanks in which all reagents are used but no DNA was added. An extraction blank using extraction buffer exposed to all procedures is always included. If there is ever evidence of contamination, all samples are discarded and steps are taken to eliminate further contamination.

Extractions took place in a dedicated clean room with positive pressure in a UV equipped laminar-flow hood using equipment that is cleaned with 10% bleach, followed by 100% ethanol and UV irradiation for 24 hours before each use. PCR took place in a room designated for

this purpose alone with dedicated equipment cleaned with 10% bleach. Gel electrophoresis and sequencing took place in two other separate rooms.

Mitochondrial DNA Typing

Both modern and ancient Amerindian populations have mtDNA patterns that typically fall into five clusters. Each cluster, or haplotype, can be characterized by specific mtDNA markers (Table 2-1). This study used a subset of these markers to determine the haplotype classification for each individual.

Although the hair shafts are only approximately 100 years old, the mtDNA was partially degraded, leaving segments of 100 to 300 bp in

Table 2-1: Primers for mitochondrial DNA typing

Grp	Position	Primer Sequence (5' to 3')	Bp*	Reference
A	L635 H708	TGAAAATGTTTAGACGGCCTCACATC TAGAGGGTGAACTCACTGGAAC	121	Handt et al. (1996) Handt et al. (1996)
B	L8215 H8297	ACAGTTTCATGCCCATCGTC ATGCTAAGTTAGCTTTACAG	119	Wrischnik et al. (1987) Wrischnik et al. (1987)
C	L13257 H13393	AATCGTAGCCTTCTCCACTTCA TCCTATTTTTCGAATATCTTGTTC	182	Handt et al. (1996) Ward et al. (1991)
D	L5127 H5189	ACTACCGCATTCCTACTACTCA GGGTGGATGGAATTAAGGGTGT	106	Handt et al. (1996) Handt et al. (1996)
X	L1689 H1782	GCTAAACCTAGCCCCAAA TTTCATCTTTCCCTTGCGGTAC	113	Green et al. (2000) Green et al. (2000)

*Number of base pairs of the amplified product

length, similar to previous reports (Paabo, 1989; Paabo et. al., 1988). Therefore, due to the small quantity and the degraded nature of hair shaft DNA, these samples were treated as ancient material and segments of less than 250 bp were amplified.

Previous ancient DNA (aDNA) studies report a standardized set of parameters that meet the needs of all primer sets used for restriction analysis (Stone and Stoneking, 1998; Kaestle and Smith, 2001). However, perhaps because of inhibitors found in hair, PCR conditions in this study required adjustment. Annealing temperatures, primer and $MgCl_2$ concentration as well as cycle number for each primer set included in the analysis had to be optimized. Additionally, heat activated AmpliTaq Gold (PE Applied Biosystems, Foster, CA) is always used in order to increase the cycle numbers, precluding the necessity of nested PCR (Baker, 2001).

For haplotype A restriction site Hae III (np663), primers L635 and H708 are used (Handt et al., 1996). The 50 μL PCR reaction contained PCR buffer [10 mM Tris (pH 8.3)], 1.7 mM $MgCl_2$, 200 μM each dNTP, 10 pmol of each primer, 2 units of AmpliTaq Gold (PE Applied Biosystems, Foster, CA) and 1 μL of extracted DNA. The DNA is amplified for 40 cycles in a Perkin Elmer 9600 programmable thermocycler. A 95°C 4 min. denaturation/enzyme activation step proceeded the cycling steps of 94°C for 1 min., 58°C for 1 min., and 72°C for 1 minute.

For the 9-bp deletion indicative of haplotype B, primers L8215 and H8297 are used (Wrischnik et al., 1987). The 25 μL PCR reaction contained PCR buffer [10 mM Tris (pH 8.3)], 1.7 mM $MgCl_2$, 200 μM each dNTP, 10 pmol of each primer, 2 units of AmpliTaq Gold (PE Applied Biosystems, Foster, CA) and 1 μL of extracted DNA. The DNA is amplified for 40 cycles in a Perkin Elmer 9600 programmable thermocycler. A 95° C 4 min. denaturation/enzyme activation step preceded the cycling steps of 94°C for 1 min., 52°C for 1 min., and 72°C for 1 minute.

Haplotype C, characterized by a HincII site at nt 13259, is amplified using primers L13257 and H13393 (Handt et al., 1996). The 50 μL PCR reaction contained PCR buffer [10 mM Tris (pH 8.3)], 1.7 mM $MgCl_2$, 200 μM each dNTP, 20 pmol of each primer, 2 units of AmpliTaq Gold (PE Applied Biosystems, Foster, CA) and 1 μL of extracted DNA. The DNA is amplified for 50 cycles in a Perkin Elmer 9600 programmable thermocycler. A 95°C 4 min. denaturation/enzyme activation step preceded the cycling steps of 94°C for 1 min., 50°C for 1 min., and 72°C for 1 minute.

The conditions for Haplotype D (AluI site nt 5176), primers L5127 and H5189, (Handt et al., 1996) and Haplotype X (DdeI site nt 1715), primers L1689 and H1728, (Green et al., 2000) are the same. The 50 μL PCR reaction contained PCR buffer [10 mM Tris (pH 8.3)], 1.9 mM $MgCl_2$, 200 μM each dNTP, 20 pmol of each primer, 2 units of AmpliTaq Gold

(PE Applied Biosystems, Foster, CA) and 1 μL of extracted DNA. The DNA is amplified for 50 cycles in a Perkin Elmer 9600 programmable thermocycler. A 95°C 4 min. denaturation/enzyme activation step preceded the cycling steps of 94°C for 1 min., 50°C for 1 min., and 72°C for 1 minute.

The PCR products are electrophoresed in a 12% acrylamide gel, stained with ethidium bromide and visualized on a UV transilluminator to estimate the DNA quantity and to check for PCR contamination. Positive controls for each haplotype are amplified at a different time to reduce the risk of contamination of the hair DNA samples. Subsequently, the samples are digested with the appropriate restriction enzyme (Hae III, HincII, AluI, or DdeI).

Restriction Digests

For restriction digests, 10-25 μL of DNA (depending on the amount of DNA obtained from the amplification) was incubated with 5 units of enzyme at 37°C for 20-24 hours and analyzed on a 12% acrylamide gel, stained with ethidium bromide and visualized on a UV transilluminator.

Mitochondrial DNA Hypervariable Region I Amplification

PCR is performed using four previously published primer sets that amplify 4 overlapping fragments encompassing 455 bp (15976-16430)

from hypervariable region I (Table 2-2). The PCR products range from 129 to 173 bases in length. Initially, larger amplicon products were attempted but for the most part they were unsuccessful due to the highly degraded nature of the DNA.

Table 2-2: Mitochondrial DNA hypervariable region I (HVR-I) primers

Set	Position	Primer Sequence	Bp*	Reference
1	L15996 H16139	CTGTCCTTGTAGTATAAACT ATACTTGACCACCTGTAGTA	173	This study Handt et al. 1996
2	L16131 H16218	CACCATGAATATTGTACGGT CAACCCTCAACTATCACACA	129	Handt et al. 1996 Handt et al. 1996
3	L16209 H16303	CCCCATGCTTACAAGCAAGT GTACATAGTACATAAAGCCA	134	Handt et al. 1996 Handt et al. 1996
4	L16287 H16410	CACTAGGATACCAACAAACC CCGTGAAATCAATATCCCGC	163	Handt et al. 1996 Handt et al. 1996

*Number of base pairs of the amplified product

The 25 μL PCR reaction contained PCR buffer [10 mM Tris (pH 8.3)], 1.7 mM $MgCl_2$, 200 μM each dNTP, 10 pmol of each primer, 2 units of AmpliTaq Gold (PE Applied Biosystems, Foster, CA) and 1 μL of extracted DNA. The DNA is amplified for 60 cycles in a Perkin Elmer 9600 programmable thermocycler. A 95°C 4 min. denaturation/enzyme activation step preceded the cycling steps of 94°C for 1 min., 55°C for 1 min., and 72°C for 1 minute.

20 μL of the PCR products were electrophoresed in a 12% acrylamide gel, stained with ethidium bromide and visualized on a UV transilluminator. The amplified product is then gel purified using the

bandstab technique described by Wilton et al. (1998). Briefly, each target band is stabbed with a sterile, filtered pipette tip. The tip is then put into a microcentrifuge tube containing PCR mix (same concentrations as before). The solution is subjected to the same PCR procedures previously described for 20 cycles. The PCR product is purified using a QIAquick purification kit (QIAGEN, Valencia, CA) according to the manufacturer's specifications before sequencing.

PCR blanks containing all reagents except for DNA and the extraction blanks are included for every PCR to detect any contamination. If there is any sign of contamination in the PCR blanks, the samples are repeated from the beginning. If the extraction blanks show signs of contamination, new extractions are performed.

DNA Sequencing

Automated DNA sequencing was performed at the UTK Molecular Biology Resource Facility with the ABI Prism Dye Terminator Cycle Sequencing reaction kit on either an ABI 373 DNA sequencer or an ABI 310 Fragment Analyzer (Perkin-Elmer Inc., Foster City, CA). The initial sequence data text files were edited following comparison with the same data displayed in four-color electropherograms before they were analyzed further.

Each of the four segments of HVR I were sequenced for both the light (L) and heavy (H) strand for each individual. The light and the heavy strands were amplified in two separate reactions in order to look for nucleotide mis-incorporations due to the PCR process. If any discrepancies occurred, the sample was amplified and sequenced again. The use of a high fidelity Taq (AmpliTaq Gold) helped to reduce mis-incorporations. In addition, several samples were extracted and sequenced multiple times for verification of sequence data.

DNA Analysis

DNA analysis was performed using several different programs available free on the world wide web. ARLEQUIN© (Schneider et al., 2000) software for population genetics data analysis was used to perform an analysis of molecular variance (AMOVA) (Excoffier et al., 1992) and to determine mismatch distributions. The AMOVA was used to determine the significance of the differences among samples creating a distance matrix between samples. This analysis determines the genetic structure of the population from which the samples are drawn. There are few assumptions about the properties of the data. AMOVA can separate and test for genetic diversity among groups of populations, diversity among the populations within groups and diversity among the individuals within

a population. AMOVA uses permutational methods so no distributional form is assumed.

Additionally, MOMA (Molecular Online Mitochondrial Analysis) (Baker, 2001), an online searchable database, was used to store and evaluate variable nucleotide positions between individuals and groups.

Samples

Only Samples that could be completely sequenced and successfully amplified for haplotyping were included in the present study. Several other samples were either partially sequenced or partially haplotyped depending on successful amplification.

The Mississippi Choctaw individuals in the present study (Appendix B) included 6 females and 13 males all claiming full Choctaw blood quantum. Their ages range from 18-68 with a mean age of 32. These individuals were either from Newton (32%) or Neshoba (68%) counties in Mississippi. The Menominee sample (Appendix B) included 15 females and 2 males. All individuals except for one (Menominee 97) claimed full Menominee blood quantum. Menominee 97 claimed 0.5 Menominee blood quantum with her father as English therefore the mtDNA is Menominee. The ages range from 5-16 with a mean age of 9.5. Although all individuals participated in the study resided in Keshina, Wisconsin, their birthplaces were varied throughout Wisconsin.

Chapter 3: Analysis of five mitochondrial DNA markers

Five Primary MtDNA Lineage Clusters in Native Americans

To date, five primary founding groups for mitochondrial DNA lineages have been established for Native Americans both past and present (Wallace et al., 1985, Schurr et al., 1990; Torroni et al., 1992; Horai et al., 1993; Scozzari et al., 1997; Brown et al. 1998; Stone and Stoneking, 1998; Kaestle and Smith, 2001). Haplogroups are determined by the presence or absence of particular polymorphisms in the coding region of mtDNA. For Native Americans these haplogroups are called A, B, C, D and X (Table 3-1).

Haplogroup A is determined by a Hae III restriction site gain at nt 663, and Haplogroup B is determined by a 9-bp deletion in region V of the mtDNA genome. Haplogroup C is ascertained by a Hinc II restriction site loss at nt 13259, an Alu I site gain at nt 10397, a Dde I site gain at 13394 and an Alu I site gain at 13161. Haplogroup D has an Alu I restriction site loss at nt 5176 and Haplogroup X is discriminated by a Dde I site loss at nt 1715 and an Acc I site gain at nt 14465. In the present analysis a subset of these markers are used: Hae III nt 663 (Haplogroup A), the 9-bp deletion (Haplogroup B), Hinc II nt 13259 (Haplogroup C), Alu I nt 5176 (Haplogroup D), and Dde I nt 1715 (Haplogroup X).

Table 3-1: Choctaw and Menominee haplogroup classifications

	Haplotype Group							Haplotype Group					
	A	**B**	**C**	**D**	**X**	**Other**		**A**	**B**	**C**	**D**	**X**	**Other**
Choc1	--	--	--	--	--	Yes	Menom14	A	.	.	.	.	.
Choc10	A	.	.	.	.	.	Menom33	A	.	.	.	.	.
Choc16	A	.	.	.	.	.	Menom57	A	.	.	.	.	.
Choc22	A	.	.	.	.	.	Menom95	A	.	.	.	.	.
Choc28	--	--	--	--	--	Yes	Menom97	A	.	.	.	.	.
Choc30	--	--	--	--	--	Yes	Menom98	A	.	.	.	.	.
Choc41	--	--	--	--	--	Yes	Menom111	A	.	.	.	.	.
Choc46	A	.	.	.	.	.	Menom113	A	.	.	.	.	.
Choc48	--	--	--	--	--	Yes	Menom114	A	.	.	.	.	.
Choc57	A	.	.	.	.	.	Menom118	A	.	.	.	.	.
Choc64	A	.	.	.	.	.	Menom124	--	--	--	--	--	Yes
Choc68	.	.	C	.	.	.	Menom128	A	.	.	.	.	.
Choc70	.	.	C	.	.	.	Menom131	.	.	C	.	.	.
Choc85	A	.	.	.	.	.	Menom139	A	.	.	.	.	.
Choc90	A	.	.	.	.	.	Menom151	A	.	.	.	.	.
Choc137	A	.	.	.	.	.	Menom154	A	.	.	.	.	.
Choc139	A	.	.	.	.	.	Menom155	A	.	.	.	.	.
Choc148	A	.	.	.	.	.							
Choc152	A	.	.	.	.	.							

