



---

## **Report Information from ProQuest**

November 28 2014 18:06

---

## Table of contents

1. Distribution of sequence variation in the mtDNA control region of Native North Americans.....	1
--	---

## Distribution of sequence variation in the mtDNA control region of Native North Americans

**Author:** Lorenz, Joseph G; Smith, David Glenn

[ProQuest document link](#)

**Abstract (Abstract):** Lorenz and Smith used the distributions of mtDNA diversity among North American haplogroups, language groups, and tribes to characterize the process of tribalization that followed the colonization of the New World. Differences between tribes and language groups were primarily due to differences in the distribution of three major haplogroups that evolved before settlement of the New World.

**Abstract:** The distributions of mtDNA diversity within and/or among North American haplogroups, language groups, and tribes were used to characterize the process of tribalization that followed the colonization of the New World. Approximately 400 bp from the mtDNA control region of 1 Na-Dene and 33 Amerind individuals representing a wide variety of languages and geographic origins were sequenced. With the inclusion of data from previous studies, 225 native North American (284 bp) sequences representing 85 distinct mtDNA lineages were analyzed. Mean pairwise sequence differences between (and within) tribes and language groups were primarily due to differences in the distribution of three of the four major haplogroups that evolved before settlement of the New World. Pairwise sequence differences within each of these three haplogroups were more similar than previous studies based on restriction enzyme analysis have indicated. The mean of pairwise sequence differences between Amerind members of haplogroup A, the most common of the four haplogroups in North America, was only slightly higher than that for the Eskimo, providing no evidence of separate ancestry, but was about two-thirds higher than that for the Na-Dene. However, analysis of pairwise sequence divergence between only tribal-specific lineages, unweighted for sample size, suggests that random evolutionary processes have reduced sequence diversity within the Na-Dene and that members of all three language groups possess approximately equally diverse mtDNA lineages. Comparisons of diversity within and between specific ethnic groups with the largest sample size were also consistent with this outcome. These data are not consistent with the hypothesis that the New World was settled by more than a single migration. Because lineages tended not to cluster by tribe and because lineage sharing among linguistically unrelated groups was restricted to geographically proximate groups, the tribalization process probably did not occur soon after settlement of the New World, and/or considerable admixture has occurred among daughter populations.

**Links:** [Check for full text in other National University Library databases](#)

### Full text: Headnote

JOSEPH G. LORENZ AND DAVID GLENN SMI1\*

### Headnote

Abstract The distributions of mtDNA diversity within and/or among North American haplogroups, language groups, and tribes were used to characterize the process of tribalization that followed the colonization of the New World. Approximately 400 bp from the mtDNA control region of 1 Na-Dene and 33 Amerind individuals representing a wide variety of languages and geographic origins were sequenced. With the inclusion of data from previous studies, 225 native North American (284 bp) sequences representing 85 distinct mtDNA lineages were analyzed. Mean pairwise sequence differences between (and within) tribes and language groups were primarily due to differences in the distribution of three of the four major haplogroups that evolved before settlement of the New World. Pairwise sequence differences within each of these three haplogroups were more similar than previous studies based on restriction enzyme analysis have indicated. The mean of pairwise sequence differences between Amerind members of haplogroup A, the most common of the four haplogroups in North America, was only slightly higher than that for the Eskimo, providing no evidence of separate ancestry,

but was about two-thirds higher than that for the Na-Dene. However, analysis of pairwise sequence divergence between only tribal-specific lineages, unweighted for sample size, suggests that random evolutionary processes have reduced sequence diversity within the Na-Dene and that members of all three language groups possess approximately equally diverse mtDNA lineages. Comparisons of diversity within and between specific ethnic groups with the largest sample size were also consistent with this outcome. These data are not consistent with the hypothesis that the New World was settled by more than a single migration. Because lineages tended not to cluster by tribe and because lineage sharing among linguistically unrelated groups was restricted to geographically proximate groups, the tribalization process probably did not occur soon after settlement of the New World, and/or considerable admixture has occurred among daughter populations.

#### **Headnote**

**KEY WORDS:** PCR, D LOOP, HAPLOGROUPS,, PHYLOGENY,, DIVERGENCE,, LANGUAGE GROUP

The earliest study of mtDNA diversity among native Americans, using six restriction enzymes, concluded that the New World was founded by a small number of females and that these founders established groups in the Americas that remained isolated from one another (Wallace et al. 1985). Later studies using 14 restriction enzymes (Schurr et al. 1990; Torroni, Schurr et al. 1993; Torroni, Chen et al. 1994) supported the hypothesis that the New World was founded by a limited number of mtDNA haplogroups, of which four remain common, and that "tribalization of the Amerinds was a relatively rapid process and that it was followed by extensive tribal isolation" (Torroni, Schurr et al. 1993, p. 585). If tribalization in the New World occurred according to this scenario, then one would expect to observe a high degree of correspondence between genetic and linguistic differences among groups.

Restriction digest analysis of the whole mtDNA molecule suggests and sequence data from the mtDNA control region subsequently confirmed (Ward et al. 1991; Horai et al. 1993) that the mtDNA lineages of most living native Americans fall into one of four Asian-derived haplogroups (designated A, B, C, and D) defined by the presence or absence of certain restriction enzyme recognition sites (Schurr et al. 1990; Torroni, Schurr et al. 1993; Lorenz and Smith 1996). The frequencies of the four major haplogroups vary significantly among tribes and among the three major language taxa (Eskimo-Aleut, NaDene, and Amerind) into which Greenberg (1987) classified all native American languages (Lorenz and Smith 1996; Merriwether et al. 1995). Specifically, haplogroups B, C, and D are rarely, if ever, found in Na-Dene groups, and haplogroup B is absent and haplogroup C is rare in most Eskimo groups.

Studies of ancient DNA have shown that at least in some geographic regions, such as the Western Great Basin, all four haplogroups are very old (predating 4500 years B.P.) and that all other haplogroups, including those that might have become extinct since, were rare (Kaestle 1996). It has been argued that at least one of these haplogroups (i.e., haplogroup B) was introduced to the New World by separate migrations from Asia [e.g., Horai et al. (1993) and Torroni et al. (1992)] and that certain haplogroups were represented in more than one migration, but this view has been contested (Merriwether et al. 1995). Bailliet et al. (1994) suggested that an HaeIII restriction site at np 16517 defines two subgroups each of haplogroups A, C, and D [and perhaps also B; see Easton et al. (1996)] that are found in both the New World and Asia and might be additional founding lineages, but others find it more likely that this HaeIII site is a hypermutable hot spot (Torroni, Sukernik et al. 1993). If the founder population included more than a single member of each haplogroup, divergence within each surviving haplogroup might predate settlement of the New World (Ward et al. 1993; Horai et al. 1993; Merriwether et al. 1995).

Ward et al. (1991) sequenced the first variable segment of the mtDNA control region of 63 individuals from a single tribe, the Nuu-Chah-Nulth (Nootka) of the Northwest coast. They found that the 63 individuals belonged to 28 mtDNA lineages. Because of the amount of intercluster sequence diversity and the time required for this diversity to accumulate (estimated to be 41,000/78,000 years, or longer than the New World is thought to have been settled), they argued that a substantial amount of mtDNA diversity in America was introduced across the

Bering land bridge with little or no bottleneck effect. A phylogenetic analysis that included sequence data from two other Northwest coast groups (the Bella Coola and the Haida) showed no clustering by language affiliation (Ward et al. 1993). Ward interpreted this as evidence for a relatively recent split among the three populations, directly contrary to the hypothesis put forth by Wallace et al. (1985).

