Native American Mitochondrial DNA Analysis Indicates That the Amerind and the Nadene Populations Were Founded by Two Independent Migrations

Antonio Torroni,* Theodore G. Schurr,* Chi-Chuan Yang,* Emőke J. E. Szathmary,** Robert C. Williams, Moses S. Schanfield,† Gary A. Troup,** William C. Knowler,†† Dale N. Lawrence,‡‡ Kenneth M. Weiss‡‡ and Douglas C. Wallace*

*Center for Genetics and Molecular Medicine and Departments of Biochemistry and Anthropology, Emory University School of Medicine, Atlanta, Georgia 30322, †Office of the Dean, Faculty of Social Science, University of Western Ontario, London, Ontario, Canada N6A 5C2, ‡Department of Anthropology, Arizona State University, Tempe, Arizona 85281, ‡‡Analytical Genetic Testing Center, Inc., Denver, Colorado 80231, **Department of Pathology, University of New Mexico, Albuquerque, New Mexico 87131, ††Diabetes and Arthritis Epidemiology Section, National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), Phoenix, Arizona 85014, ‡‡Division of Host Factors, Center for Infectious Diseases, Centers for Disease Control, Atlanta, Georgia 30333, and ™Department of Anthropology and Graduate Program in Genetics, Pennsylvania State University, University Park, Pennsylvania 16802

Manuscript received April 19, 1991
Accepted for publication September 13, 1991

ABSTRACT

Mitochondrial DNAs (mtDNAs) from 167 American Indians including 87 Amerind-speakers (Amerinds) and 80 Nadene-speakers (Nadene) were surveyed for sequence variation by detailed restriction analysis. All Native American mtDNAs clustered into one of four distinct lineages, defined by the restriction site variants: HincII site loss at np 13,259, AluI site loss at np 5,176, 9-base pair (9-bp) COII-tRNA<sup>15</sup> intergenic deletion and HaeIII site gain at np 663. The HincII np 13,259 and AluI np 5,176 lineages were observed exclusively in Amerinds and were shared by all such tribal groups analyzed, thus demonstrating that North, Central and South American Amerinds originated from a common ancestral genetic stock. The 9-bp deletion and HaeIII np 663 lineages were found in both the Amerinds and Nadene but the Nadene HaeIII np 663 lineage had a unique sublineage defined by an RsaI site loss at np 16,329. The amount of sequence variation accumulated in the Amerind HincII np 13,259 and AluI np 5,176 lineages and that in the Amerind portion of the HaeIII np 663 lineage all gave divergence times in the order of 20,000 years before present. The divergence time for the Nadene portion of the HaeIII np 663 lineage was about 6,000–10,000 years. Hence, the ancestral Nadene migrated from Asia independently and considerably more recently than the progenitors of the Amerinds. The divergence times of both the Amerind and Nadene branches of the COII-tRNA<sup>15</sup> deletion lineage were intermediate between the Amerind and Nadene specific lineages, raising the possibility of a third source of mtDNA in American Indians.

The considerable cultural diversity, linguistic complexity (Spencer et al. 1977) and biological variation (Neel 1978; Neel and Thompson 1978; Szathmary 1984) of Native Americans has been the subject of much speculation. The observed diversity has at times been attributed to multiple migrations of Asiatic peoples to the Americas, a few migrations of ethnically distinct peoples, or in situ differentiation of ancestral Native Americans occurring after the initial colonization of the New World (Laughlin 1988; Szathmary 1991). A recent hypothesis proposes that three waves of migration populated the Americas (Williams et al. 1985; Greenberg, Turner and Zegura 1986), corresponding to the tripartite division of Native American languages [Amerind, Nadene and Eskaleut] (Greenberg 1987). However, the common origin of the numerous Amerind languages and the time depth required to develop this plurality of tongues is disputed (Diamond 1990; Lewin 1990).

