Am. J. Hum. Genet. 62:1153–1170, 1998

mtDNA Analysis of a Prehistoric Oneota Population: Implications for the Peopling of the New World

Anne C. Stone and Mark Stoneking

Department of Anthropology, Pennsylvania State University, University Park

Summary

mtDNA was successfully extracted from 108 individuals from the Norris Farms Oneota, a prehistoric Native American population, to compare the mtDNA diversity from a pre-Columbian population with contemporary Native American and Asian mtDNA lineages and to examine hypotheses about the peopling of the New World. Haplogroup and hypervariable region I sequence data indicate that the lineages from haplogroups A, B, C, and D are the most common among Native Americans but that they were not the only lineages brought into the New World from Asia. The mtDNA evidence does not support the three-wave hypothesis of migration into the New World but rather suggests a single "wave" of people with considerable mtDNA diversity that exhibits a signature of expansion 23,000–37,000 years ago.

Introduction

Because of its relatively rapid rate of mutation and lack of recombination, the mitochondrial genome has been examined extensively, to investigate human population history throughout the world (Vigilant et al. 1989; Di Rienzo and Wilson 1991; Lum et al. 1994; Horai et al. 1996; Richards et al. 1996). In the New World, such studies have included both restriction-enzyme surveys and sequencing of the hypervariable regions adjacent to the origin of replication. These data have been used as evidence for hypotheses about the timing, origins, and number of migrations from Asia into the Americas, about the size of the migrant population, and about local demographic history.

Received December 22, 1997; accepted for publication March 18, 1998; electronically published April 10, 1998.

Present affiliation and address for correspondence and reprints: Dr. Anne Stone, Laboratory of Molecular Systematics and Evolution, Department of Ecology and Evolutionary Biology, 310 Biological Sciences West, University of Arizona, Tucson, AZ 85721. E-mail: acstone@u.arizona.edu

0002-9297/98/6205-0022$02.00

From restriction-enzyme studies, it was found that the mtDNA types from contemporary Amerindian populations typically fall into four primary clusters, each of related lineages. Each cluster can be characterized by a specific mtDNA marker: the 9-bp deletion in the COII/tRNA^{lys} region, a *Hae*III restriction site gain at nucleotide position 663 of the reference sequence (Anderson et al. 1981), a *Hin*cII restriction site loss at nucleotide 13259, and an *Alu*I restriction site loss at nucleotide 5176 (Wallace et al. 1985; Schurr et al. 1990; Torroni et al. 1992; Wallace and Torroni 1992). Sequence data indicate a correspondence between each marker and particular hypervariable region I (HVI) mutations (Ginther et al. 1993; Horai et al. 1993; Bailliet et al. 1994; present study). Bailliet et al. (1994) suggested a fifth cluster of lineages that has a unique HVI mutation at nucleotide 16278 and does not possess any of the characteristic markers.

To examine hypotheses about the peopling of the New World and to compare the mtDNA diversity from a pre-Columbian population with contemporary Native American and Asian mtDNA lineages, DNA was successfully extracted from 108 individuals (of 152 individuals sampled) from the Norris Farms cemetery. The Norris Farms cemetery population is ~700 years old and is classified by archaeologists as belonging to the Oneota culture. The cemetery was used for only one or two generations (Santure et al. 1990). In a previous study, we screened the markers for the four major groups of Native American mtDNA lineages in 100 adults, although we determined the haplogroup designation for only 50 of these (Stone and Stoneking 1993). In the present study, another 52 individuals were screened, including juveniles and adults, and some members of the previous sample were reexamined. In addition, the HVI was sequenced in 52 individuals who were selected to include members of the four primary clusters as well as those who did not appear to fall into one of these groups, in numbers roughly proportional to the presence of these groups in the cemetery. The genetic analysis of this cemetery population is the first extensive population survey of mtDNA diversity in a single prehistoric society from a fairly short-term occupation.

Subjects and Methods

Population Samples

The Norris Farms 36 cemetery is located in west-central Illinois, on a bluff above the Illinois River, and dates to A.D. ~1300. The cemetery and adjacent village site were variants of the Oneota cultural tradition and were used for a short time—probably only a few generations (Santure et al. 1990). The cemetery and a few other sites represent an abrupt departure from the previous cultural sequence. They appear to indicate a "frontier" population that expanded into this area from the northwest.

The cemetery was completely excavated in 1984–85 by the Illinois State Museum, and ~260 well-preserved Oneota skeletons were recovered (Santure et al. 1990, pp. 66–74). Most of the graves were located in a low mound in slightly alkaline loess. The age and sex distributions of skeletons correspond to expected human mortality patterns in traditional societies (Milner et al. 1989). These findings suggest that most, if not all, community members were buried in this cemetery.

The modern Native American HVI sequence data analyzed in this study included data from all published Native American populations with a sample $\geqslant 25$. Such samples included 11 Amerind groups: Nuu-Chah-Nulth (Ward et al. 1991); Bella Coola (Ward et al. 1993); Yakima (Shields et al. 1993); Ngobe (Kolman et al. 1995); Kuna (Batista et al. 1995); Huetar (Santos et al. 1994); Yanomami (Easton et al. 1996); Xavante, Gaviao, Zoro (Ward et al. 1996); and Mapuche (Ginther et al. 1993); and one Na-Dene group: Haida (Ward et al. 1993). These mtDNA data were also compared with Mongolian mtDNA data (Kolman et al. 1996).

DNA Extraction

DNA was extracted primarily from rib bones (usually the 11th or 12th ribs) that were free of pathological lesions. These particular specimens were chosen for analysis to minimize the destruction of parts of the skeleton frequently used in standard osteological studies. In the cases in which ribs were not available or in which they were of particular osteological interest, samples were taken from another bone, such as a phalanx or patella, or from a tooth root. To minimize the potential for contamination, bone and tooth samples were prepared for DNA extraction in the bioarchaeology laboratory, separate from the main genetics and ancient DNA laboratories. Gloves, a face shield, and a laboratory coat were worn while handling the bones. The outer layer of bone was removed with a sterile razor blade or a rotary tool with a grinding stone bit (Sears Craftsman) to prevent contamination from previous handling. When teeth were used, the whole tooth was soaked in 10% bleach for ~10 min and then rinsed with 70% ethanol. The tooth cap was then removed, with the rotary tool, at the enamel-dentin boundary, and the root was used for DNA extraction.

The bone or tooth root was ground to a fine powder by means of a bone mill (B. Braun Biotech), a modified paint shaker (Red Devil), or an electric coffee grinder (Mr. Coffee). The coffee grinder and the alumina ceramic vials that were used with the bone mill or paint shaker were washed with H_2O, soaked in 1 N HCl or bleach for $\geqslant 15$ min, rinsed with double-distilled water, and UV irradiated between uses. After the samples were ground into dust, they were stored in sterile 15-ml tubes (Falcon) at room temperature. Approximately 0.25 g of bone was used for each extraction.

The procedure used for extracting DNA from bones from the Norris Farms cemetery was initially a proteinase K digestion followed by a phenol chloroform extraction, as modified from previous protocols (Vigilant 1990; Hagelberg and Clegg 1991). This protocol was used to extract DNA from the ribs of 100 adults from the Norris Farms cemetery. These extractions were used in the analysis of the four markers that delineate Native American mtDNA lineage clusters (Stone and Stoneking 1993). Subsequent extractions were performed by use of the silica and guanidine isothiocyanate method described in Höss and Pääbo (1993), which is a modification of the protocol developed by Boom et al. (1990). A blank containing only reagents was included, with each bone sample, in the extraction procedure.

Mitochondrial DNA Typing and Sequencing

Because ancient DNA is normally degraded to fragments 100–200 bp in length (Pääbo et al. 1988, Pääbo 1989), the PCR was used to amplify specific segments of mtDNA <250 bp long that contained the characteristic Native American markers. The initial DNA amplifications of these mtDNA fragments and the primers used are described in Stone and Stoneking (1993). Subsequent PCRs were carried out in a 50-μl volume in which a wax-mediated hot start (Chou et al. 1992) was performed as described in Stone et al. (1996), except that different *Taq* polymerase (Boehringer Mannheim), another BSA (special quality for molecular biology from Boehringer Mannheim), and new primers listed in table 1 were used. The newer primers amplified smaller fragments and were more sensitive than the old primers. Forty cycles of amplification were carried out, with each cycle consisting of denaturation at 94°C for 1 min, annealing at 55°C for 1 min, and elongation at 72°C for 1 min. To detect contamination by modern DNA, PCR blanks containing all reagents except for DNA and extraction blanks were included in every PCR set. The PCR products were visualized with ethidium bromide in a 2.8% NuSieve agarose gel (FMC). Restriction digests

Table 1

Primers for mtDNA Typing

mtDNA Site and Primer	Primer Sequence (5′ to 3′)	Reference
*Hae*III site (nt 663):		
L635	TGAAAATGTTTAGACGGCCTCACATC	Handt et al. (1996)
H708	TAGAGGGTGAACTCACTGGAAC	Handt et al. (1996)
9-bp deletion:		
L8215	ACAGTTTCATGCCCATCGTC	Wrischnik et al. (1987)
H8297	ATGCTAAGTTAGCTTTACAG	Wrischnik et al. (1987)
*Hin*cII site (nt 13259):		
L13257	AATCGTAGCCTTCTCCACTTCA	Handt et al. (1996)
H13393	TCCTATTTTTCGAATATCTTGTTC	Ward et al. (1991)
*Alu*I site (nt 5176):		
L5054	TAGGATGAATAATAGCAGCTCTACCG	Present study
H5189	GGGTGGATGGAATTAAGGGTGT	Handt et al. (1996)
HVI:[a]		
L16055[a]	GAAGCAGATTTGGGTACCAC	Handt et al. (1996)
H16139	TACTACAGGTGGTCAACTAT	Handt et al. (1996)
H16142[a]	ATGTACTACAGGTGGTCAAG	Present study
L16131[a]	CACCATGAATATTGTACGGT	Handt et al. (1996)
H16218[a]	TGTGTGATAGTTGAGGGTTG	Handt et al. (1996)
L16209[a]	CCCCATGCTTACAAGCAAGT	Handt et al. (1996)
H16303[a]	TGGCTTTATGTACTATGTAC	Handt et al. (1996)
H16271	GTGGGTAGGTTTGTTGGTATCCTA	Krings et al. (1997)
L16287[a]	CACTAGGATACCAACAAACC	Handt et al. (1996)
H16356[a]	GTCATCCATGGGGACGAGAA	Handt et al. (1996)
H16379	CAAGGGACCCCTATCTGAGG	Handt (1995)
L16313	CCCTTAACAGTACATAGTAC	Handt (1995)
L16347[a]	CGTACATAGCACATTACAGT	Handt et al. (1996)
H16410[a]	GCGGGATATTGATTTCACGG	Handt et al. (1996)

[a] Main primers used for HVI PCR.

were performed as described in Stone and Stoneking (1993).

