Native American mtDNA Prehistory in the American Southwest

Ripan S. Malhi,[1]* Holly M. Mortensen,[2] Jason A. Eshleman,[3] Brian M. Kemp,[3] Joseph G. Lorenz,[4] Frederika A. Kaestle,[5] John R. Johnson,[6] Clara Gorodezky,[7] and David Glenn Smith[3,8]

[1]*Department of Human Genetics, University of Michigan, Ann Arbor, Michigan 48109*
[2]*Department of Biology, University of Maryland, College Park, Maryland 20742*
[3]*Department of Anthropology, University of California, Davis, California 95616*
[4]*Coriell Institute for Medical Research, Camden, New Jersey 08103*
[5]*Department of Anthropology, Indiana University, Bloomington, Indiana 47401*
[6]*Santa Barbara Museum of Natural History, Santa Barbara, California 93106*
[7]*Department of Immunogenetics, INDRE, Mexico City, Mexico 77600*
[8]*California Regional Primate Research Center, University of California, Davis, California 95616*

KEY WORDS admixture; migration; Uto-Aztecan; Athapaskan; Hohokam; Anasazi; haplotype

ABSTRACT This study examines the mtDNA diversity of the proposed descendants of the multiethnic Hohokam and Anasazi cultural traditions, as well as Uto-Aztecan and Southern-Athapaskan groups, to investigate hypothesized migrations associated with the Southwest region. The mtDNA haplogroups of 117 Native Americans from southwestern North America were determined. The hypervariable segment I (HVSI) portion of the control region of 53 of these individuals was sequenced, and the within-haplogroup diversity of 18 Native American populations from North, Central, and South America was analyzed. Within North America, populations in the West contain higher amounts of diversity than in other regions, probably due to a population expansion and high levels of gene flow among subpopulations in this region throughout prehistory. The distribution of haplogroups in the Southwest is structured more by archaeological tradition than by language. Yumans and Pimans exhibit substantially greater genetic diversity than the Jemez and Zuni, probably due to admixture and genetic isolation, respectively. We find no evidence of a movement of mtDNA lineages northward into the Southwest from Central Mexico, which, in combination with evidence from nuclear markers, suggests that the spread of Uto-Aztecan was facilitated by predominantly male migration. Southern Athapaskans probably experienced a bottleneck followed by extensive admixture during the migration to their current homeland in the Southwest. Am J Phys Anthropol 120: 108–124, 2003.

Stretching from Baja California to New Mexico and from Utah and Colorado south to the regions of Sonora and Chihuahua, the southwestern region of North America is characterized by diversity in landscape and culture. The people indigenous to this region include speakers of the Yuman, Seri, Piman, and Southern Athapaskan (Na-Dene) languages, as well as the culturally defined Pueblo groups. The languages and cultures of these five groups differ markedly, and the five are presumed to have experienced separate origins and prehistories. The study of prehistory of the southwestern region of North America is dominated by evidence of geographically widespread archaeological cultures practiced by multiethnic groups exhibiting marked language diversity. Ironically, in contrast to the great linguistic diversity, the region is genetically characterized by a remarkably homogenous and high frequency of mitochondrial DNA (mtDNA) haplogroup B (Lorenz and Smith, 1996). Using a larger and more representative sample of populations and mtDNA sequence data, this study examines the genetic structure of the descendants of the multiethnic Hohokam and Anasazi cultural traditions as well as the nature of the hypothesized Uto-Aztecan and Southern Athapaskan migrations into or from the Southwest region.

Grant sponsor: National Institute of Health; Grant numbers: RR00169, RR05090; Grant sponsor: Regents of the University of California; Grant number: GER9255683; Grant sponsor: National Science Foundation; Grant number: SBR9630926.

*Correspondence to: Ripan S. Malhi, Department of Human Genetics, University of Michigan, Ann Arbor, MI 48109.
E-mail: malhi@umich.edu

Received 13 June 2001; accepted 15 May 2002.

DOI 10.1002/ajpa.10138
Published online in Wiley InterScience (www.interscience.wiley.com).

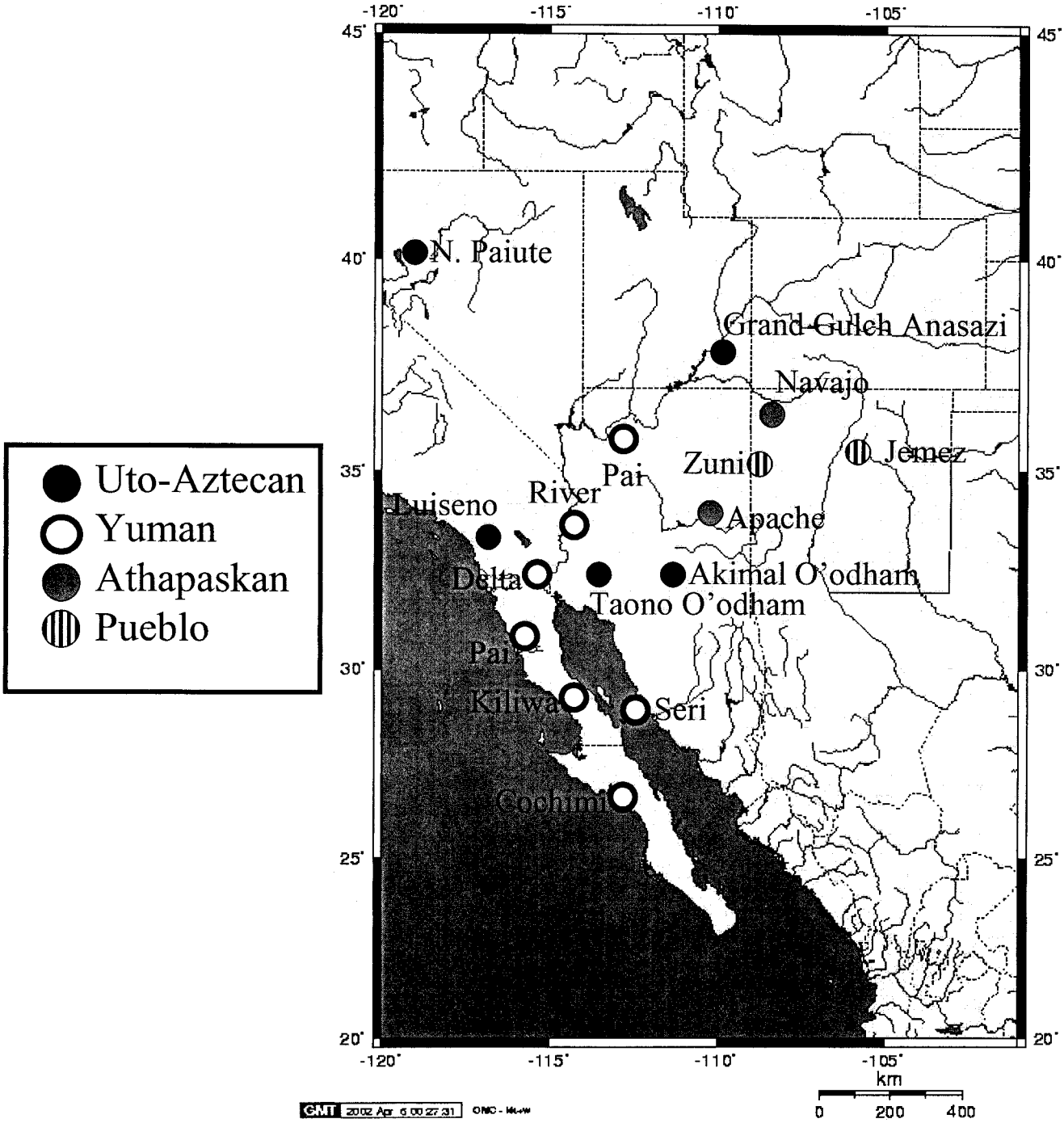

Fig. 1. Geographic location of populations analyzed in this study.

BIOCULTURAL CONTEXT

The Yumans inhabit the western end of the Southwest (Fig 1). Yuman languages have been divided into four major branches: 1) Kiliwa, consisting only of the Kiliwa language, located in Baja California; 2) Pai, located in Baja California and Arizona; 3) River Yuman, along the Colorado River, in southern California, and northern Baja California; and 4) Delta Yuman, within the Colorado Delta (Kendall, 1983). However, minor differences among Yuman languages indicate that the divisions within the language family are not very ancient, but instead represent a continuum of very closely related languages across geographic space (Kendall, 1983). Kiliwa is the most divergent of the four Yuman branches. Linguistically, the Cochimi represent the closest relative outside of the Yuman language family, and has often been classified within this family in the past (Goddard, 1996). Based on the pre-European-contact homeland of the Kiliwa and Cochimi, proto-Yuman

Fig. 2. Geographic locations of populations in the Americas analyzed for diversity in haplogroup B. 1, Nuu-Chah-Nulth (Ward et al., 1991); 2, Yakima (Shields et al., 1993); 3, Washo (Kaestle, 1998); 4, Yuman (this study); 5, Pueblo (this study); 6, Piman (this study); 7, Norris Farms (Stone and Stoneking, 1998); 8, Choctaw (Weiss, 2001); 9, Chickasaw (Weiss, 2001); 10, Cherokee (Malhi et al., 2001); 11, Ngobe (Kolman et al., 1995); 12, Kuna (Batista et al., 1995); 13, Cayapa (Rickards et al., 1999); 14, Yanomama (Merriwether et al., 2000); 15, Xavante (Ward et al., 1996); 16, Pehuenche (Moraga et al., 2000); 17, Mapuche (Chile) (Moraga et al., 2000); 18, Mapuche (Argentina) (Ginther et al., 1993).

is believed to have originated in Baja California and to have begun diversifying and expanding northward approximately 1,000 years before present (BP) (Hale and Harris, 1983). The material culture of the ancestors of Yuman speakers is presumed to be the Hakataya (a part of which is referred to as Patayan) archaeological tradition (Schroeder, 1963; Cordell, 1997). This tradition was centered in the Colorado River valley and extended Southwestward into southern California and Baja California (Schroeder, 1963).

The Seris live across the Sea of Cortez from Baja California in Sonora, Mexico and speak an isolate language (Fig.2). The geographic proximity of the Seri to Yuman speakers suggests the potential for a recent admixture. In addition, Kroeber (1915) presented a strong case for including Seri in the Hokan language superfamily together with Yuman languages and languages surrounding California's central valley. Currently, the exact relationship between the Yuman and Seri is undefined (Goddard, 1996).

The upper Pimans, consisting of the Akimal O'odham (Pima) and the Taono O'odham (Papago) peoples, are located in southeastern Arizona and the Mexican state of Sonora and are part of the Tepiman languages, which extend from Jalisco to Arizona. While the Akimal O'odham have admixed with other nearby groups, such as the River Yuman, the Taono O'odham have remained highly endogamous (Smith, 1981). Thus, a comparison between the Akimal O'odham and Taono O'odham could reveal genetic traits acquired by the Akimal O'odham through admixture (Brown et al., 1958).

