Am. J. Hum. Genet. 65:519–530, 1999

mtDNA History of the Cayapa Amerinds of Ecuador: Detection of Additional Founding Lineages for the Native American Populations

O. Rickards,[1] C. Martínez-Labarga,[1] J. K. Lum,[2] G. F. De Stefano,[1] and R. L. Cann[3]

[1]Dipartimento di Biologia, Università di Roma "Tor Vergata," Rome; [2]The Institute of Statistical Mathematics, Minato-ku, Tokyo; and [3]Department of Genetics and Molecular Biology, University of Hawaii at Manoa, Manoa

Summary

mtDNA variation in the Cayapa, an Ecuadorian Amerindian tribe belonging to the Chibcha-Paezan linguistic branch, was analyzed by use of hypervariable control regions I and II along with two linked regions undergoing insertion/deletion mutations. Three major maternal lineage clusters fit into the A, B, and C founding groups first described by Schurr and colleagues in 1990, whereas a fourth lineage, apparently unique to the Cayapa, has ambiguous affinity to known clusters. The time of divergence from a common maternal ancestor of the four lineage groups is of sufficient age that it indicates an origin in Asia and supports the hypothesis that the degree of variability carried by the Asian ancestral populations into the New World was rather high. Spatial autocorrelation analysis points out (*a*) statistically significant nonrandom distributions of the founding lineages in the Americas, because of north-south population movements that have occurred since the first Asian migrants spread through Beringia into the Americas, and (*b*) an unusual pattern associated with the D lineage cluster. The values of haplotype and nucleotide diversity that are displayed by the Cayapa appear to differ from those observed in other Chibchan populations but match those calculated for South American groups belonging to various linguistic stocks. These data, together with the results of phylogenetic analysis performed with the Amerinds of Central and South America, highlight the difficulty in the identification of clear coevolutionary patterns between linguistic and genetic relationships in particular human populations.

Received February 24, 1997; accepted for publication May 20, 1999; electronically published June 18, 1999.

Address for correspondence and reprints: Dr. Olga Rickards, Dipartimento di Biologia, Università di Roma "Tor Vergata," Via della Ricerca Scientifica, 1, 00133, Roma, Italia. E-mail: rickards@uniroma2.it

© 1999 by The American Society of Human Genetics. All rights reserved.
0002-9297/99/6502-0027$02.00

Introduction

During the past decade, demonstrated variations in human mtDNA sequence have motivated the creation of distinct, population-based research programs that are focused on the origin, dispersal, and extinction of genetic lineages and that use maternally inherited sequences (Slatkin and Hudson 1991; Rogers and Harpending 1992). By explicitly rejecting a 17th-century concept of race, which too often organizes the sampling regime of many modern biologists, mitochondrial genetics has helped to promote a wider appreciation of the global distribution of human biological diversity. However, the ability of these studies to yield statistically accurate information concerning the geographic origin of genetic founders, long-term effective population sizes, dates of population expansion, episodes of actual expansion based on shared cultural innovations, or even past episodes of natural, sexual, and artificial selection remains subject to differing interpretations (Harpending et al. 1998, Cavalli-Sforza 1998).

At present, an understanding of the last major continental expansion of our species, the movement of humans into the unoccupied tracts of the Americas, is approaching resolution, because of archaeological, linguistic, and genetic data (Bonatto and Salzano 1997; Turner 1998). The migration scenario of nomadic Arctic hunters from a now-submerged, central Beringian plain moving into North America 35,000–14,000 years before the present (ybp) has been reconstructed as a series of contrasting models chiefly differing in the number of genetically distinct Native American populations that migrated. Variations include the time scale of their dispersal and expansion, identification of a specific geographic homeland, and their linguistic roots. Stimulated by new archaeological and genetic information, specialists have moved beyond a conservative 12,000-ybp colonization date (Taylor et al. 1985; Nelson et al. 1986), to a fuller appreciation that modern Siberians may not represent the Far Eastern fringe of the relict Beringian population center (Starikovskaya et al. 1998).

However, compared with continental African, European, or Asian maternal haplotypes, a relatively small

number of aboriginal Americans have been analyzed at the DNA sequence level. This numerical bias has led to an underappreciation of the total genetic diversity of indigenous people in the Americas (Szathmary 1993), especially when the winnowing effects that epidemic diseases and random drift have on genetic diversity in geographic isolates are considered. Genetic analysis of the first real Paleoindian samples from Florida and Illinois (Hauswirth et al. 1994; Stone and Stoneking 1998) have clearly demonstrated that greater genetic complexity existed in North American only 700 ybp, compared with that found today in donors from the same tribal area.

One (Bonatto and Salzano 1997), three (Greenberg et al. 1996), or even four (Horai et al. 1993) waves of immigrants entered the Americas from Beringia, sometime during the past 30,000 years, carrying four (Torroni et al. 1993), five (Bailliet et al. 1994), six (Forster et al. 1996), seven (Easton et al. 1996), or nine (Merriwether et al. 1996) founding haplotype clusters. However, inclusion of data from South American tribal groups in these models continues to be problematic, owing to either differences in frequency or even complete absence of proposed ancestral maternal haplotypes shared by North and South Americans and either Siberians (Stone and Stoneking 1998) or Mongolians (Merriwether et al. 1996). Confusion remains with regard to absolute chronologies and relative dating of some skeletal samples and archaeological sites in western and eastern Beringia (West 1996), along with questions of admixture, recurrent mutation, contamination, and provenience (Torroni et al. 1993).

