Spatial and Temporal Stability of mtDNA Haplogroup Frequencies in Native North America

D.H. O'ROURKE,[1] M.G. HAYES,[1] AND S.W. CARLYLE[1]

Abstract Mitochondrial DNA lineage frequencies in prehistoric Aleut, eastern Utah Fremont, Southwestern Anasazi, Pyramid Lake, and Stillwater Marsh skeletal samples from northwest Nevada and the Oneota of western Illinois are compared with those in 41 contemporary aboriginal populations of North America. The ancient samples range in age from 300 years to over 6,000 years. The results indicate that the prehistoric inhabitants of North America exhibit the same level of mtDNA variability as contemporary populations of the continent. Variation in modern mtDNA haplogroup frequencies is highly geographically structured, and the prehistoric samples exhibit the same geographic pattern of variation. This indicates that differentiation of regional patterns of mtDNA lineage variation occurred early in North American prehistory (much more than 2,000 years B.P.), has remained relatively stable since its origin, and was little influenced by the disruptions hypothesized for other genetic systems as a result of population declines and relocations at contact.

Aboriginal American populations have been characterized genetically as having reduced variation relative to other world populations. The reason for the reduction in genetic variation in native American populations relative to, primarily, European populations has been attributed to acertainment bias of the studied markers, restriction of variation in small numbers of original colonists to the Americas, persistent small population size, and severe population bottleneck at contact [see reviews by Szathmary (1993) and Crawford (1998)]. At the same time, it has been clearly demonstrated that at least for native North American populations genetic variation is geographically structured (Suarez et al. 1985; O'Rourke et al. 1992; Shields et al. 1992; Lorenz and Smith 1996). All these observations and inferences are predicated on the distribution of genetic markers in contemporary populations of Amerindians. Until recently, there has been little direct evidence of the distribution of genetic variation in prehistoric native America against which to evaluate inter-

[1] Laboratory of Biological Anthropology, University of Utah, Salt Lake City, UT 84112.

Human Biology, February 2000, v. 72, no. 1, pp. 15–34.

KEY WORDS: ANCIENT DNA, NORTH AMERICAN INDIANS, MTDNA HAPLOGROUPS

pretations regarding the origin and maintenance of patterns of genetic variation in modern populations.

With the advent of ancient DNA (aDNA) analysis it has become possible to evaluate the extent and patterning of genetic variation in prehistoric populations in a fashion similar to the genetic study of modern groups. In this review we evaluate mtDNA variation in 6 ancient population samples of North America relative to 41 modern North American populations distributed throughout the continent. Analysis of the modern populations confirms substantial geographic structuring of mtDNA haplogroup frequencies (Shields et al. 1992; Lorenz and Smith 1996) in accord with earlier work on classical marker distributions in native North American populations (Suarez et al. 1985; O'Rourke et al. 1992). The ancient samples show haplogroup profiles that are most similar to those of modern populations inhabiting the same geographic areas today. This result indicates a surprising stability of mtDNA haplogroup profiles over the past 2,000+ years of the Holocene in indigenous populations of the Americas.

Materials and Methods

The Samples. Six ancient North American population samples are included in the analysis. Three of these ancient samples have been genetically characterized in our laboratory (Parr et al. 1996; Hayes 1998; Carlyle et al. 1999; O'Rourke et al. 1999), and an additional 3 have been analyzed and published by Kaestle (1997, 1998) and Stone and Stoneking (1993, 1998). Methodological details of nucleic acid extraction, amplification, and screening in these prehistoric samples can be found in the original reports.

Northern Fremont skeletal samples were excavated from the eastern margin of the Great Salt Lake (Simms 1999; Simms et al. 1991) and range in age from 500 years B.P. to over 1,500 years B.P. (Parr et al. 1996; O'Rourke et al. 1999). The Anasazi material derives from museum specimens curated at the American Museum of Natural History. It comes primarily from southeast Utah and Canyon del Muerto in northern Arizona and dates between 1,010 and 2,010 years B.P. with a mean age of nearly 1,700 years B.P. (Carlyle et al. 1999). The final ancient series characterized molecularly by us is a sample of prehistoric Aleut (Hayes 1998). These skeletal samples were excavated from the Chaluka midden on Umnak Island and burial caves on Shiprock and Kagamil islands by Hrdlicka (1945) in the 1930s and have been curated at the Smithsonian Institution since that time. These materials have yet to be directly dated, but archeological context suggests an age of approximately 2,000–4,000 years B.P. for the larger series from Chaluka and younger ages, perhaps only a few hundred years, for the burial cave material.

Molecular characterization of the other 3 ancient samples is taken from published reports. The Pyramid Lake and Stillwater Marsh samples were studied and reported by Kaestle (1997, 1998). These western Great Basin samples

have been curated at the Nevada State Museum since their archeological recovery and span an exceptionally large time range of 300 years B.P. to nearly 6,000 years B.P. The Oneota sample from western Illinois reported by Stone and Stoneking (1993, 1998) is substantially younger than the other ancient samples included in the analysis, with an age of only 700 years B.P. (Stone and Stoneking 1998).

Mitochondrial DNA diversity among modern populations was tabulated from recent research reports. Samples were included for consideration that had been screened for the suite of discrete markers that define native American haplogroups. To optimize comparability, we did not include samples whose haplogroup profile had been inferred from mtDNA hypervariable sequence data. The populations, sample sizes, and sources for all analyzed samples are given in Table 1. Sample size is of concern not only for ancient samples but also for estimation of frequencies in modern samples. Although most of the samples have sample sizes greater than 20, a few have smaller sample sizes. A few samples, notably those from Baja California, are characterized by sample sizes of less than 15, and the Paiute/Shoshone sample is based on only 9 individuals (see discussion later).

mtDNA Haplogroups. Native American mtDNA haplogroups are defined by the presence or absence of specific restriction sites and a deletion marker in the mtDNA molecule (Schurr et al. 1990; Torroni et al. 1992; Torroni, Schurr et al. 1993; Torroni, Sukernik et al. 1993; Forster et al. 1996). The relationship between these markers and haplogroup assignment is given in Table 2. Sequence data can also be used to assess haplogroup membership, although sequence support for haplogroup D is weak (Torroni, Schurr et al. 1993; Ward et al. 1991, 1993; cf. Forster et al. 1996). To maximize comparability, we have limited our attention to those groups whose haplogroup profile has been assessed using the suite of discrete markers. It is now clear that there are more than just 4 primary haplogroups defining Amerindian populations (Forster et al. 1996; Brown et al. 1998; Stone and Stoneking 1998; Smith et al. 1999).