A total of 353 bp (positions 16056–16409) were sequenced with five sets of primers (table 1) that amplified overlapping fragments of the mtDNA HVI. In some samples, more than one set of primers was used to amplify the same segment. The length range of PCR products was 102–136 bp. In general, the shorter fragments were easier to amplify, and the sequences obtained from these fragments contained fewer ambiguities, particularly in samples that contained smaller amounts of DNA.

PCR was performed as described above. For each PCR, one primer from each pair was biotinylated. PCR product bands were excised from a 2.8% NuSieve gel and placed in 100 μl of Tris-EDTA (pH 7.5). After melting at 65°C, 3 μl of the excised gel slice were used for a 25-cycle reamplification during which the annealing temperature was increased to 57°C. Magnetic beads (Dynal) were used to select and purify the biotinylated strand of the amplified DNA, for direct sequencing, as described by Redd et al. (1995). DNA was obtained from a minimum of two independent extractions per individual, and each segment of the HVI was amplified and sequenced once in each direction by use of one extract for the light strand and another for the heavy strand. If the two sequences did not agree or if they were ambiguous at particular sites, the PCR and sequencing of that fragment were repeated.

All DNA extractions and PCR preparations involving the skeletal samples were performed in a room dedicated to ancient DNA research that is physically separated from the main genetics laboratory and the bioarchaeology laboratory. Dedicated reagents and equipment, including γ-sterilized filter pipette tips (VWR), were used. Gloves, a face shield, and a laboratory coat were worn at all times in the ancient DNA lab. New primers and other PCR reagents were tested for contamination by means of three PCR blanks. After a PCR was set up, the tubes were sealed and taken upstairs to the main genetics laboratory, where they were placed in a PCR machine. PCR products were never brought into the ancient DNA laboratory.

DNA Analyses

Phylogenetic analyses of the samples included both the maximum-likelihood and neighbor-joining tree-building methods. For the maximum-likelihood method, the γ distance for the Tamura-Nei (1993) model was calculated by use of the quartet-puzzling method (Strim-

mer and von Haeseler 1996) as performed by the computer program PUZZLE (Strimmer and von Haeseler 1997). PUZZLE calculates the majority-rule consensus tree from all possible quartets that can be made from the 138 sequences. A total of 1,000 puzzling steps were performed. The γ distribution parameter α, which controls the shape of the distribution, was calculated from the data, by PUZZLE. For the phylogenetic trees constructed by the neighbor-joining method (Saitou and Nei 1987), the computer programs MEGA (Kumar et al. 1993) and MEBoot (Tamura 1997) were used, and 1,000 bootstrap replications were performed. γ distances were calculated, by use of the Tamura-Nei (1993) model and the value for α previously estimated by PUZZLE, on the basis of pairwise sequence comparison. Insertions associated with the T → C mutation at nucleotide 16189 were not included in the analyses because they are not independent events (Bendall and Sykes 1995). *Pan troglodytes* mtDNA sequences (Morin et al. 1994) were used as outgroups in the phylogenetic analyses. A median network (Bandelt et al. 1995) was also constructed, by hand, to depict the relationships of the Norris Farms sequences.

Mismatch distributions were determined for the Norris Farms Oneota, 12 Native American tribes, and Mongolians. Intermatch distributions were computed between all pairs of populations. Mismatch distributions were also determined for lineages within each haplogroup for which sufficient data were available. Mismatch distributions are histograms of the number of pairwise differences within a population, whereas intermatch distributions are the mismatch distributions between populations. Such distributions should maintain a record of past population expansions and separations (Rogers and Harpending 1992; Harpending et al. 1993; Sherry et al. 1994). The timing of these population expansions and separations was estimated as Rogers's *J* statistic (Rogers 1995) with a program provided by H. Harpending. Two substitution rates of 10.3% and 20.5% per site, per million years, for the 328 bp of the HVI common to all populations (Sherry et al. 1994; Horai et al. 1995) were used. The two rates are a slow rate and very fast rate, with standard errors of 1.35% and 3% per million years, respectively (Sherry et al. 1994; Bonatto and Salzano 1997). The 95% confidence interval (CI) was calculated as described by Sherry et al. (1994).

An analysis of molecular variance (AMOVA) (Excoffier et al. 1992) was used to determine the significance of the differences among samples, by use of the computer package Arlequin (Schneider et al. 1997). The sums of squared sequence differences were calculated between all pairs of individuals within a population, between individuals in the same group, and between groups. The variances were tested against the distributions of the theoretical values according to the method described by Excoffier et al. (1992).

Results

Of the 152 individuals from the Norris Farms cemetery for which mtDNA analysis was attempted, 102 (67.1%) could be assigned to one of the four primary haplogroups, whereas 6 individuals (3.9%) did not possess any of the characteristic markers and were designated as belonging to a group designated "other" (table 2); the remaining 44 samples (29%) did not contain sufficient DNA for analysis. For a few samples that yielded only a faint band after PCR of a particular marker, a reamplification of 25 cycles was performed with DNA from a gel slice containing the band. This was most often necessary for the *Hin*cII site primer set (both old and new) and for the old *Alu*I site primer set.

The 9-bp deletion and the *Alu*I restriction-site loss were further verified by asymmetric PCR and direct sequencing of the PCR products in five individuals with the 9-bp deletion (38.5 % of the total with the deletion) and in five individuals with the *Alu*I site loss (55.6% of the total) as described elsewhere (Stone and Stoneking 1993). The sequences confirmed the PCR and restriction map results, and no other nucleotide substitutions were observed. The *Hin*cII 13259 site loss was further verified by *Alu*I digestion, as the A → G substitution at nucleotide 13263 also creates a new *Alu*I site at nucleotide 13262. Every mtDNA with the *Hin*cII site loss also had the corresponding *Alu*I site gain.

The HVI was sequenced in 52 individuals (34%) from the total sample. These individuals included 12 with mtDNA classified as group A, 7 as group B, 25 as group C, 5 as group D, and 3 as "other." Twenty-five distinct lineages were found (table 3). Two of these lineages (2 and 25) were excluded from further analyses, as likely cases of contamination (despite multiple independent extractions of these samples). The sequence of lineage 25 (from burial 200) matched the sequence of the primary author (A.S.). The lineage 2 sequence was identical to one found in two Finnish individuals (Lahermo et al. 1996) and segregates with the reference sequence in phylogenetic analyses (data not shown). Although this sequence does not match the author's sequence or that of the primary osteologist involved with the sample, this lineage may be the result of contamination from an unknown source. Additional investigation of the mtDNA sequences of individuals involved with the skeletal collection may reveal the possible source for this sequence. Lineage 25 was from the "other" group (as is the primary author), whereas extracts from the individual with lineage 2 did have the gain of the *Hae*III site that characterizes haplogroup A in Native Americans; however,

Table 2

Frequencies of the Four Primary mtDNA Clusters A–D

Population (*n*)	mtDNA Lineage Cluster					*h*[a]	Reference
	A	B	C	D	Other		
Contemporary Na-Dene:							
Dogrib (154)	90.9	0	2	0	7.1	.17	Torroni et al. (1993), Merriwether et al. (1995)
Haida (25)	96	0	0	4	0	.08	Torroni et al. (1993)
Navajo (48)	58.3	37.5	0	0	4.2	.53	Torroni et al. (1992)
Apache (25)	64	16	12	8	0	.57	Torroni et al. (1993)
Contemporary Amerind:							
Bella Coola (25)	60	8	8	20	4	.61	Torroni et al. (1993)
Nuu-Chah-Nulth (15)	40	6.7	13.3	26.7	13.3	.78	Torroni et al. (1993)
Chippewa (15)	26.7	13.3	33.3	0	26.7	.78	Torroni et al. (1993)
Mohawk (18)	46.4	10.5	13.8	.6	28.7	.71	Merriwether et al. (1995)
Ojibwa (28)	64.3	3.6	7.1	0	25	.54	Torroni et al. (1993)
Oklahoma Muskoke (71)	36.6	15.5	9.9	38	0	.70	Merriwether et al. (1995)
Pima (30)	6.7	50	43.3	0	0	.58	Schurr et al. (1990), Torroni et al. (1992, 1993)
Zapotec (15)	33.3	33.3	33.3	0	0	.71	Torroni et al. (1994*a*)
Maya (27)	51.9	22.2	14.8	7.4	3.7	.68	Torroni et al. (1992)
Guatuso (20)	85	15	0	0	0	.27	Torroni et al. (1994*c*)
Kuna (79)	77.2	22.8	0	0	0	.36	Torroni et al. (1993), Batista et al. (1995)
Ngobe (62)	67.7	32.3	0	0	0	.44	Torroni et al. (1993), Kolman et al. (1996)
Teribe (20)	80	20	0	0	0	.34	Torroni et al. (1994*c*)
Bribri/Cabecar (24)	54.2	45.8	0	0	0	.52	Torroni et al. (1993)
Ticuna (28)	17.9	0	32.1	50	0	.64	Schurr et al. (1990), Torroni et al. (1992)
Mataco (28)	10.7	35.7	0	53.6	0	.60	Torroni et al. (1993)
Yanomama (24)	0	16.7	54.2	29.2	0	.62	Torroni et al. (1993)
Quechua (19)	26.3	36.8	5.3	31.6	0	.73	Merriwether et al. (1995)
Aymara (172)	6.4	67.4	12.2	14	0	.51	Merriwether et al. (1995)
Atacamenos (63)	14.3	71.4	9.5	4.8	0	.47	Balliet et al. (1994), Merriwether et al. (1995)
Huilliches (38)	5.3	28.9	18.4	47.4	0	.67	Balliet et al. (1994)
Mapuches (58)	5.3	31	20.6	29.3	10.3	.78	Baillet et al. (1994)
Peneunche (100)	2	9	37	52	0	.59	Merriwether et al. (1995)
Ancient groups:							
Great Salt Lake Fremont (32)	0	73	13	7	7	.45	Parr et al. (1996)
Kaweskar (19)	0	0	15.8	84.2	0	.28	Fox (1996)
Aonikenk (15)	0	0	26.7	73.3	0	.42	Fox (1996)
Yamana (11)	0	0	90.9	9.1	0	.18	Fox (1996)
Selknam (13)	0	0	46.2	46.2	7.7	.61	Fox (1996)
Norris Farms Oneota (108)	31.5	12	42.6	8.3	5.6	.70	Present study

[a] Calculated as $(1 - S[xi)2]N/(N - 1)$.

this mutation has also been found in two Caucasians (Cann et al. 1987).