Estimates of the time of arrival of the first migrants have also varied widely (Marshall 1990; Morell 1990). Dating of skeletal remains and Clovis lithic artifacts yield values of 13–14,000 years before present (YBP) (Taylor et al. 1985; Nelson et al. 1986) while archaeological studies have reported North and South American sites which date to more than 33,000 YBP (Guidon and Delibrias 1986; Dillehay and Collins 1988), with the best North American sites starting at 20,000–25,000 YBP (Adovasio et al. 1983).

To further investigate the populating of the Americas, we have examined the mtDNA variation of Native American populations. Because of its maternal inheritance (Giles et al. 1980), the mtDNA accumulates sequential mutations along radiating female lineages. The high mutation rate of the mtDNA
A random sampling of individuals within each respective of 30 Dogrib (NW Canada), 48 Navajo (Arizona and New Mexico) or buffy coats were obtained from 167 subjects including 80 Nadene and 87 Amerinds.

Accordingly the Dogrib form a recognizable genetic unit among other Athapaskan speakers, as shown by genetic distance analyses (SZATHMARY 1983). The maximum European admixture in the group was 8.7%, based on blood group, serum protein and red cell enzyme genes (SZATHMARY, FERRELL and GERSHOWITZ 1983).

The Navajo are southern Athapaskans who are closely related genetically (SZATHMARY 1983) and linguistically to the northern Athapaskan Indians living in Alaska and Canada. Their ancestors were hunting-gathering people from northwestern Canada who migrated to the southwest US after 1000 AD (HASKELL 1987). Having increased to the present 150,000 from a total of no more than 15,000 people in 1868 (SPENCER et al. 1977), the Navajo are now dispersed throughout the 30,000 square mile reservation in relatively isolated family units (WILLIAMS et al. 1981; TROUP et al. 1982).

The Pima are southern Athapaskans who are closely related genetically and linguistically to the northern Athapaskan Indians living in Alaska and Canada. Their ancestors were hunting-gathering people from northwestern Canada who migrated to the southwest US after 1000 AD (HASKELL 1987). Having increased to the present 150,000 from a total of no more than 15,000 people in 1868 (SPENCER et al. 1977), the Navajo are now dispersed throughout the 30,000 square mile reservation in relatively isolated family units (WILLIAMS et al. 1981; TROUP et al. 1982). The Navajo were one of the tribes represented in the study. These people speak Yucatec, one of many languages which includes the Papago, Hopi and several northern Mexican tribes (MATSON et al. 1968; HAURY 1976). Upon entering and settling the area, they are believed not to have intermingled with the Puebloan peoples who already occupied the region, but may have mixed considerably with other Southern Indians (MATSON et al. 1968). In 1977, the Pima numbered some 8,000 people (SPENCER et al. 1977), Gm allotype studies indicated that Caucasian admixture in the Pima was 1.0% (WILLIAMS et al. 1985, 1986). Bloods were collected at the Gila River Indian Community in Sacaton, Arizona, by the NIDDK as part of a diabetes study (KNOWLER et al. 1978).

The lowland Maya occupy parts of Mexico, Guatemala and Honduras, and are known to have inhabited the area continuously for at least 5,000 years (MACNEISH 1985). Blood samples were drawn from individuals living in an isolated village in the central Yucatan Peninsula. Because of the small size and remoteness of the village, most of its inhabitants were related at least at the second cousin level, however, only one member per family cluster was analyzed in this study. These people speak Yucatec, one of many Maya languages belonging to the Mexican Penutian subgroup (GREENBERG 1987). An analysis of blood types and serum proteins in nearby villages revealed approximately 10% European admixture in this population (our unpublished data).

The Ticuna are a linguistically distinct and geographically isolated tribe living in the Amazonian rain forest of western Brazil (MESTRINER, SIMOES and SALZANO 1980; NEEL et al. 1980). A study of blood group markers indicated only 2.2% non-native admixture in this tribe (NEEL et al. 1980). Since 1942 the Brazilian Ticuna population has increased from 2,000 to 11,000 persons (SALZANO, CALLEGARI-JACQUES and NEEL 1980). This rapid population growth has been accompanied by migrations from jungle to river settlements (LAWRENCE, BODMER and BODMER 1980; NEEL et al. 1980; SALZANO, CALLEGARI-JACQUES and NEEL 1981; TROUP et al. 1982). Blood samples were taken from individuals living in three villages located along the Rio Solimoes (LAWRENCE, BODMER and BODMER 1980).