The purpose of this study is to estimate the extent and distribution of mtDNA diversity in North America and to determine the process by which the distribution of tribal groups in the New World arrived at its present form. Earlier studies present problems when attempting to understand the connection, if any, between the degree of similarity between mtDNA and language among native North American tribes. Ward et al. (1991, 1993) drew samples from three groups in close geographic proximity (i.e., the Pacific Northwest), two of which (the Bella Coola and Nuu-Chah-Nulth) are also closely related linguistically (Greenberg 1987). On the other hand, Amerind samples from North America studied by Torroni, Schurr et al. (1993) included only two groups (the Ojibwa and Pima), which are neither geographically proximate nor linguistically related.

We have sequenced approximately 400 bp from the mtDNA control region of 1 Na-Dene and 33 Amerind individuals representing a wide range of language taxa and geographic origins. These individuals come from geographically contiguous groups (e.g., Yumans and Uto-Aztecs) whose languages are not closely related and from groups who speak closely related languages (e.g., Nahuatl from Mexico and Uto-Aztecs from California) but are geographically discontinuous. These samples were chosen to augment tribal and linguistic heterogeneity represented by published control region sequences and therefore to minimize bias in sampling levels of diversity within each mtDNA haplogroup. If tribalization and subsequent isolation occurred soon after the settlement of the New World, then phylogenetic analysis should create clusters of linguistically related groups irrespective of geographic proximity, and mean pairwise differences among linguistically related groups should be lower than differences among linguistically unrelated groups. If tribalization or retribalization following periodic declines in population occurred relatively late, lineages from the same tribe or linguistically related tribes should not cluster together phylogenetically and mean pairwise sequence differences among all tribes should be low (Ward et al. 1993). If tribalization was a continuous process, correlations between language and genes, the phylogenetic clustering of lineages from the same tribe or language group, and the mean pairwise sequence differences among all tribes should show marked geographic variation.

#### Materials and Methods

All sequence analysis was performed on DNA that had been extracted for earlier studies (Lorenz and Smith 1994, 1996) and assigned to haplogroups A, B, or C. Members of haplogroup D, whose distribution in North America is more limited, were not included in the present study because they lack unique markers in the control region (Torroni, Schurr et al. 1993) and therefore cannot be identified by their published control region sequence. The ethnic affiliations of the 34 individuals whose mtDNAs were sequenced in this study are as follows: 5 Nahuatl, 5 Chumash, 2 Bella Coola, 1 Choctaw, 2 Washo, 2 Ojibwa (Chippewa), 1 Dogrib, 1 Hopi, 1 Salinan, 4 California Uto-Aztecs (represented by 1 individual each from the Cahuilla, Fernandeno, Vanyume, and Sereno tribes), and 10 Yuman speakers from Baja, Mexico (represented by 5 Kumiai, 2 Paipai, and 1 individual each from the Kiliwa, Cucapa, and Cochimi tribes).

**Samples.** The ethnic affiliations and sources of the samples analyzed in the present study are given in Table 1. In addition to the 34 samples that we sequenced, 191 published mtDNA control region sequences of other native North Americans (Ward et al. 1991, 1993; Torroni, Schurr et al. 1993; Shields et al. 1993) were also included in the analysis. Sequences reported by Shields et al. (1993) that were generated from samples collected from various Athapaskan tribes of Alaska were treated as a single ethnic group. Samples from other studies were identified as members of one of the three haplogroups (A, B, or C) by the presence of uniquely identifying markers in the control region, identified by the boxes in Table 2 (Ward et al. 1991; Torroni, Schurr et al. 1993). Haplogroup A was identified by the presence of at least two of the following three mutations: a C - T transition at np 16290, a G - A transition at np 16319 (Torroni, Schurr et al. 1993), and a T - C transition at np

16362. Haplogroup C was also identified by the presence of at least two of three mutations: a T - C transition at np 16298, a C - T transition at np 16327 (Torroni, Schurr et al. 1993), and a T - C transition at np 16325. Although none of the analyzed samples, except those from Torroni, Schurr et al. (1993) and our own, had been assigned to one of the three haplogroups by restriction enzyme analysis, all had been screened for the presence of the 9-bp deletion in region V, a marker for haplogroup B. Membership in haplogroup B, assigned based on the presence of the region V deletion, was confirmed by the presence of T - C transitions at np 16189 and np 16217 (Torroni, Schurr et al. 1993).

**Polymerase Chain Reaction.** Nested polymerase chain reaction (PCR) amplifications were performed using external primers L15926 (5'-TCAAAGCTTACACCAGTCTTGTAAC-3') and H16498 (5'-CCTGAAGTAGGAACCAGATG-3') and internal primers L15997 (5'-AAAGAATTCCTCCACCATTAGCACCCAAAGC-3') and H16401 (5'-AAATCTAGATGATTTACGGAGGATGGTG-3') (Ward et al. 1991). The internal primers have restriction sites (EcoRI and XbaI, respectively) appended to their 5' ends to facilitate cloning the fragment into the M13 vector. Five microliters of DNA extracted from serum or 1  $\mu$ l of DNA extracted from buffy coat or hair was added to a reaction mixture containing 2.5  $\mu$ l of 10 X reaction buffer, 0.2 mM of each dNTP, 1  $\mu$ l of each external primer (50 #CM, 0.5  $\mu$ l final concentration), 0.5  $\mu$ g/ $\mu$ l BSA, and 0.5 unit of Taq polymerase (Perkin Elmer) in a final volume of 25( $\mu$ )l. The PCR cycling profile was as follows: initial denaturing at 94°C for 1 min, 40 cycles at 94°C for 10 s, 40°C for 10 s, 45°C for 10 s, and 72°C for 15 s, and a final extension step at 72°C for 1 min (1605 Thermal Cycler, Idaho Technology). Internal amplifications were performed using 1  $\mu$ l of external amplification product and the same PCR conditions as before, except that the annealing temperature was raised to 55°C. Internal PCR products were electrophoresed on a 2% NuSieve gel (FMC) and stained with ethidium bromide. Fragments were visualized using ultraviolet light, and the appropriate band (~450 bp) was excised with a razor blade. The DNA was extracted from the excised portion of gel using a Qiaquick Gel Extraction Kit (Qiagen) and eluted into 50  $\mu$ l of ddH<sub>2</sub>O. The restriction digest was performed in a final volume of 200  $\mu$ l. The fragment, along with 500 ng M13mpl9, was digested with 5  $\mu$ l of XbaI, 5  $\mu$ l of EcoRI, and 20  $\mu$ l of buffer H (Boehringer Mannheim) at 37°C for 1-2 hr. The restriction enzymes were inactivated by heating to 65°C for 15 min and were removed with the Qiaquick PCR Purification Kit (Qiagen).

**Cloning.** The eluate containing the digested PCR fragment and M13 vector was ligated overnight at 16°C with T4 DNA ligase (USB). One hundred microliters of competent E. coli cells (strain JM 109, Clontech) were transformed with the 50- $\mu$ l ligation mixture for 30 min on ice. The transfected cells were heat-shocked at 42°C for 45 s in LB top agar containing 8  $\mu$ l XGal (100 mg/ml) and 4 #cl IPTG (200 mg/ml). Two hundred microliters of exponentially growing JM 109 cells were added to the heat-shocked competent cells. The mixture was plated on LB plates, which were then incubated overnight at 37°C. Clear plaques were inoculated into 1.5 ml of LB media and incubated at 37°C for 6-8 hr with vigorous shaking. The cells were then stored at 4°C overnight. Single-stranded M13mpl9 vector containing the PCR fragment was isolated using the ssPhage Isolation Kit (Bio 101).