Of the remaining 23 Norris Farms mtDNA lineages, 29 polymorphic sites were discovered in 353 bp sequenced. Of the polymorphic sites, 28 were transitions, and one was an A↔C transversion. The most common lineage occurred in 13 individuals, although most lineages (17) were found in only one person. Each individual with one of the four primary clusters defined by marker analysis also had corresponding HVI mutations (table

Table 3

mtDNA Lineages at Norris Farms Cemetery

	Changes, from Reference Sequence (in Parentheses, below Nucleotide Positions)																																
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
	0	0	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3
	9	9	1	1	2	8	8	8	8	9	0	1	1	2	2	6	6	7	7	7	9	9	9	1	1	2	2	4	5	6	8		
	3	4	1	4	6	4	6	8	9	2	9	6	7	3	7	1	5	0	4	8	0	5	8	1	9	5	7	3	7	2	0		
Group[a]	(T	T	C	C	T	C	C	C	T	C	T	A	T	C	A	C	A	C	G	C	C	C	T	T	G	T	C	A	T	T	C)		*n*
A:																																	
1	.	.	T	.	.	.	.	.	.	T	.	.	.	T	.	.	.	.	.	.	T	.	.	.	A	.	.	.	.	C	.		4
2	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	T	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.		1
3*	.	C	T	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	T	.	.	.	A	.	.	.	.	C	.		1
4	.	.	T	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	T	.	.	.	A	.	.	.	.	C	.		1
5	.	.	T	.	.	.	.	.	.	.	C	.	.	T	.	.	.	.	.	.	T	.	.	.	A	.	.	.	.	C	.		1
6*	.	.	T	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	A	.	T	.	.	.	A	.	.	.	.	C	.		1
7*	.	.	.	.	.	T	.	.	.	.	.	.	.	T	.	T	.	.	.	.	T	.	.	.	A	.	.	.	.	C	.		1
8*	.	.	T	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	T	T	.	.	A	.	.	.	.	C	.		1
9*	.	.	T	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	T	.	.	.	A	C	.	.	.	C	.		1
B:																																	
10*	.	.	.	.	.	.	.	.	C	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.		1
11	.	.	.	.	.	.	.	.	C	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		4
12	.	.	T	.	.	.	.	.	C	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		1
13*	.	.	.	.	C	.	T	.	C	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		1
C:																																	
14*	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	C	.	.	C	T	.	.	C	.		1
15	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	C	.	.	C	T	.	.	.	.		13
16	.	.	.	.	C	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	C	C	.	C	T	.	.	.	.		1
17	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	C	.	.	.	T	.	.	.	.		7
18*	.	.	.	.	.	.	.	T	.	.	.	.	.	T	.	.	.	.	.	.	.	.	C	.	.	.	T	.	.	.	.		1
19*	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	C	.	.	.	.	.	C	.	.	.	T	.	.	.	.		1
20*	.	.	.	T	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	C	.	.	C	T	.	.	.	.		1
D:																																	
21*	.	.	T	.	.	T	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	C	.		1
22	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	C	T		3
23	.	.	.	.	.	.	.	.	C	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	C	.		1
Other:																																	
24*	C	.	.	.	.	.	.	.	C	.	.	.	.	T	G	.	.	.	.	T	.	.	.	.	.	.	.	.	C	.	.		2
25	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.		1

Note.—Polymorphic nucleotide sites are numbered according to the report by Anderson et al. (1981).

[a] Asterisks denote lineages not present in other Native Americans.

Table 4

Sample Size, Haplotype Diversity, Sequence Diversity, Tajima's *D* Values, and Percentage of Single Lineages for the Prehistoric Oneota and 12 Contemporary Native American Groups

Population (*n*)	*h*	π	Tajima's *D*	Percentage of Single Lineages
Oneota (50)	.91	.015	−.7315	73.9
Nuu-Chah-Nulth (63)	.95	.016	−.1142	46.4
Bella Coola (40)	.9	.015	.0536	9.1
Yakima (42)	.89	.015	−1.1159	80
Ngobe (46)	.77	.013	1.6835	53.3
Kuna (63)	.59	.01	1.5187	42.9
Huetar (27)	.71	.011	.4132	54.5
Yanomami (50)	.83	.011	−1.3603	75
Xavante (25)	.68	.009	.4386	25
Gaviao (27)	.87	.013	.1071	14.3
Zoro (30)	.77	.012	−.1973	44.4
Mapuche (39)	.92	.018	.355	47.4
Haida (40)	.68	.007	−1.1179	55.6

3). Ten of the lineages from the Norris Farms population have also been discovered in these modern Native Americans; 13 lineages (denoted by an asterisk in table 3) were previously unreported.

Haplotype diversity (*h*), nucleotide diversity (π), and Tajima's *D* (Tajima 1989) were calculated for the Norris Farms population and for 12 other Native American populations (table 4). For the Yanomami lineages, ambiguous sites were not included when the number of HVI lineages was determined. None of the populations had statistically significant Tajima's *D* values. The Norris Farms population had a high percentage of rare or single lineages (73.9%), whereas the average percentage in modern populations is 45.7%.

A network (Bandelt et al. 1995) of the Norris Farms sequences portrays the relationships between lineages and helps resolve the parallel and reverse mutations at hypervariable sites (fig. 1). Five distinct groups of lineages are indicated that correspond to the haplogroup designations. This network illustrates the hypervariable nature of sites 16111, 16189, and 16325 (three parallel mutations) and sites 16126, 16184, and 16362 (two parallel mutations), which were also found to be hypervariable by Hasegawa et al. (1993). The network also illustrates the distinct position of lineage 24 compared with the lineages from the four major groups.

Phylogenetic trees were constructed to examine the relationships between mtDNA lineages found in ancient and contemporary Native Americans as well as between lineages in Native Americans and Mongolians. Haplogroup designations from the marker analysis were superposed upon the tree, to examine the robustness of these groups. Data from modern Native Americans included all 12 tribes listed in table 4. For 3 of these populations (Mapuche, Bella Coola, and Haida), only HVI sequence data were available. In 5 additional populations, only a limited number of the markers characteristic of Native American mtDNA haplogroups were tested. The Nuu-Chah-Nulth were tested for the 9-bp deletion and the *Hin*cII site loss, whereas the Yakima, Xavante, Zoro, and Gaviao were surveyed only for the 9-bp deletion.

Mongolians have been identified, by several studies, as the possible source population for migrants to the New World (Turner 1984; Kolman et al. 1996; Merriwether et al. 1996). Mongolian mtDNA haplotypes include the four primary Native American haplogroups, which are typically rare in Asia, as well as additional, reportedly Asian-specific, clusters of lineages. Tibetan and central Chinese groups also possess a high percentage of the four common Native American mtDNA groups (Ballinger et al. 1992; Torroni et al. 1994*b*); however, sequence data were not reported from these populations. Both mtDNA sequence and haplotype marker data were gathered from 103 Mongolians by Kolman et al. (1996). The haplotype data in this population were obtained by analyzing seven RFLP sites (including three sites that define Native American haplogroups) and the 9-bp deletion.

A total of 328 bp (nucleotides 16056–16383) of the HVI common to all sequences were used for these analyses. Using the quartet-puzzling method to relate Native American and Mongolian sequences with maximum likelihood resulted in a poorly resolved tree. Of 14,463,090 quartets, 49.3% were unresolved, which indicates that these data are not good for this type of analysis. The γ rate heterogeneity parameter α was estimated from the data, giving $\alpha = .30$. This value of α indicates a very high level of heterogeneity in the rate of substitution between sites, with certain sites being