For convenience, in some of our analyses we have grouped populations linguistically. Thus, the Amerinds include the Pima, Maya, Ticuna, Pomo and Hopi, and the Nadene include the Dogrib, Tlingit and Navajo.

Materials and Methods:

Samples: Blood cells in the form of lymphoblasts, platelets oruffy coats were obtained from 167 subjects including 80 Nadene and 87 Amerinds. The Navajo were comprised of 30 Dogrib (NW Canada), 48 Navajo (Arizona and New Mexico) and 2 Tlingits (Alaska), and their bloods represent a random sampling of individuals within each respective tribal group. The Amerinds were comprised of 30 Pima (Arizona), 1 Hopi (Arizona), 1 Pomo (California), 27 Maya (Mexico) and 28 Ticuna (Brazil). The Pima and Ticuna individuals from whom blood was taken were known to be unrelated for at least three generations, whereas the Maya represent a random sample of that tribal group.

The Dogrib, speakers of a northern Athapaskan language, are the largest Indian group in the Canadian Northwest Territories. Their traditional subsistence economy was based on hunting and fishing in the boreal forest and adjacent barrenlands. Aboriginally they were subdivided into several regional bands, i.e., socioterritorial groups that moved within particular regions for much of the year (HELM 1981). The bloods used in this study were obtained from adult members of the Rae Band (SZATHMARY, RITENBAUGH and GOODBY 1987). In 1979 about 1600 people (75% of all Dogrib) were members of this band. Although traditionally nomadic, by 1960 the majority of the Rae Band had settled into permanent communities located within regional band areas. Consanguineal and affinal ties continue to link people in different settlements. Accordingly the Dogrib form a recognizable genetic unit among other Athapaskan-speakers, as shown by genetic distance analyses (SZATHMARY 1983).
procedures in SCHURR et al. 1990 or a modification of those. In the modified method, buffy coats or platelet pellets were incubated in a 500-μl digestion solution containing 100 mM NaCl, 10 mM Tris-Cl, pH 7.4, 5 mM Na₂EDTA, pH 8.0, 0.5% sodium dodecyl sulfate, and 500 μg/ml proteinase K, at 55° for 10-16 hr, mixed with 250 μl cold 5 M KOAc with vigorous shaking for 5 min, then incubated at 0° for 10 min, and the lysed cell debris pelleted through microcentrifugation. The supernatant was then twice extracted with phenol-chloroform, and genomic DNA within the aqueous phase ethanol precipitated.

All mtDNAs were polymerase chain reaction (PCR) amplified (SAIKI et al. 1985) in nine overlapping segments which encompassed the entire mtDNA genome (APPENDIX A). Each PCR segment was independently digested with 14 restriction endonucleases (AluI, AvaII, DdeI, HaeIII, HhaI, HinfI, HpaII, MboI, Rsal, TaqI, BamHI, HaeII, HincII), and the resulting fragments resolved by electrophoresis in 1.0-2.5% NuSieve + 1.0% SeaKem agarose (FMC BioProducts) gels and detected by ethidium bromide fluorescence. The fragments observed for each PCR segment were restriction mapped by the sequence comparison method (JOHNSON et al. 1983; CANN, BROWN and WILSON 1984), the composite restriction maps of these sets of segments representing individual haplotypes is given in APPENDIX B. The distribution and frequency of haplotypes among tribal groups is shown in Table 1.