Primer Evaluation

Of the primer sets used to amplify sections for haplotyping, those for haplogroup B and X were the most robust. The primers used for Haplogroup A often yielded the expected product along with extraneous unwanted bands; therefore, the annealing temperature had to be adjusted often and many samples were re-run to find an optimal temperature to alleviate the extra bands. Haplogroup C primers worked intermediately well when the primer concentration and $MgCl_2$ concentrations were doubled. Haplogroup D primers were the most difficult with which to obtain amplification products. Even when samples successfully amplified, the amounts of amplified product were low.

The Choctaw and Menominee

The individuals included in the analysis were sampled opportunistically for the World Columbian Exposition and are not a random sample of individuals. The Mississippi Choctaw individuals in the present study (Appendix A) included 6 females and 13 males all claiming full Choctaw blood quantum. Their ages range from 18-68 with a mean age of 32. These individuals were either from Newton (32%) or Neshoba (68%) counties in Mississippi.

The Menominee sample (Appendix A) included 15 females and 2 males. All individuals except for one (Menominee 97) claimed full Menominee blood quantum. The ages range from 5-16 with a mean age of 9.5. Although all individuals participated in the study resided in Keshina, Wisconsin, their birthplaces were varied throughout Wisconsin.

Haplogroups Mississippi Choctaw

Only two of the five haplogroups were found in the Mississippi Choctaw individuals (Table 3-1). Twelve of the 19 Choctaw individuals (63%) had the Hae III nt 663 site gain (Haplogroup A). Two individuals (11%) had the Hinc II 13259 nt site loss (Haplogroup C). Five individuals (26%) did not possess any of the 5 screened polymorphisms that would classify them into one of the accepted Native American haplogroups and were excluded from further statistical haplogroup analysis. None of the Choctaw individuals included in the present study possessed the characteristic mutations of Haplogroups B, D, or X. The Mississippi Choctaw have the lowest heterozygosity, 0.260, of all of the southeastern groups (Table 3-2).

Table 3-2: Haplogroup frequencies

Population	Language Family	n	A	B	C	D	X	h	References
Menominee	Algonquian	16	0.938	0.000	0.063	0.000	0.000	0.125	This Study
Micmac	Algonquian	6	0.333	0.000	0.167	0.000	0.500	0.733	Lorenz and Smith, 1996; Smith et al, 1999; Malhi et al., 2001
Island Ojibwa	Algonquian	33	0.325	0.097	0.269	0.040	0.269	0.754	Scozzari et al., 1997; Torroni et al., 1993
N. Ontario Ojibwa	Algonquian	26	0.643	0.036	0.071	0.000	0.250	0.538	Scozzari et al., 1997; Torroni et al., 1993
Turtle Mt Chippewa	Algonquian	28	0.571	0.179	0.179	0.000	0.071	0.627	Malhi et al., 2001
WI Chippewa	Algonquian	62	0.275	0.048	0.355	0.032	0.290	0.723	Malhi et al., 2001
Cheyenne/Arapaho	Algonquian	35	0.343	0.114	0.343	0.143	0.057	0.749	Lorenz and Smith, 1996; Smith et al, 1999; Malhi et al., 2001
MS Choctaw	Muskogean	14	0.857	0.000	0.143	0.000	0.000	0.260	This Study
OK Choctaw	Muskogean	26	0.732	0.192	0.038	0.038	0.000	0.443	Lorenz and Smith, 1996; Weiss and Smith, submitted
Chickasaw	Muskogean	8	0.125	0.750	0.125	0.000	0.000	0.464	Lorenz and Smith, 1996; Weiss and Smith, submitted
Creek	Muskogean	35	0.371	0.114	0.229	0.286	0.000	0.736	Lorenz and Smith, 1996; Weiss and Smith, submitted
Muskoke	Muskogean	71	0.366	0.155	0.099	0.380	0.000	0.697	Merriwether and Ferrell, 1996
Seminole	Muskogean	38	0.605	0.263	0.079	0.053	0.000	0.566	Huoponen et al., 1997; Weiss and Smith, submitted
Combined Choctaw	Muskogean	40	0.775	0.125	0.075	0.025	0.000	0.387	Weiss and Smith, submitted; This Study

The low heterozygosity of the Choctaw can be attributed to several factors. The Mississippi Choctaw lost more than four-fifths of its population in 1831 when the population was divided and most individuals were forcibly removed to the Indian Territory in Oklahoma. A reduction of this size could cause this reduction in haplogroup diversity and even the complete absence of three of the five haplogroups. Additionally, sampling error could play a role in the amount of diversity observed in the sample. Examination of the population data collected indicates that the individuals included were the ones that not only agreed to participate in the anthropometric portion of the study, but also agreed to have a portion of their hair collected. Out of the roughly 1,600 Choctaw living in Mississippi only 329 participated in the anthropometric study and of those only 152 contributed samples of their hair and of these only 19 of these are included in this study. With such a small percentage of the population studied (1.2%) it is difficult to know if this is a representative sample and there is surely sampling bias.

Comparison with Oklahoma Choctaw and other Muskogean Groups

In order to evaluate the haplogroup diversity of the Mississippi Choctaw from the Boas collection (1893), the results were compared with five other populations (Figure 3-1). The Chickasaw, Choctaw, Creek, Muskoke, and Seminole populations are all members of the Muskogean

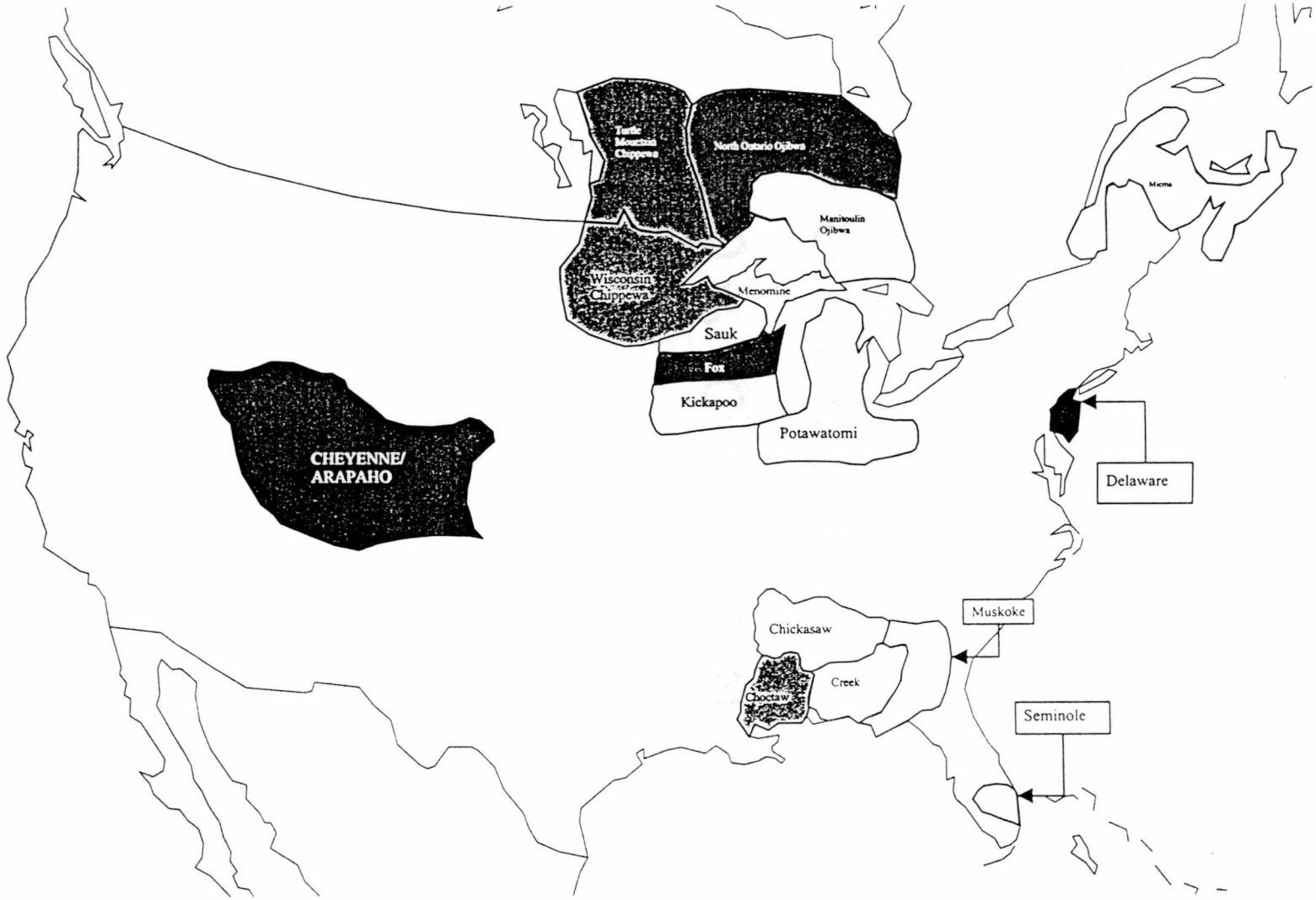

Figure 3-1: Map illustrating the areas occupied by each population.

language family that have been analyzed for DNA haplogroups in modern samples (Merriwether and Ferrell, 1996; Lorenz and Smith, 1996; Houponen et al., 1997; Weiss and Smith, submitted 2001). Each group inhabited the southeastern United States prior to forced relocation to Oklahoma in the 1830s.

The haplogroup frequencies of the Choctaw represented in the Boas collection were compared to those of the modern Choctaw residing in Oklahoma (Lorenz and Smith, 1996; Weiss and Smith, submitted 2001) (Table 6). For both of the Choctaw samples, A is the most common haplogroup, X is absent, and B, C, and D are at low frequencies or absent altogether. An analysis of molecular variance (AMOVA) was performed using ARLEQUIN (Schnieder, 2000) to see if the haplogroups of the two groups differed significantly from one another. The results showed an among group variation of 3% and a within group variation of 97%; therefore, only 3% of the variation is accounted for by differences between the two Choctaw groups. The Fst value of 0.0297 is not significant at the 0.05 level so the two groups are not significantly different from one another. Although the two groups are separated in time by more than 100 years, that has not been enough time to change their mtDNA haplogroup frequencies enough to identify them as separate populations. It also shows that the haplogroup frequencies of the Mississippi Choctaw are similar to those of the Oklahoma Choctaw

despite the population size differential due to removal. For all subsequent haplogroup analyses the Mississippi Choctaw and the Oklahoma Choctaw are combined as a single group of Choctaw. The analysis shows that the haplogroup diversity seen in the Mississippi Choctaw is similar to that of the Oklahoma Choctaw and may measure the true amount of diversity for this population.

The haplogroup frequencies of the Mississippi Choctaw combined with the Oklahoma Choctaw were then compared with the haplogroup frequencies of the Chickasaw, Creek and Seminoles (Table 3-2). Haplogroup frequencies vary greatly from group to group among these Muskogean speakers. Haplogroup A is the most frequent haplogroup for the Seminole, Choctaw, and the Creek, but for the Chickasaw haplogroup B has the highest frequency at 0.750. For the Muskoke the most frequent haplogroup is divided between A, 0.366 and D, 0.380. AMOVA results show that most of the Muskogean variation (87%) can be attributed to variation within each of the groups, while 13% of the variation is due to differences among the populations. A significant Fst value of 0.1273 was obtained indicating a substantial degree of divergence between the tribes within the southeastern Muskogean speakers.

Menominee Haplogroups

Haplogroup analysis of 17 Wisconsin Menominee represented in the Boas collection was also performed. Only two of the five haplogroups were represented in this sample of 17 individuals (Table 5). Fifteen of the 17 Menominee were haplogroup A, one individual was haplogroup C and one could not be typed as A, B, C, D or X. The single individual that was not typed as being part of one of the characteristic haplogroups was excluded from further statistical haplogroup analysis.

The Menominee haplogroup frequencies were compared with the previously reported Micmac, Island Ojibwa, Northern Ontario Ojibwa, Turtle Mountain Chippewa, Wisconsin Chippewa and the Cheyenne/Arapaho (Lorenz and Smith, 1996; Smith et al., 1999; Malhi et al., 2000; Scozzari et al., 1997), all representatives of the Algonquian language family (Figure 3-1). The Menominee have the lowest heterozygosity (0.125) of all of the groups analyzed, which together have an average heterozygosity of 0.6873 (Table 3-2).

An AMOVA performed on the haplogroup frequencies for Algonquian groups showed that 91% of the variation of these groups occurs within the population, and 9% occurs between populations. In addition, a significant Fst value of 0.089 was obtained, indicating that the populations are somewhat divergent.

Haplogroup Comparison of Muskogean and Algonquian Groups

A comparison was done between the above mentioned Muskogean and Algonquian groups to determine if the amount of among group variation is greater between language families (Figure 3-1). Once again, the majority of variation comes within the populations (86%). Four percent of the variation observed can be explained by among group differences and another 10% can be explained by among group variation within the groups.

Chapter 4: Sequence Analysis of Hypervariable Region I

Significance of Hypervariable Region I

Mitochondrial DNA (mtDNA) variation has been extensively explored in an attempt to better understand the history of human populations (Di Rienzo and Wilson, 1991; Horai et al, 1996; Richards et al., 1996; Lum et al., 1994; Vigilant et al., 1989). The high mutation rate of the hypervariable regions has made this section of mtDNA extremely useful in this effort. Many populations from around the world have been sequenced, primarily for hypervariable region I (HVR-I), in an attempt to address issues concerning population origins, expansions, bottlenecks, migrations and divergence times.

The Choctaw and the Menominee hair shaft samples from the Boas collection were sequenced for HVR-I in order to determine if archived hair shafts would provide an adequate source of mtDNA for analysis. For the present study, 410 base pairs (16000-16409) were analyzed for the Choctaw and Menominee samples. These sequences were then compared to the sequences of several other Native American groups belonging to the same language families in order to address broader questions dealing with population history. In the comparisons of the Choctaw and Menominee to other groups, a total of 346 bases were

analyzed (16064-16049) in order to have the data correspond with previously published data from other laboratories.