Haplogroup X is an additional ancestral Amerindian mtDNA lineage and is evident upon screening with additional markers and through sequence analysis. Because most of the samples available for analysis have not been screened for these additional markers, it is not possible to identify haplogroup X in all these data. Accordingly, we simply group all haplotypes that do not correspond to haplogroups A, B, C, or D as "other." It should be recognized that some of these other haplogroups may well be haplogroup X and therefore may represent an ancestral aboriginal North American mtDNA lineage. Other haplogroups may also be attributed to nonnative North American admixture (Torroni, Schurr et al. 1993; Smith et al. 1999).

Although haplogroup assignment is straightforward in modern samples, it is not always so easy for ancient samples (O'Rourke et al. 1999; Carlyle

Table 1. Native American mtDNA Haplogroup Frequencies

			Haplogroup Frequency						
Locale and Group	*Code*	*n*	*A*	*B*	*C*	*D*	*Other*	*h*	*Reference*
Arctic	ARC								
Aleut (ancient)	AAL	17	0.353	0.000	0.000	0.647	0.000	0.49	Hayes (1998); this study
St. Paul Aleut	ALT	72	0.250	0.000	0.014	0.667	0.069	0.49	Merriwether et al. (1995)
Gambell Eskimo	GAM	50	0.580	0.000	0.140	0.260	0.020	0.59	Merriwether et al. (1995)
Old Harbor Eskimo	OHB	115	0.617	0.035	0.000	0.348	0.000	0.50	Merriwether et al. (1995)
Ouzinkie Eskimo	OUZ	41	0.732	0.000	0.049	0.146	0.073	0.45	Merriwether et al. (1995)
Savoonga Eskimo	SAV	49	0.939	0.000	0.000	0.020	0.041	0.12	Merriwether et al. (1995)
Inuit	INT	30	0.967	0.000	0.000	0.033	0.000	0.07	Lorenz and Smith (1996)
Siberian Eskimos	ESK	129	0.783	0.000	0.016	0.202	0.000	0.35	Torroni, Sukernik et al. (1993); Starikovskaya et al. (1998)
Subarctic	SUB								
Dogrib	DOG	166	0.916	0.000	0.018	0.000	0.066	0.16	Lorenz and Smith (1996); Merriwether et al. (1995)
Northwest Coast	NWC								
Bella Coola	BC	36	0.500	0.056	0.140	0.250	0.056	0.68	Lorenz and Smith (1996); Torroni, Schurr et al. (1993)
Haida	HAI	29	0.966	0.000	0.000	0.034	0.000	0.07	Lorenz and Smith (1996); Torroni, Schurr et al. (1993)
Nuu-Chah-Nulth	NCN	15	0.400	0.067	0.133	0.267	0.133	0.78	Torroni, Schurr et al. (1993)
Northeast	NE								
Mohawk	MOH	18	0.464	0.105	0.138	0.006	0.287	0.71	Merriwether et al. (1995)
Ojibwa/Chippewa	OJB	43	0.512	0.070	0.163	0.000	0.256	0.66	Torroni, Schurr et al. (1993)
Chippewa/Kickapoo	CHP	62	0.484	0.113	0.194	0.000	0.210	0.68	Lorenz and Smith (1996); Torroni, Schurr et al. (1993)
Oneota (ancient)	ONT	108	0.315	0.120	0.426	0.083	0.056	0.70	Stone and Stoneking (1998)
Plains	PLN								
Cheyenne/Arapahoe	CHY	26	0.308	0.115	0.346	0.154	0.077	0.77	Lorenz and Smith (1996)
Siouan	SIO	34	0.529	0.176	0.147	0.059	0.088	0.68	Lorenz and Smith (1996)

Southeast	SE								
Cherokee	CHR	16	0.000	0.313	0.313	0.000	0.375	0.71	Lorenz and Smith (1996)
Chickasaw/Choctaw	CHC	27	0.667	0.222	0.074	0.000	0.037	0.52	Lorenz and Smith (1996)
Creek/Seminole	SEM	18	0.389	0.111	0.167	0.167	0.167	0.80	Lorenz and Smith (1996)
Muskoke	MSK	71	0.366	0.155	0.099	0.380	0.000	0.70	Merriwether et al. (1995)
California	CAL								
Californian Penutian	PEN	17	0.118	0.412	0.059	0.412	0.000	0.68	Lorenz and Smith (1996)
Salinan/Chumash	SAL	11	0.455	0.182	0.091	0.273	0.000	0.74	Lorenz and Smith (1996)
Californian Uto-Aztecan	UTO	14	0.000	0.286	0.429	0.286	0.000	0.70	Lorenz and Smith (1996)
Great Basin	GB								
Washo	WAS	28	0.000	0.536	0.357	0.107	0.000	0.60	Lorenz and Smith (1996)
Paiute/Shoshone	PAI	9	0.000	0.222	0.222	0.444	0.111	0.78	Lorenz and Smith (1996)
Pyramid Lake (ancient)	PYR	19	0.105	0.316	0.000	0.526	0.053	0.64	Kaestle (1998)
Stillwater Marsh (ancient)	STM	22	0.045	0.364	0.000	0.545	0.045	0.59	Kaestle (1998)
Fremont (ancient)	FRE	34	0.000	0.735	0.118	0.059	0.088	0.45	Parr et al. (1996); O'Rourke et al. (1999)
Southwest	SW								
Navajo	NAV	58	0.517	0.414	0.034	0.000	0.034	0.57	Lorenz and Smith (1996); Torroni et al. (1992)
Apache	APA	29	0.621	0.172	0.138	0.069	0.000	0.58	Lorenz and Smith (1996); Torroni, Schurr et al. (1993)
Jemez/Taos/San Idelfonso	JEM	36	0.000	0.861	0.028	0.028	0.083	0.26	Lorenz and Smith (1996)
Zuni	ZUN	22	0.182	0.636	0.091	0.000	0.091	0.57	Lorenz and Smith (1996)
Havasupai/Hualapai/ Yavapai/Mojave	HAV	18	0.111	0.500	0.389	0.000	0.000	0.62	Lorenz and Smith (1996)
Pima	PIM	37	0.054	0.568	0.378	0.000	0.000	0.55	Lorenz and Smith (1996); Schurr et al. (1990); Torroni et al. (1992)
Anasazi (ancient)	ANZ	22	0.227	0.591	0.091	0.000	0.091	0.61	Carlyle et al. (1999); this study
Quechan/Cocopa	QUE	23	0.000	0.652	0.304	0.000	0.043	0.50	Lorenz and Smith (1996)

Table 1. Continued

Locale and Group	*Code*	*n*	*Haplogroup Frequency*					*h*	*Reference*
			A	*B*	*C*	*D*	*Other*		
Baja	BAJ								
Cochimi	COC	13	0.077	0.462	0.462	0.000	0.000	0.61	Lorenz and Smith (1996)
Kiliwa/Paipai	KIL	11	0.000	0.909	0.091	0.000	0.000	0.18	Lorenz and Smith (1996)
Kumiai	KUM	16	0.000	0.625	0.375	0.000	0.000	0.50	Lorenz and Smith (1996)
Mexico	MEX								
Mixe	MIX	16	0.625	0.313	0.063	0.000	0.000	0.54	Torroni et al. (1994)
Alta Mixtec	MXA	15	0.733	0.133	0.133	0.000	0.000	0.46	Torroni et al. (1994)
Baja Mixtec	MXB	14	0.929	0.071	0.000	0.000	0.000	0.14	Torroni et al. (1994)
Nahua/Cora	NAH	32	0.531	0.344	0.063	0.000	0.063	0.61	Lorenz and Smith (1996)
Zapotec	ZAP	15	0.333	0.333	0.333	0.000	0.000	0.71	Torroni et al. (1994)
Maya	MAY	27	0.519	0.222	0.148	0.074	0.037	0.68	Schurr et al. (1990); Torroni et al. (1992).