**Sequence divergences:** Mean intra- and intergroup sequence divergences were estimated with NEI and TAJIMA's (1983) maximum likelihood procedure using the computer program DREST (graciously provided by L. JIN). This procedure considers the ratio of shared sites to the total number of sites between two haplotypes, and the mean length of the restriction enzyme recognition sequences to calculate an initial estimate of θ (the probability that the two mtDNAs have different nucleotides at a given nucleotide position). Using this initial estimate, θ is solved iteratively using equation 28 and the sequence divergence (δ) is estimated by equation 21 (NEI and TAJIMA 1983).

Sequence divergence values within and among the tribal groups are presented in Table 2. Divergence values shown in standard type represent interpopulational comparisons, those underlined represent intrapopulational comparisons, and those in bold type represent interpopulational comparisons, those in bold type represent interpopulational variations corrected for intrapopulational variation, δ = δi - 0.5(δj + δk), where δi is the mean pairwise divergence between individuals within a single population (X), δj is the corresponding value for a second population (Y), and δk is the mean pairwise divergence between individuals belonging to the two different populations (X and Y) (NEI and TAJIMA 1983). The Tlingit, Hopi and Pomo samples were too small in number to be used as separate units for divergence calculations among tribal groups, but were included in estimates of overall Nadene and Amerind divergences.

**Phylogenetic analysis:** The evolutionary relationships among the 50 Native American haplotypes were inferred using parsimony analysis (PAUP 3.0; SWOFFORD 1989). All but two haplotypes (28 and 29) segregated into four well defined clusters, A, B, C and D, each of which was delineated by a distinctive genetic marker (Table 1; Figure 1). The dendrogram presented in Figure 1 is rooted from HYPANC, "hypothetical ancestor a" (CANN, STUERDKING and WILSON 1987), and is 77 mutational steps in length with a consistency index of 0.833. When mtDNA haplotypes from modern Africans (our unpublished data) were substituted from HYPANC and used as outgroup, the branching order did not change. This dendrogram represents a tree generated by the subtree pruning and regrafting (SPR) branch
A. Torroni et al.

TABLE 2

mtDNA divergences of Amerinds and Nadene

<table>
<thead>
<tr>
<th>Population</th>
<th>Dogrib</th>
<th>Navajo</th>
<th>Pima</th>
<th>Maya</th>
<th>Ticuna</th>
<th>Nadene</th>
<th>Amerinds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dogrib</td>
<td>1.6 ± 1.5</td>
<td>5.2 ± 3.5</td>
<td>13.3 ± 6.5</td>
<td>8.0 ± 4.7</td>
<td>17.4 ± 7.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Navajo</td>
<td>1.3 ± 0.9</td>
<td>6.2 ± 3.8</td>
<td>13.2 ± 6.3</td>
<td>9.0 ± 5.0</td>
<td>17.9 ± 7.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pima</td>
<td>6.0 ± 2.3</td>
<td>3.6 ± 1.4</td>
<td>13.0 ± 6.0</td>
<td>14.9 ± 6.8</td>
<td>16.8 ± 7.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maya</td>
<td>2.0 ± 1.4</td>
<td>0.7 ± 0.3</td>
<td>3.2 ± 1.3</td>
<td>10.4 ± 5.1</td>
<td>17.9 ± 7.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ticuna</td>
<td>9.7 ± 3.6</td>
<td>5.8 ± 2.3</td>
<td>3.4 ± 1.2</td>
<td>5.8 ± 1.9</td>
<td>13.8 ± 6.2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Nadene: 5.0 ± 3.3 13.3 ± 6.3 3.3 ± 1.3 15.1 ± 6.7

Intrapopulation divergences are along the diagonal (underlined), interpopulation divergence above the diagonal in standard type, and interpopulation divergences corrected for intrapopulation variation (Nei and Tajima 1983) below the diagonal in bold type. Divergence values and standard errors are presented ×10^4.

RESULTS

Restriction analysis: A total of 67 polymorphic sites and the 9-base pair (9-bp) COI/1rRNA* intergenic deletion (Cann and Wilson 1983; Horai and Matsunaga 1986; Wrischnik et al. 1987; Hertzberg et al. 1989; Schurr et al. 1990) were observed in this analysis. These defined 50 haplotypes (Table 3).
Mitochondrial DNA Analysis

1.  The 14 restriction endonucleases used in this survey screened an average of 373 sites genome or over 10% of the mtDNA sequence per individual.