General Sequencing Results

A HVR-I section of 455 base pairs in length (15976-16430) was sequenced in four overlapping segments for each sample and 410 base pairs (16000-16409) were analyzed. Each segment was sequenced in both directions from two independent amplifications in order to identify PCR artifacts and eliminate any ambiguity from the analysis. There were many sequences that had to be repeated in order to clarify discrepancies between the light and heavy strand sequences. This occurred more often when larger amplicons were attempted; therefore, the four overlapping segments described were the final choice for the project.

The sample sequences were compared to that of the author as an additional check for contamination. The author has sequence variation at 16069 and 16126 and this combination was not seen in any of the sequences from the Choctaw or the Menominee in the sample. However, in the development of the extraction protocol, one sample was contaminated by the author and additional precautions were added to try to eliminate this in the future. All sequencing was performed prior to haplotype analysis to evaluate the possibility of contamination by the author.

Analysis of Choctaw Sequences

The HVR-I section (16000-16409) was analyzed for 19 Choctaw individuals. Each individual belonged to a unique mtDNA lineage, meaning that none of the sequences were shared between individuals. In other words, each individual represents a different maternal line in the population. Despite the lack of haplogroup diversity in the Choctaw sampled, there is a great amount of sequence diversity. The five individuals (Choctaw 28, 30, 41 and 48) that could not be classified as belonging to haplogroups A, B, C, D or X were included in the sequencing results despite their lack of Native American haplogroup classification.

The nucleotide diversity is 0.013 +/- 0.008, with a mean number of pairwise differences of 5.470 +/- 2.754. There were 22 observed transitions and 2 transversions, leading to 24 substitutions in 24 polymorphic sites. No sites share more than two different nucleotides.

The mismatch distribution of the number of pairwise differences within the Choctaw was calculated by Arlequin 2.000 and plotted (Figure 4-1). Harpending's raggedness index of 0.031 is low and the Choctaw's mismatch has a unimodal distribution indicative of a recent expansion sometime during the Pleistocene.

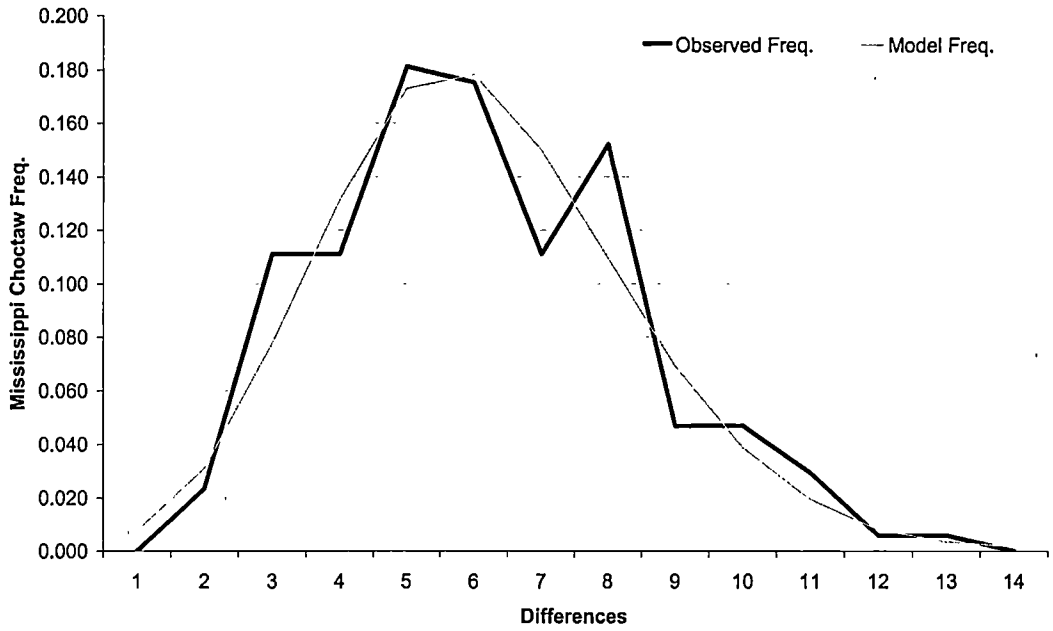

Figure 4-1: Mismatch distribution for the Mississippi Choctaw sample.
The X-axis is the pairwise differences in units of mutational time. The Y-axis is the frequency of each pairwise difference.

Comparison of Mississippi Choctaw to Oklahoma Choctaw

The sequences from these 19 Mississippi Choctaw were then compared to the sequences from 16 modern Oklahoma Choctaw (Weiss and Smith, submitted 2001). The two groups now living in two different states were part of the same group prior to the 1830s. In order to compare the sequences from the two groups, 346 bases (16064-16409) were chosen because this section was sequenced in both populations.

The individuals represented in the two groups do not share any mtDNA lineages. When the Mississippi Choctaw data is reduced to the 346 base pair comparison, Choctaw 41 and Choctaw 90 become the same sequence. These two sequences are differentiated from one another by a mutation at 16051 A-G (Anderson et al., 1981); however, because this lies out of the reduced sequence for comparison, these individuals become indistinguishable. The Oklahoma Choctaw have seven distinct mtDNA lineages that are shared by the sixteen individuals. One of the seven lineages is shared by six individuals, two of the lineages each have three individuals of the same type, and the remaining four lineages are represented by only one individual each.

An analysis of molecular variance was performed (AMOVA) using Arlequin 2.000 (Schnieder and Excoffier, 2000). Distances were computed using the Tamura-Nei model (Tamura and Nei, 1993) and the gamma distribution parameter alpha (a) was set to a = 0.4 (Excoffier,

1999). The results show that the sequences of the Mississippi Choctaw are significantly different from the Oklahoma Choctaw with an Fst of 0.200. Only 80% of the sequence variation is explained by within population differences while 20% of the variation occurs among the populations. For the subsequent analyses involving other Muskogean groups, the Mississippi Choctaw and the Oklahoma Choctaw are not combined due to their significant variations.

The mismatch distribution was also calculated for the Oklahoma Choctaw for comparison with the Mississippi Choctaw. Harpending's raggedness index for the Oklahoma Choctaw was 0.097, almost twice as much as that of the Mississippi Choctaw. The graphical representation of the distribution is bimodal and ragged (Figure 4-2).

Sequence Comparison with other Muskogean Groups

The Choctaw sequences were then compared on a larger scale by including other southeastern groups also belonging to the Muskogean language family. Sequences from 10 Chickasaw, 18 Creek, and the 16 Oklahoma Choctaw (Weiss and Smith, submitted 2001) were analyzed with the 19 Mississippi Choctaw sequences.

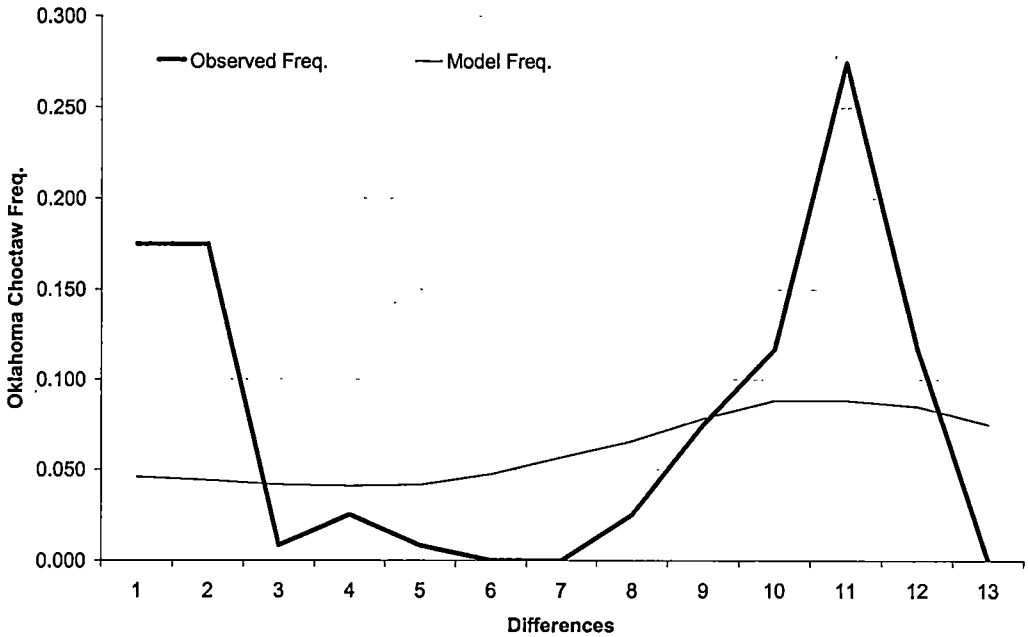

Figure 4-2: Mismatch distribution for the Oklahoma Choctaw sample.
The X-axis is the pairwise differences in units of mutational time. The Y-axis is the frequency of each pairwise difference.

There are 20 different sequence haplotypes for the Chickasaw, Creek and Oklahoma Choctaw. The 10 Chickasaw belong to seven different mtDNA lineages (or haplotypes) and the 18 Creek belonged to 12 different haplotypes. The 16 Oklahoma Choctaw described earlier represent seven different maternal haplotypes. Of these, the Oklahoma Choctaw and the Chickasaw share three haplotypes, the Chickasaw and the Creek share one and all three groups share one haplotype. The Mississippi Choctaw did not share any sequence haplotypes with any of the other Muskogean groups.

The results of an AMOVA for the four groups show that there is substantial divergence from one another. 84% of the variation occurs within the populations themselves but 16% of the variation is among the populations. A significant Fst value of 0.159 was obtained. Additionally, the Chickasaw mismatch distribution was plotted (Figure 4-3) and is multi-modal and very ragged as indicated by Harpending's raggedness index of 0.111. The Creek's raggedness index is lower (0.040) and the mismatch distribution is essentially bi-modal (Figure 4-4).

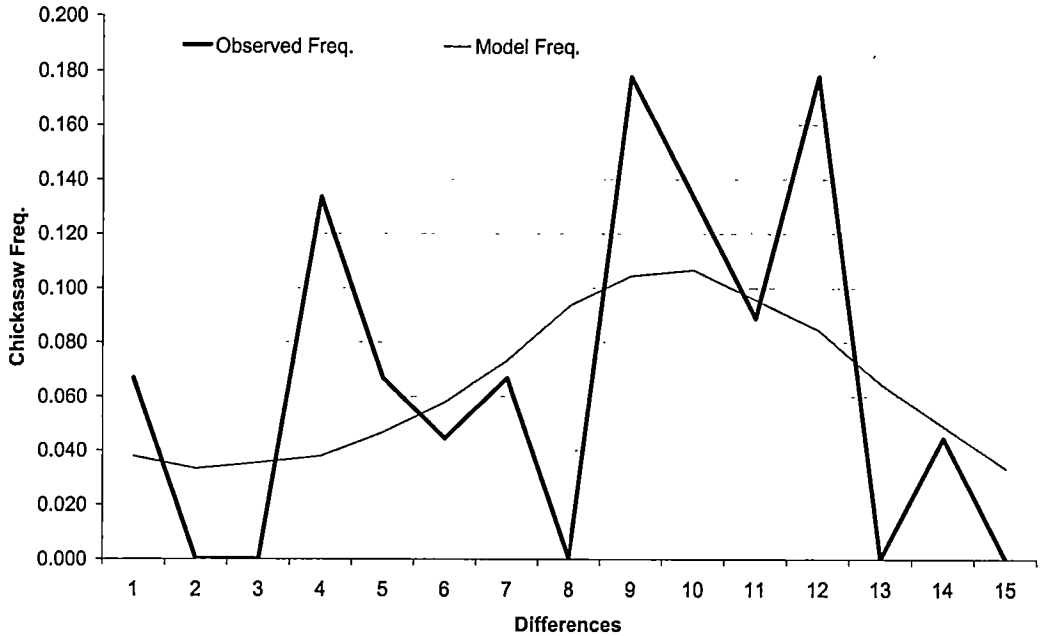

Figure 4-3: Mismatch distribution for the Chickasaw sample.
The X-axis is the pairwise differences in units of mutational time. The Y-axis is the frequency of each pairwise difference.

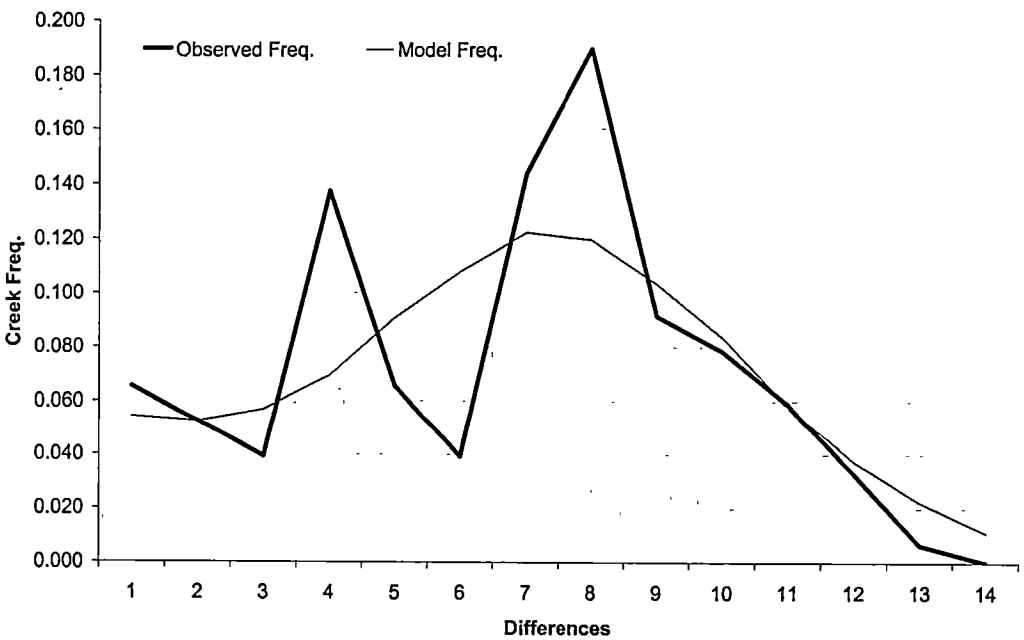

Figure 4-4: Mismatch distribution for the Creek sample.
The X-axis is the pairwise differences in units of mutational time. The Y-axis is the frequency of each pairwise difference.