The Tepiman languages are part of the Uto-Aztecan language family, whose distribution extends from Central America to the northern peripheries of the Great Basin. The origin of the Uto-Aztecan language family is disputed. The largest amount of diversity among Uto-Aztecan languages is located in Southern California, suggesting that the greatest antiquity and, therefore, the homeland for Uto-Aztecan is in Southern California approximately 5000 BP (Miller, 1983). However, Hill (2001) argued that Uto-Aztecan originated in Central Mexico, later spreading into the Southwest, Southern California, and the Great Basin. The spread of Uto-Aztecan languages northward might have been driven by the development of maize cultivation and a related population expansion in Central Mexico, which Hill (2001) believes is the source of agriculture-related terms in Hopi.

The intrusion of a Mesoamerican influence, including agriculture, into the Southwest about 3500–2500 BP coincides with the fluorescence of the Hohokam cultural tradition, while the decline of that tradition roughly temporally correlates with the diversification of proto-Yuman. Schroeder (1963) argued that the Hohokam tradition emerged from a Mesoamerican cultural influence on the Hakataya tradition, widely considered to have been practiced by ancestors of modern Yuman-speaking tribes, but others regard Hohokam as an introduction of Mesoamerican emigrants (DiPeso, 1956). Whether the expansion of Mesoamerican influence in the American Southwest was demic (i.e., the result of the migration of peoples) or simply the expansion of cultural interaction spheres, both proto-Yumans and proto-Pimans are hypothesized to have participated in the Hohokam tradition. Linguistic evidence (Shaul and Hill, 1998) and evidence from burial practices (Shaul and Anderson, 1989) suggest that Hohokam encompassed a multiethnic community that consisted of ancestors of Yumans and Pimans, and perhaps also the Zuni later in time.

Pueblo groups are geographically confined to the northern region of the Southwest and include a linguistically diverse set of people who share common cultural traits, including living in compact permanent settlements, a common ceremonial system, and similar world-view (Eggan, 1950). The antecedents of this multilingual group consisted of four different language families, Uto-Aztecan, Zuni, Keresan, and Kiowa-Tanoan, whose speakers share relative genetic homogeneity (Brown et al., 1958; Workman et al., 1974), and whose ancestors are presumed to be the people of the Anasazi tradition, dating as early

as 3000 BP (Fagan, 2000). The connection between the Anasazi and modern Pueblo groups is strengthened by a continuous culture chronology between the two. In addition, Carlyle et al. (2000) showed that the mitochondrial haplogroup frequency distributions of Anasazi people from Grand Gulch, dating to 2000 BP, are not significantly different from those of modern Pueblo groups, establishing biological as well as cultural continuity.

The Southern-Athapaskan speakers, the Navajo and Apache, are widely dispersed throughout the central region of the Southwest. Archaeologists and linguists agree that their ancestors arrived in the Southwest from a homeland to the north relatively recently, approximately 500 BP (Basso, 1983), and quickly adapted in diverse ways to their new homeland. The Navajo adopted a Pueblo lifestyle displaying the cultural patterns similar to those seen in the Hopi, Zuni, and other Pueblo groups. The Apache, however, maintained a nomadic lifestyle.

MATERIALS AND METHODS

Populations studied

The locations of the populations studied in the Southwest are shown in Figure 1. The sources for serum samples from the Zuni, Jemez, Akimal O'odham, Northern Paiute, Nahua, Pai Yuman, River Yuman, Delta Yuman, Kiliwa, Cochimi, Navajo, and Apache are described in Smith et al. (2000). The Seri samples were obtained from Clara Gorodezky (Department of Immunogenetics, INDRE, Mexico City, Mexico), and the Taono O'odham samples from Moses Schanfield (Analytical Genetics Testing Center, Denver, CO). The Zuni and Jemez are Pueblo groups. The upper Piman speakers (Akimal O'odham and Taono O'odham), the Northern Paiute of the Great Basin, and the Nahua from Cuetzlalan, Mexico all speak Uto-Aztecan languages. The Yavapai, Paipai, Kumeyaay (Diegueno), and Kiliwa represent the four main branches of the Yuman language group, and the Cochimi speak the most closely related language outside the Yuman group. Additional samples from the literature and from Lorenz et al. (unpublished) were included in analyses of group diversity (Ward et al., 1991, 1996; Ginther et al., 1993; Shields et al., 1993; Batista et al., 1995; Kolman et al., 1995; Kaestle, 1998; Rickards et al., 1999; Merriwether et al., 2000; Moraga et al., 2000; Malhi et al., 2001; Weiss, 2001).

The haplogroups of a total of 117 Native Americans were determined by restriction fragment length polymorphism (RFLP). A subset of 53 samples was sequenced from nucleotide positions (np) 16055–16548 in this study and analyzed together with an additional 29 samples that had been previously sequenced (Table 1).

TABLE 1. Samples used for dna sequence analysis[1]

Population	N	References
Alaskan Athapaskan	5	Shields et al., 1993
Tlingit	1	Torroni et al., 1993
Apache	8	This study (7); Torroni et al., 1993
Navajo	7	This study (6); Torroni et al., 1993
Jemez	8	This study
Zuni	5	This study
Akimal O'odham	7	This study (6); Torroni et al., 1993
Taono O'odham	3	This study
Northern Paiute	6	Kaestle, 1998
Nahua	5	This study
Seri	8	This study
Pai	5	This study (3); Lorenz and Smith, 1997
Cochimi	1	This study
Cocopa	2	Lorenz and Smith, 1997
Kiliwa	3	This study (1); Lorenz and Smith, 1997
Kumeyaay	4	This study (1); Lorenz and Smith, 1997
Luiseno	1	Lorenz et al., unpublished
Tubatulabel	1	Lorenz et al., unpublished
Opata	1	Lorenz et al., unpublished
Gabrielino	1	Lorenz et al., unpublished

[1] Numbers in parentheses indicate number of samples analyzed in this study.

DNA extraction and typing

DNA was extracted from 200 μl of serum using the Qiagen Blood Amp Kit. Amplification reactions were carried out in a 25-μl volume with 1–3 μl of DNA template, 50 μM of each primer, 10X Buffer (50 μM Tris, pH 8.4, 1.5 μM $MgCl_2$, 20 μM NaCl, and 500 mg/ml BSA), 1.5 units of Platinum Taq (Gibco), 200 μM of each dNTP, and 13.3 μl of ddH_2O. After an initial 4-min denaturation step at 95°C, 40 cycles were performed consisting of a denaturing at 95°C for 30 sec, an annealing step at 52–55°C for 30 sec, and an extending step at 72°C for 30 sec, followed by a final 3-min extension at 72°C. A 5-μl portion of amplification product was electrophoresed on a 6% polyacrylamide gel and stained with ethidium bromide to confirm the presence of PCR product. To assess the presence or absence of diagnostic restriction sites, the remaining 20 μl were incubated with 10 units of the appropriate restriction enzyme overnight at 37°C. Primers used for amplification of these segments are described in Smith et al. (1999).

Hypervariable segment I (HVSI) of the control region was amplified using primers described in Smith et al. (1999). The PCR products were filtered using a Microcon 100 filter unit (Millipore) and then submitted for sequencing to the DBS Automated DNA sequencing facility at the University of California at Davis. Both the heavy and light strands were sequenced to preclude sequencing errors. All sequences generated for this study can be found in the Appendix.

DNA haplogroup and sequence analysis

Altogether, 479 individuals, including those from 362 additional samples previously studied or reported in the literature, were analyzed in this study. Any individuals in a given sample determined not to

belong to haplogroups A, B, C, D, or X were assumed to represent non-Native American admixture (Smith et al., 1999) and were excluded from analysis. Treating the five Native American haplogroups as alternate alleles at a single locus, gene (haplogroup) diversity was estimated as:

$$h = \left(1 - \sum_{i=1}^{k} p_i^2\right) \Big/ n - 1$$

(Nei, 1987), where n is the number of gene copies in the sample, k is the number of haplogroups, and p_i is the sample frequency of the i-th haplogroup.

Pairwise comparisons and tests for homogeneity of haplogroup frequency distributions were made between all populations and groups using Fisher's exact probability (Weir, 1990), bootstrapping each comparison with 1,000 iterations using the Genepop software program (Raymond and Rousset, 1995). The Kiliwa and Seri were excluded from this part of the analysis due to their extremely small sample sizes (7 and 8, respectively). Genetic distances were calculated between all pairs of populations in the Southwest, using the chord distance measurement of Cavalli-Sforza and Edwards (1967) in GENDIST, and phylogenetic trees were constructed by the neighbor-joining method using NEIGHBOR and DRAWTREE in the PHYLIP 3.572 software package (Felsenstein, 1993). A consensus tree was constructed, with 100 iterations, using SEQBOOT and CONSENSUS in the PHYLIP software package. A principal coordinates analysis was performed for all groups inhabiting the Southwest. The coordinates were calculated in Genstat for Windows, using genetic similarity between populations ($1 - F_{ST}$), calculated from F_{ST} values determined in Arlequin version 2.000 (Schneider et al., 2000), and the first two coordinates are reported. Finally, an analysis of molecular variance (AMOVA) was performed, using the Arlequin package (Schneider et al., 2000), to determine whether gene flow in populations in the Southwest was structured more strongly by language boundaries or by shared archaeological traditions.

Due to the polyphyletic lineage history of Native Americans (Schurr et al., 1990), we limited our analyses to within-haplogroup comparisons. By excluding interhaplogroup comparisons, we preclude most influence of prehistoric population events that occurred in Asia prior to settlement of the Americas. Due to sampling and variation in haplogroup frequencies of Native American groups, some haplogroups are better suited for answering specific questions about population prehistory than others. Haplogroups A, B, and C are high in frequency in northern Athapaskans, Southwest populations, and most Uto-Aztecan groups, respectively. Therefore, in this study, haplogroup A was used to study the Southern Athapaskan migration, haplogroup B to study genetic relationships among Southwest populations, and haplogroup C to investigate the spread of the Uto-Aztecan languages. Haplogroups D and X are nearly absent from Southwest populations, and therefore were excluded from analysis. Haplotype median-joining networks were constructed using the Bandelt Network Program (Bandelt et al., 1999). Nucleotide positions 16182–16183 were excluded from analysis of haplogroup B haplotypes, since polymorphism at these sites appears to be hypervariable and is neither informative of phylogenetic relationships nor reported in a consistent manner by different authors. Nucleotide position 16519 was also excluded from the analysis because it is hypervariable.

Theta (θ_S), an estimate of genetic diversity, was calculated as:

$$\theta_S = \frac{S}{\sum_{i-1}^{n=1} \frac{1}{i}}$$

(Watterson, 1975), where S is the number of segregating sites and n is the sample size.