Genetic variation within Native American populations: Intra- and intergroup sequence divergences divide the populations into two groups that parallel the postulated linguistic divisions (Table 2). The mean intragroup divergence value for the Nadene is 0.050%, while that of the Amerinds is three times higher at 0.151%. This result confirms the general genetic cohesiveness of the Nadene observed in other studies (Száthmary and Ossenberg 1978; Szathmary 1979, 1983; Williams et al. 1985) and suggests that the Amerinds have radiated longer than the Nadene.

Within the Nadene, the intragroup divergence of the Navajo is much higher (0.062%) than that of the Dogrib (0.016%) (Table 2). This difference results from 37.5% of the Navajo having haplotypes associated with the 9-bp deletion which is not seen in those of the Dogrib or Tlingits. In fact, 6 out of the 10 Navajo haplotypes (haplotypes 13–18) are associated with the 9-bp deletion. However, all non-deletion haplotypes observed in the Navajo (1, 5 and 9) are shared with the Dogrib and Tlingits (Table 1).

Among the Amerinds, the Pima and Ticuna have similar intragroup divergences (0.132–0.138%) whereas the Maya are somewhat less divergent (0.104%).

The interpopulational divergence values revealed intriguing relationships between tribal groups. The South American Ticuna have similarly high divergence values relative to both the Nadene and the other Amerind groups (0.168–0.179%). The Pima also have similar divergence values relative to all other groups analyzed (0.132–0.168%), although they are less divergent from the Navajo and Dogrib than are the Ticuna. On the other hand, the Maya appear to be closer to the Nadene (0.080–0.090%) than they are to the Pima (0.149%) and Ticuna (0.179%). The reduced divergence between the Maya and the Nadene groups is primarily due to their sharing haplotype 1. This haplotype was present in 38.8% of the Nadene and 18.5% of the Maya but is not found in either the Pima or the Ticuna. This result could be explained by either gene flow from the Nadene to some Amerind populations, the loss of haplotype 1 in the Pima and Ticuna or admixture with a third migratory group.

Evolutionary relationships: Analysis of the phylogenetic relationships among Native American haplotypes revealed four well-defined mtDNA lineages, designated clusters A–D (Figure 1).

The haplotypes in cluster A are associated with the HaeIII np 663 site. This site has previously been observed at low frequencies in mtDNAs of East Asians [Han Chinese and Koreans (Ballinger et al. 1992)] and Chicanos4 (Cann, Stoneking and Wilson 1987). Cluster A is divided into two subclusters radiating from haplotypes 1 and 9. Haplotype 1 was found in 38.8% of the Nadene, 18.5% of the Maya and 5.0% of Taiwanese Han (haplotype 56 in the Asian phylogeny of Ballinger et al. 1992). Haplotype 9 differs from haplotype 1 by the presence of a HaeII np 16,517 site and was detected in 8.8% of the Nadene, 7.4% of the Maya, 7.1% of the Ticuna. This haplotype is only one mutational step away from the Asian haplotypes 28 and 103 found in 7.7% of Koreans, 7.1% of Han Chinese and 5.0% of Taiwanese Han (Ballinger et al. 1992). Therefore, haplotypes 1 and 9 are likely to have been the founding haplotypes of this lineage since they are nodal within the cluster, the most prevalent cluster A haplotypes in both the Nadene and Amerinds and the only ones found in Asia. Within cluster A two haplotypes, 5 and 6, have lost the Rsai np 16,329 site, a polymorphism not previously observed. This mutation occurs in all Nadene populations examined (50.0% of the Tlingits, 26.7% of the Dogrib, and 27.1% of the Navajo), but not in Amerinds. Therefore, the Rsai np 16,329 site loss appears to be a specific genetic marker for the Nadene.