**Sequencing.** Recombinants were sequenced using the standard dideoxy chain termination sequencing protocol from the Sequenase Kit (USB). Sequencing reactions (2-3  $\mu$ l) were electrophoresed at 1700 V for 1.5-5.5 hr on 6% polyacrylamide gels (AMRESCO) containing 7 M urea. Gels were dried and exposed to Biomax film (Kodak) overnight. Sequence data were obtained from two separate clear plaques for each individual to verify that detected differences were authentic mutations rather than cloning artifacts. Autoradiographs were read using a semi-automatic gel reader (International Biotechnologies Inc.), and the data were entered into a MacVector file (International Biotechnologies Inc.). The sequence data for each individual were imported into Assemblyalign (International Biotechnologies Inc.) and compared to the reference sequence (Anderson et al. 1981) to identify mutational sites. **Phylogenetic Analysis.** Because the different studies providing data included in the present analysis sequenced slightly different portions of the control region, we limited our analysis to the 284 bp

between nucleotide sites 16086 and 16370 to minimize the number of sequences with missing data. Phylogenetic analysis was conducted on the lineages identified using PAUP (version 3.1.1; Swofford 1993). Maximum parsimony trees were generated through the simple addition of sequences using the tree bisection and reconnection algorithm. A maximum of 2000 trees were saved for each search, and a 50% consensus tree was constructed.

**Table 1.** Distribution of mtDNA Lineages among Native North Americans

<i>Language Group</i>	<i>Ethnic Group</i>	<i>Number of Individuals per Haplogroup</i>				<i>Number of Tribal-Specific Lineages per Haplogroup</i>			
		<i>A</i>	<i>B</i>	<i>C</i>	<i>Total</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>Total</i>
Eskimo-Aleut	Inupiaq <sup>b</sup>	5	0	0	5	1	0	0	1
	Inuit <sup>b</sup>	4	0	0	4	4	0	0	4
	West Greenland Eskimos <sup>b</sup>	12	0	0	12	0	0	0	0
Na-Dene	Athapaskans <sup>b</sup>	18	0	1	19	6	0	1	7
	Dogrib <sup>a,c</sup>	3	0	0	3	0	0	0	0
	Haida <sup>c</sup>	36	0	3	39	5	0	0	5
	Tlingit <sup>c</sup>	2	0	0	2	1	0	0	1
	Navaho <sup>c</sup>	2	0	0	2	0	0	0	0
Amerind	Nuu Chah Nulth <sup>d</sup>	28	2	12	42	9	1	3	13
	Bella Coola <sup>a,c,e</sup>	27	3	5	35	5	1	2	8
	Yakima <sup>b</sup>	5	15	2	22	3	4	2	9
	Ojibwa <sup>a,c</sup>	2	1	2	5	2	1	1	4
	Washo <sup>a</sup>	0	2	0	2	0	1	0	1
	Chumash <sup>a</sup>	3	1	1	5	3	0	0	3
	Salinan <sup>a</sup>	0	1	0	1	0	0	0	0
	California Uto-Aztecs <sup>a</sup>	0	2	2	4	0	1	2	3
	Pima <sup>a</sup>	1	1	1	3	0	1	1	2
	Hopi <sup>a</sup>	0	1	0	1	0	0	0	0
	Baja Yumans <sup>a</sup>	0	8	2	10	0	3	1	4
	Choctaw <sup>a</sup>	1	0	0	1	1	0	0	1
	Maya <sup>c</sup>	1	1	1	3	1	0	1	2
	Nahuatl <sup>a</sup>	3	0	2	5	3	0	2	5
	Total	153	38	34	225	44	13	16	73

a. Present study.

b. Shields et al. (1993).

c. Torroni, Schurr et al. (1993).

d. Ward et al. (1991).

e. Ward et al. (1993).

**Table 2.** mtDNA Control Region Sequences from Native North Americans by Ethnic Group

[illegible]



Sequence<sup>b</sup>[illegible]

a. Numbers in parentheses indicate number of individuals in a particular lineage if there is more than one.  
b. A dot in the column indicates identity with the reference sequence (Anderson et al. 1981); an X indicates missing data.

Historical events leading to admixture and/or stochastic effects on the size distribution of lineages within a language group or tribe can distort estimates of pairwise sequence divergence from their expected distribution and linearity in time (Marjoram and Donnelly 1994; Rogers and Harpending 1992; Slatkin and Hudson 1991). For example, recent gene flow can pose as shared ancestry between two groups, whereas subsequent rapid increases in the size of groups or any other lineages as a result of variance in reproductive success in groups of limited size can obscure close common ancestry to other groups. Because several of the studied groups have inhabited regions through which continuous migration probably occurred, contact among unrelated or very

distantly related groups might have led to admixture, causing overestimates of genetic similarities among dissimilar language groups. The potential impact of these influences were evaluated by estimating pairwise sequence differences, in a separate analysis, using only one member of each tribalspecific or language-specific lineage represented in each group. Thus, only the variety, not the frequency distribution, of the lineages in each group could influence the differences in pairwise divergence, and the influence of population substructure on divergence estimates was minimized.

The modalities of the distributions of pairwise differences were also studied to evaluate the plausibility of panmixia and founder effects, both of which are associated with unimodal distributions of pairwise sequence differences. Population expansion leads to the increased retention of lineages, causing a starlike phylogeny and a Poisson distribution of pairwise sequence differences (Rogers and Harpending 1992). We also analyzed the rate of coalescence of lineages over time, measured as the number (or percentage) of gene substitutions, to assess whether or not expansion was constant (Nee et al. 1996), as assumed when gene divergence is calibrated to time.

## Results

The ethnic affiliations of the 225 samples are shown in Table 1 for each of the three haplogroups. The 85 distinct mtDNA lineages represented by these individuals were defined by 68 polymorphic sites, all of which were transitions; these are shown in Table 2. Of the 34 individuals that we sequenced in this study, 22 belonged to 1 of 18 previously undescribed lineages. Three additional individuals belonged to two lineages that have not been described previously in North America but have been found elsewhere: 2 Kumiai belonged to lineage 120, which has also been found among Indonesians, Papua New Guineans, Samoans, Hawaiians, Malaysians, and Chinese (Lum et al. 1994; Redd et al. 1995), and 1 Nahuatl individual belonged to lineage 171, which is also found among the Mapuche of South America (Ginther et al. 1993). The remaining nine individuals belonged to two previously described North American lineages (lineages 89 and 180).

Seventy-three of the 85 lineages were tribal specific. Approximately equal portions (85-88%) of the lineages within each of the three haplogroups were tribal specific. The 12 lineages that were shared between 2 or more tribes (identified as "multiple" in Table 2) are listed in Table 3. Six of these 12 lineages were shared by tribal groups within the same language family, as defined by Greenberg (1987) (i.e., Eskimo, Na-Dene, Amerind), with 2 of these 6 lineages (lineages 63 and 64) being shared between 2 of the 3 Eskimo groups, 1 being shared between 2 of the 5 Na-Dene groups (lineage 30), and 3 being shared among 2 or more of the 14 Amerind groups (lineages 86, 89, and 180). The most common of these 3 (lineage 89) was shared by 7 of the 14 different Amerind groups. Of the six remaining lineages that were shared by two of the language groups (none was shared by all three of the language groups), two (lineages 66 and 67) were shared between at least one of the two Eskimo groups and one Na-Dene group (the Alaskan Athapaskans), one (lineage 65) was shared between one Eskimo and one Amerind group, and three (lineages 15, 18, and 195) were shared between at least one Na-Dene (including Haida in all three cases) and at least one (but including Bella Coola in each case) Amerind group.