We focus this study on the Cayapa (or Chachi) of Ecuador, a group of South American Amerind language–family speakers who have been known to western anthropologists for almost 75 years (Barrett 1925). Their history suggests a stable population that has adapted to a number of catastrophic changes during the past 500 years. They are of distinct interest to geneticists, for at least three reasons. First, their oral traditions recount a migration from the western Amazon basin into the Andes (Barriga-Lopez 1987; Carrasco 1988) and then a flight from Ibarra into the lowland jungles to escape first Inca and finally Spanish enslavement (Barrett 1925). These movements suggest that historical events and forced acculturation have not significantly distorted their genetic identity. Second, although they have been exposed to disease epidemics brought by foreign contact, their population appears to have been stable during the past 250 years of colonial rule and has been slowly increasing to its current size of ~3,600 individuals, without significant admixture from either European- or African-derived settlements adjacent to tribal communities (Pepe et al. 1994; Rickards et al. 1994; Scacchi et al. 1994; Titus-Trachtenberg et al. 1994). Thus, they represent the opportunity to investigate genotype survival in an indigenous population that has had a relatively stable census size during the past 10 generations. Understanding patterns of allelic-lineage survival under this regime may contribute to resolution of the debate, between molecular and physical anthropologists (Turner 1998), with regard to molecular estimates of time scales and expansion waves, which has characterized discussion of diversity in the Yanomama, who totally lack mitochondrial haplogroup A (Easton et al. 1996). Third, the Cayapa are members of an important linguistic subfamily (Chibcha-Paezan; Greenberg 1987) that spans Central and South America, offering molecular anthropologists the opportunity to study the coevolution of genes and languages in spatially isolated replicates that differ in the degree of contact with their neighbors, with consequences for the lateral transmission of shared elements.

We report here sequence variation in hypervariable mtDNA control regions 1 and 2 (HV1 and HV2) and length polymorphisms within noncoding sequences previously shown to distinguish Amerindian populations from others. We compare these data with those from other Native American groups and consider their implications for both the reconstruction of the history of the Cayapa and the founding of the Americas, as deduced from mtDNA evidence.

Material and Methods

Blood samples were collected by one of us (G.F.D.S.), along the Cayapas River (approximately 0° latitude, 79°E longitude; see fig. 1), from 204 individuals who

Figure 1 Map of Esmeraldas province (northwestern Ecuador), showing geographic location of the Cayapa.

gave their informed consent. Efforts were made to include only unmixed Cayapa in this sample. Each donor was asked to supply name, birthplace, language, and ethnicity for 3 generations, in order to allow us to determine the extent of recent admixture. Blood was drawn into sterile tubes containing acid citrate–dextrose anticoagulant, was frozen in liquid nitrogen, and then was stored at −80°C until DNA was extracted. Total genomic DNA was then prepared from 5–10 ml of whole blood (Miller et al. 1988).

Three regions of human mtDNA were amplified by PCR using specific primers. HV1 and HV2 contain most of the sequence variation useful for phylogenetic analysis of human populations and were produced according to standard methods (Vigilant et al. 1989). Individual PCR-amplification products were checked for purity by electrophoresis on a 0.8% agarose/Tris-acetate EDTA electrophoretic gel. A small noncoding region showing ±9-bp length variation in the COII-$tRNA^{Lys}$ junction (Wrischnik et al. 1987) was also amplified, under published conditions that have been described by Rickards (1995). The deletion of one of the 9-bp tandem repeats was verified by direct sequencing of the double-stranded PCR products of several individuals, by use of an internal primer (Wrischnik et al.1987). The Huetar 6-bp deletion in HV2 (Santos et. al. 1994) was screened by direct sequencing of samples from 30 individuals. On the basis of absence of the deletion in this initial sample, we adopted the fast procedure (Santos and Barrantes 1994*a*) for detection of this deletion, with slight modifications. A 7-μl aliquot of HV2 amplification product was digested at 37°C for 16 h with 1.2 U of the enzyme *Bsi*HKAI (Biolabs). The resulting fragments were separated on a 4% Nusieve agarose gel run at 90 V for 2 h.

For manual sequencing, the product of an asymmetric PCR (Erlich 1989) was purified either by a Qiaquick PCR spin column (Qiagen) or ethanol precipitation. Overlapping sequencing reactions of both strands were performed with the dideoxy chain–termination procedure of Sanger et al. (1977), by use of a Sequenase 2.0 reagent kit (U.S. Biochemicals). Automated sequencing of double-strand PCR products by dideoxy terminator–cycle sequencing was performed with a GeneAmp PCR System 2400 (Perkin-Elmer) and the recommended sequencing kit protocols of the manufacturer, followed by resolution of fluorescent dye–labeled templates on an ABI Prism 310 DNA sequencer (Perkin-Elmer).

mtDNA data were used to infer both the demographic history and the genetic history of the Cayapa, by a variety of statistical indices and phylogenetic estimation procedures. We first calculated intrapopulation mean pairwise sequence differences; genetic distances, by use of Kimura's (1980) two-parameter method; and haplotype diversity (denoted as "*H*" in the work of Nei and Roychoudhury [1974]). Two measures of nucleotide diversity were then calculated: the average number of nucleotide differences, per site, between any two randomly chosen sequences in the current generation, π (Nei and Tajima 1981), is a standard metric, and $E\nu$, the long-term value (Watterson 1975), is an alternative metric that focuses directly on polymorphic sites and sample size. Mismatch distributions were computed by programs provided by G. Gorla and O. Rickards.

Genetic distances were used to estimate phylogenetic relationships between individual maternal lineages, via programs in Phylip 3.5 (Felsenstein 1993), with the corrected transition:transversion ratio of 10:1. Consensus maximum likelihood (ML), unweighted-pair-group method of analysis, and neighbor-joining (NJ) trees were obtained, and all trees were rooted by use of a highly divergent sub-Saharan mtDNA sequence (Kocher and Wilson 1991) as an outgroup. Parsimony analysis using the PAUP 3.0s package (Swofford 1993) was also performed, using heuristic searches, tree-bisection reconnection, and nearest-neighbor interchange to produce 100 minimal trees per replication. Topologies of equally most parsimonious trees were summarized as a majority-rule consensus tree. All topologies were similar to the NJ tree obtained.

Data from 378 additional Native American published sequences (Dipartimento di Biologia dell'Università–Tor Vergata on-line database of Native American mtDNA) were included in the sample, for both phylogenetic estimation and comparison with results of our analysis of lineage diversity detected in the Cayapa alone.

To avoid the possibility of distortion of the ancestral structure via recent parallelism, only those sequences found in more than two individuals were included, with use of parsimony-based methods. This was done because, under neutrality, in a young expanding population, the more frequent haplotypes tend to be older (Donnelly and Tavarè 1986). Haplotype diversity and nucleotide diversity were also calculated for 10 indigenous populations (see table 4), by use of the 186 HV1 nucleotide positions common to all studies (base pairs 16185–16370).