The haplotypes in Cluster B are defined by the 9-bp deletion. The HaeIII site gain at np 16,517 was always found in association with the deletion. Deleted mtDNAs were observed in 24.0% of the Native American mtDNAs analyzed, with haplotype 13 representing 52.5% of the mtDNAs within the cluster. This

---

1. When using the term “lineage” and “haplotype,” we define “lineage” as a set of mtDNAs related to each other by shared common and unique mutations not observed in other mtDNAs. We define “haplotype” as a distinct association of restriction endonuclease site polymorphisms for 14 enzymes within one mtDNA genome.

2. The occurrence of the HaeII np 663 site in Chicanos mtDNAs (Cann, Stoneking and Wilson 1987) indicates Native American admixture in this population; for this reason it cannot be considered a “Caucasian” marker.
haplotype is also positioned at the node of the cluster, is the only one of the Amerindian deleted haplotypes which is found in East Asians [2.9–7.1% (haplotype 54 in BALLINGER et al. 1992)], and is the only cluster B haplotype shared by more than one population (Navajo, Pima and Maya). Consequently, haplotype 13 is the most likely ancestral mtDNA of this cluster.

The haplotypes of cluster C are characterized by an A to G transition at np 13,263 which simultaneously eliminates a HincII site at np 13,259 and creates an AluI site at np 13,262 (morph-6, WALLACE, GARRISON and KNOWLER 1985; SCHURR et al. 1990). This mutation was always found in association with a Ddel site gain at np 10,394 and an AluI site gain at np 10,397. Haplotypes within this cluster are absent from the Nadene but are present in all three Amerind groups (43.3% of the Pima, 14.8% of the Maya and 32.1% of the Ticuna) at an overall frequency of 29.9%. This mutation is also found at low frequencies in Asian mtDNAs, being observed in 1.8% of Japanese (HORAI, GOJOBORI and MATSUNAGA 1984), 1.6% of Orientals (BLANC et al. 1983) and 1 out of 153 East Asians (the only population samples analyzed with the same set of enzymes, BALLINGER et al. 1992). The phylogenetic analysis indicated haplotype 45 as being central to the radiation of the cluster, and since this haplotype is identical to the one observed in East Asians (haplotype 65) it is likely to be the founding mtDNA for this lineage.

The haplotypes of cluster D are distinguished by a loss of the AluI site at np 5,176 and this mutation is virtually always associated with a Ddel np 10,394 and AluI np 10,397 site gains. Haplotypes within this cluster are found exclusively among the Amerinds (Pomo, Maya, Ticuna), comprising 19.5% of the samples analyzed. This mutation has also been observed in mtDNAs of 23.1% of Koreans, 14.3% of Chinese Han and 10.0% of Taiwanese Han (BALLINGER et al. 1992). The phylogenetic analysis positioned haplotype 44 at the node of cluster D. This haplotype is the only one shared by two tribal groups (Pomo and Ticuna) and is only a single mutational step away from Asian haplotype 25 of BALLINGER et al. 1992. Hence, it is probably the founding haplotype of this lineage.

Caucasian admixture: Haplotypes 28 and 29 of the Maya and the Navajo, respectively, lack characteristic Native American mutations and do not cluster in any of the four Native American lineages (Figure 1). Haplotype 28 has an AluI site loss at np 7,025 found predominantly in Caucasian mtDNAs (ANDERSON et al. 1981; CANN, STONEKING and WILSON 1987; SHOFFNER et al. 1990; BROWN et al. 1992), and haplotype 29 has a Ddel site loss at np 7,715 found in about 10% of Caucasian mtDNAs (SHOFFNER et al. 1990; BROWN et al. 1992). Consequently, these haplotypes were probably introduced by European admixture.