**Table 3.** Number of Individuals in Lineages Shared by Two or More Ethnic Groups

Lineage	West Greenland							Bella	
	Inupiaq	Eskimo	Tlingit	Haida	Athapaskans	Dogrib	Navaho	Nuu-Chah-Nulth	Coala
15				1				2	9
18			1	21	6	2	1	5	3
30						1	1		
63	1	2							
64	1	1							
65		1							
66	2	1			3				
67		7			2				
86								1	
89									1
180								7	
195				3					3
Total number	4	12	1	25	11	3	2	15	16

Lineage 18 was represented by the largest number of all individuals in this study and was shared by all five of the Na-Dene groups and all three of the Northwest coast Amerind groups. Most Amerind groups that share lineages with Eskimo or Na-Dene groups are located adjacent to or nearby Eskimo and/or Na-Dene groups. Moreover, a greater proportion of the Eskimo/ Aleut and Na-Dene lineages (30%) than of the Amerind lineages (11%) were shared with another tribe, even though approximately equal average numbers of individuals (10.75 and 9.9, respectively) were sampled per tribe.

A 50% majority rule consensus tree showed that all lineages (lineage numbers shown by haplogroup in Table 2 correspond to those in Figure 1) clustered into one of three groups that correspond to three of the four haplogroups (A, B, and C) defined by Torroni, Schurr et al. (1993). Although 83 of the 85 lineages clustered with other members of their haplogroup, lineages 177 and 178 did not cluster with the other haplogroup C lineages. Lineages 177 and 178 had been included in this study and assigned to haplogroup C based on the restriction enzyme data. Our sequence analysis revealed the presence in lineage 177 of the T - C transition at np 16325, which is characteristic of haplogroup C, but this lineage lacked both of the other two transitions usually associated with this haplogroup. This circumstance is typically seen only in members of haplogroup D. Lineage 178 had two of the three mutations typical of haplogroup C but was one of only three (of 18) lineages assigned to haplogroup C (lineage 177 was another of these three) that lacked a T - C transition at np 16298 and one of two lacking a C - T transition at np 16223, traits that are also common to haplogroup D. Samples that were assigned to haplogroup C had tested negative for haplogroups B and A by restriction enzyme analysis, but after testing positive for haplogroup C, they were not tested for presence of the A1 site loss at np 5176 that is diagnostic of haplogroup D. Although it is possible for similar patterns of nucleotide substitutions in the rapidly evolving control region to emerge in different mtDNA haplogroups, lineages 177 and 178 might actually have the compound haplogroup C/D.

<i>Ojibwa</i>	<i>Yakima</i>	<i>Washo</i>	<i>Chumash</i>	<i>Salinan</i>	<i>California Uto-Aztecan</i>	<i>Baja Yuman</i>	<i>Pima</i>	<i>Hopi</i>	<i>Maya</i>	<i>Total</i>
							1			13
	1									40
										2
										3
										2
	1									2
										6
										9
									1	2
		1	1	1	1	2		1		8
1			1							9
										6
1	2	1	2	1	1	2	1	1	1	102

Pairwise Sequence Differences. Of the three major language taxa, the Eskimo and Na-Dene exhibited comparable values of average within-group pairwise diversity (2.68 1.33 and 2.69 2.87, respectively), both of which were less than half that for Amerinds (6.32 3.14) (Table 4, complete North American data set). The mean of the pairwise sequence differences between only Amerind lineages from haplogroup A (Table 4, haplogroup A only) was much lower (and those of Na-Dene groups were somewhat lower) than that for the data set containing all three haplogroups (Table 4, complete data set). The distributions about these two means are given in Figure 2 (panels a and d, respectively). In addition, although the distribution of pairwise sequence differences between all Amerinds was bimodal, with modes falling at 3 and at 8 (Figure 2d), the distribution between only Amerinds of haplogroup A was unimodal (Figure 2a). Thus approximately half of the paired sequence differences between Amerind individuals (or tribes) and one-third of the differences between Na-Dene individuals are due to haplogroup affiliation (or distribution). This multimodality of the distribution of pairwise sequence differences in Figure 2d probably reflects the fact that the sample was genetically subdivided by haplogroups before settlement of the New World.

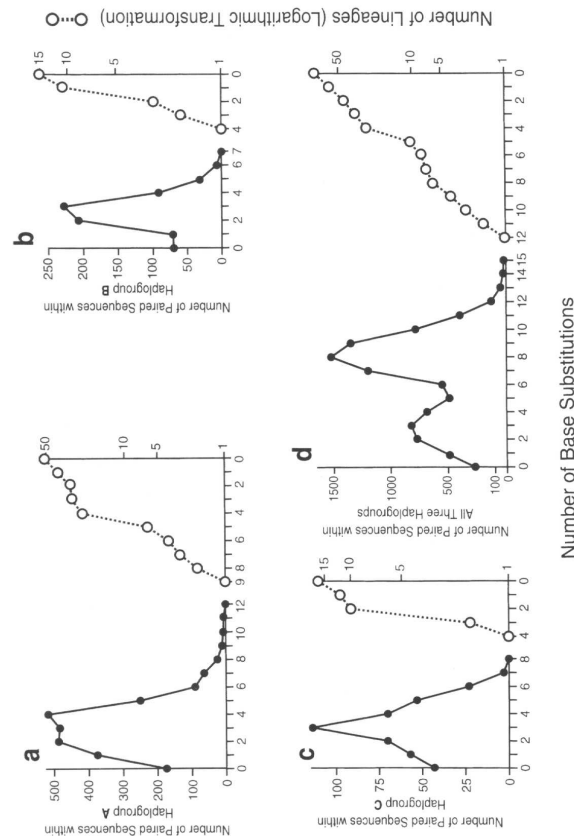
The unimodality of the mean pairwise sequence differences for haplogroup A lineages could reflect either panmixia among descendants of a common ancestor and/or recent admixture among some groups that are more geographically and culturally related than genetically related. The absence of Eskimo-specific control region mutations (e.g., T - C transitions at np 16093 or np 16311, a C - T transition at np 16173, and an A - G transition at np 16265) in Amerinds or Athapaskans and the absence of the Eskimo/Athapaskan-specific control region C - T transition at np 16192 in Amerinds (Forster et al. 1996) suggest that recent admixture is not responsible for the unimodality of the sequence differences. These observations are also consistent with the unique or rare presence of other genetic markers in Amerinds (e.g., the DI\*A red cell surface antigen, the TF\*D-Chi serum allele, and the absence of the immunoglobulin haplotype GM\*I 11,13), Athapaskans (the np 16329 RsaI site loss in mtDNA), and Eskimo/Aleuts (absence of the GM,2 21 immunoglobulin haplotype and of albumin Naskapi).

Because the pairwise differences were still greater within the Amerind than within the Eskimo and the Na-Dene language groups, their direction is consistent with the divergence estimates of Shields et al. (1992, 1993) that were based on mtDNA control region sequences and with the hypothesis, based on those estimates, that Amerind is the oldest of these three language families (Greenberg et al. 1986). However, pairwise differences in the NaDene were lower than those in the Eskimo, and those for Eskimos and Amerinds were too similar to support the hypothesis that members of the two language groups descended from different founders (i.e., migrations). Both estimates of mean pairwise intergroup differences that involved the Amerind were greater, although only marginally so, than those that did not. Moreover, the intergroup pairwise divergence between the Amerind and the Na-Dene was lower than that within the Amerind, and the divergence between the NaDene and the Eskimo was lower than that within the Eskimo.