All calculations were performed using the ARLEQUIN package (Schneider et al., 2000) and Microsoft Excel. Theta (θ_S), which is similar to the estimator E(v) as described by Excoffier and Laganey (1989), reflects the diversity in a population due to long-term history and is less influenced by generational and sampling effects. Estimates of genetic diversity within haplogroup C were excluded from the analysis, because sample sizes within groups were too small to provide an accurate estimate of diversity.

RESULTS

Haplogroup frequency distribution

As reported in previous studies (Lorenz and Smith, 1994, 1996; Carlyle et al., 2000; O'Rourke et al., 2000; Smith et al., 2000) and illustrated in Table 2, which gives the distribution of haplogroups by group, lineage B is the predominant haplogroup in the American Southwest region, reaching a maximum frequency in the Jemez Pueblo (89%) and a minimum among the Western Apache (13.2%). The frequency of haplogroup C is more uniform than that of haplogroup B across populations of the Southwest, reaching a maximum in the Cochimi and Delta Yuman populations (46.2% and 43.5%, respectively). Consequently, gene diversity (h) is lower than 60% for all but one of the 14 populations studied, because most individuals belong to either haplogroup B or C. Although most populations in the Southwest are characterized by relatively high frequencies of haplogroups B and C, the fixation of haplogroup B in the Kiliwa and the near fixation of haplogroup C in the Seri are unusual. These findings probably reflect sampling errors due to small sample size, or perhaps intense genetic drift in extremely small and/or isolated populations (Infante et al., 1999).

TABLE 2. Haplogroup frequency distribution and haplogroup diversity of native american populations[1]

Population	Language	N	A	B	C	D	X	h	References
Zuni	Zuni	26	0.154	0.769	0.077	0.000	0.000	0.394	Lorenz and Smith, 1996; this study (5)
Jemez	Tanoan	36	0.000	0.889	0.028	0.000	0.083	0.208	Lorenz and Smith, 1996; Smith et al., 1999; this study (3)
Akimal O'odham	Uto-Aztecan	43	0.047	0.535	0.395	0.000	0.023	0.568	Torroni et al., 1993; Lorenz and Smith, 1896; this study (6)
Taono O'odham	Uto-Aztecan	37	0.000	0.568	0.378	0.054	0.000	0.546	This study
N. Paiute/Shoshoni	Uto-Aztecan	94	0.000	0.426	0.096	0.479	0.000	0.586	Kaestle and Smith, 2001
Nahua	Uto-Aztecan	31	0.613	0.323	0.065	0.000	0.000	0.533	Lorenz and Smith, 1996; this study (2)
Pai Yuman	Yuman	27	0.074	0.667	0.259	0.000	0.000	0.501	Lorenz and Smith, 1996; Smith et al., 2000; this study (11)
River Yuman	Yuman	22	0.000	0.636	0.364	0.000	0.000	0.485	Lorenz and Smith, 1996; Smith et al., 2000; this study (1)
Delta Yuman	Yuman	23	0.000	0.565	0.435	0.000	0.000	0.515	Lorenz and Smith, 1996; Smith et al., 2000; this study (20)
Kiliwa	Yuman	7	0.000	1.000	0.000	0.000	0.000	0.000	Lorenz and Smith, 1996; Smith et al., 2000; this study (4)
Cochimi	Yuman	13	0.077	0.462	0.462	0.000	0.000	0.614	Lorenz and Smith, 1996; Smith et al., 2000; this study (3)
Seri	Yuman	8	0.000	0.125	0.875	0.000	0.000	0.250	This study
Navajo	Athapaskan	64	0.516	0.406	0.047	0.000	0.031	0.575	Torroni et al., 1993; Lorenz and Smith, 1996; this study (8)
Apache	Athapaskan	38	0.632	0.132	0.184	0.053	0.000	0.561	Torroni et al., 1993; Lorenz and Smith, 1996; this study (9)

[1] Numbers in parentheses indicate number of samples analyzed in this study.

Haplogroup A is extremely rare or absent in most of the populations studied except the Southern Athapaskans, i.e., the Navajo and Apache, in whom its frequency reaches 51.6% and 63.2%, respectively, and the Nahua, in whom its frequency reaches 61.3%. The Apache and Taono O'odham are the only populations in the Southwest that exhibit haplogroup D (approximately 5% in each population). The rarity of haplogroup D in the Southwest is in stark contrast to the high (48%) frequency of haplogroup D among speakers of Uto-Aztecan languages in the adjacent Great Basin (Paiute-Shoshone). Haplogroup X is present in low frequency in the Jemez Pueblo, Navajo, and Akimal O'odham (8.3%, 3.1%, and 2.3%, respectively). The presence of haplogroup X in the Akimal O'odham represents the first reported incidence of this lineage in a Uto-Aztecan-speaking population

Overall, Pueblo groups display a high frequency of haplogroup B, Yuman and Piman groups exhibit moderate frequencies of both haplogroups B and C, and Athapaskan groups share high frequencies of haplogroup A. In comparison to other Southwest populations, the Zuni and Jemez Pueblo exhibit low levels of gene diversity (0.394 and 0.208, respectively), due to very high frequencies of haplogroup B. This paucity of gene diversity could reflect matrilocal residence and the lack of female gene flow from neighboring or invading groups, consistent with the prehistoric lifeways of these people (Steward, 1937; Workman et al., 1974). The three Uto-Aztecan groups studied here (the Paiute/Shoshone of the Great Basin, the Akimal O'odham and Taono O'odham of the arid Southwest, and the Nahua of central Mexico) differ markedly from each other because of their uniquely high frequencies of haplogroups D, B, and A, respectively.

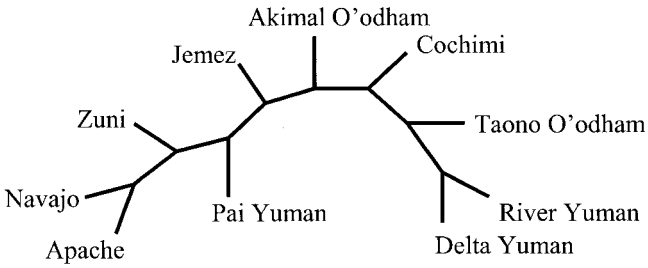

Fig. 3. Consensus tree of southwestern tribes, using chord distance measures based on haplogroup frequency distributions.

The Athapaskan groups cluster together in the consensus tree (Fig. 3), as do the Delta and River Yuman groups, probably due to common ancestry. The results of Fisher's exact test between all pairs of populations reveal no significant difference among the Yuman, Cochimi, and the Piman groups (P = 0.7067, SE = 0.016). However, haplogroup distributions of the Navajo and Apache are significantly different from each other (P = 0.00, SE = 0.00) as well as from all other groups in the Southwest. The haplogroup distribution of the Zuni and Jemez Pueblo are also significantly different from each other (P = 0.014, SE = 0.00), but the haplogroup distribution of the Zuni Pueblo is not statistically significantly different from that of the Pai Yuman (P = 0.22, SE = 0.00). The Pai and Zuni Pueblo are neighboring groups, and recent admixture could explain the similarity between them. However, the two geographically distant Pai groups were pooled in this analysis because they were not statistically in-

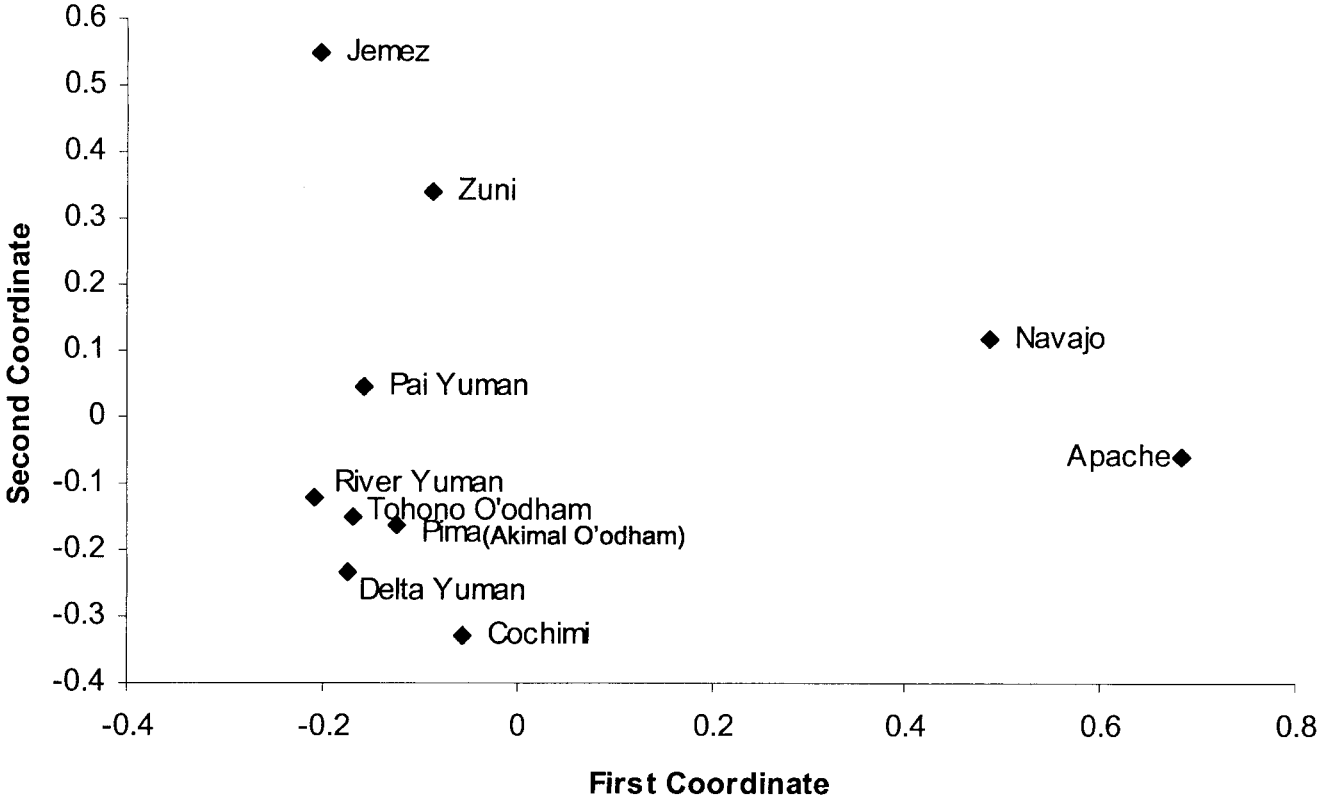

Fig. 4. Principal coordinates analysis.

distinguishable from each other, suggesting that recent admixture does not fully explain this pattern.