Spatial arrangements using the frequencies of mtDNA founding lineages previously suggested for 51 Native American populations (Dipartimento di Biologia dell'Università–Tor Vergata on-line database of Native American mtDNA) were examined by spatial autocorrelation (SA) analysis (Sokal and Oden 1978). SA statistics were calculated by the Spatial Autocorrelation Analysis Program, version 4.3, of Watenberg (1989). Pairs of populations were placed in one of five arbitrary distance classes. An SA coefficient, Moran's *I*, was then computed for each lineage and distance class. Values of this statistic range from +1 (indicative of genetic simi-

larity in the frequencies found in a given class) to −1 (indicative of dissimilarity).

Results

Patterns of mtDNA Variability

Insertion/deletion mutations (*INDELS*)*: COII-tRNA* ^{Lys} (*9 bp*)*.*—In a sample of the 204 Cayapa tested, 39% had the 9-bp deletion. We characterize this as the X.I type of deletion, caCCCCTCTA (Lum and Cann 1998). This mutation is found in the gene pool from which both Austronesian speakers in the Pacific and some Amerinds of North and South America arose. The frequency of the deletion is in close agreement with observations with regard to other Amerind groups from South America—such as the Mapuche and Mataco of Argentina, who possess it at a frequency of 36% (Ginther et al. 1993; Torroni et al. 1993; Bailliet et al. 1994), and the Mexican Zapotec, who possess it at a frequency of 33% (Torroni et al. 1994).

Amerinds from other areas generally possess this deletion in highly variable frequencies, and it tends to be most concentrated in populations settled along the coastline on both sides of the Pacific Rim. This deletion is known, on the basis of studies of ancient donors, to have been present in the North American Rockies 8,000 ybp (Stone and Stoneking 1996), but it shows no obvious correlation with linguistic characteristics in the Americas, and it is absent in circum-Arctic groups today (Shields et al. 1993).

The 9-bp deletion of the X.I type (Lum and Cann 1998), when examined by direct sequencing, occurs on the same broad group of control-region haplotypes belonging to the B-haplotype cluster in every Pacific Rim population sampled thus far, suggesting that the deletion is probably monophyletic for the Americas. The mutation event that created this deletion is inferred to be different from that observed in African and European maternal genealogies, because entirely different control-region haplotypes characterize those groups of donors (O. Rickards and C. Martínez-Labarga, unpublished data).

INDELS: HV2 Huetar (*6-bp*) *deletion.*—In contrast, a second deletion—the Huetar deletion—previously observed to occur in Panama, Costa Rica, and Chile, among speakers of the Chibchan- and Andean-language subfamilies, has a much more limited distribution. It was not seen in any of the 30 individual Cayapa who were initially assayed by direct sequencing. The remaining samples were then tested by the indirect method described in the Material and Methods section, and no individuals were judged to contain this particular marker.

Unlike the 9-bp deletion, which is always seen on the B group of haplotypes in the Americas, this length mutation is not associated with a particular subset of ancestral maternal lineages. It also occurs sporadically among non-Chibchan Amerind speakers in the Americas who belong to the A and D haplotype clusters (Merriwether et al. 1995). We suggest that this deletion most likely represents a new, independent mutation event in the same highly mutable region of hypervariable sequence, in multiple haplotype groups, and is not a particular characteristic of one founding haplotype family or linguistic subgroup. All other polymorphisms detected in the hypervariable regions, with the exception of insertions of one or two C's at position 303 and/or one C at position 315, were due to single base substitutions.

Single-Nucleotide Polymorphisms Associated with the Hypervariable Region

Thirty randomly chosen individuals were subjected to extensive sequencing for both HV1 and HV2 (871 bp). Table 1 summarizes information about the general patterns of sequence variation associated with the Cayapa lineages, for these portions of the control region. Thirty-four polymorphic sites, along with the 9-bp deletion, define 10 distinct mtDNA haplotypes (C 1–C10) in 30 individuals randomly chosen for the most extensive analysis. Haplotypes C1 and C2 are associated with the 9-bp deletion and belong with lineages assigned to the B group. The Cayapa and the Huetar from Costa Rica share 3 of the 10 identified haplotypes (Santos et al. 1994), and others have been identified in other groups as well (Ginther et al. 1993; Santos et al. 1994; Batista et al. 1995; Kolman et al. 1995; Easton et al. 1996). Statistical analysis showed a nonrandom ($P \ll .001$) distribution of substitution patterns for transitions, for HV2.

One hundred twenty additional randomly chosen Cayapa were examined for the most highly polymorphic region of HV1, covering 186 nucleotides (table 2). Only transition mutations were noted in this region. With just this limited sequence block, 13 different haplotypes were identified, and 10 appear to associate with three previously identified lineage clusters (A—C) defined on the basis of more-extensive information (Torroni et al. 1992; Horai et al. 1993). Of these 13 haplotypes, 3 (CAYAPA 11–CAYAPA13) appear frequently (22%) and may belong to a lineage that is specific to the Cayapa. Eight of the 13 haplotypes occurred more than once, and 6 occurred more than four times.

To determine whether the Cayapa lineage may be the possible consequence of admixture with surrounding populations of African Americans or Mestizos, mtDNA from 20 individuals from the black community of the Rio Cayapa, 20 Spaniards, and 40 Africans from eastern

Table 1

Polymorphic mtDNA D-Loop Sites in HV1 and HV2

	No.	Polymorphic Nucleotides	Region V Deletion
		111111111111111111	
		666666666666666666	
		1112222222222111222222222233333333	
		6674550001346991881247899900122456	
		3436230475593011397311701812957262	
CR Sequence[a]		TCATTAATGAAAAAACATTCATCCCTCAGTCTTT	
Haplotypes:[b]			
C1	5	..G........G...CCC...............	Yes
C2	1	..GC...CA...G...CCC..............	Yes
C3	5	.TGC.G....G.G..T...T...T....A....C	No
C4	1	CTGC.G....G.G..T...T...T....A...CC	No
C5	3	.TGC.G....G.G..TCC.T..TT....A....C	No
C6	1	.TGC.G...GG.G..T...T...T....A....C	No
C7	1	..G........-G--....T.....C...CT...	No
C8	2	..G...G....-G--....T.....C.G.CT...	No
C9	2	..G........-G--....T.C...C...CT...	No
C10	9	..G.C.......G......TG...T.T....C.C	No
Total	30		