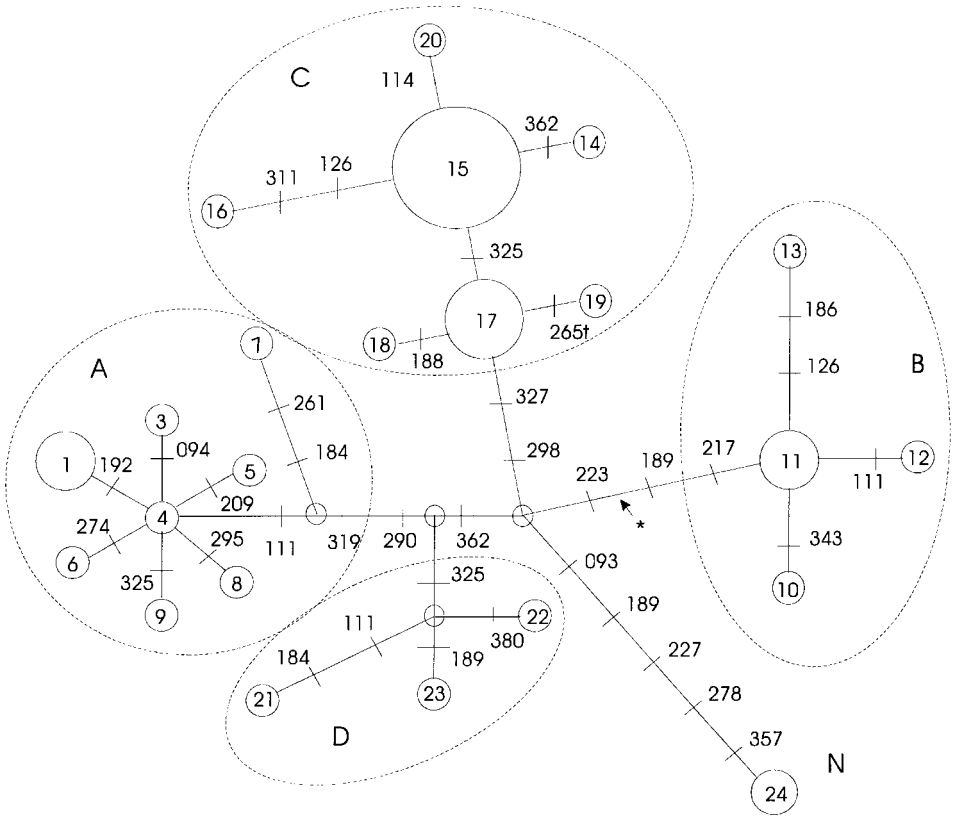

Figure 1 Median network of Norris Farms mtDNA HVI lineages. The Cambridge reference sequence is marked with an asterisk. Each mutation away from the reference sequence is represented by a tick mark, with the location of the substitution (−16,000) listed. Circles with numbers correspond to the sequences given in table 3, with larger circles representing sequences that are more frequent in the population. The haplogroup designation, based on restriction sites and the 9-bp deletion, is superposed on the network.

highly variable. For these data, 187 (57%) of 328 sites were invariant. Wakeley (1993) calculated $\alpha = .47$ for 250 bp (nucleotides 16130–16379) of HVI in humans.

Bootstrap values for the neighbor-joining trees relating these sequences were also low. In the neighbor-joining tree, A and B lineages form fairly distinct groups (fig. 2). Some Mongolian group G and F sequences are included in the deep peripheral branches of the C cluster, as is one of the Yanamami X7 lineages. Typically, group D lineages are rather dispersed, with very low bootstrap support. This results from the characteristic group D mtDNA HVI mutations that include those at nucleotide 16223 (also found in group A and C lineages) and at nucleotides 16325 and 16362, which fall into the highly variable class of sites (Hasegawa et al. 1993) found in many other lineage clusters. The Yanomami haplogroups X6 and X7, identified by Easton et al. (1996) as new, independent Native American haplogroups, are generally interspersed among the group D sequences, often sharing the same branch with group D sequences.

Lineages 1–4 from the Nuu-Chah-Nulth typically associate with three Mongolian lineages classified as group F lineages by Kolman et al. (1996). Norris Farms lineage 24 clusters with four different Mongolian lineages that are also classed as group F. In some trees, lineage 24 did group with the Nuu-Chah-Nulth lineages (data not shown). In table 5, the sequences of these Native American lineages are given, together with those from associated Mongolian lineages. The Nuu-Chah-Nulth lineages 1–4, three Mongolian lineages, and Norris Farms lineage 24 all share substitutions at nucleotides 16223 and 16278. Norris Farms lineage 24 also shares substitutions at nucleotides 16093, 16223, and 16357 with two other Mongolian lineages.

None of the above lineages appear to have much in common with the Yanomami X6 and X7 lineages. The X6 and X7 lineages share a substitution at nucleotide 16325. They tend to associate with group D in phylogenetic trees, sharing (to varying degrees) characteristic substitutions as well as private substitutions at nucleotides 16104 and 16294. According to the limited restriction sites surveyed, X6 lineages would also fall into

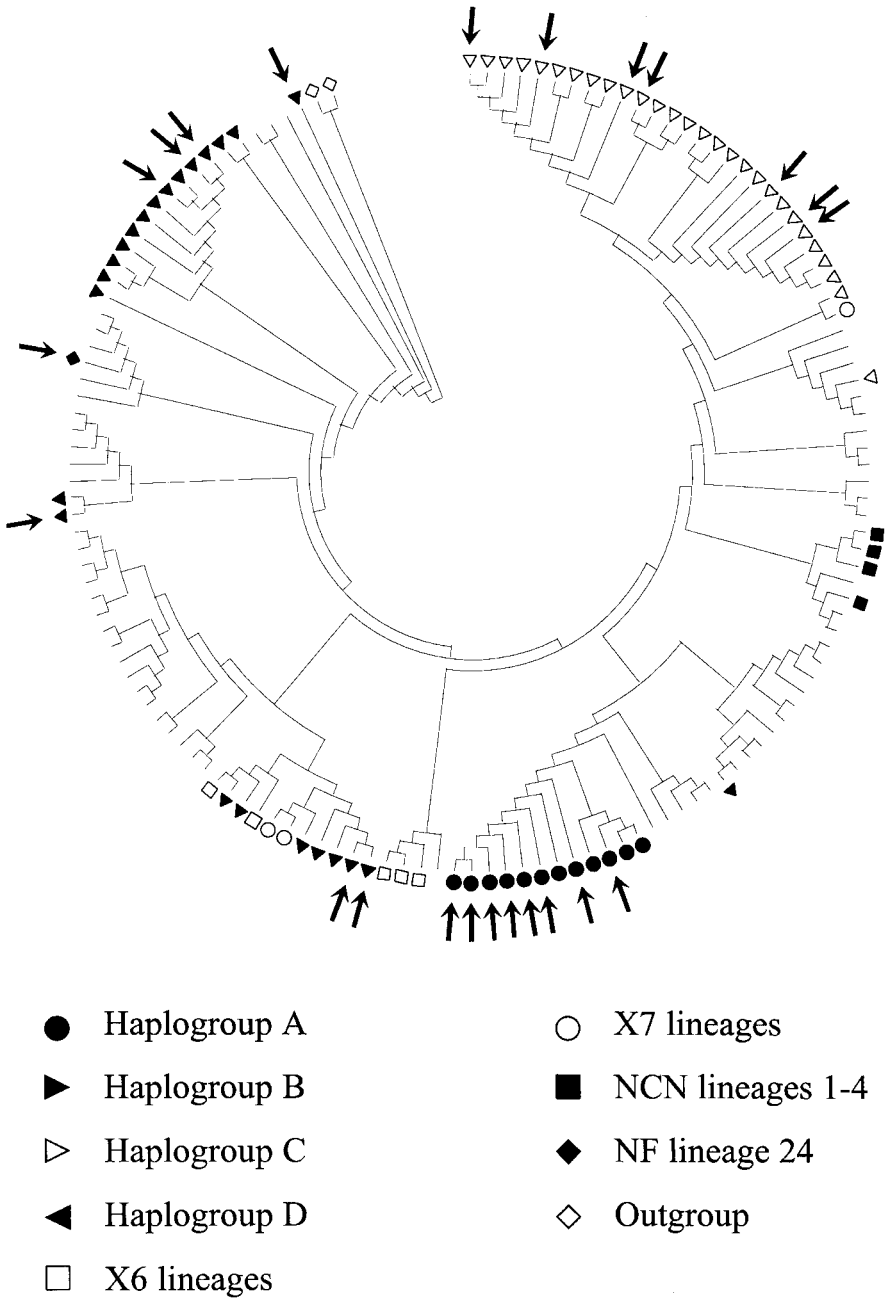

Figure 2 Neighbor-joining tree of Native American and Mongolian lineages, constructed with Tamura-Nei distance and the program MEBoot. The haplogroup designation of lineages is indicated by symbols. Unlabeled branches are Mongolian lineages that do not fall in haplogroups A–D, according to restriction analysis. Norris Farms lineages are indicated by arrows. NCN lineages 1–4 are found in the Nuu-Chah-Nulth (Ward et al. 1991), and the X6 and X7 lineages are found in the Yanomami (Easton et al. 1996).

group F, whereas X7 lineages would fall into group G of the Mongolian lineages. Lineages in group F do not cluster but are scattered throughout the tree. Mongolian groups F and G, as defined by Kolman et al. (1996), do not correspond with Asian groups F and G as defined by Ballinger et al. (1992) or with Tibetan groups F and G as defined by Torroni et al. (1994*b*). Because the full complement of restriction sites was not surveyed, it cannot be determined into which Asian haplotype groupings the Mongolian F and G lineages fall. It is likely that they fall into several of the Ballinger et al. (1992) haplotype groups, since group F and group G as defined by Kolman

Table 5

Native American mtDNA Lineages that Do Not Appear to Belong to Haplotype Groups A–D and Associated Mongolian mtDNA Lineages

Population and Lineage	Changes from Reference Sequence (in Parentheses)																													Haplotype Group
	16092 (T	16093 T	16104 C	16127 A	16129 G	16145 G	16189 T	16193 C	16209 T	16213 G	16223 C	16227 A	16234 C	16256 C	16278 C	16294 C	16295 C	16297 T	16298 T	16309 A	16311 T	16316 A	16325 T	16330 T	16337 C	16347 T	16353 C	16357 T	16362 T)	
Nuu-Chah-Nulth:																														
1	.	.	.	.	.	.	C	.	.	A	T	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?
2	.	.	.	.	.	.	.	.	.	A	T	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?
3	.	.	.	.	.	.	.	.	.	.	T	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	?
4	.	.	.	.	.	.	.	.	.	.	T	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	C	?
Mongolia:																														
3n4	.	.	.	.	.	A	.	.	.	.	T	G	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	C	F
8n22	.	C	.	.	.	.	.	.	.	.	T	.	.	.	T	.	.	.	.	.	.	.	C	.	.	.	.	.	C	F
12n1	.	.	.	.	.	.	.	.	.	.	T	G	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	C	F
17n4	.	.	.	.	A	.	C	.	.	.	T	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	F
13n1	.	.	.	.	A	.	C	.	C	.	T	.	.	.	.	.	T	.	.	.	C	.	.	.	.	.	.	C	.	F
17n10	.	C	.	.	.	.	.	T	.	.	T	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	C	.	F
17n26	.	C	.	.	A	.	.	T	.	.	T	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	C	.	F
Norris Farms:																														
24	.	C	.	.	.	.	C	.	.	.	T	G	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	C	.	Other