The AMOVA of haplogroup frequency distributions for populations in the Southwest assigned the majority (74%) of haplogroup variation to differences within populations. Differences between descendants of different archaeological (cultural) traditions (26.17%) account for a greater proportion of the total variation than do differences between language families (21.78%). This result suggests that participation in common prehistoric lifestyles and/or geography were more instrumental in structuring gene flow than was language in the Southwest.

Fifty-three percent of the variation in the principal coordinates analysis, shown in Figure 4, is accounted for by the first coordinate (X-axis), and 39% is explained by the second coordinate (Y-axis). This analysis revealed one main cluster that includes the Yuman groups, the linguistically related Cochimi, and the Akimal O'odham (Fig. 4), in agreement with the results of Fisher's exact test. The Zuni are located equidistant from this main cluster and the Jemez Pueblo. Finally, the Navajo and Apache group at a distance from the main cluster. The high frequency of haplogroup A, which the Navajo and Apache share, is almost certainly due to common ancestry, as haplogroup A approaches fixation in other Athapaskan groups in Alaska, such as the Dogrib (Torroni et al., 1993; Merriwether et al., 2000; Lorenz and Smith, 1996), and it was probably nearly fixed in the unadmixed Athapaskans who founded the Apachean populations in the Southwest.

Deletions/insertions

Two of three Nahua members of haplogroup A (assessed through the presence of the *Hae*III restriction site gain at np 663 and the presence of diagnostic HVSI mutations at np 16223, 16290, 16319, and 16362) also possessed the COII-tRNA^{lys} intergenic 9-bp deletion. This is consistent with previous conclusions that the 9-bp deletion has occurred more than once in several different continents, and suggests that the deletion has multiple origins in the Americas, a pattern previously seen in Africa (Soodyall et al., 1996) and India (Watkins et al., 1999). The haplogroup A/9-bp deletion motif has been reported in one Boruca individual (Torroni et al., 1993), one Maya individual (Schurr et al., 1990), three Baja Mixtec (Torroni et al., 1994), and 10 individuals from the Northern Mexican cities of Juárez and Ojinaga (Green et al., 2000). Recently, this derived form of haplogroup A was discovered in three pre-Columbian Aztec individuals from Tlateloco, whose remains date to approximately 500–700 BP (Kemp et al., 2002). The only report of this type outside of Mexico or Central America is in one individual from Puerto Rico (Martínez-Cruzado et al., 2001). It has been suggested that the haplogroup A/9-bp deletion type was brought to Puerto Rico via pre-Columbian slave trade between the Caribbean

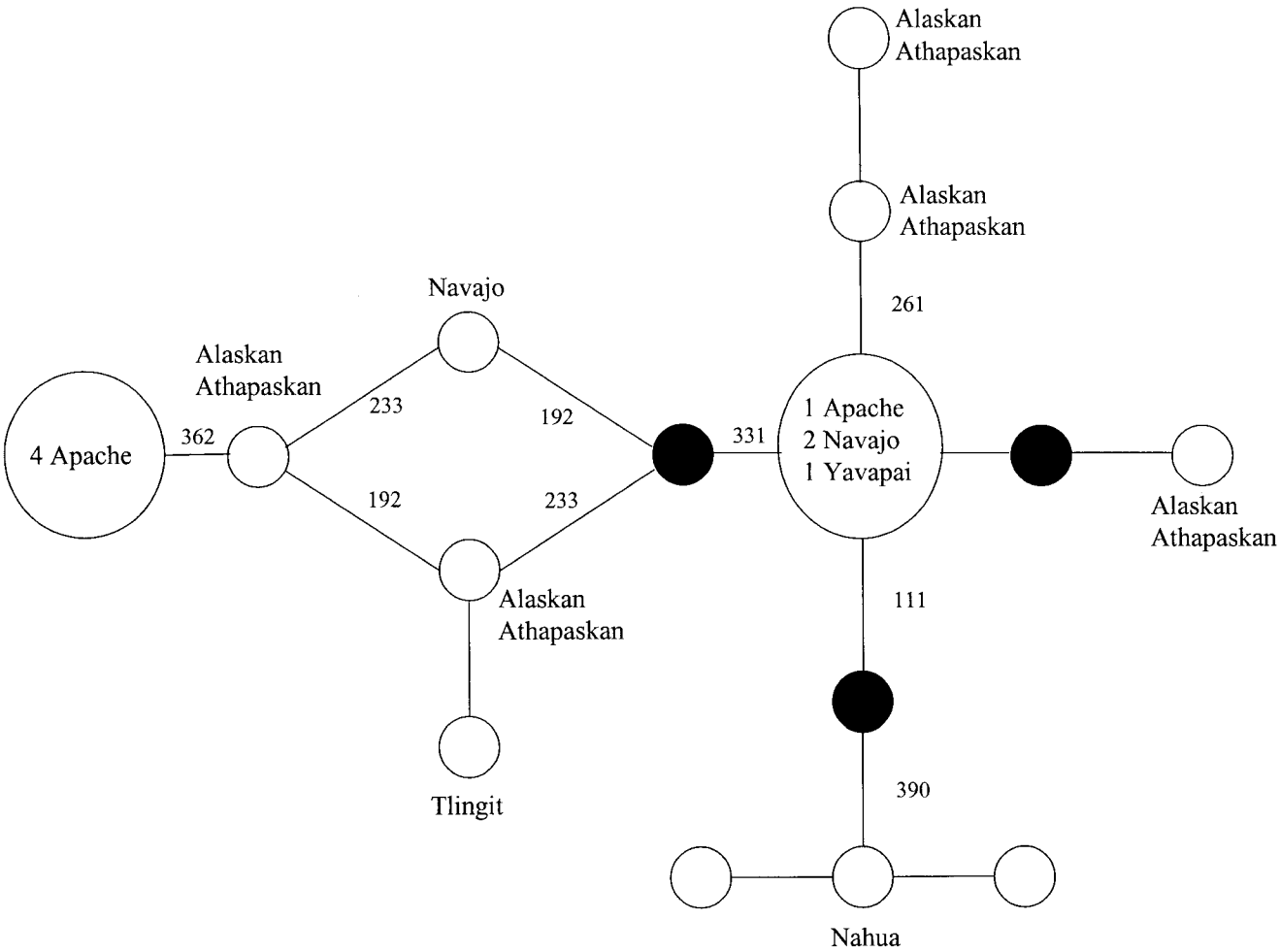

Fig. 5. Haplogroup A network. Numbers correspond to last three digits of nucleotide position, and indicate defining mutations for each clade. Size of circle and numbers preceeding names correspond to number of individuals found with that haplotype. Solid circles represent hypothetical haplotypes not found in our sample.

and theYucatan Peninsula (Martínez-Cruzado et al., 2001).

Two of three Taono O'odham samples assigned to haplogroup B whose HVSI regions were sequenced exhibited a CC insertion between np16193 and np16194. This dinucleotide insertion was previously unreported in Native North American populations and might represent a private polymorphism in the Taono O'odham population. However, this insertion has also been reported in the Kuna, a Central American population, that is not linguistically closely related to the Taono O'odham (Batista et al., 1995). It is unclear whether this dinucleotide insertion is hypervariable and developed independently in the Taono O'odham and the Kuna, or if this reflects a distant common ancestry. Further analysis of additional HVSI sequences from members of haplogroup B in the Americas is needed to address this issue.

DNA sequence analysis

The mtDNA haplotype networks based on HVSI sequences are given in Figures 5–7 for haplogroups A, B, and C, with sample sizes of 18, 36, and 29, respectively. The haplogroup A network contains Alaskan Athapaskan, Nahua, and Southwest haplotypes. The Southern Athapaskans are found in three haplotypes in the A network. Three Southern Athapaskan samples are found together with one Yavapai haplotype, in what Forster et al. (1996) described as the A1 founding haplotype. The remaining five Southern Athapaskan haplotypes are found together in a clade (also containing two Alaskan Athapaskans and a Tlingit) defined by a mutation at np 16331. The Nahua haplotypes cluster together, but do not cluster with the Athapaskan haplotypes, due to mutations at np 16111 and np 16390.

The haplogroup B network contains three central shared haplotypes. However, the haplotype defined by Forster et al. (1996) as the founding B lineage is not shared. This latter haplotype is found only in four Jemez samples. This is rather surprising, as founding haplotypes are generally found in wider distributions than derivative ones (Forster et al., 1996). The most common haplotype, shared by the Navajo, Zuni, Jemez, and Seri, is characterized by

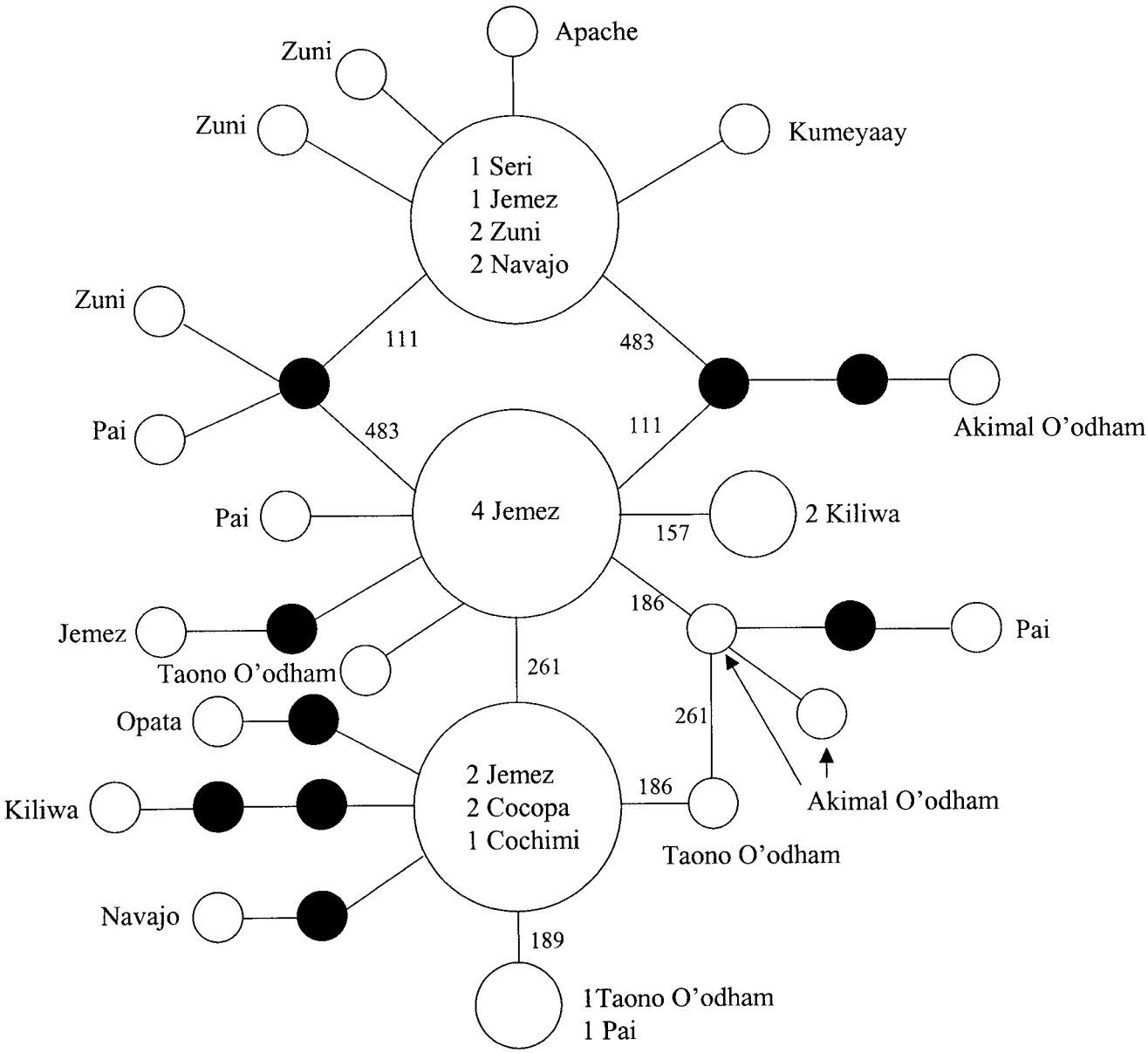

Fig. 6. Haplogroup B network. Numbers correspond to last three digits of nucleotide position, and indicate defining mutations for each clade. Size of circle and numbers preceeding names correspond to number of individuals found with that haplotype. Solid circles represent hypothetical haplotypes not found in our sample.