The comparisons of individual ethnic groups were consistent with the linguistic results. Of the ethnic groups compared in Table 5, the Haida, a NaDene tribe, are the least divergent and the Bella Coola, an Amerind tribe, are the most divergent. The two differ by a factor of 2 [1.22 ( 1.09) and 2.82 ( 1.59), respectively]. The values for the West Greenland Eskimo, the Alaskan Athapaskan, and the Nuuchah-Nulth are intermediate and similar. The average divergence between the Alaskan Athapaskan and the West Greenland Eskimo (1.87 1.45) is almost as low as the average divergence within both of these groups; the average paired divergence between the Haida and the Bella Coola (2.37 1.30), whose ancestors have been said to descend from different founders, was actually lower than the value of paired divergence within the Bella Coola (2.82 1.59). The Bella Coola were the most divergent of the five ethnic groups, including the Nootka, whose language has been argued to be remotely related to the Bella Coola's (Drucker 1965; Greenberg 1987).

**Figure 1.** Maximum parsimony tree generated using PAUP (Swofford 1993) using the tree bisection and reconstruction swapping method and rooted using the Anderson (1981) sequence. Tree length is 123; consistency index = 0.553; retention index = 0.820. The numbers above each node give the proportion of generated trees that retain that branch.

	<i>Number of Lineages</i>	<i>Number of Individuals</i>	<i>Eskimo</i>	<i>Na-Dene</i>	<i>Amerind</i>
Complete North American set					
Eskimo	10	21	2.68 (1.33)		
Na-Dene	19	65	2.92 (2.38)	2.69 (2.87)	
Amerind	62	139	5.96 (3.02)	5.68 (3.33)	6.32 (3.14)
Haplogroup A only					
Eskimo	10	21	2.68 (1.33)		
Na-Dene	17	61	2.46 (1.49)	1.81 (1.48)	
Amerind	30	71	3.35 (1.56)	2.76 (1.76)	3.02 (1.79)



**Figure 2.** Distribution (left-hand y axis) of the number of base-pair differences (reading left to right on the x axis) between pairs of Amerind sequences and plot (right-hand y axis) of the logarithmic transformation of the number of lineages (reading right to left on the x axis) by time (from present) to coalescence (measured in number of base substitutions) within all native American lineages of (a) haplogroup A, (b) haplogroup B, and (c) haplogroup C and (d) within all three haplogroups. Lineages 177 and 178, whose haplogroup affiliations are uncertain, were excluded from the plot of the number of lineages against time in part c.

**Table 5.** Mean Pairwise Sequence Differences ( $\pm$  SD) between Members of Haplogroup A by Ethnic Group<sup>a</sup>

Ethnic Group	N	n	West			
			Greenland Eskimo	Alaskan Athapaskan	Haida	Bella Coola Nuu-Chah-Nulth
West Greenland						
Eskimo	12	5	1.79 (1.76)			
Alaskan						
Athapaskan	18	9	1.87 (1.45)	1.94 (1.25)		
Haida	36	7	2.19 (1.36)	2.24 (1.62)	1.22 (1.09)	
Bella Coola	27	7	3.44 (1.59)	3.44 (1.80)	2.37 (1.30)	2.82 (1.59)
Nuu-Chah-Nulth	28	11	2.53 (1.29)	2.56 (1.50)	2.06 (1.44)	2.98 (1.57)
						1.95 (1.28)

N, number of individuals; n, number of lineages.

a. Within-group differences are italicized. Eighteen of the sequences reported by Shields et al. (1993) obtained from various Athapaskan tribes in Alaska were treated as a single population sample.

As shown in Table 6, the average numbers of pairwise sequence differences between any two individuals within haplogroups A, B, and C were 2.71 ( 1.73), 2.44 ( 1.29), and 3.04 ( + 1.70), respectively. The average numbers of pairwise sequence differences between any two members of only the Amerind language group within haplogroups A, B, and C were 3.02 ( 1.79), 2.44 ( 1.29), and 2.88 ( 1.68), respectively. The distributions of pairwise sequence differences within each of these haplogroups were unimodal in both cases and approximated Poisson distributions (see Figure 2), suggesting panmictic populations experiencing population growth.

The estimates of average pairwise divergence for haplogroup A within and between the three language groups and tribes using only one representative of each tribal-specific lineage are given in Table 7. The within-group divergence estimates for Na-Dene (Table 7, language group) and for both Alaskan Athapaskans and the Haida (Table 7, tribe) were much larger than those in Tables 4 and 5, respectively. The heterogeneous (i.e.,

multitribal) sampling of Athapaskans (Shields et al. 1993) was in part responsible for this. The within-group estimates for all three language groups and for all tribes other than the Haida were similar, differing by less than 10%, with the Haida showing a low value. Likewise, all between-group estimates were similar, differing by less than one-third, with the exceptions being those between Alaskan Athapaskans and both the Haida and the Bella Coola, which showed high values.

## Discussion

**Table 6.** Mean Pairwise Sequence Divergence within Haplogroups<sup>a</sup>

<i>Haplogroup</i>	<i>Study</i>	<i>Number of Lineages</i>	<i>Number of Individuals</i>	<i>Mean Pairwise Sequence Differences</i>	<i>Sequence Divergence (284 Sites) (%)</i>
A	All groups (present study)	52	153	2.71	0.953
	Amerind groups only				
	Present study				
	All individuals	30	71	3.02	1.064
	Amerind-specific lineages	27	27	4.23	1.49
B	Torrioni, Neel et al. (1994)	24	131		0.075
	All groups (present study)	15	38	2.44	0.859
	Amerind groups only				
	Present study				
	All individuals	15	38	2.44	0.859
C	Amerind-specific lineages	15	15	3.24	1.14
	Torrioni, Neel et al. (1994)	19	83		0.034
	All groups (present study)	18	34	3.04	1.067
	Amerind groups only				
	Present study				
	All individuals	17	30	2.88	1.013
	Amerind-specific lineages	16	16	3.92	1.38
	Torrioni, Neel et al. (1994)	25	18		0.096

a. The criteria used by Torrioni, Neel et al. (1994) for assigning haplotypes to specific haplogroups is not precisely the same as that used in the present study, but this difference does not seriously influence the comparisons between the two studies.

Shields et al. (1993) found lower pairwise sequence divergence among circumarctic populations, which include the Inupiaq, the Inuit, the West Greenland Eskimo, and Athapaskans, and greater pairwise sequence divergence among members of their Amerind sample, which consisted of the Yakima, Bella Coola, and Haida. They concluded that the geographically widespread circumarctic populations form a cohesive biological entity that diverged relatively recently, whereas the Amerinds diverged much earlier. The mean pairwise sequence differences within and between the Eskimo, NaDene, and Amerind language taxa for our own complete data set of 225 individuals (see Table 4) are, in fact, consistent with their conclusion. The intragroup mean pairwise difference in the Amerind group is more than twice that in the Eskimo and Na-Dene taxa. Moreover, the estimate of mean paired divergence between the Eskimo and Na-Dene taxa is nearly half that between either group and the Amerind group.

Table 7. Mean Pairwise Sequence Differences ( $\pm$  SD) by Language Group and Tribe for the Haplogroup A Lineages<sup>a</sup>

Language group	Number of Lineages	Language Group			Tribe			
		Eskimo	Na-Dene	Amerind	Alaskan Athapaskans	Inuit	Haida	Bella Coola Nuu-Chah-Nulth
Eskimo	7	3.81 (1.47)						
Na-Dene	13	4.30 (1.40)	4.18 (1.46)					
Amerind	27	4.21 (1.47)	4.60 (1.50)	4.30 (1.70)				
Tribe								
Inuit <sup>a</sup>	4				3.17 (0.75)			
Alaskan Athapaskan	6				3.42 (1.10)			
Haida	5				4.05 (0.83)		2.20 (0.79)	
Bella Coola	5				4.35 (0.88)		3.20 (1.23)	
Nuu-Chah-Nulth	9				3.53 (1.03)		4.18 (1.05)	
								3.17 (1.21)

a. Within-group differences are italicized. Inuit replaced West Greenland Eskimo in the analysis by tribe because no tribal-specific lineages were found in the latter group.