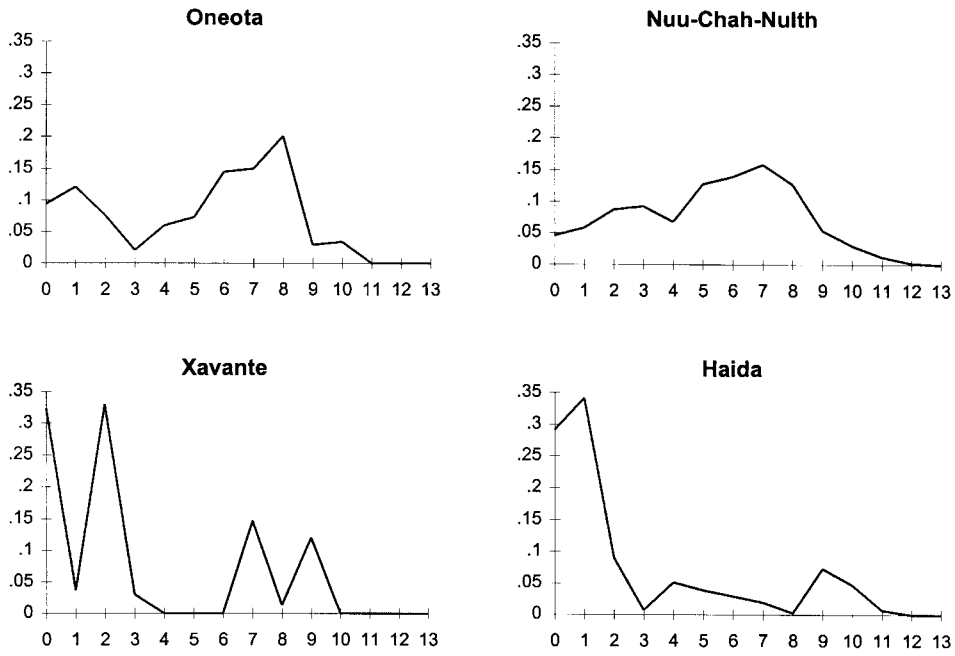

Figure 3 Distributions of pairwise differences from a selection of populations that show the different patterns present in Native American groups. The *X*-axis is the pairwise differences in units of mutational difference, and the *Y*-axis is the frequency of each pairwise difference.

et al. (1996) correspond, respectively, to groups F, B, or I and groups E or G in Ballinger et al. (1992).

AMOVA

AMOVAs of Native American mtDNA data illustrate that most of the variation is found within populations (>74%, $P = .00$) and among populations within regions (16.4%–19.2%, $P = .00$). The smallest percentage of variation was accounted for by regional differentiation. When the regions were defined as North American Amerinds, Central Americans, South Americans, and Haida, the percentage of variation among regions was 7.42% ($P = .03$), whereas when the Haida were included with North Americans, this value was 5.12% ($P = .06$). When only two regions, North America and Haida, were examined, the variation among regions (9.96%) was not significant ($P = .21$), which indicates that mtDNA sequences from the Haida, classified as Na-Dene speakers by Greenberg (1987), are not significantly different from those found in North American Amerind speakers.

Pairwise Comparisons of mtDNA Sequences

Two general patterns are evident among the mismatch distributions of Native Americans (fig. 3). Seven populations (Haida, Ngobe, Huetar, Kuna, Xavante, Gaviao, and Zoro) have more ragged bimodal distributions, apparent by visual inspection and reduced mtDNA diversity. The second distribution pattern is present in five populations, including the Norris Farms Oneota, and is essentially a unimodal distribution. Several populations also have a fairly high number of related lineages. This is most likely due to inbreeding, particularly in the Oneota population, for which related individuals were certainly included in the sample.

The intermatch distributions between Native American populations (not shown) and between Mongolians and Native American populations (such as the Oneota in fig. 4) do not lead the mismatch distributions. Instead, they resemble the distribution found when one population is subdivided and compared. Thus, according to these data, Mongolians and Native Americans look like members of the same population that began an expansion ~95,000 years before the present (B.P.) (48,000 years under the faster rate). Similar results were obtained when the Mongolian data were subdivided into Dariganga and Khalkha cultural groups and compared to the Norris Farms Oneota.

When Native American mtDNA sequences are subdivided according to haplogroup rather than by population group, additional features are revealed (fig. 5). In haplogroups A–D, the peaks of the mismatch distributions are at two to three mutational units and the estimation of expansion dates (table 6) is 23,000–37,000 years B.P. under the rate of 10.3% per million years (11,000–19,000 years under the faster rate), whereas the intermatch distribution peaks are from five to eight units with mean divergence dates estimated to be

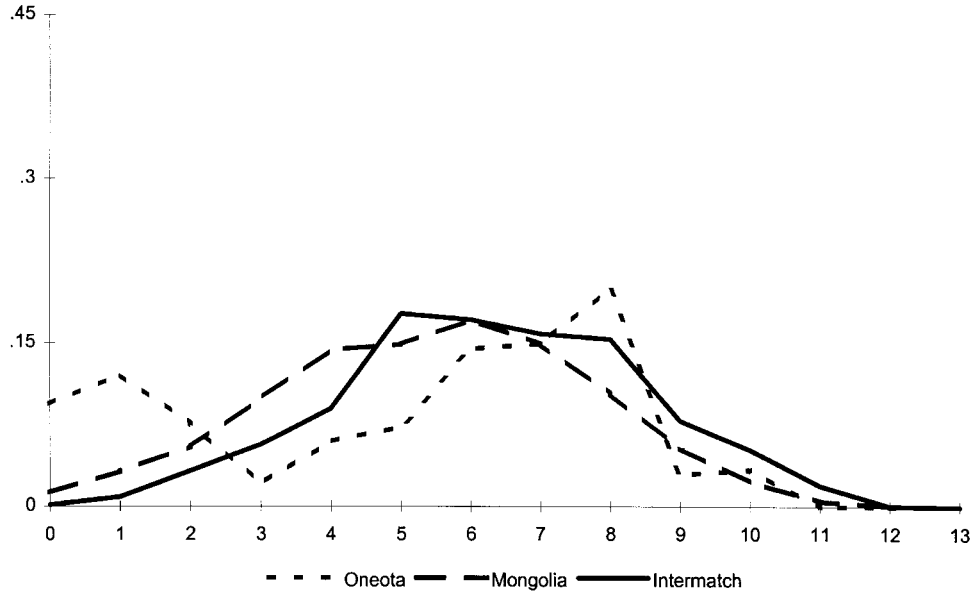

Figure 4 Mismatch and intermatch distributions between the Norris Farms Oneota and Mongolians. The *X*-axis is the pairwise differences in units of mutational difference, and the *Y*-axis is the frequency of each pairwise difference.

68,000–122,000 years B.P. (34,000–61,000 years). If haplogroup A is examined in the Haida, the only Na-Dene population, the expansion time is 11,000 years (6,000 years). When the intermatch distribution is examined between group A lineages in Amerinds and group A lineages in the Haida, it does not lead the mismatch distributions. Tests of the sequences within each of the four major haplogroups give negative Tajima's *D* values, with $P < .01$ in all cases (table 6).

Discussion

Although the sample surely includes closely related individuals, a large amount of mtDNA diversity was found in the prehistoric Norris Farms population. All four of the primary Native American mtDNA haplogroups, as well as an additional group, were found in this population. The frequencies of the four main haplogroups (table 2) and the level of mtDNA sequence diversity (table 4) in the Norris Farms population fall within the range of frequencies found in other Amerind populations, as would be expected from the geographical location of this late prehistoric site. Some modern populations, in particular the Haida, Kuna, and Xavante, exhibit much less diversity, probably as the result of founder effect and drift (Ward et al. 1993, 1996; Batista et al. 1995). It does not appear that European contact, on the whole, significantly altered patterns of Amerind mtDNA variation, despite the accompanying sudden and drastic decrease in population size (Thornton 1987; Ubelaker 1988, 1992), although some reduction in the number of rare lineages may have occurred in many populations (table 4). The Norris Farms Oneota possessed a high percentage of single lineages (73.9%) compared to most modern populations. However, this could also reflect Oneota population history. Additional precontact populations should be examined to determine whether a high number of rare lineages is a general feature.

The questions of whether a severe bottleneck occurred during the initial colonization of the Americas, how many migrations took place, and when colonization occurred have been hotly debated. In addressing the first question, Wallace and coworkers have favored the bottleneck scenario and have claimed that only four haplotypes, represented by only four or five founding variants, were present in the initial colonizers of the New World (Wallace and Torroni 1992). Sequence data also appeared consistent with the presence of only four clusters of mtDNA lineages in the Americas (Ward et al. 1991; Horai et al. 1993), although these researchers contested the hypothesis of a severe bottleneck during colonization, citing the amount of diversity present in Native Americans. Few researchers, however, have examined both the HVI sequence and all four markers in their samples, thus causing some confusion when we attempt to compare data and conclusions about the number of lineage clusters found in the Americas.