[a] Cambridge reference sequence (Anderson et al. 1981).
[b] Dots (.) indicate identity with the CR sequence; dashes (–) denote deletions.

and western Africa were included. No evidence of the Cayapa lineage was discovered in these samples (O. Rickards and C. Martínez-Labarga, unpublished results). It shares four mutations (at nucleotide positions 73, 263, 16223, and 16362) in common with the A and D lineage clusters described by Torroni et al. (1992). The latter two substitutions (i.e., those at 16223 and 16362) were determined, by Wakeley (1993), to be the most variable (they had 19 inferred changes) in HV1 and thus are not good indicators of lineage associations. The Cayapa lineage (table 3) is distinct from both A- and D-haplogroup lineages, at a number of other characteristic positions.

Estimations and Comparisons of Diversity

Haplotype and nucleotide diversities for the Cayapa, Chibchan speakers from lower Central America (the Kuna, Ngöbe, and Huetar), and linguistically distant Native American tribes from diverse parts of the continent are shown in table 4. Nucleotide-diversity values (π) for the Cayapa are low compared with those for other Chibchan speakers, but $E\nu$ values are high, an indication that Cayapa population growth during the past 2–3 generations has been slow (Carrasco 1988), compared with that of other Chibchan speakers.

The distribution of pairwise sequence differences (Slatkin and Hudson 1991; Rogers and Harpending 1992) for the Cayapa (fig. 2) shows two major peaks, one of which is at 0 (reflective of identical sequences) and one of which is at 6, a pattern noted in other tribal populations from the same language family (Batista et al. 1995; Kolman et al. 1995). This pattern may reflect the nonmonophyletic nature of the Cayapa, with the peak at 6 reflecting diversity that arose in ancestral Asian populations and with the peak at 0 (reflective of a large number of identical lineages) stemming from a recent bottleneck event that reduced the number of effective females. If the peak at 0 reflects merely the steady-state population growth of the tribe, it is unclear why the

Table 2

Polymorphic mtDNA D-Loop Sites in HV1

	Haplotype	No.	Polymorphic Nucleotides
			1111111111111111111
			6666666666666666666
			1222222222233333333
			8111247899900122455 6
			9347311701812957269 2
CR Sequence			TGCTCATCCCTCAGTCTTTT
Lineages:			
B	CAYAPA1	48	C..C................
A	CAYAPA2	22	....T...T....A.....C
	CAYAPA3	8	C...T..TT....A.....C
	CAYAPA4	1	....T..TT....A.....C
	CAYAPA5	2	.A..T...T....A.....C
	CAYAPA6	1	....T...T....A...C.C
	CAYAPA7	1	..T.T...T....A...C.C
C	CAYAPA8	1	....T.....C...CT....
	CAYAPA9	4	....T.....C.G.CT....
	CAYAPA10	6	....T.C...C...CT....
Cayapa	CAYAPA11	6	....TG...T......C..C
	CAYAPA12	19	....TG...T.T....C..C
	CAYAPA13	1	....TG...T......C.CC
Total		120	

Note.—Data are as defined in the footnotes to table 1.

Table 3

Control-Region Variable Sites Defining Major Amerindian mtDNA Clusters

	Polymorphic Nucleotides[a]
	11111111111111
	66666666666666
	1112222211222222333333
	6745534699181249990122 46
	4362359301197310181957 22
CR Sequence	CATTAAAAAACTTCACCTCGTCTT
Lineages:[b]	
A	TGC.GG.G..T..T.T...A...C
B	.G.....G...CC...........
C	.G....–G––...T...C..CT..
Cayapa	.G.C...G.....TG.T.....CC
D	.G.C...G.....T....T.C..C

Note.—Data are as defined in the footnotes to table 1.

[a] Data are from Ginther et al. (1993), Santos et al. (1994), Batista et al. (1995), Kolman et al. (1995), and the present study.

[b] According to Torroni et al. (1992), Horai et al. (1993), and the present study.

maternal lineage generating a substantial fraction of this peak (in the Cayapa lineage) is not found in adjacent groups.

Phylogenetic Analyses of Haplotypes via Genetic Distances

A consensus tree (fig. 3) for the 10 Cayapa haplotypes, discovered by use of HV1 and HV2 sequences, was produced by the neighbor-joining method used with genetic distances estimated by Kimura's (1980) two-parameter model, with an African outgroup (Kocher and Wilson 1991). Three major clusters appeared. The same results were found by the ML method and by the maximum parsimony (MP) algorithm (data not shown). We evaluated the resultant tree by the bootstrap method (Efron 1979), using 1,000 replications. The first group contains haplotypes C1 and C2 together, the second group contains C3–C6, and the third group contains C7–C9. The tenth haplotype (C10) connects to the C7-C8-C9 cluster of lineages, albeit poor (47%) bootstrap support, and lacks three deletion mutations (at nucleotide positions 249, 290, and 291) characteristic of this group. The three clusters with good bootstrap support correspond, respectively, to lineage groups B, A, and C, which previous studies have identified in Native Americans (Torroni et al. 1992; Horai et al. 1993). The C10 haplotype is present at high (30%) frequency in our sample and does not appear to have been described in previous reports.

SA

Although many individual Amerind populations lack at least one of the hypothesized ancestral lineage clusters A–D (Dipartimento di Biologia dell'Università–Tor Vergata on-line database of Native American mtDNA) and have been described informally as showing north-south geographic clines for maternal lineage diversity, table 5 shows Moran's I values that document this trend as a decreasing value of I, with distance, for lineage clusters A–C. For lineage cluster D, northern and southern populations are similar to each other but are different from Central American populations. This result emphasizes both (1) the importance of the isthmus of Panama as a geographic sink for the dispersal of some genotypes in the Americas and (2) the unique nature of isolation around this zone, which has created unusual microevolutionary forces among the Chibchan speakers residing in lower Central America (Barrantes et al. 1990).