Note added in proof: We recently discovered that the Ojibwa and Chippewa samples are not completely independent. Reanalysis of these samples with the overlap eliminated resulted in no detectable changes to the results or inferences presented. Addition of another independent sample of Ojibwa (Scozzari et al. 1997) also resulted in no appreciable change in results, inferences, or conclusions.

Table 2. Haplogroup Definitions

	Polymorphic Marker			
Haplogroup	*Hae*III *np 663*	*9-bp Deletion*	*Hin*cII *np 13259*	*Alu*I *np 5176*
A	+	−	+	+
B	−	+	+	+
C	−	−	−	+
D	−	−	+	−

+, presence of marker.
−, absence of marker.

et al. 1999). Haplogroup assignment is based on the joint occurrence of a minimum of 4 markers (Table 2). In ancient samples, where not every primer set produces a product on every sample, it is not uncommon to be able to determine only 2 or 3 of the markers listed in Table 2. This complicates haplogroup diagnosis. If the markers observed include 1 of the diagnostic markers for a haplogroup (e.g., the 9-bp deletion that defines haplogroup B), it strongly suggests that the sample represents this haplogroup, even if not all markers were observed. However, if the markers are not diagnostic (e.g., presence of the *Alu* restriction site at np 5176), haplogroup identification is indeterminate. Thus the correspondence of marker frequencies and haplogroup frequencies may not be concordant in ancient samples, depending on the failure to observe a rate for each marker. This can become problematic when comparing samples.

For example, in the ancient Fremont material (Parr et al. 1996) 47 skeletons were examined, of which 43 yielded aDNA. Of these samples, 41 could be reliably screened for the 9-bp intergenic deletion, resulting in a frequency of this marker of 0.61 (25/41). However, because other markers could not be amplified in several samples, only 34 individuals could reliably be assigned to a haplogroup. Of these, nearly 74% possessed the deletion and would be considered haplogroup B. Yet this must be a spuriously high haplogroup frequency, because in the full series only 61% possessed the deletion. This is an extreme example, but it illustrates the difficulty posed in comparative studies of haplogroup frequencies when using samples of aDNA.

In the present study we have used haplogroup frequencies rather than marker frequencies, because for many of the modern samples marker frequencies were not given. This clearly introduces some degree of error, given the difficulties of haplogroup assignment in prehistoric material just described. As a control for such difficulties we replicated analyses using marker frequencies as surrogates for haplogroup frequencies in the ancient samples. Because this resulted in no substantive differences in result or inference, the results of this analysis are not presented.

Analysis. Ancient DNA samples are not populations in the traditional sense of the term. The individual specimens that constitute aDNA samples may span several centuries and even geographic space. Thus they are the equivalent of sampling an individual every few generations to characterize a continuous population. As a result, we have been reluctant to subject aDNA data to the same types of statistical analyses commonly used for contemporary samples, because they clearly do not conform to the distributional and sampling assumptions of the methods. Here, we base our inferences on 2 simple methods of analysis. First, we evaluated visual comparisons of haplogroup profiles from aDNA samples and regionally coherent modern samples. Second, we subjected the haplogroup frequencies to a principal components analysis and plotted the samples in the significant eigenvalue space to evaluate a more quantitative assessment of population similarity and dissimilarity. The principal components analysis was performed without factor rotation using SPSS, version 8.0, on the 4 principal Amerindian haplogroup frequencies.

Results

A graphical comparison of the haplogroup frequencies in Table 1 is given in Figure 1. Two things are immediately apparent from this figure: (1) There is considerable variation in North American populations with respect to haplogroup frequencies, and (2) the variation is geographically patterned. Populations of northern North America and the eastern United States are characterized by high frequencies of haplogroup A, defined by the *Hae*III restriction site at np 663. The obvious exceptions to this pattern are the Aleut samples, both ancient and modern, which exhibit notably lower haplogroup A frequencies and dramatically elevated frequencies of haplogroup D relative to other Arctic populations. This distinctive haplogroup profile of the Aleut distinguishes these samples from contemporary Siberian populations (data not shown) and from their North American Arctic neighbors. The other exception to the general pattern is the Cherokee sample from Oklahoma.

Populations of western and southwestern North America are generally characterized by low to zero frequencies of haplogroup A, higher frequencies of haplogroup B, and variable frequencies of haplogroups C and D. In particular, populations of the greater US Southwest have low to zero frequencies of haplogroup A, and high frequencies of haplogroup B. These populations are also typically characterized by low to zero frequencies of haplogroup D. Groups from the western Great Basin and California, in contrast, exhibit moderate frequencies of haplogroup B and high frequencies of haplogroup D.

Finally, the Mexican populations, exclusive of Baja, are characterized by moderate to high frequencies of haplogroup A, moderate frequencies of haplogroups B and C, and an absence of haplogroup D. This is a haplogroup profile not observed in populations farther north. Frequencies of other hap-