mutations at np 16111 and np 16483. The next most common haplotype is shared among the Jemez, Cocopa, and Cochimi samples, and is defined by a mutation at np 16261. The third shared haplotype, the least common of the three, is shared by a Yuman (Pai) and Piman (Taono O'odham) sample. In addition, many Pimans share a mutation at np 16186, although they are further differentiated from each other by additional mutations. Another feature of this network is the large number of undetected haplotypes (unobserved intermediate haplotypes that are steps between observed haplotypes).

The haplogroup C network contains haplotypes from the Southwest as well as from Uto-Aztecan groups from Central Mexico, the Great Basin, and Southern California. Other than the founding haplotype of haplogroup C (Forster et al., 1996), shared by the Shoshone and central Uto-Aztecan groups (together with three Seri), the network displays distant genetic relationships among the three geographically distant Uto-Aztecan groups. The Seri that fall outside the founding haplotype cluster tightly together, all containing a mutation at np 16301, and the Northern Paiute cluster together due to a mutation at np 16189. A single haplotype is shared between one Delta Yuman (Kumeyaay) and one California Uto-Aztecan (Luiseño). This result likely reflects gene flow structured through geo-

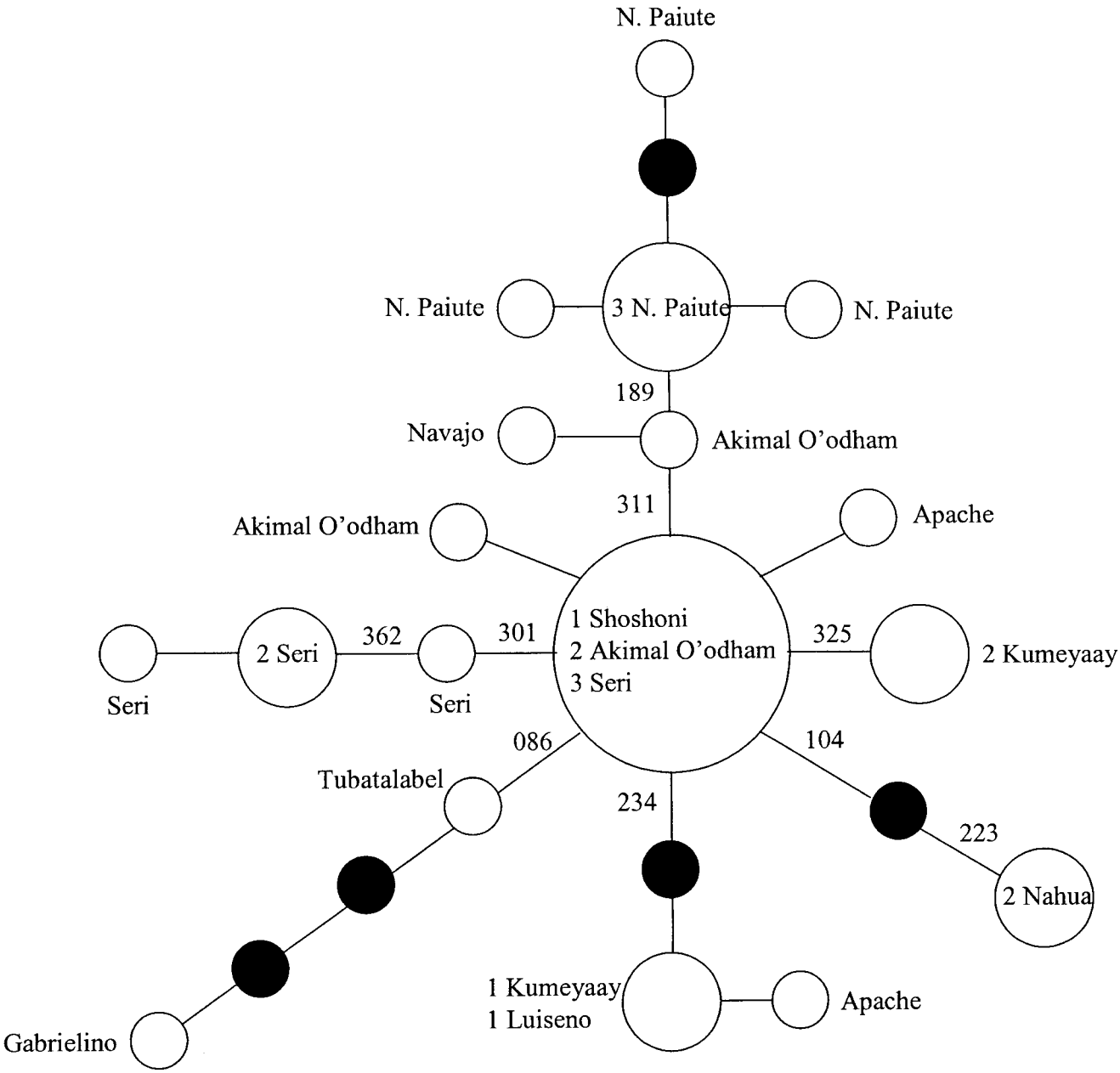

Fig. 7. Haplogroup C network. Numbers correspond to last three digits of nucleotide position, and indicate defining mutations for each clade. Size of circle and numbers preceeding names correspond to number of individuals found with that haplotype. Solid circles represent hypothetical haplotypes not found in our sample. The maternal ancestry of the Gabrielino sample is uncertain.

graphic proximity, since the Luiseño and the Kumeyaay are neighboring groups living near the border of California and Mexico. Both Apache haplotypes are a single mutational step away from Yuman or Uto-Aztecan haplotypes, suggesting that the Apache acquired them through admixture.

Table 3 shows the genetic diversity within haplogroups for Southwest populations and likely descendants of archaeological traditions. Table 4 displays the genetic diversity within haplogroup B for a large number of populations throughout the Americas (see Fig. 2 for corresponding geographical locations). Southern Athapaskans exhibit substantially less di-

TABLE 3. Diversity estimates of native american groups for haplogroups a and b

Population	N	Haplotypes	θ_S
Haplogroup A			
Alaskan Athapaskan	6	6	3.500
Southern Athapaskan	8	3	1.930
Haplogroup B			
Zuni	5	4	1.440
Jemez	8	4	1.930
Piman	6	6	3.940
Yuman	11	9	4.440
Descendants of Pueblo	13	8	1.930
Descendants of Hohokam/ Hakatayan	17	15	5.030

TABLE 4.
Diversity estimates within haplogroup B for 18 tribes from North, Central, and South America[1]

Population	N	Haplotypes	θ_S	References
Cayapa	6	1	0.000	Rickards et al., 1999
Ngobe	15	3	0.310	Kolman et al., 1995
Mapuche (Chile)	8	2	0.386	Moraga et al., 2000
Kuna	18	3	0.580	Batista et al., 1995
Xavante	21	3	0.834	Ward et al., 1996
Chikasaw	5	5	1.320	Weiss and Smith, unpublished findings
Yanomama	10	5	1.410	Merriwether et al., 2000
Choctow	5	3	1.440	Weiss, 2001
Nuu-Chah-Nulth	5	4	1.440	Torroni et al., 1993; Malhi, 2001
Yakima	15	4	1.540	Shields et al., 1993
Norris Farms	7	4	1.630	Stone and Stoneking, 1998
Cherokee	11	5	1.710	Malhi et al., 2001
Mapuche (Argentina)	15	5	1.850	Ginther et al., 1993
Pueblo	15	7	2.150	This study
Pehuenche	7	6	2.450	Moraga et al., 2000
Piman	8	6	3.090	This study
Yuman	11	8	4.100	This study
Washo	5	4	4.320	Kaestle, 1998

[1] Nucleotide positions 16092–16360 were analyzed.

versity in haplogroup A than Alaskan Athapaskans. For haplogroup B, Pueblo groups show lower diversity than Yuman and Piman groups. Estimates (based on θ_S) of such diversity are higher in Southwestern (and highest in the Washo) than in other populations in the Americas. North American and some South American populations exhibit higher diversity within haplogroup B than do Central and Northern South American populations.

DISCUSSION

Origins and pattern of haplogroup B diversity

The known emergence and expansion of archaeological traditions in the American Southwest had a significant effect on the genetic structure of native populations in this region. This pattern is apparent despite the homogenizing effects of high frequencies of haplogroup B in most Southwest populations. This high frequency of haplogroup B is accompanied by a high level of diversity within this haplogroup.

Aside from the Washo in Western North America, Southwest populations contain the highest amount of diversity within haplogroup B in the Americas (Kaestle, 1998). In contrast, populations in Central and South America exhibit a drastically reduced level of diversity within haplogroup B, as evidenced by their low value of θ_S and their high proportion of founding haplotypes and single haplotypes one mutational step away from the founding lineage (network not shown). This supports the related hypotheses that 1) these same populations underwent a population bottleneck during the peopling of Central and South America (Batista et al., 1995; Kolman et al., 1995), and that 2) a high population density was reached in Central America soon after South America was settled, inhibiting Southward migrations from North America (O'Rourke et al., 1992). Kolman and Bermingham (1997) speculated that the cultural and genetic distinctiveness of Central American populations suggests that they acted as a barrier to migration through this region. Thus, populations of South America might have experienced an extended period of very low population density, relative isolation, and genetic drift, due to a second genetic bottleneck subsequent to the founder effect associated with the earliest settlement of the Americas.