Recently, many authors have focused on the number of founding lineages to enter the New World as determined by the presence of these lineages in both Asia and

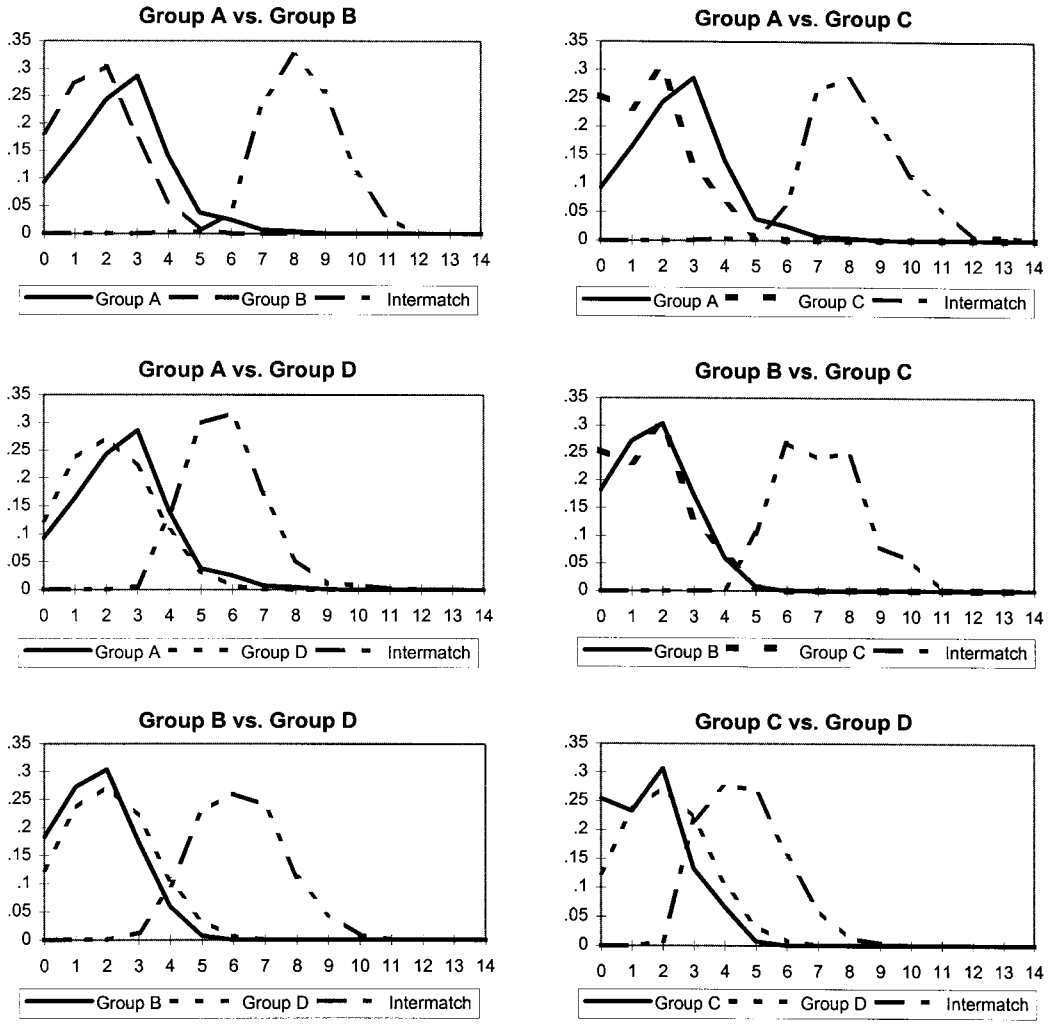

Figure 5 Mismatch and intermatch distributions between haplogroup HVI lineages. The *X*-axis is the pairwise differences in units of mutational difference, and the *Y*-axis is the frequency of each pairwise difference.

the New World, the positions of the lineages as nodal in the phylogenetic tree, and the presence of the lineages throughout the New World (Torroni et al. 1993). Bailliet et al. (1994) suggested the presence of a fifth group of related mitochondrial DNA lineages that does not possess any of the changes found in the four sites identified by Schurr et al. (1992). They suggested that this fifth lineage cluster corresponded to cluster I in the Nuu-Chah-Nulth and Yakima and is characterized by a transition at position 16278 in the HVI. This hypothesis was difficult to evaluate, however, because those lineages were not subject to complete marker analysis. Torroni and Wallace (1995) criticized the evidence for a fifth lineage by pointing out that the absence of any of the characteristic markers could also be the result of admixture or of a second mutation that disrupted the restriction site rather than an additional Asian haplotype that was brought to the New World. To clarify an "anomalous" mtDNA pattern, they recommended both high-resolution haplotype analysis and HVI sequence analysis. Bailliet et al. (1994) also suggested that haplogroups A, D, and C were represented by more than one founder lineage, on the basis of presence or absence of a *Hae*III site at nucleotide 16517. Forster et al. (1996) postulated six founding sequences for Amerinds (one lineage from groups B, C, D, and X and two from group A) by analyzing HVI sequences from Native Americans and Asians. Their group X is comprised of sequences with the characteristic mutation at nucleotide 16278. Easton et al. (1996) and Merriwether and Ferrell (1996) proposed additional founding haplogroups on the basis of the lack of one of the four markers and the presence (type X7) or absence (type X6) of the *Hae*III site at nucleotide 16517. Easton et al. (1996) performed both

Table 6

Sample Sizes, Nucleotide Diversity, and Expansion Times, Calculated by Use of Two Rates of Mutation—and Corresponding 95% CI and Tajima's *D* Value for Each Haplogroup

Haplogroup (n)	π	Time (95% CI) (Years B.P.)		Tajima's *D*
		10.3% Rate	20.5% Rate	
A (179)	.007	37,000 (25,000–57,000)	19,000 (12,000–30,000)	−1.9694
B (113)	.0052	25,000 (16,000–41,000)	12,000 (8,000–21,000)	−1.9873
C (57)	.0046	22,000 (13,000–40,000)	11,000 (6,000–21,000)	−1.8047
D (70)	.0063	31,000 (19,000–51,000)	15,000 (9–,000–27,000)	−2.1172

marker and sequence analysis and tentatively characterized these X6 and X7 groups by mutations in HVI at nucleotides 16223, 16234, and 16324 and at nucleotides 16256, 16353, and 16362, respectively.

The results from the Norris Farms population are the first from a survey of a single Native American population with highly diverse mtDNA lineages (including lineages from all four primary clusters) to include both marker data from all four sites as well as HVI sequence data. These data, as well as new comparative data from Mongolians (Kolman et al. 1996), clarify the debate concerning the number of lineage clusters present in the Americas.

The presence of a fifth lineage cluster in the Americas is supported by one lineage found in two individuals in the Norris Farms population. Lineage 24 in the Norris Farms population does not possess any of the four markers characteristic of clusters A–D as defined by Schurr et al. (1992). This lineage also does not appear to be the result of contamination (as was lineage 25 and, probably, lineage 2). Instead, lineage 24 groups with Mongolian sequences that belong to Asian haplogroup F in Kolman et al. (1996). Moreover, as noted by Bailliet et al. (1994) and Forster et al. (1996), Nuu-Chah-Nulth lineages 1–4 probably do not belong to one of the four primary Native American haplogroups. In this research, they cluster with Norris Farms lineage 24 in phylogenetic trees of Native American lineages and with group F Mongolian lineages in phylogenetic trees of Native American and Mongolian lineages (fig. 2). These lineages, along with the lineage represented by Windover sample 94 (Hauswirth et al. 1994), all possess the mutation at nucleotide 16278 that Bailliet et al. (1994) suggested is characteristic of a fifth cluster of mtDNA lineages in Native Americans. Because haplogroup F defined by Kolman et al. (1996) may correspond to more than one Asian haplogroup defined by high-resolution RFLP analysis (Ballinger et al. 1992), it is difficult to determine whether Norris Farms lineage 24 belongs to the same haplogroup as Nuu-Chah-Nulth lineages 1–4 or represents a sixth group of lineages. In addition, phylogenetic analyses sometimes place lineage 24 in a group with the Nuu-Chah-Nulth lineages (data not shown) and at other times place it separate from them (fig. 2). These Native American lineages all share mutations at nucleotides 16223 and 16278; however, Norris Farms lineage 24 also shares mutations at nucleotides 16092, 16223, and 16357 with other Mongolian lineages from cluster F (table 5).

The additional founding lineages X6 and X7 suggested by Easton et al. (1996) and Merriwether and Ferrell (1996) do not possess the mutation at nucleotide 16278 and thus do not correspond to the fifth cluster cited above. Phylogenetic analyses show a close relationship between the X6 and X7 lineages and other Native American lineages that fall into groups C and D (fig. 2). These lineages have been defined by the absence of one of the four Native American markers and by the presence or absence of a *Hae*III site at nucleotide 16517. This *Hae*III site, however, is not a particularly informative marker, because it is polymorphic in three of the four primary Native American haplogroups, as well as in other groups, and the polymorphism seems to be due to a single hypervariable nucleotide (16519), according to sequence data from Horai et al. (1993) and from Horai and Hayasaka (1990). Thus, as noted by Forster et al. (1996), the *Hae*III site should not be used alone to define any new haplogroup, and it may be questionable to assume that the site is informative about the number of founding lineages. As a result, the conclusion that X6 and X7 represent new founding lineages does not seem warranted, and it seems more likely that they are derived from C and D lineages.

In addition to the number of lineage clusters present in the Americas, the number of migrations into the New World has also been controversial. It has been suggested that the migrants to the New World came in one (Bonatto and Salzano 1997; Merriwether et al. 1995), three (Greenberg et al. 1986), or four (Horai et al. 1993) distinct waves. According to Horai et al. (1993), each lineage cluster was introduced into the Americas by a separate migration. This hypothesis is unlikely, given that all four clusters are found in tribes throughout the Americas and given the kinds of group dynamics and marriage patterns in small mobile populations. The "three waves of migration" theory of colonization was proposed by Greenberg et al. (1986) on the basis of linguistic, dental, and genetic evidence, and it posits separate waves of ancestral Amerindians, Na-Dene, and Eskimo-Aleuts. Ward et al. (1993), Torroni et al. (1993), and Torroni et al. (1992) proposed separate migrations for Na-Dene and Amerindian populations, perhaps from the same source population in Asia. Shields et al. (1993) noted that mtDNA lineages from Na-Dene and Eskimo-Aleuts,

as well as Chukchi from Siberia (together termed "Circumarctic" populations), were much less diverse within and among each other, compared with three Amerindian tribes, which led them to suggest a recent and rapid evolutionary radiation of Circumartic populations separate from a previous Amerindian migration into the Americas. Others have suggested that the mtDNA distances are not significantly different between Na-Dene and Amerinds and that the diversity differences may be the result of other demographic processes unrelated to the initial colonization of the Americas (Batista et al. 1995; Kolman et al. 1995; Merriwether et al. 1995). Forster et al. (1996) and Bonatto and Salzano (1997) concluded that one wave of migration occurred with a subsequent reexpansion, from Beringia or northwest North America, of the Eskimo-Aleut and Na-Dene speaking populations.