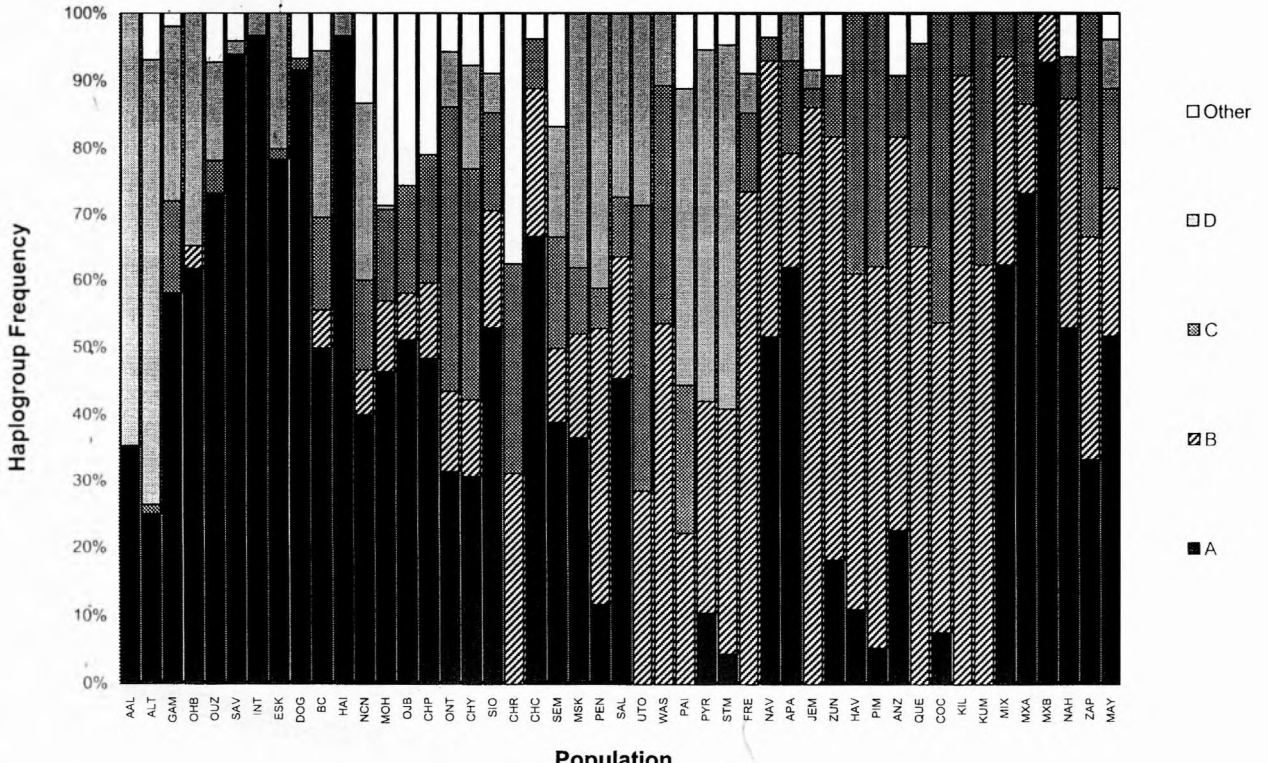

Figure 1. mtDNA profiles of native North American populations. Samples are grouped by geographic region. Sample codes are given in Table 1.

Table 3. Principal Components Analysis

Component	*Eigenvalue*	*Variance (%)*	*Cumulative %*
1	2.104	52.588	52.588
2	1.162	29.058	81.647

logroups are sufficiently low in most populations to be of little interest with respect to identifying patterns of variation.

In addition to the obvious geographic patterning in Figure 1, it is important to note that the aDNA samples exhibit haplogroup profiles that are similar to the profiles that characterize the modern populations inhabiting the same geographic regions today. This implies a temporal and geographic stability of these mtDNA haplogroup profiles that had not been anticipated. To further explore this issue, we subjected the haplogroup frequencies in Table 1 to a principal components analysis. The results are given in Table 3 and plotted in Figure 2. Two significant components (eigenvalues greater than 1) were extracted from the haplogroup frequencies and collectively account for nearly 82% of the original variation. The first component is characterized by loadings with contrasting frequencies of haplogroups A and B, whereas the second component reflects variation in haplogroups A and D.

The geographic patterning noted from a visual inspection of haplogroup frequencies (Figure 1) is apparent here as well (Figure 2), although it is not as sharply defined as in the earlier analysis. All Arctic populations are distributed on the extreme left of the component plot, and the distinctiveness of the Aleut samples is obvious. Northwest coast and subarctic populations cluster near one another along with 1 Eskimo population and a central Mexican group. Although not geographically proximate, these groups (Haida, Inuit, Savoonga Eskimo, Dogrib, and Baja Mixtec) share the characteristic of having the highest haplogroup A frequencies among all samples.

The greatest uniformity is seen among the US Southwest and Baja California populations, which plot in a tight cluster in the lower right quadrant. Curiously, the Washo of the western Great Basin, the ancient Oneota, and the ancient Fremont appear in this cluster. The Paiute/Shoshone of the Great Basin plot in the upper central portion of the component map, near the geographically proximal ancient Pyramid Lake and Stillwater Marsh samples. Two California linguistic groups, California Penutian and Uto-Aztecan, also plot in this area and are geographically not too removed from the locale of the ancient samples.

Finally, a large cluster of populations plot in the lower center of the component map and seem to share moderate frequencies of haplogroups A and B and variable frequencies of haplogroups C and D and other haplogroups.

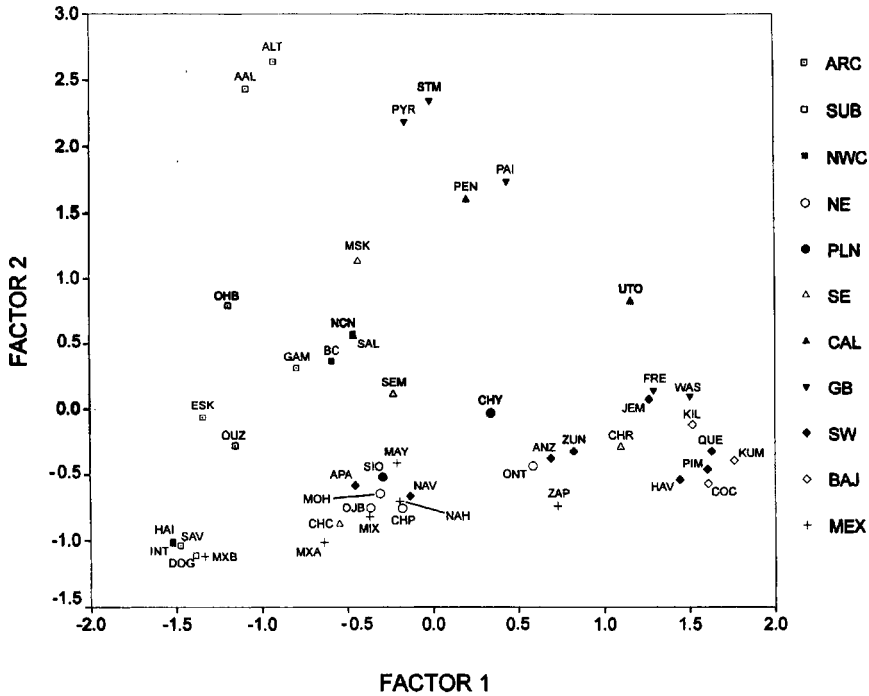

Figure 2. Plot of the component scores from the first 2 principal components of mtDNA haplogroup frequencies for modern and ancient populations of native North America. Sample codes are given in Table 1.