Discussion

A general population model of American local tribal groups undergoing drift and local lineage extinction, with maternally inherited diversity being periodically replenished by admixture with adjacent, larger source populations, seems to describe well the picture of haplotype diversity in the Cayapa. They contain most of the apparently major lineage clusters of Amerindians, with a coalescent indicating diversity that arose in Asian ancestors well before the peopling of the Americas. The Cayapa thus appear to be closer to the metapopulation concept (Hanski and Gilpin 1997) of population biology than to that which emphasizes the stability and persistence of local groups. This genetic-diversity pattern contrasts with that found in other tribal populations occupying extreme human habitation zones of North, Central, and South America—for example, the Savoonga Eskimo, Guaymi, and Xavante, all of which have lost lineages and may be geographically or culturally isolated from central source populations. Population forces that have kept the Cayapa total census size small have, at the same time, promoted structuring that yields a rather large effective population size, and recent archaeological descriptions of Amazonian foraging cultures dating as far back as 15,000 ybp suggest that fishing camps may have commonly developed along waterways during the Holocene (Roosevelt et al. 1998), providing the mechanism by which local demes were linked. Short-range gene flow under such a model would have been sufficient to reintroduce lineages that previously had been lost.

What is also striking is that this small tribal population has maintained its relative isolation for a period of time sufficiently long that it has evolved, in situ, an

apparently unique mitochondrial marker, such as the C10 haplotype discovered here. Alternatively, C10 may exist at low frequencies in other populations, which have not been adequately sampled. If this mutation is not shared by other neighboring groups, it is possible that it results from the chance inclusion of a few maternal lines with the appearance of extra high rates of substitution.

The distribution of new mitochondrial mutations—for example, those arising in the Americas, as distinct from those found among potential ancestral populations in Asia and elsewhere—is primarily dependent on both the length of time since the mutation's origin and the strength of gene flow to distribute the new mutation widely among geographically isolated groups. The combination of both cosmopolitan and unique lineages implies that the ancestors of the Cayapa not only experienced gene flow in the past but also underwent periods of extensive isolation from their neighbors (and this can be inferred from the shape of the mismatch distribution of sequence diversity, which shows two peaks). This finding is in line with oral traditions and historical records that state that the Cayapa fled the Inca, other Amerind groups, and the Spanish. Their population size, estimated to have been ~1,500 in 1734 (Carrasco 1988), remained stable until the report by Barrett (1925) in 1908–9 and has now approximately doubled, to >3,000. These numbers help to establish the parameters under which the rates of lineage extinction and lineage origination may be expected to balance each other in technically simple human societies.

Many aboriginal populations, when examined for haplotype diversity, show a reduced number of distinct lineages compared with what is seen in urban groups, supporting the idea that a population bottleneck occurred during the dispersal of humans into the Americas (Schurr et al. 1990). Stone and Stoneking (1998) report that, on the basis of hypervariable sequence diversity, the average percentage of rare- or single-occurrence lineages in modern Native American populations is 45.7%, whereas the frequencies in the Yanomama (75%) of South America and the prehistoric Oneota (74%) of North America appear to be at the extreme end for the

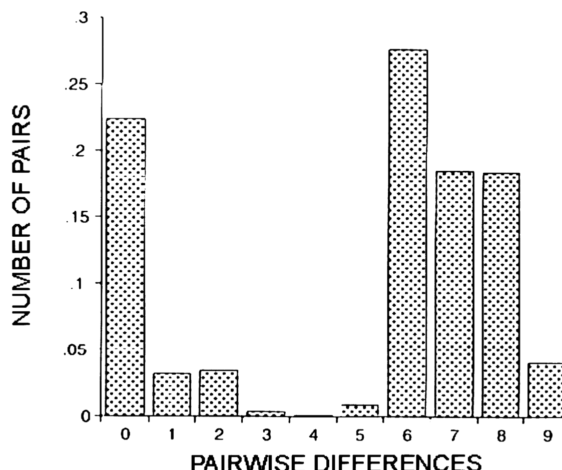

Figure 2 Distribution of pairwise sequence differences within the Cayapa. The abscissa indicates the number of base differences in a pair, and the ordinate indicates the frequency (in %) of each difference.

Table 4

mtDNA Sequence Diversity in the Cayapa and Other Native American Populations

	No. of[a]				Nucleotide Diversity	
Population	Individuals	Mitochondrial Haplotypes	Polymorphic Sites	Haplotype Diversity[b]	π[c]	$E(\nu)$[d]
Cayapa	120	13	20	.7766	.2502	3.7340
Kuna	63	5	9	.5698	.3467	1.9132
Ngöbe	46	5	9	.7429	.4141	2.0531
Huetar	27	6	9	.6724	.3065	2.3467
Mapuche	38	11	18	.8976	.2891	4.2709
Xavante	25	3	7	.6400	.3086	1.8642
Zoro	30	9	14	.7747	.2345	3.5494
Gavião	27	7	13	.8661	.2758	3.3897
Nuu-Chah-Nult	63	24	19	.9376	.2207	4.0390
Bella Coola	40	10	15	.8730	.2371	4.2447
Haida	41	8	13	.6586	.1536	4.2195

[a] Data are from Ward et al. (1991, 1993,1996), Ginther et al. (1993), Santos et al. (1994), Batista et al. (1995), Kolman et al. (1995), and the present study.

[b] *H,* according to Nei and Roychoudhury's (1974) formula.

[c] According to Nei and Tajima's (1981) formula.

[d] According to Watterson's (1975) formula.