Analysis of Menominee Sequences

A 410 bp section of HVR-I (16000-16409) was analyzed for 17 Menominee individuals. There were 14 unique haplotypes out of the 17 sequences compared. Despite the lack of haplogroup diversity in the Menominee sampled, there are many different sequence lineages. Menominee 33 and Menominee 131 share a common sequence, Menominee 95 and 97 and Menominee 111 and Menominee 139. It is interesting that Menominee 33 and 131 share identical sequences because they belong to two different haplogroups. Menominee 33 has the characteristic HaeIII nt 663 site gain indicative of haplogroup A, and Menominee 131 has the Hinc II nt 13259 site loss representative of haplogroup C. They both share sequence changes at 16111 (C-T), 16290 (C-T), 16319 (G-A) and 16362(T-C). These polymorphisms are characteristic of those typically found in individuals belonging to haplogroup A. The single individual, Menominee 124, that could not be classified as belonging to haplogroups A, B, C, D or X was included in the sequencing results despite the lack of haplogroup classification.

The nucleotide diversity for the Menominee was 0.010 +/- 0.006 with a mean pairwise difference of 4.267 +/- 2.224. There were 18 observed transitions and 1 transversions for a of total 19 substitutions at 19 polymorphic sites. No sites share more than two different nucleotides.

The mismatch distribution of the number of pairwise differences within the Menominee was calculated by Arlequin 2.000 and plotted (Figure 4-5). Harpending's raggedness index of 0.024 is very low and the Menominee's mismatch distribution is unimodal, which is indicative of a Pleistocene population expansion.

Sequence Comparison with other Algonquian Groups

A 346 bp section of HVR-I (16064-16409) was compared between the Menominee sequences and individuals from 8 different previously studied Algonquian groups: Chippewa (n = 29), Micmac (n = 4), Delaware (n = 1), Potawatomi (n = 1), Kickapoo (n = 5), Fox (n = 1), Cheyenne, and Arapaho combined (n = 4) (Malhi et al., 2001). Some populations were combined to form a single group because several were represented by a very small number of individuals. The Micmac and Delaware groups were combined, as were the Potawatomi, Kickapoo and Fox individuals and the Cheyenne and Arapaho individuals to form three different amalgamated groups.

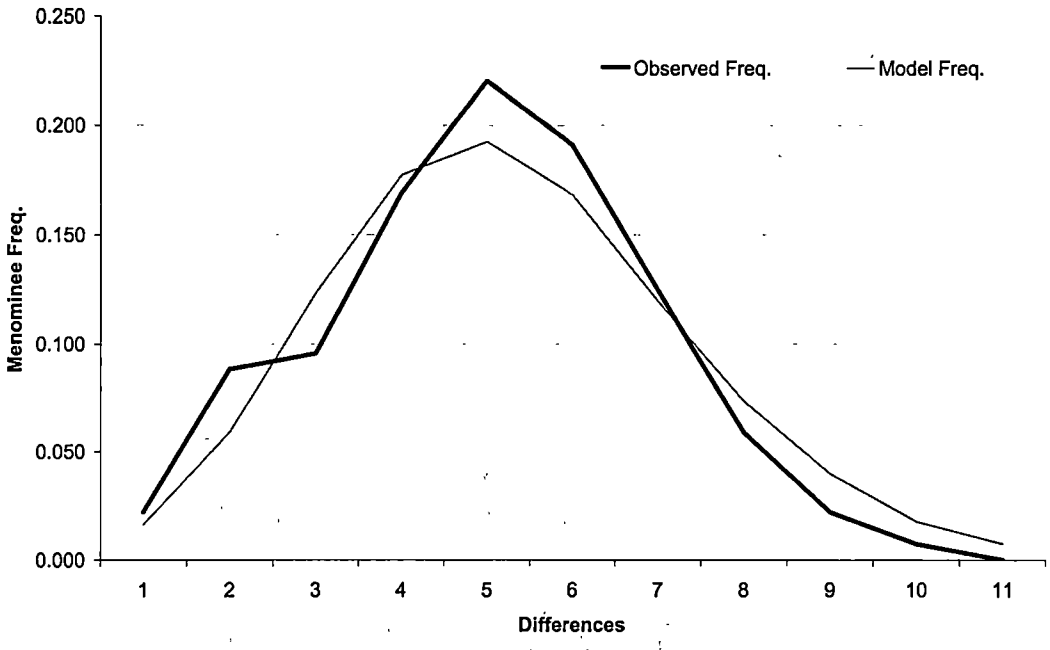

Figure 4-5: Mismatch distribution for the Menominee sample.
The X-axis is the pairwise differences in units of mutational time. The Y-axis is the frequency of each pairwise difference.

There are 36 different sequence haplotypes for the five groups including the Menominee. The 29 Chippewa have 16 different sequence haplotypes. Of these, nine individuals share one, each of which were RFLP typed as haplogroup X. The next most common sequence haplotype is shared by five individuals, also all typed as haplogroup X. There is only one other haplotype that is common among the Chippewa and it is shared by two individuals. The Chippewa have six sequence haplotypes in common with other groups. One sequence type is shared by at least one member of each of the groups.

The results of an AMOVA for the five Algonquian groups show that there is substantial divergence. Only 69% of the variation occurs within the populations themselves and 31% of the total variance is accounted for by variation among the populations. A significant Fst value of 0.308 was obtained.

The mismatch distributions for all groups were calculated and plotted. Unlike the unimodal distribution of the Menominee, the distributions of the other groups are more bi-modal and multi-modal. While the Menominee had a raggedness index of 0.024, the Chippewa had a lower index of 0.025 and a bimodal distribution (Figure 4-6). The Kickapoo/Potawatomi/Fox group had a raggedness index of 0.052 and an essentially bimodal distribution (Figure 4-7). The Micmac/Delaware

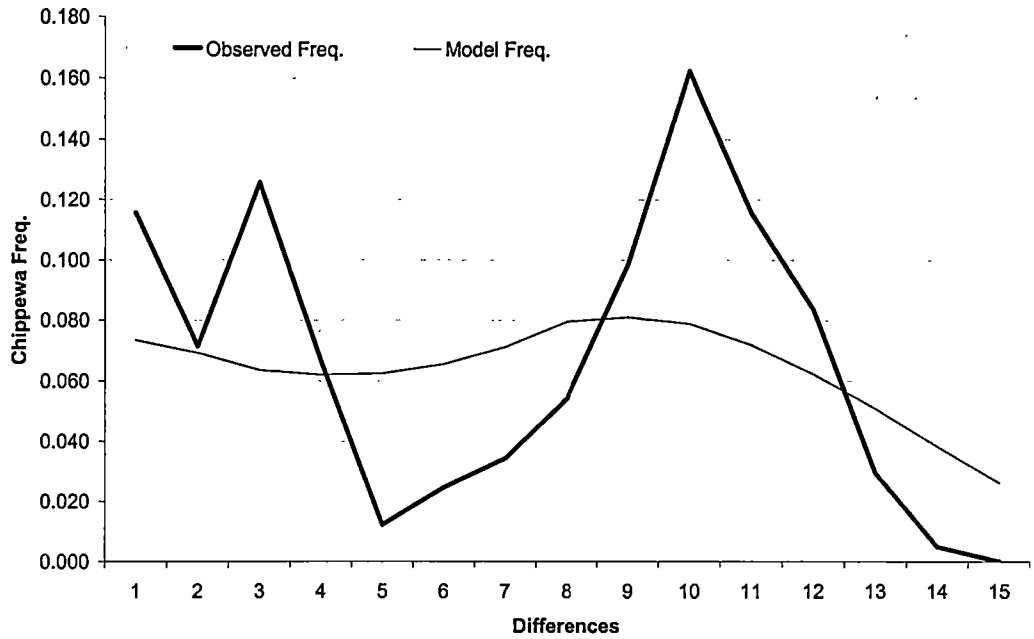

Figure 4-6: Mismatch distribution for the Chippewa sample.
The X-axis is the pairwise differences in units of mutational time. The Y-axis is the frequency of each pairwise difference.

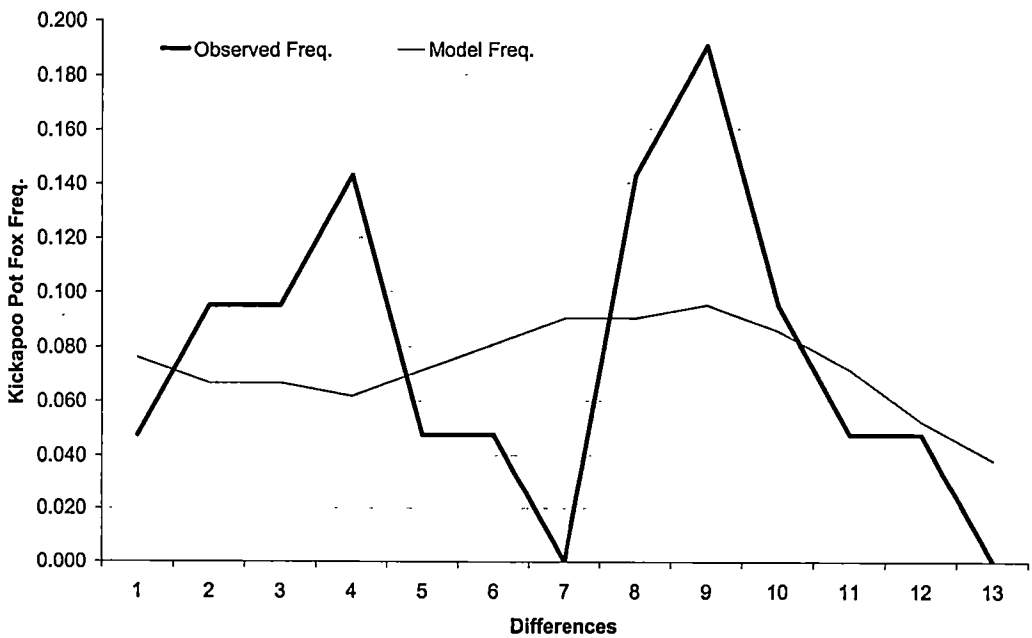

Figure 4-7: Mismatch distribution for the Kickapoo/Potawatomi/Fox sample.
The X-axis is the pairwise differences in units of mutational time. The Y-axis is the frequency of each pairwise difference.

group had a raggedness index of 0.1500 and a multi-modal distribution (Figure 4-8). The Cheyenne/Arapaho group had the highest raggedness index of 0.3333 and has two very distinct peaks in its mismatch distribution (Figure 4-9).

Comparison of the Muskogean and the Algonquian Groups

All of the groups in the analysis were collapsed into their respective language family groups of either Muskogean or Algonquian. For the Muskogean groups there were a total of 38 different sequence haplotypes for the 63 individuals represented, and for the Algonquian groups there were a total of 36 sequence haplotypes for the 62 individuals represented. The Muskogean groups and the Algonquian groups share 2 haplotypes, leading to a total of 72 different sequence haplotypes that are described in the present study. One of the shared haplotypes occurred in six Muskogean samples and seven Algonquian samples. All of the individuals for this sequence haplotype were previously characterized as belonging to haplogroup A by restriction digest analysis except for Menominee 131 and all share base changes 16111 (C-T), 16223 (C-T), 16290 (C-T), 16319 (G-A) and 16362 (T-C). The other shared haplotype is found in seven Muskogean and only one Algonquian, all described as belonging to haplogroup C by restriction digest and sharing base changes of 16223 (C-T), 16298 (T-C), 16325 (T-C) and 16327 (C-T). The mean

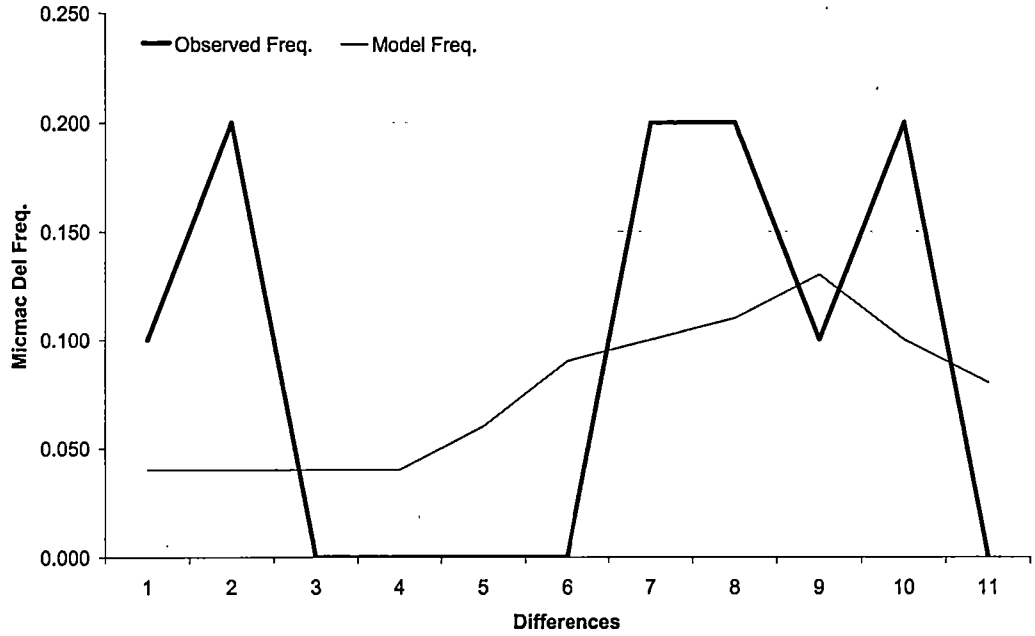

Figure 4-8: Mismatch distribution for the Micmac/Delaware sample.
The X-axis is the pairwise differences in units of mutational time. The Y-axis is the frequency of each pairwise difference.