Discussion

The patterns of variation among mtDNA haplogroups in native American groups illustrated in Figures 1 and 2 are concordant with other syntheses of mtDNA lineage variation in the Americas (Lorenz and Smith 1996; Carlyle et al. 1999). In particular, haplogroup A predominates throughout the Americas, except in the US Southwest. Here, this most common haplogroup is observed at low frequency and is often absent. Haplogroup B, which has only moderate to low frequencies in most other geographic regions of North America, has high frequency among US Southwest populations. It is of interest that among contemporary populations of western America those residing in the western Great Basin and northern California are distinguished from the more general US Southwest groups by their high frequency of haplogroup D. This lineage is rare to absent in the greater Southwest, as it is in most of Mexico.

Navajo and Apache populations are resident in the Southwest but are believed to have migrated there in the recent past from the Athapaskan home-

land in subarctic Canada (Haskell 1987). Their mtDNA profile, as given in Table 1 and Figure 1, illustrates this nicely. They do not appear at all similar to other Southwest populations as a result of the very high frequency of haplogroup A: greater than 50% in each southern Athapaskan group. This high frequency of haplogroup A more clearly aligns the Navajo and Apache populations with their Athapaskan kin living well to the north, where this haplogroup occurs at high frequencies, being nearly fixed in many northern Athapaskan populations. Interestingly, both the Navajo and Apache possess the 9-bp deletion and hence moderate frequencies of haplogroup B, although at frequencies lower than is typical for Southwest populations. This seems likely the result of admixture with neighboring Southwest populations subsequent to their migration into the region. The Apache also retain a low frequency of haplogroup D, which is commonly found in Northern Athapaskan populations not fixed for haplogroup A but which is absent in most Southwest and Mexican populations (Table 1, Figure 1).

Of particular interest is the concordance of mtDNA haplogroup profiles of the aDNA samples relative to the modern populations occupying the same geographic areas today. The aDNA samples conform to the same patterns of regional variation as the contemporary samples. The most obvious example is also perhaps the least surprising. The ancient Aleut samples are virtually identical to the single report of modern Aleut mtDNA haplogroup frequencies from the Pribilof Islands (Merriwether et al. 1995). Given the island nature of these populations and their relative isolation until the 18th century, it is perhaps not unexpected that the ancestors of the modern Aleuts are genetically similar to the modern descendents. However, the oldest of the prehistoric Aleut samples are archeologically dated to perhaps 4,000 years B.P. Given this temporal dimension and the vagaries of sampling that necessarily are attendant to aDNA samples, the degree of similarity of the prehistoric and modern Aleut samples is striking. This is especially true because both are easily distinguishable from other Eskimo, Inuit, and Indian populations of northwest North America (Figures 1 and 2) as well as from indigenous communities in northeast Asia (data not shown).

The correspondence of ancient and modern mtDNA profiles within a geographic region is evident also for the other aDNA samples examined. The Anasazi of the Four Corners region of the US Southwest are, on the basis of mtDNA haplogroup frequencies, indistinguishable from modern Native American populations in the region (Carlyle et al. 1999). All share a low frequency of haplogroups A and D and a preponderance of haplogroup B. In Figure 2 the Anasazi appear closest to modern Zuni. In similar fashion the ancient Fremont sample from northern Utah is also indistinguishable from modern US Southwest populations. This may be unexpected because the geographic region from which the ancient Fremont samples derive is part of the Great Basin geographic area, not the Southwest. However, the geographic distance separating the Anasazi and Fremont samples is scarcely 300 miles,

and there are well-known archeological similarities between these archeological traditions. The similarities are so great that, at least for the more southerly Fremont, some archeologists postulated the Fremont as a northern colonization of the southwestern Anasazi (Morss 1931; Gunnerson 1969; Berry 1975). The similarity of the Fremont to the Anasazi and modern Southwest populations led us to suggest that Fremont origins should be sought in the south and west, rather than among archeological traditions of the Great Basin (Parr et al. 1996; O'Rourke et al. 1999). The age range of the Fremont samples is 750–2,000 years B.P., whereas the Anasazi are slightly older, with an age range of 1,000–2,000 years B.P. Thus the regional Southwest pattern that these ancient samples reflect is of considerable antiquity (Carlyle et al. 1999).

In this context the placement of the Washoe in Figure 2 is intriguing. Although residing in the extreme western Great Basin, the Washoe show considerable similarity to Southwest groups but little similarity to the other western Great Basin (Paiute/Shoshone) or Californian groups (e.g., Salinan, Penutian) to which they are geographically proximate. This relationship was also noted by Lorenz and Smith (1996). The Washoe, Hokan speakers from the region surrounding Lake Tahoe, are the only Great Basin population that does not speak Numic, nor are they closely related linguistically to their neighbors to the west in California (Jacobsen 1986). The Washoe are longtime residents of the eastern Sierra region, presumably predating the hypothesized later influx of Numic speakers (Lamb 1958). Although generally no longer given much credence by comparative linguists, the Washoe language was formerly thought to be part of a larger Hokan grouping that included the Coahuiltecan branch. This linkage would imply some linguistic and perhaps phylogenetic relationship to other Southwest populations. Although not strongly supported on linguistic grounds, the mtDNA data are consistent with such an ancestral relationship. In this view the Washoe are related ancestrally to greater Southwest populations in the same fashion as we have postulated for the eastern Great Basin Fremont. That the extreme southern Californian and Baja groups are clearly aligned genetically with the Southwest populations of the Four Corners area makes this less of a stretch of an interpretation than it might otherwise appear.

The ancient Pyramid Lake and Stillwater Marsh samples also illustrate the temporal and regional continuity seen in the Southwest samples. In both Figure 1 and Figure 2 these ancient western Great Basin samples show greatest similarity to the Paiute/Shoshone of northwestern Nevada, along with the California Penutian sample. Indeed, the historical range of the Northern Paiute includes the Lahontan Basin north and east of Lake Tahoe where the Pyramid Lake and Stillwater Marsh collections were recovered. However, determination of similarity of these samples to modern populations in this geographic region is problematic. Only 2 modern western Great Basin populations are available for examination. As noted, the Washoe show strong mtDNA affinity with populations in the Southwest, well outside the western Great Basin geo-

graphic area, and are genetically distinct from their neighbors, the Northern Paiute and Shoshone. With only 2 samples from the Great Basin it is not possible to accurately evaluate a real geographic profile for this region among modern populations. It cannot be clearly determined whether one of these populations accurately represents a general Great Basin mtDNA profile while the other represents an outlier or whether the region was typified by extreme heterogeneity in mtDNA diversity.