However, because haplogroups A, B, C, and D emerged before the settlement of the New World, Amerind haplogroup A lineages are more closely related to Eskimo and Na Dene haplogroup A lineages than to Amerind haplogroup B or C lineages. In fact, the mean pairwise sequence divergence between pairs of all members of haplogroup A was actually slightly lower, not higher, than that between only Amerind members of haplogroup A. This is also evident from the binomial distribution of the number of base substitutions between paired sequences of Amerinds of all three haplogroups in Figure 2d. Although sampling might have influenced this outcome (and this should be reevaluated with more distant groups, such as those from South America), our results suggest that the language families as defined by Greenberg (1987) account for little of the genetic differences among native Americans.

The maximum parsimony tree confirms that North American mtDNA lineages as defined by sequencing do cluster into discrete groups that correspond to the haplogroups defined by restriction enzyme analysis (see Figure 1). Moreover, the unimodal distribution of pairwise sequence differences of each of the three Amerind haplogroups, which when combined, exhibit a multimodal distribution (Figure 2), suggests that the three Amerind haplogroups include panmictic portions of a subdivided population, each of which exhibits a founder effect. Therefore a true measure of the relative time depths of Amerinds and non-Amerinds requires estimating divergence times for each haplogroup separately if differences in haplogroup frequencies between groups occurred. For example, because most Amerind populations retain at least three of the four haplogroups, whereas haplogroup A overwhelmingly predominates in Eskimo and Na-Dene groups, Shield et al.'s (1993) estimates of sequence divergence were strongly influenced by distributional differences in haplogroup frequencies. Comparisons of divergence values therefore must be restricted to a common haplogroup. Haplogroup A, which is more commonly shared among different language and geographic groups than other haplogroups, is most useful in this regard because it is absent only in native American populations from the southwestern United States. Moreover, the unimodal distribution of paired sequence differences within



haplogroup A suggests that its membership is panmictic and not derived from more than a single or a few closely related founding lineages.

The pairwise sequence divergences for members of haplogroup A Amerind lineages (3.02 1.79) but not those of other groups (Table 4, haplogroup A only) were less than half that for the entire data set (6.32 3.14). This is, of course, due to the elimination from the Amerind sample of that group of lineages least related to those lineages (almost) exclusively found in Eskimo and Na-Dene individuals. As a result, the mean pairwise within-group sequence divergence between members of haplogroup A lineages, relative to their standard deviations, are much more similar for the Eskimo, NaDene, and Amerind taxa than when all haplogroups are combined. The smaller differences in within-group sequence divergence among the three language groups were neither consistent with the conclusions of Shields et al. (1993) nor sufficiently large to support or in the direction predicted by the three-migration hypothesis (Greenberg et al. 1986). The lower value for the Haida (31 of 39 of whom belong to 1 of only 2 lineages) than for the Alaskan Athapaskan in Table 5 results in part because the Alaskan Athapaskan group includes members of numerous tribes and in part, perhaps, because the Haida underrepresent sequence diversity within the Na-Dene group because they have experienced greater genetic drift as a result of their geographic isolation. With a level of heterogeneity of tribal representation comparable to that for the Eskimo, diversity within the Na-Dene might well have been lower. However, even though only two Alaskan Athapaskans shared one of the six tribalspecific Alaskan Athapaskan lineages, four of these lineages, compared to four of five Haida-specific lineages, clustered together in the maximum parsimony tree (Figure 1), reflecting the genetic homogeneity of that "tribal" sample. Clearly, because sampling strongly influences estimates of sequence divergence within haplogroups and geographic and language groups, comparisons of divergence estimates should be made cautiously.

Ward et al. (1993) compared the amount of sequence divergence within and among the Haida, Bella Coola, and Nuu-Chah-Nulth. As in the present study, the conclusions of which are based on diversity within members of only haplogroup A, they found that the mean intratribal pairwise sequence divergence was lowest for the Haida and that the Haida were less divergent from the other two groups, especially the Bella Coola, than either of them were from each other, despite their apparent linguistic similarity (see Tables 5 and 7). Rather than attribute this lack of correspondence to gene flow, Ward postulated that the population divergence leading to the three ethnic groups, especially the Haida, was relatively recent.

In our own study the Haida lineages did not form a distinct clade but rather clustered with the Bella Coola lineages, and in the three shared lineages that were found among the Haida (lineages 15, 18, and 195) 25 individuals had identical sequences to 9 Bella Coola individuals. Because the Bella Coola are geographically close to the Haida, whose oral traditions place their immediate ancestors in the Skeena River valley, just north of the Bella Coola (Drucker 1965), and because coastal Salishan-speaking peoples are probably recent emigrants from the interior of British Columbia, recent gene flow could be more responsible for genetic similarities between the Haida and both the Bella Coola and the Nuu-Chah-Nulth than recent divergence from a common ancestor.

The frequent lineage sharing between Eskimos and Alaskan Athapaskans and the restriction of lineage sharing between Amerinds and nonAmerinds to adjacent tribes also suggest the importance of recent admixture. When pairwise sequence divergence is based only on single representatives of tribal-specific lineages, the pairwise divergence between the Haida and the Nuu-Chah-Nulth increases dramatically, suggesting that admixture, not recent ancestry, is responsible for the lineage sharing between the two tribes. Because the lowest between-group estimate for any tribal pair is that between the Haida and the Bella Coola, recent divergence cannot be discounted as a plausible explanation for their genetic similarity. That the divergence between Alaskan Athapaskans and the Haida is higher than all other intertribal estimates except that between the Bella Coola and Alaskan Athapaskans suggests that language adoption might be responsible.

The amount of sequence divergence within haplogroups A, B, and C, based on 284 nt sequences of the first

hypervariable segment of the control region from 225 individuals, was more than an order of magnitude greater than that obtained by Torroni, Schurr et al. (1993) and Torroni, Neel et al. (1994) based on restriction enzyme analysis of the whole mtDNA molecule. This is because the rate of evolution of mtDNA, which varies from one region of the molecule to another (Vigilant et al. 1991), is greater in the noncoding control region than elsewhere and because mutations in that region are less subject to functional constraints. The date for the peopling of the New World has been estimated by calibrating the amount of divergence between paired sequences to time. This calibration assumes that divergence and the rate of lineage extinction, which is inversely proportional to the rate of population growth, proceed at a constant rate. The plots of the logarithmic transformations of the number of lineages at each node of the tree in Figure 1 (righthand y axes in Figure 2) against time to coalescence to a single lineage (measured as the number of base substitutions and read from right to left on the x axis,) reflect early and late periods of rapid population growth (shallow, convex portions of the curves) separated by an interim period of slower growth (concave midsection of the curves). The distributions of the numbers of base substitutions (read left to right on the x axis) by which paired members of Amerind haplogroups A, B, and C differ (left-hand y axis in Figure 2) are Poisson-like, with their variances (3.22, 1.65, and 2.80, respectively) being of similar value to their means (3.02, 2.44, and 2.88, respectively). Such distributions are expected in exponentially expanding populations (Rogers and Harpending 1992). The longer time to coalescence for members of haplogroup A in Figure 2 probably represents greater heterogeneity within the founders of that lineage rather than a more ancient presence of haplogroup A in the New World, because the average pairwise sequence divergences for all three haplogroups are similar. The epidemic transformation (Nee et al. 1996) of the left-hand y axis in Figure 2 generates slopes that are nearly identical to those generated by their logarithmic transformations, with no pronounced and systematic change over time, a pattern expected of a population whose exponential growth has occurred at a relatively constant rate. Thus the linear calibration of sequence divergence to time given in Table 6 is valid if the appropriate rate is chosen.