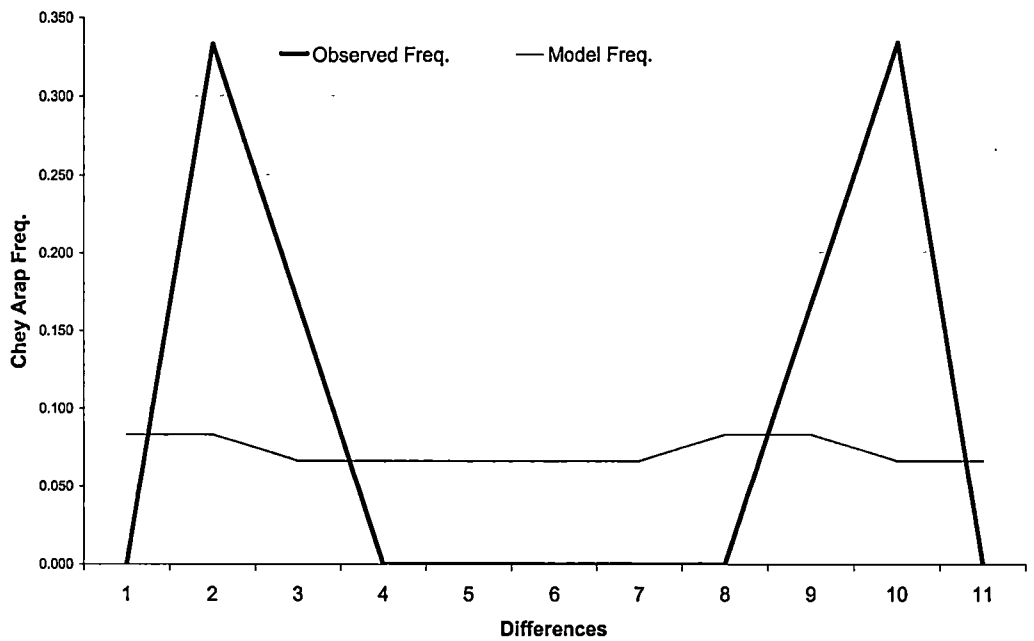

Figure 4-9: Mismatch distribution for the Cheyenne/Arapaho sample.
The X-axis is the pairwise differences in units of mutational time. The Y-axis is the frequency of each pairwise difference.

number of pairwise differences for the Muskogean group is 7.411 +/- 3.512 and the mean number of pairwise differences for the Algonquian group is 7.759 +/- 3.663.

An AMOVA was performed to determine the amount of variation that occurs between language families. The amount of variation that can be attributed to the differences between the Muskogean and the Algonquian groups represented in this study is only 6%. It is interesting that the lowest amount of among population variation is seen in this comparison. The rest of the variation, 94%, is due to differences within the populations. A significant Fst of 0.060 was obtained.

Although the actual sequence variation in the form of haplotypes is different for the two groups, the amount of variation is similar for the groups. The Muskogean group had a total of 47 substitutions including 43 transitions and 4 transversions. The Algonquian group had a total of 46 substitutions including 42 transitions and 4 transversions.

The mismatch distributions and Harpending's raggedness index for the two language groups were calculated and compared. The raggedness values for the two groups are very similar and both are low; 0.009 for the Muskogean group and 0.008 for the Algonquian group. The mismatch distribution for the Muskogean group is unimodal (Figure 4-10) and the distribution for the Algonquian group is slightly bimodal (Figure 4-11).

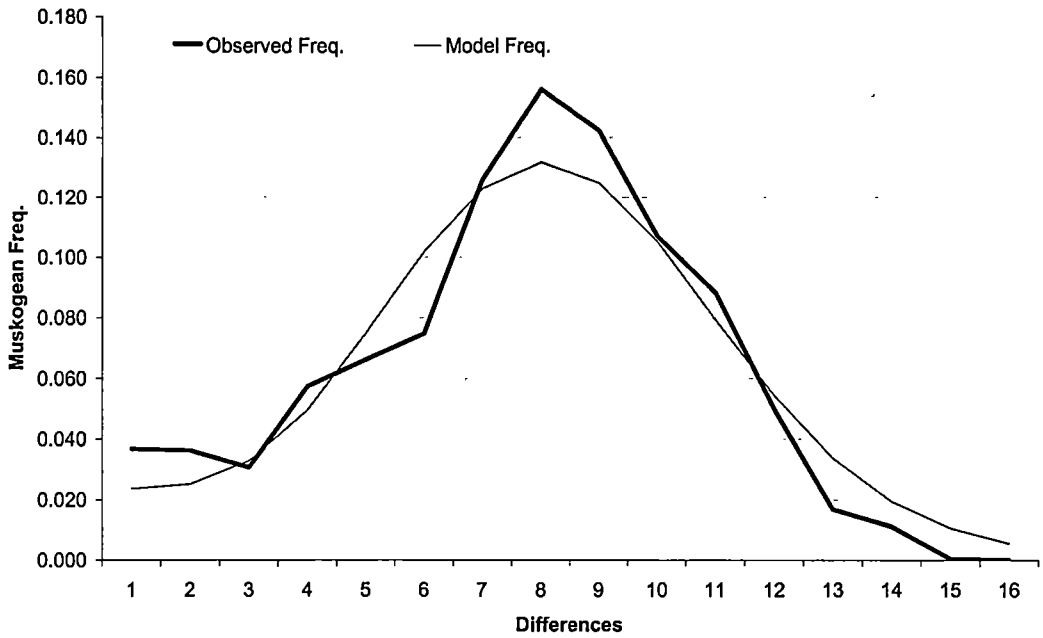

Figure 4-10: Mismatch distribution for the Muskogean sample.
The X-axis is the pairwise differences in units of mutational time. The Y-axis is the frequency of each pairwise difference.

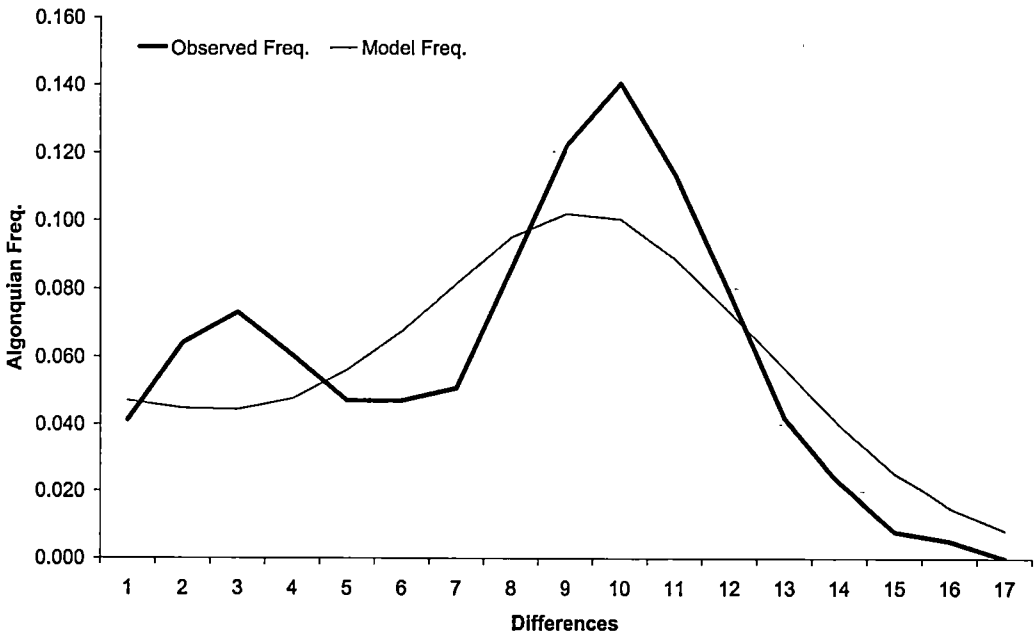

Figure 4-11: Mismatch distribution for the Algonquian sample.
The X-axis is the pairwise differences in units of mutational time. The Y-axis is the frequency of each pairwise difference.

This indicates that the groups as a whole have gone through a recent expansion (Rogers, 1995; Rogers and Harpending, 1993; Harpending et al, 1993; Sherry et al., 1994).

Chapter 5: Discussion and Conclusions

Discussion

In an attempt to better understand the genetic diversity of the Mississippi Choctaw and the Wisconsin Menominee, the research presented herein analyzes Native American hair samples collected under the supervision of Franz Boas. This is the first study to examine population differentiation by use of DNA from historic hair shaft samples, and it includes the analysis of both mtDNA haplotype and control region sequence data.

The specific objectives of this research are threefold. First, to develop procedures to successfully extract and analyze mitochondrial DNA from historic hair shaft samples. Secondly, to show that analyzed DNA from aged hair shafts can be used to address questions of the genetic diversity in distinct human populations at a fixed point in time. Lastly, this research is an attempt to describe the utility of archived collections, such as the Boas collection, to fill in temporal and geographical gaps in the genetic history of Native American groups.

MtDNA Haplotyping and Sequencing

Choctaw

The Mississippi Choctaw represented in this sample only have individuals from two of the five Native American haplogroups, resulting in low heterozygosity (0.260). There are several possible reasons why 3 of the 5 haplogroups are not present in the Mississippi Choctaw sampled. First, sampling error may play a role in the described diversity. Of the 1,600 Choctaw individuals residing in Mississippi in 1893 only 21% participated in the anthropometric study, and only 1.2% of them were represented in the present study. These issues are compounded by the fact that only Choctaw individuals who claimed full Choctaw ancestry were considered for inclusion in the present study. Because neither the original sample collection process nor those employed in this genetic study adhered to a true random sampling method, it is difficult to know if the present study is a measure of the true diversity of the Choctaw.

In addition to sampling error, the original Choctaw population underwent substantial changes prior to the acquisition of data for the Columbian Exposition. The Mississippi Choctaw lost more than four-fifths of its population in 1831 due to a forced relocation that left only a small number of Choctaw living in their ancestral homeland of Mississippi. It is conceivable that the reduction in haplogroup diversity

is due to this reduction in population size. In order to address this possibility, the Mississippi Choctaw were compared to a modern sample of Oklahoma Choctaw (Lorenz and Smith, 1996; Weiss and Smith, submitted 2001) (Table 3-2). The Oklahoma sample contains 4 of the 5 haplogroups, with a heterozygosity of 0.443. Although the heterozygosity of the Oklahoma sample is quite a bit higher than that of the Mississippi group, with the representation of four of the five haplogroups, the frequencies of haplogroups C and D for the Oklahoma sample are very low. The heterogeneity of the groups was further assessed by means of analysis of molecular variance (AMOVA) (Excoffier et al., 1992). The sums of squared sequence differences between individuals are compared within populations and between populations. Significance is calculated by a nonparametric randomization approach for the observed values. Only 3% of the total variation within the Choctaw is due to differences between the two groups. The results show that both groups of Choctaw individuals, although separated by over one hundred years and over 1,000 miles, can still be considered a single population. The effects of the population division and years of separation can not yet be appreciated by mtDNA haplogroup analysis.

Because the mutation rate is slower for the coding region of mtDNA than that of the hypervariable region, a separate analysis of the two Choctaw groups was performed to determine if the amount of

sequence variation is the same for the two data sets. Sequence variation from HVR-I was compared for the 19 Mississippi Choctaw from the Boas collection and 16 modern Oklahoma Choctaw (Weiss and Smith, submitted 2001). The two Choctaw groups were significantly different when sequences of HVR-I were compared. In contrast to the 3% difference between the groups for haplogroup data, there is a 20% difference between the group's sequence data.

Of the 19 Mississippi Choctaw, there were five samples that could not be classified as A, B, C, D or X. These five samples account for 26% of the individuals in the analysis. There are two reasons that non-A, B, C, D or X haplogroups exists in a Native American population. The first reason may be that these low frequency haplotypes are founding haplogroups, with distributions and frequencies in Native American populations small enough to preclude their acceptance as a standard haplogroup. The most likely reason, however, is the introduction of various haplotypes by non-Native American admixture. A Choctaw census from 1906 (Table 5-1) provides almost the same percentage (22%) of admixture as seen in the mtDNA data (26%) (McKee and Schlenker, 1980; Debo, 1961). These close admixture estimates suggest that the non-A, B, C, D or X samples may, in fact, be due to admixture. The same 1906 census estimates the admixture for the Oklahoma Choctaw as being 59% of the total population. Because most individuals only

Table 5-1: Choctaw Census, 1906

Oklahoma Choctaws		
Full bloods	7,067	
Mixed bloods, ¾ or more	706	
Mixed bloods, ½ to ¾	1,636	
Less than ½, including whites	9,563	
Total		18,981
Freedmen		5,994
Mississippi Choctaws		
Full bloods	1,344	
Mixed bloods, ¾ or more	85	
Mixed bloods, ½ to ¾	27	
Less than ½, including whites	183	
Total		1,639
Grand Total		26,614

recall information of ancestral makeup to three or four generations in the past, it is important to take past census data into consideration when examining modern populations.

Muskogean Comparisons

The Choctaw were compared to several other Muskogean groups in order to understand how the mtDNA variation of these historic individuals fit into the variation of southeastern Native American groups. For haplogroup frequency analysis, the Mississippi Choctaw (present study) were combined with the Oklahoma Choctaw (Weiss and Smith, submitted 2001) because of their lack of divergence. The Choctaw group (now n = 40) was compared to 8 Chickasaw, 35 Creek, 38 Seminole (Lorenz and Smith, 1996; Weiss and Smith, submitted 2001), and 71 Muskoke (Merriwether and Ferrell, 1996) individuals from each group. The haplogroup frequencies vary between the groups. The AMOVA results indicate that 13% of the Muskogean haplogroup frequency variation is due to among group differences.