The issue is complicated by the very small size ($n = 9$) of the Paiute/Shoshone sample. However, Kaestle (1998) reported haplogroup frequencies based on mitochondrial HVS1 sequence data for a sample of 33 Northern Paiute and Shoshone and haplogroup frequencies based on discrete marker data for 98 Northern Paiute (Kaestle 1997) embedded in a larger series of 116 northern Uto-Aztecan speakers. These additional data are not radically different from those used here (Lorenz and Smith 1996). The sequence data indicate a higher frequency of haplogroup B (0.39) and an absence of other haplogroups, whereas the large northern Uto-Aztecan series also indicates a higher frequency of haplogroup B (0.42) and a lower frequency of haplogroup C (0.15). These frequencies led Kaestle (1998) to conclude that the ancient western Great Basin samples had greater similarity to California Penutian groups than to either a combined group of Numic speakers or a combined sample of populations from throughout the Great Basin. We also find the ancient Nevada samples to be similar to Californian Penutian samples (Figures 1 and 2). The similarity of these ancient samples to the California Penutian sample in Figure 2 is determined by the shared high frequency of haplogroup D, a characteristic of many populations of northern California and the Columbia Plateau (see Figure 1) and also the western Great Basin. The California Penutian sample is also characterized by a low frequency of haplogroup A, enhancing the similarity to the ancient northwest Nevada samples.

Although the larger series of northern Uto-Aztecan speakers studied by Kaestle (1997, 1998) has a higher frequency of haplogroup B than the Pyramid Lake and Stillwater Marsh samples, the frequency is not as high as is typical in Southwest populations (Table 1). In addition, the lower frequency of haplogroup C in the northern Uto-Aztecan series coupled with the high frequency of haplogroup D is reminiscent of the haplogroup profile seen in the ancient western Great Basin samples. It is worth noting that the California Penutian sample represents populations from the geographic region immediately west of the Lake Tahoe area. Thus they are geographically proximal to the area of northwestern Nevada that yielded the ancient remains from Pyramid Lake and Stillwater Marsh, separated only by the small area occupied by the Washoe. Finally, Kaestle (1998) noted linguistic arguments that place the origin of some California Penutian speakers in northwestern Nevada. Given the extremely long temporal range of the ancient samples from the western Great Basin, it is perhaps not surprising that they show affinity to

more than 1 modern population sample but only those in the general geographic area or those in adjacent geographic areas with reasonable links to the area prehistorically.

Finally, the prehistoric Oneota of western Illinois are considerably younger than the other ancient samples examined. Dating to only 700 years B.P., they represent a much more restricted time frame than the other ancient samples. This skeletal series initially appears to be an exception to the concordance of haplogroup profiles between ancient and modern populations in a geographic region. They plot adjacent to the Anasazi and other Southwest groups in Figure 2. Nevertheless, they also plot near the Cheyenne and Zapotec in this component map and are on the edge of the Southwest cluster of groups in the direction of other upper plains and Mexican populations, which share moderate frequencies of haplogroup A and variable frequencies of the other markers. It is clear from Figure 1 that the Oneota are virtually identical in haplogroup profile to the Cheyenne and are similar to several Algonquian-speaking groups of the northeast (Mohawk, Ojibwa, Chippewa) with respect to the frequencies of haplogroups A, B, and C. These modern groups are the geographically closest contemporary Amerindian populations to which the Oneota can be compared. There are, unfortunately, no data on modern Amerindians in the western Illinois region from which the Oneota skeletal series was excavated for comparison.

The principal difference between the prehistoric Oneota and these modern groups is the substantially higher frequency of other haplogroups in the modern groups. Stone and Stoneking (1998) and Brown et al. (1998) recently demonstrated that haplogroup X is present in these prehistoric samples. This confirms the presence of haplogroup X in pre-Columbian samples and indicates that the other haplogroup category here represents the fifth Amerindian lineage. Moreover, modern Ojibwa populations exhibit substantial frequencies of haplogroup X (Forster et al. 1996; Scozzari et al. 1997), although this was initially ascribed to European admixture by Torroni, Schurr et al. (1993). The presence of haplogroup X in modern Ojibwa and the ancient Oneota clearly indicates its presence in ancestral Amerindian groups and suggests its presence in other North American populations where other (i.e., non-A, B, C, or D) haplogroups have been observed. In any event the haplogroup profiles illustrated in Figure 1 suggest that the Oneota present haplogroup frequencies that are consistent with those expected of populations that would have inhabited the area in the recent past.

The nature of aDNA samples raises substantive questions regarding their use as populations in analyses such as those described here. Indeed, this is 1 reason simple and straightforward analyses are used when analyzing ancient samples in a population context. In many instances ancient samples do not meet the expectations and assumptions of statistical manipulations. Given the temporal distribution of most ancient samples, we might question whether they

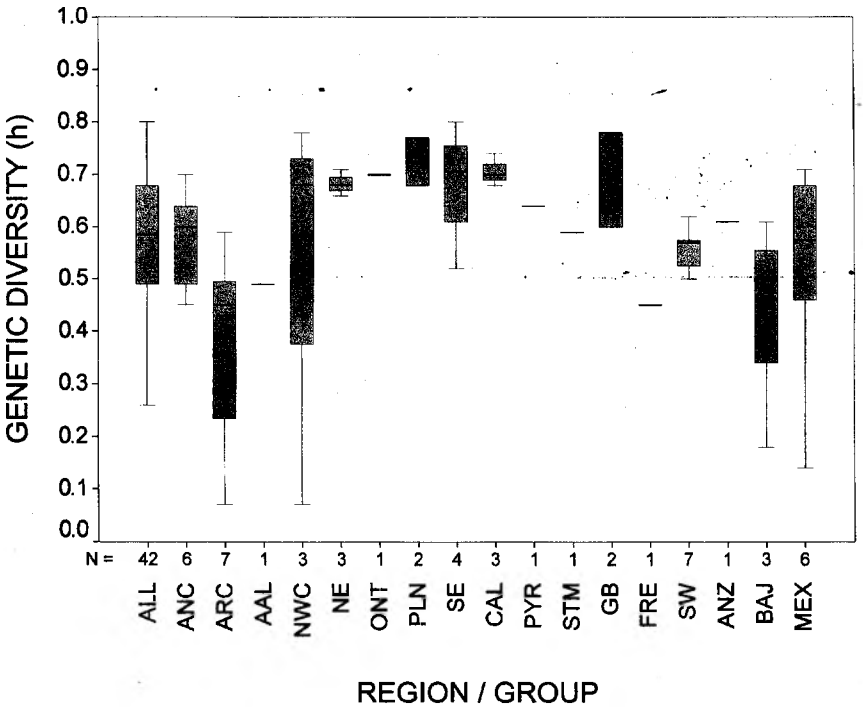

Figure 3. Distribution of genetic diversity *h* estimates by geographic region. See text for discussion and explanation. Sample codes are given in Table 1.