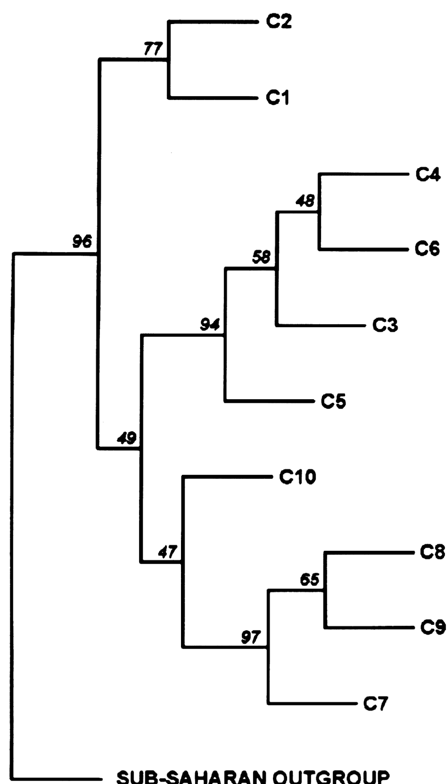

Figure 3 Phylogenetic tree inferred from sequence data reported in table 1, by use of Kimura's (1980) two-parameter distance and NJ method. This is a 50%-majority-rule consensus tree generated by the TBR branch-swapping algorithm. Numbers at the nodes indicate the percentage of bootstrap replications supporting each node. The same topology was also obtained by the MP and ML methods (data not shown). The labels at the tips of the branches indicate the different mtDNA haplotypes. The tree was generated by 1,000 resamplings of the 30 Cayapa sequences, along with a sub-Saharan African sequence as outgroup (Kocher and Wilson 1991).

range. The value that we calculate for the Cayapa, 38.5%, is approximately half of these estimates and refers specifically to the 120 individuals in whom the HV1 region of the D loop has been sequenced. This measure in the Cayapa is very close to that reported for the Kuna of Central America (43%), another tribe in the Chibchan-language family, which is estimated to be ~10 times as large as the Cayapa and is known to have undergone rapid growth during the past 50–60 years (Gjording 1991; Kolman et al. 1995).

The similar frequency of the 9-bp deletion in the Cayapa, compared to other South American Amerind populations, and its tight association with B haplogroup lineages emphasizes the monophyletic nature of this deletion mutation on a particular ancestral haplotype in the Americas. This B lineage group constitutes a large proportion of the surviving indigenous South Americans, and its tight association with indigenous populations of remote Pacific islands indicates that it arose in the same source population of Asia that contributed to a dispersal event in two different directions at two different times. The mutation associated with this deletion has to be as old as the first settlement, judging from its broad distribution among Amerinds and Pacific Islanders.

In contrast, the 6-bp Huetar deletion is lacking in the Cayapa, although it is present in the Boruca, Bribri, Cabecar, Huetar, Guaymi, Kuna, and Teribe, who are Chibchan speakers in Central America (Santos and Barrantes 1994*b*, 1995). By other measures of diversity, the Cayapa appear closer to non-Chibchan speakers of South America than to representatives of the same language family in Central America, a classic case of discordance between genetic and linguistic data, as is to be expected among complex polyphyletic populations such as these (Rickards et al. 1994). This deletion may have arisen recently in Central America, before the lineage has had time to spread widely. On the other hand, the number of Cayapa whom we sampled was insufficient to allow us to detect this low-frequency haplotype. Finally, if it is assumed that the mutation arose in the ancestral proto-Chibchan population (Santos and Barrantes 1994*b*), there remains the possibility that it was simply lost in the founder Cayapa population of Chibchan speakers who presumably crossed the isthmus of Panama. Since this deletion is found in tribal Chilean populations outside the same language family (Merriwether et al. 1995), more analysis is needed to determine whether (*a*) it had recent multiple origins because of hot spots for variation in HV2 (Parsons et al. 1997; Jazin et al. 1998) or (*b*) its distribution simply represents the ancient assimilation of individuals from different language groups into what now looks like a single, lin-

Table 5

Moran's *I* Computed for Pairs of Populations, for Each Native American mtDNA Founding Lineage and for Each Distance Class

	Moran's *I* for[a]				
Lineage	Class 1	Class 2	Class 3	Class 4	Class 5
A	.57**	.45**	.05*	−.16**	−.39**
B	.37**	−.01	.04*	−.06	−.21**
C	.52**	−.03	−.07	.01	−.13**
D	.39**	.12*	.02	−.24**	−.02
Average	.46	.13	.01	−.11	−.19

[a] Distance classes are defined as follows: class 1, all pairs of populations 0–1,000 km apart (87 pairs); class 2, all pairs of populations 1,001–2,000 km apart (122 pairs); class 3, all pairs of populations 2,001–5,500 km apart (476 pairs); class 4, all pairs of populations 5,501–10,000 km apart (326 pairs); and class 5, all pairs of populations 10,001–18,128 km apart (264 pairs). The upper limit of class 5 corresponds to the maximum distance found among all possible pairs of populations included in the analysis.

* $P < .05$.

** $P < .01$.

guistically homogeneous subfamily. Unique polymorphisms identified in the Chibchan speakers of Central America contribute to the view that there has been a substantial period of genetic isolation between Central, South, and North American Amerinds, perhaps as long as 7,000 years (Barrantes et al. 1990).

Whatever the antiquity of the Cayapa as a defined linguistic group, their demographic history, as inferred from mtDNA sequence data and their mismatch distribution (see fig. 2), supports no prolonged population bottleneck stemming from the differential reproductive effort of only a few founders. Even though the population has persisted in modest numbers for the past 250 years, extensive maternal genetic diversity has been maintained. This observation is important in the consideration of how small Beringian populations of Arctic hunters were able to procure new mates and avoid inbreeding depression during extended periods in which they were cut off from major population centers, either by being trapped in North America on one side of the ice-free MacKenzie corridor or by being isolated in eastern Beringia for several thousand years as the central Beringian plain oscillated in size 26,000–10,000 ybp (West 1981, 1996). Fluctuating population sizes should result in the random extinction of lineages via genetic drift, and kinship systems that encourage exogamy might be predicted to offset this trend. Multiple patterns of coevolution between genes and language should therefore be expected, depending on how resources have been partitioned when and if societies developed from mobile groups of patrilocal hunters into the more sedentary practice of farming.