While the Mapuche of Argentina and the Pehuenche of Chile exhibit unusually large amounts of diversity within haplogroup B relative to other populations of South America, the Mapuche of Chile (from Huapi Island) exhibit significantly different haplogroup distributions than the Mapuche of Argentina (Moraga et al., 2000). A similar pattern is exhibited in the Yanomama of Brazil and Venezuela (Merriwether et al., 2000), and is consistent with events leading to strong genetic drift. The high levels of diversity in some groups are probably due to recent admixture with neighboring populations (O'Rourke et al., 1992). The increased level of isolation among tribes is probably attributable to higher levels of habitat and linguistic diversity in South America than in North or Central America (Mace and Pagel, 1995).

The high frequency and diversity of haplogroup B in the Southwest are probably due to an early colonization of this region by populations that contained or developed a high frequency of haplogroup B, followed by a rapid population expansion later in time. Currently, the oldest archaeological site in the Southwest (the Aubery site), at 11,550 rcBP (13,400 BP), contains Clovis technology (Fiedel, 1999) whose users probably experienced the effects of very low population densities. One possible explanation for the predominance of haplogroup B in these early Southwest populations is that early inhabitants of the Southwest were big game hunters who experienced genetic drift due to an initial small population

size, causing haplogroup B to become the predominant haplogroup in this region. Alternatively, haplogroup B might represent an independent migration to the Americas, due to its absence in Siberia and curious level of genetic diversity (Starikovskaya et al., 1998; Schurr et al., 1999).

The introduction of maize agriculture from Central Mexico (approximately 3500 BP; Smith, 1995) probably contributed to the eventual expansion of haplogroup B. Even though agriculture spread from Central Mexico to North and South America at about the same time (Smith, 1995), this population expansion associated with agriculture in North America was far more limited in South America, as evidenced by the lack of genetic homogeneity over a large geographic area usually observed with a population expansion. The limited influence of maize agriculture on populations in South America was probably due to the barrier of Central American populations to an expansion southward as well as high levels of ecological diversity in South America.

Southwest populations also exhibit relatively high frequencies of the B haplotype with T at np 16,261 also found in Mongolia (Kolman et al., 1996) and South China (Yao et al., 2000). The occurrence of this haplotype in the Nuu-Chah-Nulth, located in the Pacific Northwest (Malhi, 2001), as well as in Southwest populations, suggests that this haplotype might be a founding lineage in colonizing populations (Malhi et al., 2002). Previous studies attempting to determine the number of founding haplotypes for Native Americans were based on relatively few samples representing a large geographic region, and probably resulted in an oversimplified view of the peopling process. Forster et al. (1996) were only able to identify a single founding B haplotype from their general survey of populations throughout the Americas. The present study shows that an extensive survey of mitochondrial DNA variation within regions that exhibit high frequencies of a certain haplogroup, in this case haplogroup B, can reveal previously unknown potential founding Native American haplotypes. Detailed studies of populations in the Northeast and the interior West of North America might identify additional founding haplotypes for haplogroups C and D, respectively, which are found in high frequencies in these regions (Lorenz and Smith, 1996). Due to the high frequency of lineage extinctions in populations over time (Avise, 2000), it is possible that additional founding haplotypes do not survive in modern Native American populations. In that event, analysis of ancient populations in North America holds the greatest potential for discovering additional Native American founding haplotypes.

Anasazi and Hohokam

The Zuni share a main haplotype with the Jemez, suggesting, along with the archaeological record, a common ancestry located in the heart of the Southwest perhaps as long as 3000 BP. Based on the distribution of haplogroup frequencies, the Pueblo groups are not statistically genetically different from the prehistoric Anasazi from Grand Gulch (Carlyle et al., 2000). The low levels of diversity within haplogroup B for the Jemez and Zuni suggest that these populations experienced similar histories. Workman et al. (1974) showed that Zuni and Taos Pueblo groups share an unusually high level of blood group B and overall lack genetic diversity, likely due to isolation of the Zuni and other Pueblo groups. The archaeological record of the ancient Anasazi and Pueblo traditions reveals a large-scale abandonment of village sites, followed by aggregation into compact isolated communities (Fagan, 2000). This pattern suggests the possibility of high levels of lineage extinction due to a genetic bottleneck in the Anasazi and Pueblo groups that might have resulted in the low diversity of haplogroup B found among their descendants.

The Jemez, Cocopa, and Cochimi also share a central haplotype within haplogroup B that is rare outside of these Southwest groups. Ancestors of the Yumans probably had close contact with the ancestral Jemez. Therefore, contrary to conclusions based on the linguistic data, genetic data point to a Yuman homeland in the Arizona/New Mexico region of the Southwest rather than in Baja California. The Yumans may then have expanded to Southern California and Baja California later, as evidenced by the distribution of the Hakatayan culture.

The nearly identical distribution of haplogroups in both Yumans and Pimans is consistent with blood group data (Brown et al., 1958), while the appearance of the Albumin*Mexico variant in all tribes representing both of these groups suggests either extensive admixture between these groups or common ancestry for them. The Zuni also contain the Albumin*Mexico variant (Schell and Blumberg, 1977), and their haplogroup distribution is statistically indistinguishable from that of the Pai Yuman. These results are in accordance with linguistic evidence presented by Shaul and Hill (1998) suggesting that ancestors of these three groups participated in the Hohokam culture. It is interesting to note that no Piman haplotypes are represented in the three main shared haplotypes of the B network, and that many of the Piman haplotypes contain unique mutations at np 16186 and np 16317. It is possible that the ancestors of Pimans are not native to the American Southwest but migrated to the American Southwest from the region now identified as the Mexican state of Sonora, approximately 1500 BP. Pimans probably extensively admixed with, and introduced Albumin*Mexico to, Yumans upon entry into the American Southwest, during the emergence of the Hohokam cultural period. If the south to north movement of the ancestors of Pimans coincides with the spread of the Tepiman languages, this is in disagreement with the conclusions of Shaul and Hill (1998), who provided multiple lines of linguistic

evidence that suggest a north to south spread of the Tepiman languages. Perhaps the split and southward spread of Tepiman languages postdated the movement of Piman ancestors into the American Southwest. It is also possible that analysis of nuclear and Y-chromosome markers will show a different genetic pattern, since the distribution of mtDNA haplotypes is biased by female movement.

Uto-Aztecan migration

Due to the independent origin of the 9-bp deletion in members of haplogroup A in the Americas, it is possible that samples identified as haplogroup B in Mesoamerica, using 9-bp deletion detection and not confirmed by mtDNA control region sequence analysis or tested for the presence of the *Hae*III site gain at np 663, are actually members of haplogroup A. Therefore, the Nahua might be even less similar to the Pimans and Northern Paiute than previously reported (Smith et al., 2000), due to an overestimate of the frequency of haplogroup B in Mesoamerica. The large differences in haplogroup frequency distributions among populations of the main branches of Uto-Aztecan, along with the distribution of haplotypes in the haplogroup C network, suggest that the spread of Uto-Aztecan was not the result of a population expansion northward caused by the development of maize cultivation in Mesoamerica. A population expansion caused by the development of agriculture would have likely involved the movement of women; therefore, the distribution of Uto-Aztecan was caused either by a language/culture spread that did not involve the movement of people, or by the migration of predominantly males, perhaps merchants engaged in trade activity along the Tepiman corridor.

The latter hypothesis is more consistent with the distribution of Albumin*Mexico and GM haplotypes (Callegari-Jacques et al., 1993). Anthony (1990) described the behavior of migration as typically performed by defined groups. He described Julius Caesar's documentation of the migration of the Helvetii in 58 BC as a movement inspired by the ideology of "glory-seeking young men." The major interpretation of the linguistic evidence suggests that proto-Uto-Aztecan diversified and spread southward from the American Southwest approximately 5500 BP (Miller, 1983). The direction of this movement agrees with Aztec legends of their descent from Chichimec barbarians from the north who invaded Central Mexico approximately 700 BP (Fagan, 1984). However, this interpretation is in disagreement with the pattern of distribution of Albumin*Mexico. The distribution of Albumin*Mexico is in equilibrium with mtDNA haplogroups in the Pimans but not in the Yumans (Smith et al., 2000). Disequilibrium between the Albumin and mtDNA loci in Yumans suggests they acquired Albumin*Mexico from the Pimans relatively recently. That Albumin*Mexico is widely dispersed among groups speaking a variety of Uto-Aztecan and non-Uto-Aztecan languages in Mexico, but is limited to Southwestern groups whose ancestors participated in the Hohokam cultural tradition, suggests that the mutation was introduced into the Southwest by immigrants from Mesoamerica. It is possible that the pattern and distribution of Albumin*Mexico in the American Southwest is the result of male Piman ancestors moving north, approximately 1500 BP, long after the initial spread of the Uto-Aztecan languages. Studies of nuclear DNA, especially Y-chromosome markers, should provide insight into the origins and spread of Uto-Aztecan languages.

Athapaskan migration

The lower sequence variation in the Southern Athapaskans compared to Northern Athapaskan groups suggests that Southern Athapaskans experienced a founder effect/bottleneck during and/or after their migration to the Southwest. This founder effect probably explains the high frequency of the otherwise rare haplotypes with mutations at np 16331 and np 16233 in Southern Athapaskans. Since this haplotype is not seen in the Nahua (Uto-Aztecan) of Mexico or in Native American haplotypes from North-Central Mexico (Green et al., 2000), a majority of haplogroup A present in the Southwest must have arrived with Athapaskans migrating from the North rather than from Mexico. However, this founder effect does not explain the disparity in haplogroup frequencies between Northern and Southern Athapaskans. If a founder effect were responsible for the higher frequencies of haplogroups B, C, and D in Southern Athapaskans, the Navajo and the Apache should exhibit similar frequencies of these haplogroups, because linguistic evidence suggests that the Navajo and Apache migrated to the Southwest as a single group (Hoijer, 1956).

However, the Navajo and Apache display significantly different haplogroup frequencies. The Navajo share a major nonfounding haplogroup B haplotype with the Zuni and Jemez, while the Apache share a haplogroup C haplotype with the Yavapai. Brown et al. (1958) also observed that Navajo and Apache groups share blood group phenotypes with those groups in closest geographic proximity to them. This result suggests that the Southern Athapaskans obtained haplogroups other than A solely through admixture. Specifically, the Navajo admixed with the Pueblo groups and the Apache admixed with the Yuman and Piman groups. This result is consistent with the high frequency of Albumin*Mexico in the Apache and the low frequency of Albumin*Mexico in the Navajo, since the Yumans and Pimans possess Albumin*Mexico and the Pueblo groups do not (Smith et al., 2000).