The pairwise comparison of sequences from Native Americans and Mongolians sheds some light on the debate over the number and diversity of migrant populations. The data indicate that Mongolian and Native American populations, including the Haida, have not been isolated from one another for a sufficiently long period of time to generate the mutations needed to result in a leading intermatch distribution. These intermatch distributions resemble the distribution generated when two populations diverge and then expand at approximately the same time (Harpending et al. 1993).

When the intermatch distribution is examined between group A lineages in Amerinds and group A lineages in the Haida (where 88% of lineages are group A), it does not lead the mismatch distributions. In addition, the expansion time of haplogroup A lineages in the Haida is estimated to be ~11,000 years (6,000 years) compared with the expansion time of Amerind haplogroup lineages of 37,000 years (table 6). These results are in agreement with Bonatto and Salzano's (1997) estimation of ~41,000 years for the expansion time of Amerind haplogroup A lineages. Their estimation of the expansion time of haplogroup A lineages within the Na-Dene was ~42,000 years. The difference in estimated times probably reflects the limited sample (one population) analyzed in this paper. The AMOVA analysis between the Haida and other North American populations also does not indicate that the Haida are significantly different from other Native Americans. These data thus suggest that the ancestors of the Haida were included in the initial colonization of the Americas and not the product of a later separate migration from Asia.

The Haida do exhibit reduced diversity and a ragged mismatch distribution, which may indicate recent population restriction or long-term constant size (Slatkin and Hudson 1991; Rogers and Harpending 1992). This pattern is also found in the Xavante, who have been isolated for some time, and in Chibchan populations. Among Chibchans, this pattern is likely to be the result of European contact or Chibchan ethnogenesis (Santos et al. 1994; Batista et al. 1995; Kolman et al. 1995). Watson et al. (1996) have suggested that a ragged mismatch distribution in African populations reflects "relic" populations with hunting and gathering rather than agricultural histories. These Native American populations seem to be exceptions to this, since Chibchan populations are agriculturalists and the Haida share food procurement strategies as well as geographic location with the Nuu-Chah-Nulth, who have a much smoother mismatch distribution and a high amount of sequence diversity.

The timing of New World colonization has long been contentious for archaeologists. They are divided between those who argue for a late migration of populations ~12,000–20,000 years B.P. and those who contend that it occurred earlier, >20,000 years B.P. (Szathmary 1993). Genetic data have also been used to estimate the date of colonization. In their analysis of the Nuu-Chah-Nulth mtDNA data, Ward et al. (1991) suggested that the sequence differences within lineage clusters coalesce ~8,000–15,000 years B.P. and that many of these differences occurred within Amerindian populations. Torroni et al. (1994*c*) used intrahaplogroup sequence divergence, calculated from Chibchan RFLP data, and time of diversification, estimated from archaeological, genetic, and linguistic data, to obtain dates of 22,000–29,000 years B.P. They suggest that this favors an early migration into the New World. Forster et al. (1996) propose dates of 20,000–25,000 years B.P. for the colonization of the New World, based on sequence analysis performed with the average distance to root sequences. Bonatto and Salzano (1997) suggested a time of colonization at 22,000–55,000 years B.P. on the basis of haplogroup A diversity.

The mismatch distributions of Native American HVI sequences, subdivided by haplogroup, suggest that the expansion of the ancestral population carrying these lineages began ~23,000–37,000 years B.P. if the rate of 10.3% per million years is used. Tests of the sequences within each population do not give the significant Tajima's D values that would indicate population expansion (Tajima 1989); however, Aris-Brosou and Excoffier (1996) found that high mutation-rate heterogeneity (such as $\alpha = .3$) will affect Tajima's D values in a direction contrary to population expansion. In their analysis of the Nuu-Chah-Nulth mtDNA sequences, they suggested that the slightly negative D value and the small number of polymorphic sites may indicate both rate heterogeneity and mild expansion. Six of the 13 Native American populations in table 4, including the Nuu-Chah Nulth, the Oneota, and the Haida, appear to fit this pattern. On the other hand, tests of the sequences

within each of the four major haplogroups give highly significant negative Tajima's *D* values (table 6).

Whether this expansion began in Asia, Beringia, or North America cannot be determined from these data. Given the short evolutionary time span under discussion and the small number of mutations that will occur during that time span, the times estimated by the molecular clock are subject to a large amount of variation. Although they exclude an extremely early migration into the Americas prior to ~50,000 years B.P., these data do not rule out the hypothesis of a late migration.

In conclusion, these data indicate that although the lineages from haplogroups A, B, C, and D are indeed the most common among Native Americans, they were not the only lineages brought into the New World from Asia. The mtDNA evidence does not support the three-wave hypothesis of migration into the New World. Native American mtDNA lineages are a subset of Asian lineages, and these lineages are typically rare in Asian populations. Consequently, one would not expect to see these same lineages introduced repeatedly into the Americas. Most likely, many small groups of people wandered across Beringia as they followed game and moved from resource to resource. This single wave of people contained considerable mtDNA diversity that exhibits a signature of expansion ~23,000–37,000 years B.P.. Ultimately, some of these peoples moved south into the Americas, and the rise in sea level largely blocked further migration into the New World from Asia. mtDNA is, however, only one locus, and therefore other rapidly evolving loci should be used to evaluate these results.

These data are the first population-level survey of mitochondrial DNA hypervariable region sequences from a prehistoric Native American community and the largest such survey in an ancient population. The application of molecular DNA techniques to archaeological samples has great potential to answer both site-specific and larger questions, although several researchers have cautioned that problems with authenticating results have made such research difficult and controversial (Richards et al. 1995; Handt et al. 1996). The recovery of DNA from the Norris Farms population, though not without difficulties from contamination, illustrates that authentic ancient DNA can be successfully recovered when adequate precautions are taken.

Acknowledgments

We would like to thank the Illinois State Museum, for permission to work with the Norris Farms collection. We would like to thank S. Pääbo, G. Milner, S. Sherry, and H. Harpending, for helpful discussion; Oliva Handt, for assistance with primer design; and David Steven, for computer assistance. This research was supported by an NSF dissertation improvement grant and by a grant from the L. B. S. Leakey Foundation to A.S. During 1 year of this research, A.S. was supported by a Fulbright Fellowship, in the laboratory of Dr. S. Pääbo, University of Munich.

References

Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, Drouin J, Eperon IC, et al (1981) Sequence and organization of the human mitochondrial genome. Nature 290: 457–465

Aris-Brosou S, Excoffier L (1996) The impact of population expansion and mutation rate heterogeneity on DNA sequence polymorphism. Mol Biol Evol 13:494–504

Bailliet G, Rothhammer F, Carnese FR, Bravi CM, Bianchi NO (1994) Founder mitochondrial haplotypes in Amerindian populations. Am J Hum Genet 55:27–33

Ballinger SW, Schurr TG, Torroni A, Gan YY, Hodge JA, Hassan K, Chen K-H, et al (1992) Southeast Asian mitochondrial DNA analysis reveals genetic continuity of ancient mongoloid migrations. Genetics 130:139–152

Bandelt H-J, Forster P, Sykes BC, Richards MB (1995) Mitochondrial portraits of human populations using median networks. Genetics 141:743–753

Batista O, Kolman CJ, Bermingham E (1995) Mitochondrial DNA diversity in the Kuna Amerinds of Panama. Hum Mol Genet 4:921–929

Bendall KE, Sykes BC (1995) Length heteroplasmy in the first hypervariable segment of the human mtDNA control region. Am J Hum Genet 57:248–256

Bonatto SL, Salzano FM (1997) A single and early migration for the peopling of the Americas supported by mitochondrial DNA sequence data. Proc Natl Acad Sci USA 94:1866–1871

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

Cann RL, Stoneking M, Wilson AC (1987) Mitochondrial DNA and human evolution. Nature 325:31–36

Chou Q, Russell M, Birch DE, Raymond J, Bloch W (1992) Prevention of pre-PCR mis-priming and primer dimerization improves low-copy-number amplifications. Nucleic Acids Res 20:1717–1723

Di Rienzo A, Wilson AC (1991) Branching pattern in the evolutionary tree for human mitochondrial DNA. Proc Natl Acad Sci USA 88:1597–1601

Easton RD, Merriwether DA, Crews DE, Ferrell RE (1996) mtDNA variation in the Yanomami: evidence for additional New World founding lineages. Am J Hum Genet 59: 213–225

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplogroups: applications to human mitochondrial DNA restriction data. Genetics 131:479–491

Forster P, Harding R, Torroni A, Bandelt H-J (1996) Origin and evolution of Native American mtDNA variation: a reappraisal. Am J Hum Genet 59:935–945

Fox CL (1996) Mitochondrial DNA haplogroups in four tribes from Tierra del Fuego–Patagonia: inferences about the peopling of the Americas. Hum Biol 68:855–871

Ginther C, Corach D, Penacino GA, Rey JA, Carnese FR, Hutz

MH, Anderson A, et al (1993) Genetic variation among the Mapuche Indians from the Patagonian region of Argentina: mitochondrial DNA sequence variation and allele frequencies of several nuclear genes. In: Pena SDJ, Chakraborty R, Epplen JT, Jeffreys AJ (eds) DNA fingerprinting: state of the science. Birkhäuser Verlag, Basel, pp 211–219

Greenberg JH (1987) Language in the Americas. Stanford University Press, Stanford

Greenberg JH, Turner CG II, Zegura SL (1986) The settlement of the Americas: a comparison of the linguistic, dental, and genetic evidence. Curr Anthropol 27:477–497

Hagelberg E, Clegg JB (1991) Isolation and characterization of DNA from archaeological bone. Proc R Soc Lond B Biol Sci 244:45–50

Handt O (1995) Molekulargenetische analyse alter menschlicher DNA Biology. PhD thesis, Ludwig-Maximillians-Universitat, Munich

Handt O, Krings M, Ward RH, Pääbo S (1996) The retrieval of ancient human DNA sequences. Am J Hum Genet 59: 368–376

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

Hasegawa M, Rienzo AD, Kocher TD, Wilson AC (1993) Toward a more accurate time scale for the human mitochondrial DNA tree. J Mol Evol 37:347–354