The finding that both the nonagrarian Cayapa and the Yanomama, an Amazonian–rain-forest tribe within the same linguistic subfamily, show very similar measures of genetic diversity when parameters other than numbers of unique lineages are considered gives hope that, however disturbed these populations may now be, there are general molecular-genetic patterns reflecting a common history of the settlement of South America that remain to be discovered. At the same time, discrepancies between diversity estimates for the Cayapa compared with Chibchan speakers of Central America highlight the difficulty in the identification of a clear relationship between linguistic affiliation and genetic affinity. Mismatch distributions (see fig. 2) show the two major peaks characteristic of Chibchan speakers in both areas. Admixture between geographic neighbors and gene flow that is not recent in origin were suspected for some Central American groups, and, in most population studies, this remains a difficult problem to dissect (Wakeley and Hey 1997). However, our data show that no recent gene flow with non-Amerindian groups accounts for the mixture of maternal lineages now seen among the Cayapa. Thus, discussion of the origin of the Chibcha-Paezan language speakers from a common stock in both Central and South America versus the importance of gene flow from neighboring tribes (Torroni et al. 1993) would be furthered by inclusion, in future studies, of the Paezan-speaking Colorado from Colombia and Ecuador.

With regard to the larger question, we are of the opinion that the number of migrations and the exact time scale for this major dispersal event(s) into the Americas cannot be profitably extracted from these genetic data by current methods, because most tribal groups demonstrate the pattern of polyphyletic lineage history. No debatable archaeological evidence currently supports an early dispersal into the Americas with a sustained cultural record of tool use or habitat modification prior to ~16,000 ybp (Roosevelt et al. 1998; Turner 1998). One previous attempt that used mtDNA sequence to present a synthetic summary of genetic, linguistic, and archaeological evidence was limited to an analysis of haplogroup-A types only (Bonatto and Salzano 1997); the result allowed the authors to claim that values for divergence between any pair of Native American populations were always significantly lower than those between this pair and the combined Asian group. We suggest that such a result would not have been obtained if the other haplogroups had been included within the analysis, since the results for these haplogroups would show the Cayapa to be the closest genetic relatives of a group of non-Native Americans (to date, the Kapingamarangi, a Polynesian outlier within Micronesia), probably because of similar patterns of ancestral lineage loss in small populations. Each time that we make inappropriate assumptions about tribal-group composition and stability, we ignore quantities of data that underscore the episodic loss and renewal of human genetic diversity.

Human population-genetic models that use monophyly of lineages as an assumption (such as branching algorithms and mismatch distributions) need to be discussed as approximations—and not as ultimate solutions to intractable archeological problems, such as the geological disappearance of a major dispersal corridor. A second major synthetic study using mtDNA sequence data (Horai et al. 1993) included all four major haplogroups known before extensive sequencing revealed other distinct lineage groups, but it had such shallow structure in the estimated phylogenetic trees that little confidence could be placed in the subsequent historical reconstruction of dispersal history, including waves and dates. An updated study on ~400 Native Americans from Central and South America would only be further complicated by the addition of three to five presumably ancestral lineage clusters to the first four, because of similarly low bootstrap support for each major node (data not shown).

Human geneticists might be well advised to only modestly suggest that their suggestions with regard to the

identification of population waves for archaeological consideration are simply exercises in speculation that have little precision. Our research continues to document the unique composition of genomes in space and time, but interpretations of the exact process by which genetic diversity has accumulated should be stated with greater caution, if it is to have credibility among a broader range of disciplines. The Cayapa demonstrate that lineage polyphyly complicates every statement made about genetic population history, even in a system that considers only maternal transmission. The difficulties that attend the appropriate incorporation of information from biparentally inherited loci into the effort to reconstruct population history—an effort that is the ultimate goal of most anthropological geneticists—can be only broadly imagined on the basis of this example.

Acknowledgments

We wish to express our thanks to the Cayapa population for their voluntary collaboration in the study. This project was supported by the Comunità Economica Europea, which provided an individual fellowship to C.M.-L. (Human Capital and Mobility contract ERBCHBICT920251, allotted to O.R.); by a Ministero dell'Università e della Ricerca Scientifica e Tecnologica grant to G.F.D.S.; and by Japan Society for the Promotion of Science fellowship P 97904 to J.K.L. We thank G. Barbujani and V. Sbordoni for their advice and comments on an early version of the manuscript, R. Guderian and A. Guevara for helping in the sample collection, and T. Salviati for his invaluable help with the drawings. Two anonymous reviewers greatly contributed to the improvement of this article. Webpages are designed and maintained by F. Carnera.

Electronic-Database Information

The URL for data in this article is as follows:

Dipartimento di Biologia dell'Università–Tor Vergata, http://www.uniroma2.it/biologia/lab/anthromol/mtdna.htm (for Native American mtDNA sequences)

References

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

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

Barrantes R, Smouse PE, Mohrensweiser HW, Gershowitz H, Azofeifa J, Arias TD, Neel JV (1990) Microevolution in lower Central America: genetic characterization of the Chibcha-speaking groups of Costa Rica and Panama, and a consensus taxonomy based on genetic and linguistic affinity. Am J Hum Genet 46:63–84

Barrett SA (1925) The Cayapa Indians of Ecuador. Indian notes and monographs 40. Heye Foundation, New York

Barriga-Lopez F (1987) Cayapas o Chachis. Vol 4: Etnologia ecuatoriana. Graficas Duque. Instituto Ecuatoriano de Credito Educativo y Becas, Quito

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

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

Carrasco E (1988) El pueblo Chachi: el jeengume avanza. Abya-Yala Ediciones, Quito

Cavalli-Sforza LL (1998) The DNA revolution in population genetics. Trends Genet 14:60–65

Donnelly P, Tavarè S (1986) The ages of alleles and a coalescent. Adv Appl Prob 18:1–19

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

Efron B (1979) The jackknife, the bootstrap and other resampling plans. Society for Industrial and Applied Mathematics, Philadelphia

Erlich AH (ed) (1989) PCR technology: principles and applications for DNA amplification. Stockton Press, New York, London, Tokyo, Melbourne, Hong Kong

Felsenstein J (1993) PHYLIP (phylogeny inference package), version 3.5c. Department of Genetics, University of Washington, Seattle

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

Ginther C, Corach D, Penacino GA, Rey JA, Carnese FR, Hutz MH, Anderson A, et al (1993) Genetic variation among the Mapuche Indians from the Patagonian region of Argentina: mitochondrial DNA sequence variation and allele frequencies of several nuclear genes. In: Pena SDJ, Epplen JT, Jeffreys AJ (eds) DNA fingerprinting: state of the science. Birkhäuser-Verlag, Basel, pp 211–219