represent the same level of genetic variation observed in modern populations. One measure of variation relevant here is gene diversity *h*, given as

$$h = n(1 - \Sigma x^2)/(n - 1), \quad (1)$$

where n is sample size and x is haplogroup frequency (Nei 1987). Gene diversity estimates are given for each sample included in Table 1. It is clear that *h* for each of the aDNA samples is within the range of those observed for the modern populations, indicating that the aDNA samples do represent the same level of mtDNA variation seen in contemporary groups. This is illustrated clearly in Figure 3, where the shaded boxes indicate the interquartile range of diversity values for samples within a geographic region, the bar is the median of that distribution, and the extensions to the boxes illustrate the full range of observed *h* values. For comparison, gene diversity estimates for the aDNA samples are plotted as single bars adjacent to the geographic areas from which they derive. As can be seen, all the aDNA samples fall

within the observed ranges of the modern population estimates of h, and several are within the range, even the interquartile range, of their geographic area. The single exception to this is the Utah Fremont. This ancient sample has a lower diversity value than either the modern Great Basin groups or the modern Southwest populations, although it is closer to the latter. This is not too surprising given the absence of haplogroup A from this relatively large skeletal series. The observed diversity value for this ancient sample is, however, within the overall range of gene diversities seen in modern populations of North America.

The overall inference from the comparison of haplogroup frequencies in ancient North America with those observed in contemporary populations is one of regional and temporal stability. The regional patterning so clearly evident in mtDNA data of North Amerindians and noted by several researchers (Lorenz and Smith 1996; Shields et al. 1992; Carlyle et al. 1999; Merriwether et al. 1995) is clearly reflected in prehistoric populations as well. This rather unexpected observation has several ramifications. First, prehistoric populations of the Americas were characterized by considerable genetic variation, at least mtDNA lineage variation. Second, the geographic patterns so evident in modern populations arose early (greater than 2,000 years B.P.) and have remained relatively stable since. Third, the stability of these regional patterns indicates that the mtDNA genome was comparatively little affected by the disruption and demographic collapse that characterized many Amerindian populations at contact.

We hasten to emphasize that mtDNA data are but 1 realization of evolutionary and phylogenetic events. Mitochondrial DNA variation represents a single locus and is subject to random lineage extinction (Avise et al. 1984), especially in small populations, such as most populations of the Americas probably were for most of their history. Data on Y-chromosome and other nuclear markers would be most instructive to our reconstruction of these regional patterns and would not necessarily yield the same inferences (Jorde et al. 1995; Hammer et al. 1997). Dos Santos et al. (1999) recently reported on genetic variation in an admixed community in Amazonia. Using the same mtDNA markers employed here, they estimated that over 50% of the contemporary mtDNA genome derived from native American mitochondria. In contrast, screening for the *DYS199T* allele, which is specific for Amerindians (Underhill et al. 1996), indicated that less than 5% of the contemporary genome was contributed by indigenous men. Asymmetry in gene flow at admixture is not surprising, and similar asymmetric effects on different marker systems through mortality, morbidity, migration, and population displacement as a result of contact would also not be surprising. As additional types of markers (e.g., Y-chromosome markers, STRs, *Alus*) become routinely available for analysis in prehistoric samples, our understanding of population dynamics and history in the Americas and elsewhere will be enhanced.

Conclusions

The analysis of mtDNA haplogroup variation in both ancient and contemporary native populations of North America demonstrates that (1) native populations of North America are characterized by substantial mtDNA lineage variation, (2) the distribution of these lineages is geographically structured, and (3) the pattern of haplogroup (lineage) variation in ancient samples is concordant with the pattern observed in modern populations living in the same geographic region. (4) Thus regional differentiation in mtDNA variation in native America is relatively stable, temporally and geographically, and is of considerable antiquity (much greater than 2,000 years). (5) The stability of the mtDNA profiles in geographic regions suggests minimal disruption of this genetic system and its patterns of variability as a result of the disruptions and demographic declines associated with contact.

Acknowledgments We gratefully acknowledge the Northwestern Band of the Shoshone Nation for permission to study the material from the Great Salt Lake Wetlands burials and to the Aleut Corporation and the Chaluka Corporation for permission to analyze the Aleut material. We also appreciate permission to sample the Anasazi material from the American Museum of Natural History, New York, and the Aleut samples from the Smithsonian Institution, Washington, DC.

The analyses were supported by the National Science Foundation through grants BNS 89-20463 and OPP 99-74623 and doctoral dissertation improvement grants OPP 98-13044 and SBR 98-09609 and the Wenner-Gren Foundation for Anthropological Research. S. Carlyle and G. Hayes were supported by the University of Utah Graduate Research Fellowships, and G. Hayes also received support from a Post-Graduate Scholarship from the Natural Sciences and Engineering Research Council of Canada. Additional research support was provided by the Department of Anthropology, College of Social and Behavioral Science, and the George H. Thomas Presidential Chair Research Fund at the University of Utah.

We are pleased to thank our colleagues James O'Connell, Duncan Metcalfe, Steve Simms, Ryan Parr, Henry Harpending, Alan Rogers, Elise Eller, Steve Wooding, Rika Kaestle, and David Glenn Smith for stimulating, informative, and productive discussions and assistance.

Received 3 September 1999; revision received 16 September 1999.

Literature Cited

Avise, J.C., J.E. Neigel, and J. Arnold. 1984. Demographic influences on mitochondrial DNA lineage survivorship in animal populations. *J. Molec. Evol.* 20:99–105.

Berry, M.S. 1975. *An Archeological Survey of the Northeast Portion of Arches National Park.* Antiquities Section Selected Papers, v. 1, no. 3. Salt Lake City, UT: Utah Division of State History.

Brown, M.D., S.H. Hosseini, A. Torroni et al. 1998. mtDNA haplogroup X: An ancient link between Europe/Western Asia and North America? *Am. J. Hum. Genet.* 63:1852–1861.

Carlyle, S.W., R.L. Parr, M.G. Hayes et al. 1999. The context of maternal lineages in the greater Southwest. Unpublished.

Crawford, M.H. 1998. *The Origins of Native Americans: Evidence from Anthropological Genetics.* Cambridge, England: Cambridge University Press.

Dos Santos, S.E.B., J.D. Rodrigues, A.K.C. Ribeiro-Dos-Santos et al. 1999. Differential contribution of indigenous men and women to the formation of an urban population in the Amazon region as revealed by mtDNA and Y-DNA. *Am. J. Phys. Anthropol.* 109:175–180.

Forster, P., R. Harding, A. Torroni et al. 1996. Origin and evolution of Native American mtDNA variation: A reappraisal. *Am. J. Hum. Genet.* 59:935–945.

Gunnerson, J.H. 1969. *The Fremont Culture: A Study in Cultural Dynamics on the Northern Anasazi Frontier.* Papers of the Peabody Museum of Archeology and Ethnology, v. 59, no. 2. Cambridge, MA: Harvard University Press.

Hammer, M.F., A.B. Spurdle, T. Karafet et al. 1997. The geographic distribution of human Y-chromosome variation. *Genetics* 145:787–805.