In addition, historic records document that during the formation of the historic Navajo population, large numbers of Pueblo refugees were absorbed into Navajo populations during the Pueblo Revolt of the 1680s (Brooks, 1999). This amalgam-

APPENDIX: DNA Sequence (nucleotide positions 16055–16548)

Sample	16092	16111	16188	16189	16192	16223	16233	16290	16319	16331	16362	16390	16519	N
CRS	T	C	C	T	C	C	A	C	G	A	T	G	T	
Navajo A	.	T	.	.	.	T	.	T	A	.	C	.	.	2
Navajo A	.	T	.	C	T	T	G	T	A	G	C	.	.	1
Nahua A	C	.	.	.	.	T	.	T	A	.	C	A	C	1
Nahua A	.	.	T	.	.	T	.	T	A	.	C	A	.	1
Nahua A	.	.	.	.	.	T	.	T	A	.	C	A	.	1
Apache A	.	T	.	.	T	T	G	T	A	G	.	.	.	4

Sample	16075	16092	16111	16157	16164	16182	16183	16186	16189	16197	16217	16223	16227	16249	16261	16278	16311	16317	16325	16342	16357	16483	16519	N
CRS	T	T	C	T	A	A	A	C	T	C	T	C	A	T	C	C	T	A	T	T	T	G	T	
Zuni B	.	.	T	.	.	.	C	.	C	.	C	.	.	.	.	.	.	.	.	.	C	A	C	1
Zuni B	.	.	T	.	.	.	C	.	C	.	C	.	.	.	.	.	.	.	.	.	.	A	C	2
Zuni B	.	.	.	.	.	.	C	.	C	T	C	.	.	.	.	.	.	.	.	.	.	A	C	1
Zuni B	.	.	T	.	.	.	C	.	C	T	C	.	.	.	.	.	.	.	.	.	.	A	C	1
Jemez B	.	.	.	.	.	C	C	.	C	.	C	.	.	.	.	.	.	.	.	.	.	.	C	1
Jemez B	.	.	T	.	.	.	C	.	C	.	C	.	.	.	.	.	.	.	.	.	.	A	C	1
Jemez B	.	.	.	.	.	.	C	.	C	.	C	.	.	.	T	.	.	.	.	.	.	.	C	2
Jemez B	.	.	.	.	.	C	C	.	C	.	C	.	.	.	.	.	.	.	.	.	.	.	C	3
Jemez B	.	.	.	.	.	.	C	.	C	T	C	.	.	.	.	T	.	.	.	.	.	.	C	1
Seri B	.	.	T	.	.	C	C	.	C	.	C	.	.	.	.	.	.	.	.	.	.	A	C	1
Navajo B	.	.	T	.	.	.	C	.	C	.	C	.	.	.	.	.	.	.	.	.	.	A	C	1
Navajo B	N	N	N	.	.	.	.	.	C	.	C	.	.	.	.	.	.	.	.	.	.	A	C	1
Apache B	.	.	T	.	.	.	C	.	C	.	C	.	.	.	.	.	.	.	.	.	.	A	C	1
Pima B	.	.	.	.	.	.	C	T	C	.	C	.	.	.	.	.	.	T	.	.	.	.	C	1
Pima B	.	C	.	.	.	.	C	T	C	.	C	.	.	.	.	.	.	.	.	.	.	.	C	1
Taono O'odham B	C	.	.	.	.	.	.	.	C	.	C	.	.	.	.	.	.	T	.	.	.	.	C	1
Taono O'odham B	.	.	.	.	.	.	C	T	C	.	C	.	.	.	T	.	.	.	.	.	.	.	C	1
Taono O'odham B	.	.	.	.	.	.	C	.	.	.	C	.	.	.	T	.	.	.	.	.	.	.	C	1
Paipai B	.	.	.	.	.	.	C	.	.	.	C	.	.	.	T	.	.	.	.	.	.	.	C	1
Paipai B	.	.	.	C	.	.	.	T	C	.	C	.	.	C	.	.	.	.	.	.	.	.	C	1
Yavapai B	.	.	.	.	.	.	C	.	C	.	C	.	.	.	.	.	.	.	.	C	.	A	C	1
Kiliwa B	.	.	.	.	.	.	C	.	C	T	C	T	G	.	T	.	.	.	.	.	.	.	C	1
Kumeyaay B	.	.	T	.	.	C	C	.	C	.	.	.	.	.	.	.	.	.	.	.	.	A	C	1
Cochimi B	.	.	.	.	.	.	C	.	C	.	C	.	.	.	T	.	.	.	.	.	.	A	C	1

Sample	16104	16129	16188	16223	16234	16295	16298	16301	16311	16325	16327	16362	16385	16519	N
CRS	C	G	C	C	C	C	T	C	T	T	C	T	A	T	
Navajo C	.	.	T	T	.	.	C	.	C	C	.	.	.	.	1
Nahua C	T	.	.	.	.	.	C	.	.	C	T	.	.	.	2
Apache C	.	.	.	T	.	.	C	.	.	C	T	.	G	.	1
Apache C	.	A	.	T	T	T	C	.	.	C	T	.	.	.	1
Pima C	.	.	.	T	.	.	C	.	.	C	T	.	.	.	2
Pima C	.	.	.	T	.	.	C	.	C	C	T	.	.	.	1
Seri C	.	.	.	T	.	.	C	T	.	C	T	C	.	C	1
Seri C	.	.	.	T	.	.	C	T	C	C	T	C	.	.	1
Seri C	.	.	.	T	.	.	C	T	.	C	T	.	.	.	1
Seri C	.	.	.	T	.	.	C	.	.	C	T	.	.	C	2
Seri C	.	.	.	T	.	.	C	T	.	C	T	C	.	.	1
Seri C	.	.	.	T	.	.	C	.	.	C	.	.	.	C	1

Sample	16111	16179	16182	16183	16189	16213	16223	16278	16483	16519	N
CRS	C	C	A	A	T	G	C	C	G	T	
Pima X	T	T	C	C	C	.	.	T	A	C	1

CRS, Cambridge Reference Sequence (Anderson et al, 1981).

ation probably produced some of the similarities observed between Navajo and Pueblo groups, but it does not fully explain the genetic patterns for Southern Athapaskans described above. Since non-Athapaskan Southwestern groups do not carry significant levels of haplogroup A, it is reasonable to conclude that gene flow with Athapaskan groups was almost entirely unidirectional. Perhaps Southern-Athapaskans acquired wives through warfare or trade, a circumstance that might have been necessary for the survival of a (presumably) small immigrant group.

CONCLUSIONS

The diversity within haplogroup B in populations from western North America suggests that groups exhibiting high frequencies of haplogroup B experienced a population expansion in this region in prehistoric times. The introduction of maize cultivation into the Southwest may have contributed to this expansion. In contrast, populations in South America exhibit low diversity estimates within haplogroup B. This low amount of diversity may be a result of a population bottleneck during the peopling

of South America, or the result of relatively increased isolation among South American populations. Further investigation into the within-haplogroup diversity among many widespread North American and South American populations may clarify this issue.

The results from this study suggest that language differences played a minimal role in structuring gene flow among populations in the Southwest. Despite the high frequency of haplogroup B within Pueblo groups, they exhibit a paucity of diversity within this haplogroup. Yuman and Piman groups, however, exhibit a large amount of diversity within haplogroup B and in haplogroup frequencies. This suggests that groups in the Southwest experienced significantly different population histories, possibly as a result of their inclusion in different cultural traditions during prehistoric times.

The distribution of mtDNA haplogroups and haplotypes among Uto-Aztecan-speaking groups in the Southwest and in Central Mexico suggests that the spread of Uto-Aztecan was not the result of a population expansion northward caused by the development of maize cultivation, as suggested by Hill (2001). The distribution of nuclear markers such as Albumin*Mexico (Smith et al., 2000), however, suggests that the spread of Uto-Aztecan may have been a predominantly male-mediated event. In addition, the reduced amount of variation in haplogroup A in Southern Athapaskans compared to Northern Athapaskan groups suggests that Southern Athapaskans experienced a founder effect during their migration to the Southwest. However, the significant difference in haplogroup frequencies between the Apache and Navajo is the result of a large amount of admixture with different Southwest groups. Specifically, the Apache admixed with Yuman and Piman groups, while the Navajo admixed with Pueblo groups. Future studies, focusing on nuclear and specifically Y-chromosome variation within Southwest, Athapaskan, and Uto-Aztecan groups, will provide useful information that can be used to evaluate the existence and nature of these migrations.

ACKNOWLEDGMENTS

We are indebted to numerous personnel of Indian Health Service facilities where most of the samples studied were obtained, and to individuals who provided samples. We also thank Ann Horsburgh, David Shaul, and Victor Golla for assistance in the laboratory and/or helpful discussions.

LITERATURE CITED

Anderson A, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJH, Staen R, Young IG. 1981. Sequence and organization of the human mitochondrial genome. Nature 290:457–470.

Anthony DW. 1990. Migration in archeology—the baby and the bathwater. Am Anthropol 92:895–914.

Avise JC. 2000. Phylogeography: the history and formation of species. Cambridge, MA: Harvard University Press.

Bandelt HJ, Forster P, Rohl A. 1999. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16:37–48.

Basso KH. 1983. Western Apache. In: Ortiz A, editor. Handbook of North American Indians: Southwest 10. Washington, DC: Smithsonian Institution. p 139–152.

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

Brooks JF. 1999. Violence, justice, and state power in the New Mexican borderlands, 1780–1880. In: White R, Findley, J, editors. Power and place in the North American west. Seattle University of Washington Press.

Brown KS, Hanna BL, Dahlberg AA, Strandskov HH. 1958. The distribution of blood group alleles among Indians of Southwest North America. Am J Hum Genet 10:175–195.

Callegari-Jacques SM, Salzano FM, Constans J, Maurieres P. 1993. GM haplotype distribution in Amerindians—relationship with geography and language. Am J Phys Anthropol 90:427–444.

Carlyle SW, Parr RL, Hayes MG, O'Rourke DH. 2000. Context of maternal lineages in the Greater Southwest. Am J Phys Anthropol 113:85–101.

Cavalli-Sforza LL, Edwards AWF. 1967. Phylogenetic analysis models and estimation procedures. Am J Hum Genet 19:223–257.

Cordell LS. 1997. Archaeology of the Southwest. San Diego, CA: Academic Press.

DiPeso CC. 1956. The Upper Pima of San Cayetano del Tumacacori. Dragoon, Arizona: Amerind Foundation Publication.

Dixon EJ. 2001. Human colonization of the Americas: timing, technology and process. Q Sci Rev 20:277–299.

Eggan F. 1950. Social organization of the Western Pueblos. Chicago: University of Chicago Press.

Excoffier L, Langaney A. 1989. Origin and differentiation of human mitochondrial DNA. Am J Hum Genet 44:73–85.

Fagan BM. 1984. The Aztecs. New York: W.H. Freeman and Co.

Fagan BM. 2000. Ancient North America: the archaeology of a continent. New York: Thames & Hudson.