Hauswirth WW, Dickel CD, Rowold DJ, Hauswirth MA (1994) Inter- and intrapopulation studies of ancient humans. Experientia 50:585–591

Horai S, Hayasaka K (1990) Intraspecific nucleotide sequence differences in the major noncoding region of human mitochondrial DNA. Am J Hum Genet 46:828–842

Horai S, Hayasaka K, Kondo R, Tsugane K, Takahata N (1995) Recent African origin of modern humans revealed by complete sequences of hominoid mitochondrial DNAs. Proc Natl Acad Sci USA 92:532–536

Horai S, Kondo R, Nakagawa-Hattori Y, Hayashi S, Sonoda S, 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, Harihara S, et al (1996) mtDNA polymorphism in east Asian populations, with special reference to the peopling of Japan. Am J Hum Genet 59:579–590

Höss M, Pääbo S (1993) DNA extraction from Pleistocene bones by a silica-based purification method. Nucleic Acids Res 21:3913–3914

Kolman CJ, Bermingham E, Cooke R, Ward RH, Arias TD, Guionneau-Sinclair F (1995) Reduced mtDNA diversity in the Ngöbé Amerinds of Panama. Genetics 140:275–283

Kolman CJ, Sambuughin N, Bermingham E (1996) Mitochondrial DNA analysis of Mongolian populations and implications for the origin of New World founders. Genetics 142:1321–1334

Krings M, Stone A, Schmitz RW, Krainitzki H, Stoneking M, Pääbo S (1997) Neanderthal DNA sequences and the origin of modern humans. Cell 90:19–30

Kumar S, Tamura K, Nei M (1993) MEGA: Molecular Evolutionary Genetics Analysis, version 1.01. Pennsylvania State University, University Park

Lahermo P, Sajantila A, Sistonen P, Lukka M, Aula P, Peltonen L, Savontaus M-L (1996) The genetic relationships between the Finns and the Finnish Saami (Lapps): analysis of nuclear DNA and mtDNA. Am J Hum Genet 58:1309–1322

Lum JK, Rickards O, Ching C, Cann RL (1994) Polynesian mitochondrial DNAs reveal three deep maternal lineage clusters. Hum Biol 66:567–590

Merriwether DA, Ferrell RE (1996) The four founding lineage hypothesis for the New World: a critical reevaluation. Mol Phylogenet Evol 5:241–246

Merriwether DA, Hall WW, Vahlne A, Ferrell RE (1996) mtDNA variation indicates Mongolia may have been the source for the founding population for the New World. Am J Hum Genet 59:204–212

Merriwether DA, Rothhammer F, Ferrell RE (1995) Distribution of the four founding lineage haplotypes in Native Americans suggests a single wave of migration for the New World. Am J Phys Anthropol 98:411–430

Milner GR, Humpf DA, Harpending HC (1989) Pattern matching of age-at-death distributions in paleodemographic analysis. Am J Phys Anthropol 80:49–58

Morin PA, Moore JJ, Chakraborty R, Jin L, Goodall J, Woodruff DS (1994) Kin selection, social structure, gene flow, and the evolution of chimpanzees. Science 265:1193–1201

Pääbo S (1989) Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. Proc Natl Acad Sci USA 86:1939–1943

Pääbo S, Gifford JA, Wilson AC (1988) Mitochondrial DNA sequences from a 7000-year-old brain. Nucleic Acids Res 16:9775–9787

Parr RL, Carlyle SW, O'Rourke D (1996) Ancient DNA analysis of Fremont Amerinidians of the Great Salt Lake wetlands. Am J Phys Anthropol 99:507–518

Redd AJ, Takezaki N, Sherry ST, McGarvey ST, Sofro ASM, Stoneking M (1995) Evolutionary history of the COII/tRNALys intergenic 9 base pair deletion in human mitochondrial DNAs from the Pacific. Mol Biol Evol 12:604–615

Richards M, Côrte-Real H, Forster P, Macaulay V, Wilkinson-Herbots H, Demaine A, Papiha S, et al (1996) Paleolithic and Neolithic lineages in the European mitochondrial gene pool. Am J Hum Genet 59:185–203

Richards MB, Sykes BC, Hedges REM (1995) Authenticating DNA extracted from ancient skeletal remains. J Archaeol Sci 22:291–299

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

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

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Santos M, Ward RH, Barrantes R (1994) mtDNA variation in the Chibcha Amerindian Huetar from Costa Rica. Hum Biol 66:963–977

Santure SK, Harn AD, Esarey D (eds) (1990) Archaeological investigations at the Morton Village and Norris Farms 36 cemetery: reports of investigations, no 45. Illinois State Museum, Springfield

Schneider S, Kueffer J-M, Roessli D, Excoffier L (1997) Ar-

lequin: an exploratory population genetics software environment. University of Geneva, Geneva

Schurr TG, Ballinger SW, Gan Y-Y, Hodge JA, Merriwether DA, Lawrence DN, Knowler WC, et al (1990) Amerindian mitochondrial DNAs have rare Asian mutations at high frequencies, suggesting they derived from four primary lineages. Am J Hum Genet 46:613–623

Sherry S, Rogers A, Harpending H, Soodyall H, Jenkins T, Stoneking M (1994) Mismatch distributions of mtDNA reveal recent human population expansions. Hum Biol 66: 761–775

Shields GF, Schmiechen AM, Frazier BL, Redd A, Voevoda MI, Reed JK, Ward RH (1993) mtDNA sequences suggest a recent evolutionary divergence for Beringian and northern North American populations. Am J Hum Genet 53:549–562

Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. Genetics 129:555–562

Stone A, Milner G, Pääbo S, Stoneking M (1996) Sex determination of ancient human skeletons using DNA. Am J Phys Anthropol 99:231–238

Stone A, Stoneking M (1993) Ancient DNA from a pre-Columbian Amerindian population. Am J Phys Anthropol 92: 463–471

Strimmer K, von Haeseler A (1996) Quartet puzzling: a quartet maximum-likelihood method for reconstructing tree topologies. Mol Biol Evol 13:964–969

——— (1997) PUZZLE: maximum likelihood analysis for nucleotide and amino acid alignments. Ludwig-Maximillian Universitat, Munich

Szathmary EJE (1993) Genetics of aboriginal North Americans. Evol Anthropol 1:202–220

Tajima F (1989) Statistical method for testing the neutral muation hypothesis by DNA polymorphism. Genetics 123: 585–595

Tamura K (1997) MEBoot. Tokyo Metropolitan University, Tokyo

Tamura K, 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

Thornton R (1987) American Indian holocaust and survival: a population history since 1492. University of Oklahoma Press, Norman

Torroni A, Chen Y-S, Semino O, Santachiara-Beneceretti AS, Scott CR, Lott MT, Winter M, et al (1994*a*) mtDNA and Y-chromosome polymorphisms in four Native American populations from southern Mexico. Am J Hum Genet 54: 303–318

Torroni A, Miller JA, Moore LG, Zamudio S, Zhuang J, Droma T, Wallace DC (1994*b*) Mitochondrial DNA analysis in Tibet: implications for the origin of the Tibetan population and its adaptation to high altitude. Am J Phys Anthropol 93:189–199

Torroni A, Neel JV, Barrantes R, Schurr TG, Wallace DC (1994*c*) Mitochondrial DNA "clock" for the Amerinds and its implications for timing their entry into North America. Proc Natl Acad Sci USA 91:1158–1162

Torroni A, Schurr TG, Cabell MF, Brown MD, Neel JV, Larsen M, Smith DG, et al (1993) Asian affinities and continental radiation of the four founding Native American mtDNAs. Am J Hum Genet 53:563–590

Torroni A, Schurr TG, Yang C-C, Szathmary E, Williams RC, Schanfield MS, Troup GA, et al (1992) Native American mitochondrial DNA analysis indicates that the Amerindian and the Nadene populations were founded by two independent migrations. Genetics 130:153–162

Torroni A, Wallace DC (1995) mtDNA haplogroups in Native Americans. Am J Hum Genet 56:1234–1236

Turner CG (1984) Advances in the dental search for Native American origins. Acta Anthropogenetica 8:23–78

Ubelaker DH (1988) North American Indian population size, A.D. 1500 to 1985. Am J Phys Anthropol 77:289–294

Ubelaker DH (1992) Patterns of demographic change in the Americas. Hum Biol 64:361–379

Vigilant L (1990) Control region sequences from African populations and the evolution of human mitochondrial DNA Biochemistry. PhD thesis, University of California (Berkeley), Berkeley

Vigilant L, Pennington R, Harpending H, Kocher TD, Wilson AC (1989) Mitochondrial DNA sequences in single hairs from a southern African population. Proc Natl Acad Sci USA 86:9350–9354

Wakeley J (1993) Substitution rate variation among sites in hypervariable region 1 of human mitochondrial DNA. J Mol Evol 37:613–623

Wallace DC, Garrison K, Knowler WC (1985) Dramatic founder effects in Amerindian mitochondrial DNAs. Am J Phys Anthropol 68:149–155

Wallace DC, Torroni A (1992) American Indian prehistory as written in the mitchondrial DNA: a review. Hum Biol 64: 403–416

Ward RH, Frazier B, Dew-Jager K, Pääbo S (1991) Extensive mitochondrial diversity within a single Amerindian tribe. Proc Natl Acad Sci USA 88:8720–8724

Ward RH, Redd A, Valencia D, Frazier B, Pääbo S (1993) Genetic and linguistic differentiation in the Americas. Proc Natl Acad Sci USA 90:10663–10667

Ward RH, Salzano FM, Bonatto SL, Hutz MH, Coimbra CEA Jr, Santos RV (1996) Mitochondrial DNA polymorphism in three Brazilian Indian tribes. Am J Hum Biol 8:317–323

Watson E, Bauer K, Aman R, Weiss G, von Haeseler A, Pääbo S (1996) mtDNA sequence diversity in Africa. Am J Hum Genet 59:437–444

Wrischnik LA, Higuchi RG, Stoneking M, Erlich HA, Arnheim N, Wilson AC (1987) Length mutations in human mitochondrial DNA: direct sequencing of enzymatically amplified DNA. Nucleic Acids Res 15:529–542