Haskell, J.L. 1987. *Southern Athapaskan Migration A.D. 200–1750.* Tsaile, AZ: Navajo Community College Press.

Hayes, M.G. 1998. Mitochondrial DNA variation of ancient Aleuts. *Am. J. Phys. Anthropol.* (suppl. 26):95.

Hrdlicka, A. 1945. *The Aleutian and Commander Islands and Their Inhabitants.* Philadelphia, PA: Wistar Institute of Anatomy and Biology.

Jacobsen, W.H., Jr. 1986. Washoe language. In *Handbook of North American Indians,* v. 11, *Great Basin,* L. D'Azevedo, ed. Washington, DC: Smithsonian Institution, 107–112.

Jorde, L.B., M.J. Bamshad, W.S. Watkins et al. 1995. Origins and affinities of modern humans: A comparison of mitochondrial and nuclear genetic data. *Am. J. Hum. Genet.* 57:523–538.

Kaestle, F. 1997. Molecular archeology: An analysis of ancient Native American DNA from western Nevada. *Nev. Hist. Soc. Q.* 40:85–96.

Kaestle, F.A. 1998. Molecular evidence for prehistoric Native American population movement: The Numic expansion. Ph.D. dissertation, University of California, Davis.

Lamb, S.M. 1958. Linguistic prehistory in the Great Basin. *Int. J. Am. Ling.* 24:95–100.

Lorenz, J.G., and D.G. Smith. 1996. Distribution of four founding mtDNA haplogroups among native North Americans. *Am. J. Phys. Anthropol.* 101:307–323.

Merriwether, D.A., F. Rothhammer, and R.E. Ferrell. 1995. Distribution of the four founding haplotypes in Native Americans suggest a single-wave migration for the New World. *Am. J. Phys. Anthropol.* 98:411–430.

Morss, N. 1931. *The Ancient Culture of the Fremont River in Utah.* Papers of the Peabody Museum of Archeology and Ethnology, v. 12, no. 3. Cambridge, MA: Harvard University Press.

Nei, M. 1987. *Molecular Evolutionary Genetics.* New York: Columbia University Press.

O'Rourke, D.H., A. Mobarry, and B.K. Suarez. 1992. Patterns of genetic variation in native America. *Hum. Biol.* 64:417–434.

O'Rourke, D.H., R.L. Parr, and S.W. Carlyle. 1999. Molecular genetic variation in prehistoric inhabitants of the eastern Great Basin. In *Understanding Prehistoric Lifeways in the Great Basin Wetlands: Bioarcheological Reconstruction and Interpretation,* B.E. Hemphill and C.S. Larsen, eds. Salt Lake City, UT: University Press of Utah, 84–102.

Parr, R.L., S.W. Carlyle, and D.H. O'Rourke. 1996. Ancient DNA analysis of Fremont Amerindians of the Great Salt Lake Wetlands. *Am. J. Phys. Anthropol.* 99:507–518.

Schurr, T.G., S.W. Ballinger, Y.Y. Gan et al. 1990. Amerindian mitochondrial DNAs have rare Asian mutations at high frequencies, suggesting they derived from four primary maternal lineages. *Am. J. Hum. Genet.* 46:613–623.

Scozzari, R., F. Cruciani, P. Santolamazza et al. 1997. mtDNA and Y-chromosome-specific polymorphisms in modern Ojibwa: Implications about the origin of their gene pool. *Am. J. Hum. Genet.* 60:241–244.

Shields, G.F., K. Hecker, M.I., Voevoda et al. 1992. Absence of the Asian-specific region V mitochondrial marker in native Beringians. *Am. J. Hum. Genet.* 50:758–765.

Simms, S.R. 1999. Farmers, foragers, and adaptive diversity: The Great Salt Lake wetlands project. In *Understanding Prehistoric Lifeways in the Great Basin Wetlands: Bioarcheological Reconstruction and Interpretation,* B.E. Hemphill and C.S. Larsen, eds. Salt Lake City, UT: University Press of Utah, 21–54.

Simms, S.R., C.J. Loveland, and M.E. Stuart. 1991. *Prehistoric Human Skeletal Remains and the Prehistory of the Great Salt Lake Wetlands.* Utah State Contributions to Anthropology 6. Logan, UT.

Smith, D.G., R.S. Malhi, J.A. Eshleman et al. 1999. Distribution of mtDNA haplogroup X among native North Americans. *Am. J. Phys. Anthropol.* 110:271–284.

Starikovskaya, Y.B., R.I. Sukernik, T.G. Schurr et al. 1998. mtDNA diversity in Chukchi and Siberian Eskimos: Implications for the genetic history of ancient Beringia and the peopling of the New World. *Am. J. Hum. Genet.* 63:1473–1491.

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

Stone, A.C., and M. Stoneking. 1998. mtDNA analysis of prehistoric Oneota population: Implications for the peopling of the New World. *Am. J. Hum. Genet.* 62:1153–1170.

Suarez, B.K., J.D. Crouse, and D.H. O'Rourke. 1985. Genetic variation in North Amerindian populations: The geography of gene frequencies. *Am. J. Phys. Anthropol.* 67:217–232.

Szathmary, E.J.E. 1993. Genetics of aboriginal North Americans. *Evol. Anthropol.* 1:202–220.

Torroni, A., Y.-S. Chen, O. Semino et al. 1994. mtDNA and Y-chromosome polymorphisms in four Native American populations from southern Mexico. *Am. J. Hum. Genet.* 54:303–318.

Torroni, A., T.G. Schurr, M.F. Cabell et al. 1993. Asian affinities and continental radiation of the four founding Native American mtDNAs. *Am. J. Hum. Genet.* 53:563–590.

Torroni, A., T.G. Schurr, C. Yang et al. 1992. Native American mitochondrial DNA analysis indicates that the Amerind and the NaDene populations were founded by two independent migrations. *Genetics* 130:153–162.

Torroni, A., R.I. Sukernik, T.G. Schurr et al. 1993. mtDNA variation of aboriginal Siberians reveals distinct genetic affinities with Native Americans. *Am. J. Hum. Genet.* 53:591–608.

Underhill, P.A., L. Jin, R. Zemans et al. 1996. A pre-Columbian Y-chromosome-specific transition and its implications for human evolutionary history. *Proc. Natl. Acad. Sci. USA* 93:196–200.

Ward, R.H., B.L. Frazier, K. Dew-Jager et al. 1991. Extensive mitochondrial diversity within a single Amerindian tribe. *Proc. Natl. Acad. Sci. USA* 88:8720–8724.

Ward, R.H., A. Redd, D. Valencia et al. 1993. Genetic and linguistic differentiation in the Americas. *Proc. Natl. Acad. Sci. USA* 90:10,663–10,667.