Torroni, Neel et al. (1994) used a rate of 0.022-0.029% per 10,000 years, based on data from Chibcha-speaking tribes of Central America, to calibrate the molecular clock for the entire mtDNA molecule. Vigilant et al. (1991) estimated a rate of evolution for the control region of 11.5-17.3% per 1,000,000 years based on mtDNA divergence between humans and chimpanzees. This rate and the slightly higher one (just over 20% per 1,000,000 years) estimated by Horai et al. (1995) for hypervariable segment I require an unrealistically early date (before 50,000 years B.P.) for settlement of the New World to reach the levels of pairwise sequence divergences given in Table 6, unless the founders themselves exhibited considerable diversity within haplogroups A, B, and C [as Ward et al. (1991) have argued]. Given the low frequency of all four common Amerind haplogroups in most Asian populations today [but see Sukernik et al. (1996)] and the apparent antiquity of the limited number of haplogroups in the New World, the probability that founders of the same haplogroup exhibited a high level of sequence diversity seems low.

Ward et al. (1991) used a rate of 33% per 1,000,000 years, also based on human-chimpanzee comparisons but using a higher transition to transversion ratio (at least 30:1 instead of 15:1) than Vigilant et al. (1991). Because we found no transversions among the 68 mutations described in Table 6, Ward's rate might be more valid than Vigilant's rate and provides estimates of times of divergence for haplogroups A and C that are closely in accord with those published by Torroni, Neel et al. (1994). However, Torroni, Schurr et al. (1993) and Torroni, Neel et al. (1994) found that haplogroup B exhibited less than half the level of sequence divergence of haplogroup A and about one-third the level of haplogroup C. Based on these data and the absence of the 9-bp deletion among Eskimo and Na-Dene groups (thought to have settled in the New World later than the Amerind groups), Torroni and co-workers hypothesized that haplogroup B reached the New World in a migration following the arrival of the Amerinds but preceding the arrival of the Na-Dene and that haplogroup A reached the New World in more than a single migration.

In our own study haplogroup B was also less divergent than haplogroups A and C, although to a much lower

degree (10-20%), whereas haplogroups A and C exhibited equal diversity. In our earlier study of the distribution of the 9-bp region V deletion, a defining marker for haplogroup B, we found that the deletion was almost universal among North American Amerinds, most of whom exhibit at least three of the four haplogroups and thus probably did not arrive in the New World by a separate Amerind migration (Lorenz and Smith 1994). Given the low frequency of haplogroup B in most of Asia, it also seems unlikely that this haplogroup arrived in the New World through multiple migrations. The slightly younger date for haplogroup B, resulting from a slightly lower estimate of lineage divergence within haplogroup B, might occur because lineages A and C but not (or less so) lineage B [but see Easton et al. (1996)] had already begun to diverge when the New World was first settled (Bailliet et al. 1994). The lower diversity in members of haplogroup B reported by Torroni, Sukernik et al. (1993) and Torroni, Neel et al. (1994) might also result from the fact that the tribal heterogeneity of their Amerind sample was much lower than that of our own. This again underscores how crucial an influence tribal sampling imposes on comparisons of sequence divergence.

The rate of control region evolution estimated from a coalescent model (Lundstrom et al. 1992) applied to data from one of the Amerind populations in the present study (Nuu-Chah-Nulth) is much higher (i.e., 1% sequence divergence per 8,950 years, or 111.7% per 1,000,000 years). This rate, recently contested by Forster et al. (1996), is based on statistical principles and does not rely on the imprecise estimates of the time and extent of divergence between chimpanzees and humans (Ward et al. 1993; Shields et al. 1993). If the rate obtained from the coalescent model is applied to the mean intragroup pairwise sequence divergence of the three haplogroups in this study, the time of divergence within haplogroups ranges from 8,000 to 10,000 years. This range of dates (among haplogroups) is much narrower and more recent than those obtained by Torroni, Schurr et al. (1993) and Torroni, Neel et al. (1994) and is too recent to accommodate even the traditional archeological view that the "Clovis" paleo-Indian culture [e.g., see Lynch (1990) and Meltzer (1995)] and that of cave-dwelling contemporaries in the Amazon basin (Roosevelt et al. 1996) provide the earliest valid evidence for human occupation in the New World. These late dates become even less plausible if some small amount of diversity within haplogroups predates the colonization of the New World. Within-group divergence estimates based only on single representatives of each tribal-specific lineage are higher for all haplogroups, tribes (except the Haida), and language groups (between 1.1% and 1.5%) and are consistent with a date between 10,000 and 13,000 years B.P. for New World settlement. However, this leaves little time for divergence of the haplogroups before their arrival in the New World and for the full-blown emergence of such divergent cultures as those cited. Except as noted for the Alaskan Athapaskan and Haida, there appears to be little clustering of haplogroup A lineages by tribe in Figure 1. Although the five Bella Coola lineages fell into three clusters (two of which include the Haida), the Inuit, Nuu-Chah-Nulth, Yakima, and Nahuatl specific lineages were dispersed throughout the tree. Only two of the four Yakima-specific haplogroup B lineages clustered together, as did the three Nuu-Chah-Nulth haplogroup C lineages. The paucity of clustering of lineages by tribe, especially for Amerinds, is evidence that tribes in the New World are not the result of an early tribalization process in which ancestral populations split with little or no contact among daughter populations. Although intertribal admixture might have been low in general (Forster et al. 1996), local instances of language spread (Kaestle 1996) and substantial gene flow (Lorenz and Smith 1996) have obscured correlations between languages and genes, as has the decline in the number of shared derived etymologies between languages whose common ancestor predates 10,000 years B.P.

We also noted a large increase (131%) in the mean paired within-group divergence for the Na-Dene language group and lesser increases (82% and 80%, respectively) for both Na-Dene tribes (Table 7) when divergences were estimated using only one representative of each tribal-specific lineage. These findings suggest that population substructure might have strongly influenced the estimates of mean pairwise sequence divergence for those groups in Table 4. The overrepresentation of one or two lineages (especially lineage 18) in the Haida might reflect effects of genetic drift or a recent bottleneck. A lesser yet still substantial (75%) increase in the

mean pairwise divergence between the Na-Dene and Eskimo taxa might reflect the influence of admixture between those language groups that is minimized when only tribal-specific lineages are considered (i.e., when lineages 66 and 67 are excluded).

In addition, pairwise estimates involving the Inuit (Table 7) were difficult to compare with those for the West Greenland Eskimos (Table 4), because the Eskimos exhibited no tribal-specific lineages and therefore could not be included in Table 7. However, the substantial increase in pairwise divergence for comparisons between both Eskimo populations with a NaDene tribe (83% and 85%, respectively) is consistent with that for the Eskimo/ Na-Dene comparison in Table 7. Moreover, the widespread and unique sharing between Eskimos and Athapaskans, including the Navaho (Torroni, Sukernik et al. 1993) of a C - T transition at np 16192 and an A - G transition at np 16265 suggests their common ancestry. It is less clear why the lowest mean pairwise intergroup comparison in Table 7 (3.2 + 1.23) is that between the Haida and the Bella Coola, whose languages are regarded as being unrelated, whereas the second highest estimate is that between the Haida and Alaskan Athapaskans, both of whom speak related languages. Although admixture or language adoption, respectively, might be responsible for these observations, given that it has occurred far more commonly than is sometimes acknowledged [e.g., see Ward et al. (1993)], additional sequencing studies that include larger numbers of individuals from other Na-Dene tribes (studies now underway) are needed to resolve this issue and that pertaining to separate native American migrations.