Gjording CN (1991) The Guaymi Indians and mining multinationals in Panama. Smithsonian Institution Press, Washington, DC

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

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

Hanski I, Gilpin ME (1997) Metapopulation biology. Academic Press, New York

Harpending HC, Batzer MA, Gurven M, Jorde LB, Rogers AR, Sherry ST (1998) Genetic traces of ancient demography. Proc Natl Acad Sci USA 95:1961–1967

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

Horai S, Kondo R, Nakagawa-Hattori Y, Hayashi S, Sonoda S, Tajima K (1993) Peopling of the Americas, founded by four major lineages of mitochondrial DNA. Mol Biol Evol 10:23–47

Jazin E, Soodyall H, Jalonen P, Lindholm E, Stoneking M, Gyllensten U (1998) Mitocondrial mutation rate revisited: hot spot and polymorphism. Nat Genet 18:109–110

Kimura M (1980) A simple model for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–120

Kocher TD, Wilson AC (1991) Sequence evolution of mitochondrial DNA in human and chimpanzees: control region and a protein-coding region. In: Osawa S, Honjo T (eds) Evolution of life: fossils, molecules, and culture. Springer-Verlag,Tokyo, pp 391–413

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

Lum JK, Cann RL (1998) mtDNA and language support of common origin of Micronesians and Polynesians in Island South East Asia. Am J Phys Anthropol 105:109–119

Merriwether DA, Ferrell RE, Rothhammer F (1995) mtDNA D-loop 6-bp deletion found in the Chilean Aymara: not a unique marker for Chibcha-speaking Amerindians. Am J Hum Genet 56:812–813

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

Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 16:1215

Nei M, Roychoudhury AK (1974) Sampling variances of heterozygosity and genetic distance. Genetics 76:379–390

Nei M, Tajima F (1981) DNA polymorphism detectable by restriction endonucleases. Genetics 97:145–163

Nelson DE, Morlan RE, Vogel JS, Southon JR, Harington CR (1986) New dates on northern Yukon artifacts: Holocene not Upper Pleistocene. Science 232:749–751

Parsons TJ, Muniec DS, Sullivan K, Woodyatt N, Alliston-Greiner R, Wilson MR, Berry DL, et al (1997) A high observed substitution rate in the human mitochondrial DNA control region. Nat Genet 15:363–368

Pepe G, Rickards O, Buè C, Martinez-Labarga C, Tartaglia M, De Stefano GF (1994) EcoRI, RsaI, and MspI RFLPs of the COL1A2 gene (type I collagen) in the Cayapa, a Native American population of Ecuador. Hum Biol 66:979–989

Rickards O (1995) Analysis of the region V mitochondrial marker in two black communities of Ecuador, and in their parental populations. Hum Evol 10:5–16

Rickards O, Tartaglia M, Martinez-Labarga C, De Stefano GF (1994) Genetic characterization of the Cayapa Indians of Ecuador and their genetic relationships with other Native American populations. Hum Biol 66:299–322

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

Roosevelt AC, Lima da Costa M, Lopes Machado C, Michab M, Mercier N, Valladas H, Feathers J, et al (1998) Paleoindian cave dwellers in the Amazon: the peopling of the Americas. Science 272:373–384

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 74: 5463–5467

Santos M, Barrantes R (1994*a*) Direct screening of a mitochondrial DNA deletion valuable for Amerindian evolutionary research. Hum Genet 93:435–436

——— (1994*b*) D-loop mtDNA deletion as a unique marker of Chibchan Amerindians. Am J Hum Genet 55:413–414

——— (1995) Reply to Merriwether et al. Am J Hum Genet 57:195–196

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

Scacchi R, Corbo RM, Rickards O, De Stefano GF (1994) The Cayapa Indians of Ecuador: a population study of seven protein genetic polymorphisms. Ann Hum Biol 21:67–77

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

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

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

Sokal RR, Oden NL (1978) Spatial autocorrelation in biology. I. Methodology. Biol J Linnean Soc 10:199–228

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

Stone AC, Stoneking M (1996) Genetic analyses of an 8,000 year-old Native American skeleton. Ancient Biomolecules 1: 83–87

——— (1998) mtDNA analysis of a prehistoric Oneota population: implications for the peopling of the New World. Am J Hum Genet 62:1153–1170

Swofford DL (1993) PAUP: Phylogenetic analysis using parsimony, version 3.1.1. Illinois Natural History Survey, Champaign

Szathmary EJE (1993) mtDNA and the peopling of the Americas. Am J Hum Genet 53:793–799

Taylor RE, Payen LA, Prior CA, Slota PJ Jr, Gillespie R, Gowlett JAJ, Hedges REM, et al (1985) Major revisions in the Pleistocene age assignments for North American human skeletons by C-14 accelerator mass spectrometry: none older that 11,000 C-14 years BP. Am Antiquity 50:136–140

Titus-Trachtenberg EA, Rickards O, De Stefano GF, Erlich HA (1994) Analysis of HLA class II haplotypes in the Cayapa Indians of Ecuador: a novel DRB1 allele reveals evidence for convergent evolution and balancing selection at position 86. Am J Hum Genet 55:160–167

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

Torroni A, Schurr TG, Cabell MF, Brown MD, Neel JV, Larsen M, Smith DG, et al (1993) Asian affinities and continental

radiation of the four founding Native American mtDNAs. Am J Hum Genet 53:563–590

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

Turner C II (1998) Book review of *American Beginnings*. Am J Phys Anthropol 106:408–409

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

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

Wakeley J, Hey J (1997) Estimating ancestral population parameters. Genetics 145:847–855

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

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

Watenberg D (1989) SAAP: a spatial autocorrelation analysis program, version 4.3. University of Medicine and Dentistry of New Jersey, Piscataway

Watterson GA (1975) On the number of segregating sites in genetical models without recombination. Theor Popul Biol 7:256–276

West FH (1981) The archaeology of Beringia. Columbia University Press, New York

——— (1996) American beginnings. University of Chicago Press, Chicago

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