DISCUSSION

Native American mtDNA markers: Our analysis shows that all Native American mtDNAs are characterized by one of the four following mutations: HaeIII np 663 site gain; 9-bp deletion; HincII np 13,259 site loss/AluI np 13,262 site gain; and AluI np 5,176 site loss. All of these mutations have been observed at much lower frequencies in Asians (HORAI, GOJOBORI and MATSUNAGA 1984; HARIHARA, HIRAI and OMOTO 1986; HORAI and MATSUNAGA 1986; CANN, STONEKING and WILSON 1987; HARIHARA et al. 1992). The first three of these mutations have not been observed in Africans and Caucasians, whereas the latter has been observed at low frequency in Africans but not in association with the Ddel np 10,394 and AluI np 10,397 site gains as observed in Amerind cluster D haplotypes (CANN, STONEKING and WILSON 1987; BROWN et al. 1992; A. TORRONI, unpublished data).

Hence, these are highly informative genetic markers which can be used to confirm American Indian admixture in other American ethnic groups.

Founding haplotypes: This study suggests that only five haplotypes (1, 9, 13, 43 and 44) founded the Amerinds, whereas only two or three of these (1, 9 and 13) founded the Nadene. In general, these are the most prevalent among the populations, are shared between different tribes, are the only ones also detected in Asians, and are positioned at the nodes of the phylogenetic haplotype clusters. Except for the nodal haplotypes and Nadene haplotype 5, all other mtDNA haplotypes are present at low frequencies and are tribal specific ("private polymorphisms"; NEEL 1978), suggesting that they developed in situ in the Americas after the initial radiation of the founding population/s in the Americas or in Siberia.

Genetic divergence times: Calculations of intracluster divergences yielded similarly high values for Amerind clusters C and D (0.102% and 0.073%; Table 3). Likewise, the divergence value for the cluster A haplotypes in the Amerinds was 0.071% which is essentially identical to that calculated for cluster D. The sequence divergence of the Nadene cluster A was much lower than the Amerind specific clusters, being only 0.021%. The overall divergence of the cluster B deletion haplotypes in Navajo, Pima and Maya, was of 0.038% and independent cluster B calculations for the Navajo and Amerinds gave very similar values of 0.032% and 0.037%, respectively. Hence, the divergence of cluster B haplotypes is intermediate between the Amerind and Nadene specific lineages.

The similarity of the sequence divergences of the Amerind clusters A, C and D suggests that these were the founding lineages of the Paleoindians. Using the weighted mean value of 0.084%, the Amerind mtDNA lineages are estimated to have begun radiat-
The ancestors of the Navajo are believed to have migrated south in small familial groups through regions exclusively inhabited by Amerinds about 1000 A.D. (HASKELL 1987). After their movement into what is now north-central New Mexico, the Navajo became semisedentary and mixed both genetically and culturally with Puebloan Indians, borrowing both agricultural practices and ceremonial rituals (SPENCER et al. 1977). Their genetic admixture with surrounding peoples has also been confirmed by the presence of Albumin variants specific to southern Native Americans (Albumin Mexico; SCHELL and BLUMBERG 1988). However, this scenario does not explain how the Navajo acquired deletion haplotypes from the Amerinds without obtaining any haplotypes from Amerind clusters C and D (Table 1, Figure 1). Similarly, the limited amount of divergence observed for Amerind cluster B haplotypes (0.037%) can not be reconciled with the higher divergence values determined for the other Amerind clusters.

The third alternative is that the cluster B haplotypes arrived in the Americas in a separate migration that was intermediate between those of the Paleoindians and ancestral Nadene. As this migration moved south, perhaps along the continental divide, it would have mixed with some Amerind tribes. The subsequent arrival and limited southward expansion by the ancient Nadene would explain the preferential admixture of the deletion haplotype in the Navajo and the similar and intermediate divergence of cluster B haplotypes in both modern Nadene and Amerinds. Interestingly enough, in Asia the frequency of mtDNAs characterized by the 9-bp deletion reaches its highest value in coastal populations (BALLINGER et al. 1992) and deletion mtDNAs were virtually the only ones present in the migratory groups of Asian ancestry that began to colonize Polynesia around 5,000 YBP (HERTZBERG et al. 1989). Hence this could have been a separate migration. Clearly, further clarification of the origin and migrations of American Indian mtDNAs will require characterization of the nature and frequency of these mtDNA lineages in Asia, Siberia and the Pacific.