**Acknowledgments** This study was supported by the US Public Health Service through grants RR00169 and RR05090 and by a Faculty Research Grant from the Regents of the University of California. Received 18 September 1996; revision received 24 April 1997.

## References

### Literature Cited

## References

- Anderson, S., A.T. Bankier, B.G. Barrell et al. 1981. Sequence and organization of the human mitochondrial genome. *Nature* 290:457-465.
- Bailliet, G., F. Rothhammer, F.R. Carnes et al. 1994. Founder mitochondrial haplotypes in Amerindian populations. *Am. J. Hum. Genet.* 54:27-33.
- Drucker, P. 1965. *Cultures of the North Pacific Coast*. San Francisco, CA: Chandler.
- Easton, R.D., D.A. Merriwether, D.E. Crews et al. 1996. Mitochondrial DNA variation in the Yanomami: Evidence for additional New World founding lineages. *Am. J. Hum. Genet.* 59:213-225.

## References

- 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.
- Ginther, C., D. Corach, G.A. Penacino 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 *DNA Fingerprinting: State of the Science*, S.D.J. Pena, R. Chakraborty, J.T. Epplen et al., eds. Basel, Switzerland: Birkhauser Verlag, 211-219.
- Greenberg, J.H. 1987. *Language in the Americas*. Stanford, CA: Stanford University Press.
- Greenberg, J.H., C.G. Turner II, and S.L. Zegura. 1986. The settlement of the Americas: A comparison of the linguistic, dental, and genetic evidence. *Curr. Anthropol.* 27(5):477-497.

## References

- Horai, S., K. Hayasaka, R. Kondo et al. 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., R. Kondo, Y. Nakagawa-Hattori et al. 1993. Peopling of the Americas founded by four major lineages of mitochondrial DNA. *Molec. Biol. Evol.* 10:23-47.
- Kaestle, F. 1996. Molecular analysis of ancient Native American DNA from western Nevada. Paper presented at the Great Basin Anthropological Conference, Reno, Nevada.
- Lorenz, J.G., and D.G. Smith. 1994. Distribution of the 9 bp mitochondrial DNA region V deletion among North American Indians. *Hum. Biol.* 66:777-788.
- Lorenz, J.G., and D.G. Smith. 1996. Distribution of four

mtDNA haplogroups among native North Americans. *Am. J. Phys. Anthropol.* 101:307-323.

Lum, J.K., O. Rickards, C. Ching et al. 1994. Polynesian mitochondrial DNAs reveal three deep maternal lineage clusters. *Hum. Biol.* 66:567-590.

Lundstrom, R., S. Tavare, and R.H. Ward. 1992. Estimating substitution rates from molecular data using the coalescent. *Proc. Natl. Acad. Sci. USA* 89:5961-5965.

## References

Lynch, T.F. 1990. Glacial-age man in South America? A critical review. *Am. Antiq.* 55:1236.

## References

Marjoram, P., and P. Donnelly. 1994. Pairwise comparisons of mitochondrial DNA sequences in subdivided populations and implications for early human evolution. *Genetics* 136:673-683.

Meltzer, D. 1995. Clocking the first Americans. *Annu. Rev. Anthropol.* 24:21-45. Merriwether, D.A., F. Rothhammer, and R.E. Ferrell. 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.

Nee, S., E.C. Holmes, A. Rambaut et al. 1996. Inferring population history from molecular phylogenies. In *New Uses for New Phylogenies*, P.H. Harvey, A.J.L. Brown, J.M. Smith et al., eds. Oxford, England: Oxford University Press, 6680.

Nichols, J. 1990. Linguistic diversity and the settlement of the New World. *Language* 66:475-521.

## References

Redd, A.J., N. Takezaki, S.T. Sherry et al. 1995. Evolutionary history of the COII/tRNA<sup>Lys</sup> intergenic 9 base pair deletion in human mitochondrial DNAs from the Pacific. *Molec. Biol. Evol.* 12:604-615.

Rogers, A.R., and H. Harpending. 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Molec. Biol. Evol.* 9:552-569.

Roosevelt, A.C., M.L. da Costa, C.L. Machado et al. 1996. Paleoindian cave dwellers in the Amazon: The peopling of the Americas. *Science* 272:373-384.

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.

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.

Shields, G.F., A.M. Schmiechen, B.L. Frazier et al. 1993. mtDNA sequences suggest a recent evolutionary divergence for Beringian and northern North American populations. *Am. J. Hum. Genet.* 53:549-562.

Slatkin, M., and D. Hudson. 1991. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* 129:555-562.

Sukernik, R.I., T.G. Schurr, E.B. Starikovskaya et al. 1996. Mitochondrial DNA variation in native Siberians, with special reference to the evolutionary history of American Indians: Studies on restriction polymorphism. *Genetika* 32:432-439 (in Russian).

Swofford, D.L. 1993. *Phylogenetic Analysis Using Parsimony (PAUP)*, version 3.1.1. Washington, DC: Smithsonian Institution.

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.

## References

Torroni, A., J.V. Neel, R. Barrantes et al. 1994. 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., T.G. Schun, 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.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.

## References

Vigilant, L., M. Stoneking, H. Harpending et al. 1991. African populations and the evolution of human mitochondrial DNA. *Science* 253:1503-1507.

## References

Wallace, D.C., K. Garrison, and W.C. Knowler. 1985. Dramatic founder effects in Amerindian mitochondrial DNAs. *Am. J. Phys. Anthropol.* 68:149-155.

Ward, R.H., B.S. Frazier, K. Dew et al. 1991. Extensive mitochondrial diversity within a single Amerindian tribe. *Proc. Natl. Acad. Sci. USA* 88:872-8724. 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.

## AuthorAffiliation

<sup>1</sup> Analytical Genetic Testing Center, 7808 Cherry Creek S. Drive, Denver, CO 80231. <sup>2</sup> Department of Anthropology, University of California, Davis, CA 95616.

**Subject:** Deoxyribonucleic acid; DNA; Biology; Native Americans;

**MeSH:** Anthropology, Cultural, Ethnic Groups -- genetics, Genetic Variation, Humans, Sequence Analysis, DNA, Mitochondrial -- genetics (major), Genetics, Population (major), Indians, North American -- genetics (major)

**Substance:** DNA, Mitochondrial;

**Publication title:** Human Biology

**Volume:** 69

**Issue:** 6

**Pages:** 749-76

**Number of pages:** 28

**Publication year:** 1997

**Publication date:** Dec 1997

**Year:** 1997

**Publisher:** Wayne State University Press

**Place of publication:** Detroit

**Country of publication:** United States

**Publication subject:** Biology

**ISSN:** 00187143

**CODEN:** HUBIAA

**Source type:** Scholarly Journals

**Language of publication:** English

**Document type:** PERIODICAL

**Accession number:** 9353972, 03469696

**ProQuest document ID:** 224542025

**Document URL:**

<http://ezproxy.nu.edu/login?url=http://search.proquest.com/docview/224542025?accountid=25320>

**Copyright:** Copyright Wayne State University Press Dec 1997

**Last updated:** 2011-04-28

**Database:** ProQuest Central

---

**Contact ProQuest**

Copyright © 2014 ProQuest LLC. All rights reserved. - **Terms and Conditions**