Sequence variation of HVR-I in the 19 Mississippi Choctaw from the Boas collection was also compared to other Muskogean groups. The data for the Mississippi Choctaw showed a different number of lineages depending on how much of the data was included. In the independent

analysis of the Choctaw, sequence data from nt 16000 to nt 16409 was used and from this, all of the 19 samples comprised 19 separate lineages. In order to make the data correspond to the published data for other Muskogean groups, only data for nt 16064 to nt 16409 was compared. Of the 19 Mississippi Choctaw, there were 18 different mtDNA sequence patterns observed and none of these were shared by any of the individuals from the other populations. For sequence data analysis, the Mississippi Choctaw and the Oklahoma Choctaw were not combined because of their significant variation. The significant Fst value of 0.20 indicates that the sequences of the Mississippi Choctaw and the Oklahoma Choctaw are divergent. The Mississippi Choctaw sequence data was compared with 16 Oklahoma Choctaw, 18 Creek and 10 Chickasaw individuals (Weiss and Smith, submitted 2001). Each group had relatively similar nucleotide diversity around 0.02. The population history of each group was examined by plotting the mismatch distributions for each group separately. The Mississippi Choctaw have a unimodal distribution that is indicative of a recent population expansion (Rogers, 1995; Rogers and Harpending, 1993; Harpending et al., 1993; Sherry et al., 1994). The Oklahoma Choctaw and the Creek both have bimodal distributions and the Chickasaw have a very different, ragged distribution. The mismatch distributions of these populations all appear to be in equilibrium.

The AMOVA results from the sequence data comparison show that 16% of the sequence variation is due to among group differences. This is higher than the 13% among group variation seen from the haplogroup data. The higher mutation rate may account for this disparity of higher divergence in the control region data.

Algonquian Comparisons

Haplotype analysis of 17 Wisconsin Menominee represented in the Boas collection was also performed. Like the Mississippi Choctaw, the Menominee sample only had two of the five Native American haplogroups represented. Of these 17, 15 were haplogroup A. The heterozygosity for the Menominee is the lowest of all of the groups in the present study (0.125) due to the large amount of haplogroup A. The population size of the Menominee in historic records has consistently been estimated to be roughly 1,600 individuals (Ourada, 1979). In addition, the Menominee population, like the Choctaw, went through a drop in size during the 1830s. This decline was due to a small pox epidemic that reduced the population by one-fourth of its original size (Ourada, 1979). However, also like the Choctaw, the lack of haplogroup diversity in the Menominee is not reflected in their sequence diversity. Of the 17 Menominee sequenced, there were 14 different lineages or sequence types. There are three pairs that share common sequence lineages: Menominee 33 and

Menominee 131, Menominee 95 and Menominee 97, and Menominee 111 and Menominee 139. While Menominee 95, 97, 111 and 139 all share the same haplogroup (A), Menominee 33 and 131 belong to two different haplogroups, A and C respectively. Menominee 33 has the Hae III site gain at nt 663 indicative of haplogroup A and Menominee 131 has the Hinc II loss at nt 13259 indicative of haplogroup C. Graven et al. (1995) report a similar occurrence. Two individuals in their study are identical in the control region sequence yet both belong to different RFLP haplogroups. While this is not a frequent occurrence is argues against trying to estimate haplogroup designations through sequence similarities.

One Menominee individual could not be typed as A, B, C, D or X. Census data for the Menominee taken in 1916 (Table 5-2), 24 years after the hair samples were collected, show that 75% of the Menominee population was admixed by that time, therefore this could be due to non-Native American admixture. Unlike the Choctaw, this is not reflected in the mtDNA data. The Menominee do not show the same levels of sequence diversity that is seen in the Choctaw possibly due to the incorporation of less non-Native American females into the tribe.

Table 5-2: Menominee Census Information

	1916	1919	1922	1925	1927	1928	1929
Total Population	1,736	1,733	1,819	1,890	1,941	1,940	1,939
Full Bloods	434	395	380	300	300	477	463
More than ½ Meno.	868	897	900	900	900	1,463	1,476
Less than ½ Meno.	434	441	539	690	741		
% of Full Bloods	25	22.8	20.9	15.9	15.4	24.6	23.9

The Menominee haplogroup frequencies were compared with the previously reported Micmac, Island Ojibwa, Northern Ontario Ojibwa, Turtle Mountain Chippewa, Wisconsin Chippewa and the Cheyenne/Arapaho (Lorenz and Smith, 1996; Smith et al., 1999; Malhi et al., 2000; Scozzari et al., 1997), all representatives of the Algonquian language family. The Menominee have the lowest heterozygosity (0.125) of all of the groups analyzed, which together have an average heterozygosity of 0.6873 (Table 3-2). Like the other Algonquian groups, the Menominee have an absence of the D haplogroup, but, unlike other groups, they do not contain any individuals belonging to haplogroup X. Like the Micmac, the Menominee lack haplogroup B, which is found at low frequencies in the other Algonquian groups.

An AMOVA was performed on the haplogroup frequencies for Algonquian groups with results indicating that 9% of the haplogroup variation occurs between populations. This implies that the populations are somewhat divergent, but not quite as much as the Muskogean speakers in the southeast (13%).

HVR-I sequence variation was compared between the Menominee and 8 other Algonquian groups: Chippewa, Micmac, Delaware, Potawatomi, Kickapoo, Fox, Cheyene and Arapaho. There are 36 different sequence lineages for the groups including the Menominee. The

Menominee share only two sequence types with the four other combined Algonquian groups, indicating that the Menominee exhibit substantial sequence diversity. The results of an AMOVA for the five Algonquian groups show that there is considerable divergence from one another. Thirty-one percent of the sequence variation is among the Algonquian groups, which is the most diversity seen in the entire present study.

Evidence of a recent expansion is seen in the Menominee's unimodal mismatch distribution. Unlike the unimodal distribution of Menominee, the distributions of the other groups are more bi-modal and multi-modal. The Chippewa, the Kickapoo/Potawatomi/Fox group and the Cheyenne/Arapaho group had essentially bimodal distributions and the Micmac/Delaware group had a multi-modal distribution, indicating that these groups are in equilibrium.

Comparison of the Muskogean and the Algonquian Groups

All of the groups in the analysis were collapsed into their respective language family groups of either Muskogean or Algonquian.

Interestingly, only 6% of the variation of the sequence data is due to among group differences between the two language groups. This is the lowest amount of among group variation seen in this entire study. For the haplogroup data, the among group variation for the two language families is 14%. The mismatch distributions for the Muskogean group

and for the Algonquian group are essentially unimodal. This indicates that the groups as a whole have gone through a recent population expansion (Rogers, 1995; Rogers and Harpending, 1993; Harpending et al., 1993; Sherry et al., 1994).

Summary

Hair shaft DNA analysis presents additional challenges to the already daunting task of ancient DNA analysis. Due to the extremely small amounts of DNA in each hair shaft, extreme care had to be taken at all steps to avoid contamination. In addition, there are many inhibitors in hair that make enzymatic reactions difficult. There were many amplification failures until optimization for each primer set was achieved.

Heteroplasmy

In addition, heteroplasmy is an increased risk with hair DNA. The author concedes that sequence polymophisms may be due, in part, to heteroplasmy. Heteroplasmy is the existence of two or more major mtDNA sequence types in a single individual. Although heteroplasmy was first described in 1988 (Holt et al., 1988), the first occurrence of a single nucleotide heteroplasmic condition was not reported until 1994 (Gill et al., 1994). Heteroplasmy occurs in different tissues at different

rates. It also occurs at different nucleotide positions at varying rates. Because certain sites are more polymorphic than others, they are more useful in the discrimination of closely related groups, however, this increased variation of sites also suggests that some sites will prove to be heteroplasmic in individuals rendering them less useful for population studies.

Remarks

As mtDNA analysis passes from its infancy a compilation of data is being assessed from many varied regions and including individuals with different language affiliations. As this occurs, the general assessments made in the past can turn to more specific questions about particular groups.

The information and biological samples Franz Boas preserved for the more than 15,000 individuals undoubtedly will continue to provide diverse information for generations to come. Genetic analysis of this collection is significant because of its extensive documentation of demographic and kinship information. This detailed collection removes the need to estimate admixture rates due to the known biological heritage of each individual and their parents. In addition, admixture rates are less than in modern groups because of the antiquity of the sample.

As more and more groups are genetically characterized, a true picture of the relationships of world populations will be understood. Additionally, there are several laboratories that are specifically aimed at the characterization of the genetic structure of ancient populations. Ancient DNA analysis has now passed the stage of merely seeing what can be done and with work like that of Stone and Stoneking, 1998 has moved into the realm of actual population genetics. However, many technical challenges still remain. Issues of contamination, amplification optimization, and heteroplasmy need to be further addressed.

In order to better understand how populations actually relate, the use of ancient and archived samples may be the key to filling in temporal and well as geographical data. In order to help piece together the migration patterns and origins of modern populations, the analysis of historic and ancient samples will be of vital importance.

LIST OF REFERENCES

Allen M, Engstrom AS, Meyers S, Handt O, Saldeen T, von Haeseler A, Paabo S, and Gyllensten U (1998) Mitochondrial DNA sequencing of shed hairs and saliva on robbery caps: sensitivity and matching probabilities. J Forensic Sci *43:*453-64.

Anderson S, Bankier AT, Barrell BG, Bruijn MHd, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJ, Staden R, and Young IG (1981) Sequence and organization of the human mitochondrial genome. Nature *290:*457-465.

Baker L, McCormick W, and Matteson K (2001) A Silica-Based Mitochondrial DNA Extraction Method Applied to Forensic Hair Shafts and Teeth. Journal of Forensic Sciences *46:*126-130.

Baker LE (in press) DNA from Hair. In D DeGusta (ed.): Methods in Molecular Biology: Ancient DNA Techniques. Totowa, NJ: Humana Press, pp. Chapter 12.

Barbujani G, Stenico M, Excoffier L, and Nigro L (1996) Mitochondrial DNA sequence variation across linguistic and geographic boundaries in Italy. Hum Biol *68:*201-15.

Bettersworth JK (1959) Mississippi: A History. Austin: Steck Company.

Boas F (1895) Zur Anthropologie der nordamerikanischen Indianer. Z. Ethnol. *27:*366-411.

Boas F (1899) Anthropometry of the Shoshonean tribes. Am Anthropol *1:*751-758.

Boas F (1905) Huntington California expedition: Anthropometry of central California. Bull Am Mus Natur Hist *17*:347-380.

Boom R, Sol CJ, Salimans MM, Jansen CL, Wertheim-van Dillen PM, and van der Noordaa J (1990) Rapid and simple method for purification of nucleic acids. J Clin Microbiol *28*:495-503.

Brown M, Hosseini S, Torroni A, Bandelt H, Allen J, Schurr T, Scozzari R, Cruciani F, and Wallace D (1998) mtDNA haplogroup X: an ancient link between Europe/Western Asia and North America? American Journal of Human Genetics *63*:1852-1861.

Brown WM, Prager EM, Wang A, and Wilson AC (1982) Mitochondrial DNA sequences of primates: tempo and mode of evolution. J Mol Evol *18*:225-39.

Cann RL, Stoneking M, and Wilson AC (1987) Mitochondrial DNA and human evolution. Nature *325*:31-6.

Comuzzie AG, Duggirala R, Leonard WR, and Crawford MH (1995) Population relationships among historical and modern indigenous Siberians based on anthropometric characters. Hum Biol *67*:459-79.

Condit CM (1999) The Meanings of the Gene. Madison, WI: The University of Wisconsin Press.

Darnell R (1970) The development of American anthropology 1879-1920 from the Beureau of American Ethnology to Franz Boas. Ph.D. Dissertation, University of Pennsylvania, Philadelphia.

Debo A (1934) The Rise and Fall of the Choctaw Republic. Norman, OK: University of Oklahoma Press.

DiRienzo A, and Wilson A (1991) Branching pattern in the evolutionary tree for human mitochondrial DNA. Proc Natl Acad Sci U S A *88:*1597-1601.

Edwards J (1932) The Choctaw indians in the middle of the of the nineteenth century. Chronicles of Oklahoma *10:*392-425.

Excoffier L, Smouse P, and Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics *131:*479-491.

Excoffier L, and Yang Z (1999) Substitution rate variation among sites in mitochondrial hypervariable region I of humans and chimpanzees. Mol Biol Evol *16:*1357-1368.

Falsetti A, and Jantz R (1990) Anthropometry of Amerindians: An examination of scaling phenomena (abs.). American Journal of Physical Anthropology *81:*220-221.

Falsetti AB (1989) Anthropometry of Native Americans from the Artic, Subartic, Northwest Coast, Great Basin, and California: An examination of scaling phenomena. Ph.D., University of Tennessee, Knoxville.

Finkler K (2000) Experiencing the New Genetics: Family and Kinship on the Medical Frontier. Philadelphia: University of Pennsylvania Press.

Gibson A (1973) The Indians of Mississippi. In R McLemore (ed.): A History of Mississippi. Hattiesburg: University and College Press, pp. 69-89.

Giles RE, Blanc H, Cann HM, and Wallace DC (1980) Maternal inheritance of human mitochondrial DNA. Proc Natl Acad Sci U S A *77*:6715-9.

Ginther C, Issel-Tarver L, and King MC (1992) Identifying individuals by sequencing mitochondrial DNA from teeth. Nat Genet *2*:135-8.

Graven L, Passarino G, Semino O, Boursot P, Santachiara-Benerecetti S, Langaney A, and Excoffier L (1995) Evolutionary correlation between control region sequence and restriction polymorphisms in the mitochondrial genome of a large Senegalese Mandenka sample. Mol Biol Evol *12*:334-45.

Green, LD, Derr JN, and Knight A (2000) mtDNA affinities of the peoples of North-Central Mexico. Am J Hum Genet *66*(3): 989-998.

Greenberg BD, Newbold JE, and Sugino A (1983) Intraspecific nucleotide sequence variability surrounding the origin of replication in human mitochondrial DNA. Gene *21*:33-49.

Halbert H (1899) Nanih Waiya, the sacred mound of the Choctaws. Publications of the Mississippi Historical Society *2*:223-34.

Hall RL, and Hall DA (1995) Geographic variation of native people along the Pacific Coast. Hum Biol *67:*407-26.

Hall RL, and MacNair PL (1972) Multivariate analysis of anthropometric data and the classification of British Columbia natives. Am J Phys Anthropol *37:*401-409.

Handt O, Hoss M, Krings M, and Paabo S (1994) Ancient DNA: methodological challenges. Experientia *50:*524-9.

Harpending HC, Sherry ST, Rogers AR, and Stoneking M (1993) The genetic structure of human populations. Curr Anthropol *34:*483-496.