Felsenstein J. 1993. PHYLIP (phylogeny inference package). Seattle: Department of Genetics, University of Washington.

Fiedel SJ. 1999. Older than we thought: implications of corrected dates for Paleoindians. Am Antiq 64:95–115.

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

Ginther C, Corach D, Penacino GA, Rey JA, Carnese FR, Hutz MH, Anderson A, Just J, Salzano FM, King M-C. 1993. Genetic variation among the Mapuche Indians from the Patagonian region of Argentina: mitochondrial DNA sequence variation and allele frequencies of several nuclear genes. In: Pena SDJ, Chakraborty R, Epplen JT, Jeffreys AJ, editors. DNA fingerprinting: state of the science. Basel: Birkhauser Verlag. p 211–219.

Goddard I. 1996. The classification of the native languages of North America. In: Goddard I, editor. Handbook of North American Indians: languages 17. Washington, DC: Smithsonian Institution. p 290–324.

Green LD, Derr JN, Knight A. 2000. mtDNA affinities of the peoples of north-central Mexico. Am J Hum Genet 66:989–998.

Hale K, Harris D. 1983. Historical linguistics and archaeology. In: Ortiz A, editor. Handbook of North American Indians: Southwest 10. Washington, DC: Smithsonian Institution. p 170–177.

Hill JH. 2001. Proto-Uto-Aztecan: a community of cultivators in Central Mexico? Am Anthropol 103:913–934.

Hoijer H. 1956. The chronology of the Athapaskan languages. Int J Anthropol Linguist 22:219–232.

Infante E, Olivo A, Alaey C, Williams F, Middleton D, De La Rosa G, Pujol C, Duren J, Navarro L, Gorodezky C. 1999. Molecular analysis of HLA class I alleles in Mexican Seri Indians: Implications for their origin. Tissue Antigens 54:35–42.

Kaestle FA. 1998. Molecular evidence for prehistoric Native American population movement: the numic expansion. Davis: University of California.

Kaestle FA, Smith DG. 2001. Ancient mitochondrial DNA evidence for prehistoric population movement: the Numic expansion. Am J Phys Anthropol 115:1–12.

Kemp BM, Resendez A, Román-Berrelleza JA, Malhi RS, Smith DG. 2002. An analysis of ancient mtDNA from Tlateloco: pre-Columbian relations and the spread of Uto-Aztecan. In: David Reed, editor. Biomolecular archaeology: Genetic approaches to the past. Carbondale, IL.

Kendall MB. 1983. Yuman languages. In: Ortiz A, editor. Handbook of North American Indians: Southwest 10. Washington, DC: Smithsonian Institution. p 4–24.

Kolman CJ, Bermingham E. 1997. Mitochondrial and nuclear DNA diversity in the Choco and Chibcha amerinds of Panama. Genetics 147:1289–1302.

Kolman CJ, Bermingham E, Cooke R, Ward RH, Arias TD, Guionneausinclair F. 1995. Reduced mtDNA diversity in the Ngobe Amerinds of Panama. Genetics 140:275–283.

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

Kroeber AL. 1915. Serian, Tequislatacan, and Hokan. Berkeley: University of California. p 279–290.

Lorenz JG, Smith DG. 1994. Distribution of the 9-bp mitochondrial DNA region V deletion among North American Indians. Hum Biol 66:777–788.

Lorenz JG, Smith DG. 1996. Distribution of four founding mtDNA haplogroups among Native North Americans. Am J Phys Anthropol 101:307–323.

Lorenz JG, Smith DG. 1997. Distribution of sequence variation in the mtDNA control region of native North Americans. Hum Biol 69:749–776.

Mace R, Pagel M. 1995. A latitudinal gradient in the density of human languages in North America. Proc R Soc Lond [Biol] 261:117–121.

Malhi RS. 2001. Investigating prehistoric population movements in North America with ancient and modern mtDNA. Davis: University of California.

Malhi RS, Schultz BA, Smith DG. 2001. Distribution of mitochondrial DNA lineages among Native American tribes of northeastern North America. Hum Biol 73:17–55.

Malhi RS, Eshleman JA, Greenberg JA, Weiss DA, Shultz Shook BA, Kaestle FA, Lorenz JG, Kemp BM, Johnson JR, Smith DG. 2002. The structure of diversity within New World mitochondrial DNA haplogroups: implications for the prehistory of North America. Am J Hum Genet 70:905–919.

Martínez-Cruzado JC, Toro-Labrador G, Ho-Fung V, Estévez-Montero MA, Lobaina-Manzanet A, Padovani-Claudio DA, Sánchez-Cruz H, Oritz-Bermúdez P, Sánchez-Crespo A. 2001. Miotchondrial DNA analysis reveals substantial Native American ancestry in Puerto Rico. Hum Biol 73:491–511.

Merriwether DA, Kemp BM, Crew DE, Neel JV. 2000. Gene flow and genetic variation in the Yanomama as revealed by mitochondrial DNA. In: Renfrew C, editor. America past, America present: genes and languages in the Americas and beyond. Cambridge, UK: McDonald Institute for Archaeological Research. p 89–124.

Miller WR. 1983. Uto-Aztecan languages. In: Ortiz A, editor. Handbook of North American Indians: Southwest 10. Washington, DC: Smithsonian Institution. p 113–124.

Moraga ML, Rocco P, Miquel JF, Nervi F, Llop E, Chakraborty R, Rothhammer F, Carvallo P. 2000. Mitochondrial DNA polymorphisms in Chilean aboriginal populations: implications for the peopling of the southern cone of the continent. Am J Phys Anthropol 113:19–29.

Nei M. 1987. Molecular evolutionary genetics. New York: Columbia University Press.

O'Rourke DH, Mobarry A, Suarez BK. 1992. Patterns of genetic variation in Native America. Hum Biol 64:417–434.

O'Rourke DH, Hayes MG, Carlyle SW. 2000. Spatial and temporal stability of mtDNA haplogroup frequencies in native North America. Hum Biol 72:15–34.

Raymond M, Rousset F. 2000. GENEPOP (ver. 1.2): a population genetics software for exact test and ecumenicism. J Hered 86:248–249.

Rickards O, Martinez-Labarga C, Lum JK, De Stefano GF, Cann RL. 1999. mtDNA history of the Cayapa Amerinds of Ecuador: detection of additional founding lineages for the Native American populations. Am J Hum Genet 65:519–530.

Schell LM, Blumberg BS. 1977. The genetics of human serum albumin, function and uses. Oxford and New York: Pergamon Press.

Schneider S, Roessli D, Excoffier DL. 2000. Arlequin version 2.000: a software for poulation genetic data analysis. Geneva: Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Schroeder AH. 1963. Hakataya, Patayan, and Hohokam. Santa Fe: National Park Service.

Schultz BA, Malhi RS, Smith DG. 2001. Examining the proto-Algonquian migration: analysis of mtDNA. In: Nichols JD, Ogg A, editors. Proceedings of the 32nd Algonquian Conference. Ottawa: Carleton.

Schurr TG, Ballinger SW, Gan YY, Hodge JA, Merriwether DA, Lawrence DN, Knowler WC, Weiss KM, Wallace DC. 1990. Amerindian mitochondrial DNAs have rare Asian mutations at high frequencies, suggesting they derived from four primary maternal lineages. Am J Hum Genet 46:613–623.

Schurr TG, Sukernik RI, Starikovskaya YB, Wallace DC. 1999. Mitochondrial DNA variation in Koryaks and Itel'men: population replacement in the Okhotsk Sea-Bering Sea region during the Neolithic. Am J Phys Anthropol 108:1–39.

Shaul DL, Anderson JM. 1989. A case for Yuman participation in the Hohokam regional system. Kiva 54:105–126.

Shaul DL, Hill JH. 1998. Tepimans, Yumans, and other Hohokam. Am Antiq 63:375–396.

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

Smith BD. 1995. The emergence of agriculture. New York: Scientific American Library, distributed by W.H. Freeman.

Smith DG. 1981. Admixture and population replacement of the Sells Papago Indians: three strategies. Soc Biol 28:126–144.

Smith DG, Malhi RS, Eshleman J, Lorenz JG, Kaestle FA. 1999. Distribution of mtDNA haplogroup X among Native North Americans. Am J Phys Anthropol 110:271–284.

Smith DG, Lorenz J, Rolfs BK, Bettinger RL, Green B, Eshleman J, Schultz B, Malhi R. 2000. Implications of the distribution of Albumin Naskapi and Albumin Mexico for New World prehistory. Am J Phys Anthropol 111:557–572.

Soodyall H, Vigilant L, Hill AV, Stoneking M, Jenkins T. 1996. MtDNA control-region sequence variation suggests multiple independent origins of an Asian-specific 9-bp deletion in sub-Saharan Africans. Am J Hum Genet 58:595–608.

Starikovskaya YB, Sukernik RI, Schurr TG, Kogelnik AM, Wallace DC. 1998. mtDNA diversity in Chukchi and Siberian Eskimos: implications for the genetic history of ancient Beringia and the peopling of the New World. Am J Hum Genet 63:1473–1491.

Steward JH. 1937. Ecological aspects of Southwestern society. Anthropos 32:87–104.

Stone AC, Stoneking M. 1998. MtDNA analysis of a prehistoric Oneota population. Implications for the peopling of the New World. Am J Hum Gen 62:1153–1170.

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

Torroni A, Chen Y, Semino O, Santachirara-Beneceretti AS, Scott CR, Lott MT, Winter M, Wallace DC. 1994. mtDNA and Y-chromosome polymorphisms in four Native American populations from southern Mexico. Am J Hum Genet 54:303–318.

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

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

Watkins WS, Bamshad M, Dixon ME, Rao BB, Naidu JM, Reddy PG, Prasad BVR, Das PK, Reddy PC, Gai PB, Bhanu A, Kusuma YS, Lum JK, Fischer P, Jorde LB. 1999. Multiple origins of the mtDNA 9-bp deletion in populations of South India. Am J Phys Anthropol 109:147–158.

Watterson G. 1975. On the number of segregation sites in the genetical models without recombination. Theor Popul Biol 7:256–276.

Weir BS. 1990. Intraspecific differentiation. In: Hillis DM, Moritz DC, Marble BK, editors. Molecular systematics. Sunderland, MA: Sinauer Associates, Inc. p 373–410.

Weiss DA. 2001. Mitochondrial DNA diversity among Native Americans from the southeastern United States. Am J Phys Anthropol [Suppl] 30:63.

Workman PL, Niswander JD, Brown KS, Leyshon WC. 1974. Population studies on southwestern Indian tribes IV. The Zuni. Am J Phys Anthropol 41:119–132.

Yao YG, Watkins WS, Zhang YP. 2000. Evolutionary history of the mtDNA 9-bp deletion in Chinese populations and its relevance to the peopling of East and Southeast Asia. Hum Genet 107:504–512.