The authors would like to thank C. RONALD SCOTT for contributing the Tlingit samples and JAMES DERR for assistance with the PAUP bootstrap analysis. This work was supported by National Science Foundation grant BNS8718775 awarded to D.C.W.

LITERATURE CITED


BLANG, H., K. H. CHEN, M. A. D'AMORE and D. C. WALLACE,


Shoffner, J. M., IV, M. T. Lott, A. M. S. Lezza, P. Seibel, S. W. Ballinger and D. C. Wallace, 1990 Myoclonic epilepsy and ragged-red fiber disease (MERRF) is associated with a...
Mitochondrial DNA Analysis


Communicating editor: V. G. FINNERTY

APPENDIX A

Oligonucleotide primers for PCR amplifications of Native American mtDNAs are shown in Table 4.

**TABLE 4**

<table>
<thead>
<tr>
<th>Primer coordinates</th>
<th>Tm (°C)</th>
<th>Segment size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1562–1581, 3717–3701</td>
<td>51</td>
<td>2155</td>
</tr>
<tr>
<td>3007–3023, 5917–5898</td>
<td>55</td>
<td>2910</td>
</tr>
<tr>
<td>5317–5333, 7608–7588</td>
<td>57</td>
<td>2291</td>
</tr>
<tr>
<td>7392–7410, 8921–8902</td>
<td>57</td>
<td>1529</td>
</tr>
<tr>
<td>8282–8305, 10107–10088</td>
<td>57</td>
<td>1825</td>
</tr>
<tr>
<td>9911–9932, 11873–11851</td>
<td>69</td>
<td>1962</td>
</tr>
<tr>
<td>11675–11691, 13950–13932</td>
<td>57</td>
<td>2277</td>
</tr>
<tr>
<td>13914–13930, 16547–16527</td>
<td>47</td>
<td>2633</td>
</tr>
<tr>
<td>16453–16472, 1696–1677</td>
<td>61</td>
<td>1812</td>
</tr>
</tbody>
</table>

Primer pair are numbered according to ANDERSON _et al._ (1981), with the 5' → 3' coordinates before the comma corresponding to the forward primer and those after the comma to the reverse primer. The Tm used for annealing was the lowest for each primer pair as calculated from the nucleotide sequence of each primer, where Tm = 4( C+ G) + 2(T + A) − 5°. All primers were synthesized at the Emory University Microchemical Facility.

APPENDIX B

Figure 2 shows the polymorphic restriction sites observed in Native American mtDNA haplotypes.
**Figure 2.** Table presenting the polymorphic restriction sites observed in Native American mtDNA haplotypes. Haplotypes are numbered according to Table 1. A "1" indicates the presence of a site and a "0" indicates the absence of a site except for region V where "1" indicates a single copy of the 9-bp repeat (deletion) and "2" indicates two copies of the repeat. Sites are numbered from the first nucleotide of the recognition sequence according to the published sequence (Anderson et al. 1981). Bold face numbers indicate site gains relative to the published sequence and standard type numbers indicate site losses. The 14 restriction enzymes used in the analysis are designated by the following single-letter code (after Cann, Brown and Wilson 1984): a, AluI; b, AvaII; c, DdeI; e, HaeIII; f, Hhal; g, Hinfl; h, HpaI; i, HpaII; j, MboI; k, RsaI; l, TaqI; m, BamHI; n, HaeII; o, HincII. Sites separated by a diagonal line indicate either simultaneous site gains or site losses for two different enzymes or a site gain for one enzyme and a site loss for another because of a single inferred nucleotide substitution; these sites are considered to be only one restriction site polymorphism in the statistical analysis. Sites marked with an asterisk were found to be present or absent in all samples (except where polymorphic) contrary to the published sequence, and were confirmed in mtDNAs from all major ethnic subdivisions by DNA sequencing (Wallace et al. 1988; Shoffner et al. 1990; Brown et al. 1992; A. Torroni, unpublished data).