Higuchi R, von Beroldingen CH, Sensabaugh GF, and Erlich HA (1988) DNA typing from single hairs. Nature *332:*543-6.

Hoffman WJ (1970) The Menomini Indians. New York: Johnson Reprint Corporation.

Holland MM, Fisher DL, Mitchell LG, Rodriguez WC, Canik JJ, Merril CR, and Weedn VW (1993) Mitochondrial DNA sequence analysis of human skeletal remains: Identification of remains from the Vietnam War. Journal of Forensic Sciences *13:*82-92.

Horai S, Kondo R, Nakagawa-Hattori Y, Hayashi S, Sonoda S, and Tajima K (1993) Peopling of the Americas, founded by four major lineages of mitochondrial DNA. Mol Biol Evol *10:*23-47.

Horai S, Murayama K, Hayasaka K, Matsubayashi S, Hattori Y, Fucharoen G, and Harihara S (1996) mtDNA polymorphism in east

Asian populations, with special reference to the peopling of Japan. Am J Hum Genet *59:*579-590.

Hosmer BC (1999) American Indians in the Marketplace: Persistence and Innovation Among the Menominees and Metlakatlans, 1870-1920. Lawrence: University of Kansas Press.

Hoss M, and Paabo S (1993) DNA extraction from Pleistocene bones by a silica-based purification method. Nucleic Acids Research *21:*3913-3914.

Hudson P (1934) Reminiscences by Peter J. Hudson. Chronicles of Oklahoma *10:*294-304.

Huhne J, Pfeiffer H, Waterkamp K, and Brinkmann K (1999) Mitochondrial DNA in human hair shafts--existence of intra-individual differences? Int J Legal Med *112:*172-5.

Hunt D, Jantz R, and Falsetti A (1988) Anthropometric variation among Amerindians from the North American Pacific Coast (abs.). American Journal of Physical Anthropology *75:*224.

Huoponen K, Juvonen V, Iitia A, Dahlen P, Siitari H, Aula P, Nikoskelainen E, and Savontaus ML (1994) Time-resolved fluorometry in the diagnosis of Leber hereditary optic neuroretinopathy. Hum Mutat *3:*29-36.

Hutchinson CA, Newbold JE, Potter SS, and Edgell MH (1974) Maternal inheritance of mammalian mitochondrial DNA. Nature *251:*536-538.

Jantz RL (1995) Franz Boas and Native American biological variability. Hum Biol *67*:345-53.

Jantz RL, Hunt DR, Falsetti AB, and Key PJ (1992) Variation among North Amerindians: analysis of Boas's anthropometric data [see comments]. Hum Biol *64*:435-61.

Jantz RL, and Meadows L (1995) Population structure of Algonquian speakers. Hum Biol *67*:375-86.

Kaestle FA, and Smith DG (2001) Ancient mitcochondrial DNA evidence for prehistoric population movement: the Numic expansion. American Journal of Physical anthropology *115*:1-12.

Kalbe J, Kuropka R, Meyer-Stork L, Sauter S, Loss P, Henco K, Riesner D, Hocker H, and Berndt H (1988) Isolation and characterization of high-molecular mass DNA from hair shafts. Biol Chem Hoppe Seyler *369*:413-416.

Keesing FM (1987) The Menomini Indians of Wisconsin: A Study of Three Centuries of Cultural Contact and Change. Madison, WI: The University of Wisconsin Press.

Konigsberg LW, and Ousley SD (1995) Multivariate quantitative genetics of anthropometric traits from the Boas data. Hum Biol *67*:481-98.

Lorenz JG, and Smith DG (1996) Distribution of four founding mtDNA haplogroups among Native North Americans. Am J Phys Anthropol *101*:307-23.

Lum JK, Rickards O, Ching C, and Cann RL (1994) Polynesian mitochondrial DNAs reveal three deep maternal lineage clusters. Human Biology *66*:567-590.

Malhi RS, Schultz BA, and Smith DG (2001) Distribution of mitochondrial DNA lineages among native american tribes of northeastern North America. Human Biology *73*:17-55.

McKee JO, and Schlenker JA (1980) The Choctaws: Cultural Evolution of a Native American Tribe. Jackson: University Press of Mississippi.

Merriwether DA, Hall WW, Vahlne A, and Ferrell RE (1996) mtDNA variation indicates Mongolia may have been the source for the founding population for the New World. Am J Hum Genet *59*:204-12.

Morrison J (1987) The social history of the Choctaw Nation: 1865-1907. Durant, Oklahoma: Creative Informatics.

O'Rourke DH, Hayes MG, and Carlyle SW (2000) Spacial and temporal stability of mtDNA haplogroup frequencies in native North America. Human Biology *72*:15-34.

Ourada PK (1979) The Menominee Indians: A History. Norman: University of Oklahoma Press.

Ousley SD (1995) Relationships between Eskimos, Amerindians, and Aleuts: old data, new perspectives. Hum Biol *67*:427-58.

Paabo S (1989) Ancient DNA: extraction, charachterization, molecular cloning, and enzymatic amplification. Proc Natl Acad Sci U S A *86:*1939-1943.

Paabo S, Gifford JA, and Wilson AC (1988) Mitochondrial DNA sequences from a 7000-year old brain. Nucleic Acids Res *16:*9775-87.

Pfeiffer H, Huhne J, Ortmann C, Waterkamp C, and Brinkmann B (1999) Mitochondrial DNA typing from human axillary, pubic and head hair shafts-success rates and sequence comparisons. International Journal of Legal Medicine *112:*287-290.

Piercy R, Sullivan KM, Benson N, and Gill P (1993) The application of mitochondrial DNA typing to the study of white Caucasian genetic identification. Internation Journal of Legal Medicine *106:*85-90.

Prince JM (1989) Secular trands in stature in an historic Sioux population. M.A., University of Tennessee, Knoxville.

Richards M, Corte-Real H, Forster P, Macaulay V, Wilkinson-Herbots H, Demaine A, Papiha S, Hedges R, Bandelt H, and Sykes B (1996) Paleolithic and Neolithic lineages in the European mitochondrial gene pool. American Journal of Human Genetics *59:*185-203.

Rogers A, and Harpending H (1992) Population growth makes waves in the distribution of pairwaise differences. Mol Biol Evol *9:*552-569.

Rogers AR (1995) Genetic evidence for a pleistocene population explosion. Evolution *49:*608-615.

Schneider S, Roessli D, and Excoffier L (2000) Arlequin ver.2.000: a software for population genetics data analysis: Genetics and Biometry Laboratory, University of Geneva.

Schurr TG, Ballinger SW, Gan YY, Hodge JA, Merriwether DA, Lawrence DN, Knowler WC, Weiss KM, and Wallace DC (1990) Amerindian mitochondrial DNAs have rare Asian mutations at high frequencies, suggesting they derived from four primary maternal lineages. Am J Hum Genet *46:*613-23.

Scozzari R, Cruciani F, Santolamazza P, Sellitto D, Cole DE, Rubin LA, Labuda D, Marini E, Succa V, Vona G, and Torroni A (1997) mtDNA and Y chromosome-specific polymorphisms in modern Ojibwa: implications about the origin of their gene pool [letter]. Am J Hum Genet *60:*241-4.

Sherry ST, Rogers AR, Harpending H, Soodyall H, Jenkins T, and Stoneking M (1994) Mismatch distributions of mtDNA reveal recent human population expansions. Hum Biol *66:*761-75.

Smith D, Malhi R, Eshleman J, Lorenz J, and Kaestle F (1999) Distribution of mtDNA haplogroup X among native North Americans. American Journal of Physical Anthropology *110:*271-284.

Stivers D (1990) Changes in stature and health status as related to the emergence of diabetes among eastern Cherokee Indians in North Carolina. Ph.D. Dissertation, University of Tennessee, Knoxville.

Stone AC, and Stoneking M (1998) mtDNA analysis of a prehistoric Oneota population: implications for the peopling of the New World. Am J Hum Genet *62:*1153-70.

Stoneking M, Hedgecock D, Higuchi RG, Vigilant L, and Erlich HA (1991) Population variation of human mtDNA control region sequences detected by enzymatic amplification and sequence-specific oligonucleotide probes. Am J Hum Genet *48:*370-82.

Sullivan L (1920) Anthropometry of the Siouan tribes. Anthropol. Pap. Am. Mus. Natur. Hist. *23:*81-174.

Sweet DJ, and Sweet CH (1995) DNA analysis of dental pulp to link incinerated remains of homicide victim to crime scene. J Forensic Sci *40:*310-4.

Szathmary EJ (1995) Overview of the Boas anthropometric collection and its utility in understanding the biology of Native North Americans. Hum Biol *67:*337-44.

Tamura K, and Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNa in humans and chimpanzees. Mol Biol Evol *10:*512-526.

Torroni A, Schurr TG, Yang CC, Szathmary EJ, Williams RC, Schanfield MS, Troup GA, Knowler WC, Lawrence DN, Weiss KM, and et al. (1992) Native American mitochondrial DNA analysis indicates that the Amerind and the Nadene populations were founded by two independent migrations. Genetics *130:*153-62.

Uchihi R, Tamaki K, Kojima T, Yamamoto T, and Katsumata Y (1992) Deoxyribonucleic acid (DNA) typing of human leukocyte antigen (HLA)- DQA1 from single hairs in Japanese. J Forensic Sci *37:*853-9.

Vigilant L, Pennington R, Harpending H, Kocher TD, and Wilson AC (1989) Mitochondrial DNA sequences in single hairs from a southern African population. Proc Natl Acad Sci U S A *86:*9350-4.

Wallace DC, Garrison K, and Knowler WC (1985) Dramatic founder effects in Amerindian mitochondrial DNAs. Am J Phys Anthropol *68:*149-55.

Weiss DA, and Smith DG (Submitted for Publication) Evidence for a recent mtDNA bottleneck among native Americans from the southeastern United States.

Wilson MR, DiZinno JA, Polanskey D, Replogle J, and Budowle B (1995a) Validation of mitochondrial DNA sequencing for forensic casework analysis. International Journal of Legal Medicine *108:*68-74.

Wilson MR, Holland MM, Stoneking M, DiZinno JA, and Budowle B (1993) Guidelines for the use of mitochondrial DNA sequencing in forensic science. Crime Laboratory Digest *20:*68-77.

Wilson MR, Polanskey D, Butler J, DiZinno JA, Replogle J, and Budowle B (1995b) Extraction, PCR amplification and sequencing of mitochondrial DNA from human hair shafts. Biotechniques *18:*662-9.

APPENDICIES

Appendix A: Sample Information

Sample	Sex	Collection Location	Date	Age	Tribe	Purity	Birthplace	Mother Tribe	Father Tribe
Choctaw1	M	Newton Co., MS	7/7/1892	26	Choctaw	Full	Newton Co	Choc.	Choc.
Choctaw10	M	Newton Co., MS	7/8/1892	42	Choctaw	Full	Newton Co	Choc.	Choc.
Choctaw16	F	Newton Co., MS	7/9/1892	18	Choctaw	Full	Newton Co	Choc.	Choc.
Choctaw22	F	Newton Co., MS	7/11/1892	65	Choctaw	Full	Newton Co	Choc.	Choc.
Choctaw28	M	Newton Co., MS	7/12/1892	21	Choctaw	Full	Newton Co	Choc.	Choc.
Choctaw30	M	Newton Co., MS	7/12/1892	22	Choctaw	Full	Newton Co	Choc.	Choc.
Choctaw41	M	Neshoba Co., MS	7/15/1892	32	Choctaw	Full	Neshoba Co	Choc.	Choc.
Choctaw46	M	Neshoba Co., MS	7/16/1892	35	Choctaw	Full	Neshoba Co	Choc.	Choc.
Choctaw48	M	Neshoba Co., MS	7/16/1892	31	Choctaw	Full	Neshoba Co	Choc.	Choc.
Choctaw57	F	Neshoba Co., MS	7/18/1892	20	Choctaw	Full	Neshoba Co	Choc.	Choc.
Choctaw64	F	Neshoba Co., MS	7/19/1892	20	Choctaw	Full	Neshoba Co	Choc.	Choc.
Choctaw68	M	Neshoba Co., MS	7/20/1892	25	Choctaw	Full	Neshoba Co	Choc.	Choc.
Choctaw70	M	Neshoba Co., MS	7/20/1892	34	Choctaw	Full	Neshoba Co	Choc.	Choc.
Choctaw85	M	Neshoba Co., MS	7/23/1892	26	Choctaw	Full	Neshoba Co	Choc.	Choc.
Choctaw90	M	Neshoba Co., MS	7/23/1892	48	Choctaw	Full	Neshoba Co	Choc.	Choc.
Choctaw137	F	Neshoba Co., MS	7/31/1892	26	Choctaw	Full	Neshoba Co	Choc.	Choc.
Choctaw139	F	Neshoba Co., MS	7/31/1892	27	Choctaw	Full	Neshoba Co	Choc.	Choc.
Choctaw148	M	Neshoba Co., MS	7/31/1892	27	Choctaw	Full	Neshoba Co	Choc.	Choc.
Choctaw152	M	Neshoba Co., MS	8/1/1892	68	Choctaw	Full	Neshoba Co	Choc.	Choc.
Menom14	M	Keshina, Wis	2/18/1892	9	Menom.	Full		Menom	Menom
Menom33	F	Keshina, WI	2/18/1892	7	Menom.	Full		Menom.	Menom.
Menom57	M	Keshina, WI	2/19/1892	9	Menom.	Full	Menom, WI	Menom.	Menom.
Menom95	F	Keshina, WI	2/20/1892	13	Menom.	Full	Keshena	Menom.	Menom.
Menom97	F	Keshina, WI	2/20/1892	9	Menom.	½		Menom.	English
Menom98	F	Keshina, WI	2/20/1892	10	Menom.	Full	Keshena	Menom.	Menom.
Menom111	F	Keshina, WI	2/22/1892	8	Menom.	Full		Menom.	Menom.
Menom113	F	Keshina, WI	2/22/1892	16	Menom.	Full		Menom.	Menom.
Menom114	F	Keshina, WI	2/22/1892	8	Menom.	Full		Menom.	Menom.
Menom118	F	Keshina, WI	2/22/1892	5	Menom.	Full		Menom.	Menom.
Menom124	F	Keshina, WI	2/23/1892	5	Menom.	Full		Menom.	Menom.
Menom128	F	Keshina, WI	2/23/1892	7	Menom.	Full	Marinett	Menom.	Menom.
Menom131	F	Keshina, WI	2/23/1892	5	Menom.	Full	Keshena	Menom.	Menom.
Menom139	F	Keshina, WI	2/23/1892	9	Menom.	Full	Kanikua	Menom.	Menom.
Menom151	F	Keshina, WI	2/23/1892	12	Menom.	Full	Keshena	Menom.	Menom.
Menom154	F	Keshina, WI	2/23/1892	14	Menom.	Full	Keshena	Menom.	Menom.
Menom155	F	Keshina, WI	2/23/1892	14	Menom.	Full	Keshena	Menom.	Menom.

Appendix B: Observed Choctaw and Menominee mtDNA Sequence Variation

Changes from Reference Sequence

	16000	16001	16002	16003	16004	16005	16006	16007	16008	16009	16010	16011	16012	16013	16014	16015	16016	16017	16018	16019	16020	16021	16022	16023	16024	16025	16026
Reference	G	A	T	T	C	T	A	A	T	T	T	A	A	A	C	T	A	T	T	C	T	C	T	G	T	T	C
Choc1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc16	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc22	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc28	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc30	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc41	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc46	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc48	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc64	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc68	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc70	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc85	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc90	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc137	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc139	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc148	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc152	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom14	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom33	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom95	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom97	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom98	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom111	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom113	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom114	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom118	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom124	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom128	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom131	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom139	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom151	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom154	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom155	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Changes from Reference Sequence

	16027	16028	16029	16030	16031	16032	16033	16034	16035	16036	16037	16038	16039	16040	16041	16042	16043	16044	16045	16046	16047	16048	16049	16050	16051	16052	16053
Reference	T	T	T	C	A	T	G	G	G	G	A	A	G	C	A	G	A	T	T	T	G	G	G	T	A	C	C
Choc1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc16	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc22	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc28	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc30	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc41	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc46	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.
Choc48	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc64	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc68	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc70	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc85	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc90	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.
Choc137	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc139	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc148	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc152	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom14	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom33	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom95	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom97	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom98	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom111	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom113	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom114	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom118	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom124	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom128	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom131	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom139	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom151	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom154	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom155	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Changes from Reference Sequence

	16054	16055	16056	16057	16058	16059	16060	16061	16062	16063	16064	16065	16066	16067	16068	16069	16070	16071	16072	16073	16074	16075	16076	16077	16078	16079	16080
Reference	A	C	C	C	A	A	G	T	A	T	T	G	A	C	T	C	A	C	C	C	A	T	C	A	A	C	A
Choc1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc16	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc22	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc28	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc30	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc41	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc46	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc48	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc64	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc68	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc70	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc85	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc90	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc137	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc139	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc148	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc152	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom14	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom33	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom95	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom97	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom98	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom111	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom113	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom114	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom118	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom124	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom128	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom131	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom139	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom151	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom154	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom155	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Changes from Reference Sequence

	16081	16082	16083	16084	16085	16086	16087	16088	16089	16090	16091	16092	16093	16094	16095	16096	16097	16098	16099	16100	16101	16102	16103	16104	16105	16106	16107
Reference	A	C	C	G	C	T	A	T	G	T	A	T	T	T	C	G	T	A	C	A	T	T	A	C	T	G	C
Choc1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc16	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc22	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc28	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc30	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc41	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc46	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc48	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc64	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc68	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc70	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc85	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc90	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc137	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc139	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc148	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc152	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom14	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom33	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom95	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom97	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom98	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom111	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom113	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom114	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom118	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom124	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom128	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom131	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom139	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom151	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom154	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom155	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Changes from Reference Sequence

	16108	16109	16110	16111	16112	16113	16114	16115	16116	16117	16118	16119	16120	16121	16122	16123	16124	16125	16126	16127	16128	16129	16130	16131	16132	16133	16134
Reference	C	A	G	C	C	A	C	C	A	T	G	A	A	T	A	T	T	G	T	A	C	G	G	T	A	C	C
Choc1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc16	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc22	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc28	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc30	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc41	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc46	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc48	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc57	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
Choc64	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc68	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc70	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc85	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc90	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc137	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.
Choc139	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.
Choc148	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.
Choc152	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom14	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom33	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom57	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom95	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom97	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom98	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.
Menom111	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom113	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom114	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom118	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom124	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom128	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom131	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom139	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom151	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom154	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom155	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Changes from Reference Sequence

	16135	16136	16137	16138	16139	16140	16141	16142	16143	16144	16145	16146	16147	16148	16149	16150	16151	16152	16153	16154	16155	16156	16157	16158	16159	16160	16161
Reference	A	T	A	A	A	T	A	C	T	T	G	A	C	C	A	C	C	T	G	T	A	G	T	A	C	A	T
Choc1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.
Choc16	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc22	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc28	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc30	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc41	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc46	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc48	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc64	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc68	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc70	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc85	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc90	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc137	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc139	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc148	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc152	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom14	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom33	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom95	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom97	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom98	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom111	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom113	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom114	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom118	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom124	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom128	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom131	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom139	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom151	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom154	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom155	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Changes from Reference Sequence

	16162	16163	16164	16165	16166	16167	16168	16169	16170	16171	16172	16173	16174	16175	16176	16177	16178	16179	16180	16181	16182	16183	16184	16185	16186	16187	16188
Reference	A	A	A	A	A	C	C	C	A	A	T	C	C	A	C	A	T	C	A	A	A	A	C	C	C	C	C
Choc1	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc16	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc22	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc28	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.
Choc30	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc41	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc46	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.
Choc48	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc64	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc68	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc70	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc85	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.
Choc90	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc137	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc139	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc148	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc152	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom14	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom33	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom57	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom95	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom97	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom98	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom111	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom113	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom114	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom118	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.
Menom124	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.
Menom128	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom131	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom139	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom151	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom154	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom155	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.

Changes from Reference Sequence

	16189	16190	16191	16192	16193	16194	16195	16196	16197	16198	16199	16200	16201	16202	16203	16204	16205	16206	16207	16208	16209	16210	16211	16212	16213	16214	16215
Reference	T	C	C	C	C	A	T	G	C	T	T	A	C	A	A	G	C	A	A	G	T	A	C	A	G	C	A
Choc1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc16	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc22	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc28	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc30	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc41	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc46	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc48	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc64	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc68	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc70	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc85	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc90	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc137	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc139	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc148	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G
Choc152	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom14	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom33	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom95	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom97	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom98	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom111	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom113	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom114	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom118	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom124	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom128	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom131	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom139	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom151	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom154	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom155	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Changes from Reference Sequence

	16216	16217	16218	16219	16220	16221	16222	16223	16224	16225	16226	16227	16228	16229	16230	16231	16232	16233	16234	16235	16236	16237	16238	16239	16240	16241	16242
Reference	A	T	C	A	A	C	C	C	T	C	A	A	C	T	A	T	C	A	C	A	C	A	T	C	A	A	C
Choc1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc16	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc22	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc28	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc30	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc41	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc46	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Choc48	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc64	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc68	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc70	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc85	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc90	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc137	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc139	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc148	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc152	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom14	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.
Menom33	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom95	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom97	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom98	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom111	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom113	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom114	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom118	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom124	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom128	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom131	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom139	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom151	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom154	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.
Menom155	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Changes from Reference Sequence

	16243	16244	16245	16246	16247	16248	16249	16250	16251	16252	16253	16254	16255	16256	16257	16258	16259	16260	16261	16262	16263	16264	16265	16266	16267	16268	16269
Reference	T	G	C	A	A	C	T	C	C	A	A	A	G	C	C	A	C	C	C	C	T	C	A	C	C	C	A
Choc1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.
Choc16	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc22	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc28	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.
Choc30	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc41	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.
Choc46	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc48	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc64	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc68	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc70	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc85	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc90	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.
Choc137	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.
Choc139	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc148	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc152	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom14	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom33	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom95	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom97	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom98	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom111	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.
Menom113	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.
Menom114	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom118	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom124	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom128	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom131	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom139	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.
Menom151	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.
Menom154	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom155	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Changes from Reference Sequence

	16270	16271	16272	16273	16274	16275	16276	16277	16278	16279	16280	16281	16282	16283	16284	16285	16286	16287	16288	16289	16290	16291	16292	16293	16294	16295	16296
Reference	C	T	A	G	G	A	T	A	C	C	A	A	C	A	A	A	C	C	T	A	C	C	C	A	C	C	C
Choc1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc16	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc22	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc28	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc30	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc41	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc46	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc48	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc57	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc64	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc68	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc70	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc85	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc90	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc137	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc139	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc148	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc152	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom14	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom33	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.
Menom57	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom95	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom97	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom98	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom111	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.
Menom113	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom114	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom118	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom124	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.
Menom128	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom131	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.
Menom139	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.
Menom151	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.
Menom154	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	T	.	.	.	.	.	.
Menom155	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Changes from Reference Sequence

	16297	16298	16299	16300	16301	16302	16303	16304	16305	16306	16307	16308	16309	16310	16311	16312	16313	16314	16315	16316	16317	16318	16319	16320	16321	16322	16323
Reference	T	T	A	T	C	A	G	T	A	C	A	T	A	G	T	A	C	A	T	A	A	A	G	C	C	A	T
Choc1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc16	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.
Choc22	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Choc28	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc30	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc41	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Choc46	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc48	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Choc64	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Choc68	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Choc70	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc85	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Choc90	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Choc137	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc139	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Choc148	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc152	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom14	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Menom33	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Menom57	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom95	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Menom97	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Menom98	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Menom111	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Menom113	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Menom114	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Menom118	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Menom124	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom128	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Menom131	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Menom139	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Menom151	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Menom154	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Menom155	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Changes from Reference Sequence

	16324	16325	16326	16327	16328	16329	16330	16331	16332	16333	16334	16335	16336	16337	16338	16339	16340	16341	16342	16343	16344	16345	16346	16347	16348	16349	16350
Reference	T	T	A	C	C	G	T	A	C	A	T	A	G	C	A	C	A	T	T	A	C	A	G	T	C	A	A
Choc1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc16	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc22	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc28	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc30	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc41	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc46	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc48	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc64	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc68	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc70	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc85	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc90	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc137	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc139	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc148	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc152	.	C	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom14	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom33	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom95	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom97	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom98	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom111	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom113	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom114	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom118	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom124	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom128	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom131	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom139	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom151	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom154	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom155	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Changes from Reference Sequence

	16351	16352	16353	16354	16355	16356	16357	16358	16359	16360	16361	16362	16363	16364	16365	16366	16367	16368	16369	16370	16371	16372	16373	16374	16375	16376	16377
Reference	A	T	C	C	C	T	T	C	T	C	G	T	C	C	C	C	A	T	G	G	A	T	G	A	C	C	C
Choc1	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc16	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc22	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc28	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc30	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc41	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc46	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc48	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc64	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc68	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc70	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc85	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc90	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc137	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc139	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc148	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc152	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom14	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom33	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom57	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom95	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom97	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom98	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom111	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom113	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom114	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom118	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom124	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom128	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom131	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom139	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom151	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom154	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom155	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Changes from Reference Sequence

	16378	16379	16380	16381	16382	16383	16384	16385	16386	16387	16388	16389	16390	16391	16392	16393	16394	16395	16396	16397	16398	16399	16400	16401	16402	16403	16404
Reference	C	C	C	T	C	A	G	A	T	A	G	G	G	G	T	C	C	C	T	T	G	A	C	C	A	C	C
Choc1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc16	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc22	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc28	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc30	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc41	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc46	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc48	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc64	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc68	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc70	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc85	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc90	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc137	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.
Choc139	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc148	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Choc152	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom14	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom33	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom95	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom97	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom98	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom111	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom113	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom114	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom118	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom124	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom128	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom131	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom139	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom151	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom154	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Menom155	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Changes from Reference Sequence

	16405	16406	16407	16408	16409
Reference	A	T	C	C	T
Choc1	.	.	.	.	.
Choc10	.	.	.	.	.
Choc16	.	.	.	.	.
Choc22	.	.	.	.	.
Choc28	.	.	.	.	.
Choc30	.	.	.	.	.
Choc41	.	.	.	.	.
Choc46	.	.	.	.	.
Choc48	.	.	.	.	.
Choc57	.	.	.	.	.
Choc64	.	.	.	.	.
Choc68	.	.	.	.	.
Choc70	.	.	.	.	.
Choc85	.	.	.	.	.
Choc90	.	.	.	.	.
Choc137	.	.	.	.	.
Choc139	.	.	.	.	.
Choc148	.	.	.	.	.
Choc152	.	.	.	.	.
Menom14	.	.	.	.	.
Menom33	.	.	.	.	.
Menom57	.	.	.	.	.
Menom95	.	.	.	.	.
Menom97	.	.	.	.	.
Menom98	.	.	.	.	.
Menom111	.	.	.	.	.
Menom113	.	.	.	.	.
Menom114	.	.	.	.	.
Menom118	.	.	.	.	.
Menom124	.	.	.	.	.
Menom128	.	.	.	.	.
Menom131	.	.	.	.	.
Menom139	.	.	.	.	.
Menom151	.	.	.	.	.
Menom154	.	.	.	.	.
Menom155	.	.	.	.	.

VITA

Lori Elmore Baker was born in Mesquite, Texas, on September 4, 1970. She attended primary and secondary schools in Lufkin, Texas, and graduated from Lufkin High School in June 1989. Ms. Baker continued her education at Baylor University in Waco, Texas, where she received bachelor degrees in Anthropology and Sociology in May of 1993. Upon graduation, she began the Master's program at Baylor for a Master of Arts degree that she completed in August of 1994. Ms. Baker then spent the next year learning DNA extraction and analysis techniques while working at the Plant Biotechnology Center at Baylor. Ms. Baker then accepted a graduate position at the University of Tennessee, Knoxville, where she pursued a Ph.D. in Anthropology. She received the Doctor of Philosophy degree in December 2001 and will further her education as a postdoctoral fellow in the Department of Medical Genetics of the Graduate School of Medicine, Knoxville, Tennessee.