

# Mitochondrial DNA diversity in the Kuna Amerinds of Panamá

Oriana Batlsta, Connie J. Kolman<sup>1</sup> and Eldredge Bermingham<sup>1,\*</sup>

Universidad de Costa Rica and <sup>1</sup>Smithsonian Tropical Research Institute, Apartado 2072, Balboa, Panamá

Received December 27, 1994; Revised and Accepted February 15, 1995

GenBank accession nos L39319–L39326

**Mitochondrial DNA (mtDNA) haplotype diversity was determined for 63 Chibcha-speaking Kuna Amerinds sampled widely across their geographic range in eastern Panamá. The Kuna data were compared with mtDNA control region I sequences from two neighboring Chibchan groups, the Ngöbé and the Huetar; two Amerind groups located at the northern and southern extremes of Amerind distribution, the Nuu-Chah-Nulth of the Pacific Northwest and the Chilean Mapuche; and with a single Na-Dene group, the Haida of the Pacific Northwest. The Kuna exhibited low levels of mitochondrial diversity as had been reported for the other two Chibchan groups and, furthermore, carried only two of the four Amerind founding lineages first reported by Schurr and coworkers (*Am. J. Hum. Genet.* 1990; 46: 613–623). We posit that speakers of modern Chibchan languages (henceforth referred to as the Chibcha) passed through a population bottleneck caused either by ethnogenesis from a small founding population and/or subsequent European conquest and colonization. Using the approach of Harpending *et al.* (*Curr. Anthropol.* 1993; 34: 483–496), we estimated a Chibchan population bottleneck and subsequent expansion approximately 10 000 years before present, a date consistent with a bottleneck at the time of Chibchan ethnogenesis. The low mtDNA diversity of Kuna Amerinds, as opposed to the generally high levels of mtDNA variation detected in other Amerind groups, demonstrates the need for adequate sampling of cultural or racial groups when attempting to genetically characterize human populations.**

## INTRODUCTION

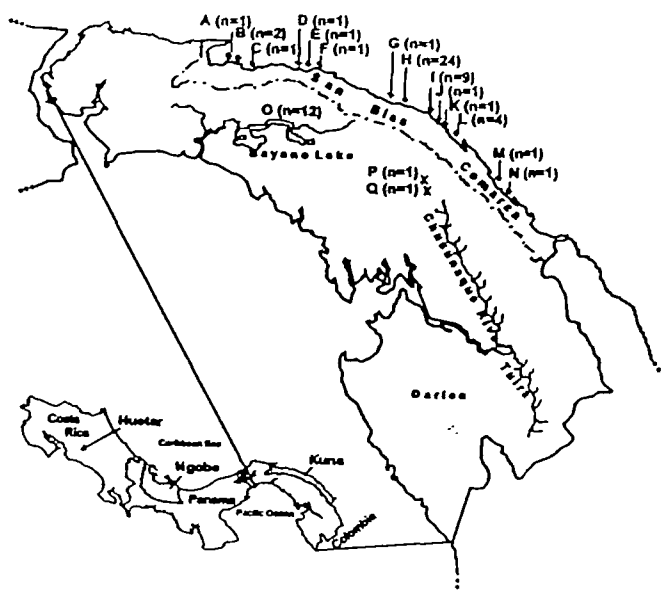
One of the most exciting arenas of human population study is the New World, which represents the most recent human colonization of a previously uninhabited continental landmass (30 000–12 000 years BP; 1). Surviving Native Americans exhibit remarkable cultural and linguistic diversity; for example, seven distinct cultural groups, speaking mutually unintelligible, but historically related languages, live today within the small country of Panamá (2). Another aspect of the

early human settlement of the New World that makes the Americas an interesting landscape against which to study human evolution is the rich archaeological record left by the prehistoric Amerinds (3,4). Panamá, with its prosperous contemporary Amerind populations, wealth of prehistoric sites and detailed documentary record for the post-Colombian period, appears particularly promising for the study of human populations.

Herein we present mitochondrial DNA (mtDNA) sequence data on the Kuna Amerinds of eastern Panamá. The Kuna language is assigned to the Paya-Chibchan linguistic phylum, a family of languages spoken by Amerind groups currently distributed from Nicaragua to Colombia (2). The majority of Kuna inhabit the islands of the San Blas Comarca, which stretch along the Caribbean coast between Colon and Colombia (Fig. 1). Smaller Kuna settlements are found on the Pacific slopes of eastern Panamá and in the westernmost part of Colombia.

The origin of the Kunas is not well understood. Some historians suggest that the Kuna are descendants of the 'Cueva', a group of Amerind populations of unknown linguistic and cultural affiliation. The Cueva were first identified by the Spaniards in the 1500s and inhabited the Pacific coast of Panamá until the late 1500s when it is thought that they were extirpated by the Spaniards and disease. Other investigators propose that the Kuna immigrated from Colombia and claimed the territory vacated by the Cueva. The scarcity of linguistic data on the Cueva makes it difficult to establish a firm relationship between the two groups leading Romoli (5) to support the second scenario. Romoli proposed that the Cueva-speaking peoples disappeared before the middle of the sixteenth century, roughly 70 years before the arrival of the Kuna (5). It has also been proposed that the Kuna may not represent a distinct Amerind group, but may instead encompass an agglomeration of remnant Amerind groups decimated during the Spanish conquest (8). Howe (9) further suggested that these remnant groups were located in the immediate contact area between Colombia and Panamá, which was home to Choc-speaking groups at the time of the conquest, leading to the possibility of Chocoan origination or admixture in the Kuna. Although the historical record is ambiguous concerning Kuna residence on one or both of the Panamanian coasts prior to the 1800s, it is known that the Kuna moved en masse from the Pacific to the Caribbean coast and islands beginning in the mid-1800s due to pressure from Choc-speaking groups immigrating from Colombia and the development of new sources of trade and subsistence on the Caribbean coast (6,7).

\*To whom correspondence should be addressed at: Smithsonian Tropical Research Institute, Unit 0948, APO AA 34002-0948, USA



**Figure 1.** Geographic distribution of the three Chibchan groups of lower Central America (see insert) and location of the Kuna collection sites. The collection sites are represented by letters and the sample size is given for each site. A: Isla Carti-Tupile, B: Isla Carti-Suitupo, C: Isla Río Sidra, D: Isla Río Azúcar, E: Isla Narganá, F: Isla Tigre, G: Isla Playon Chico, H: Isla Tupile, I: Isla Ailgandi, J: Isla Achutupu, K: Isla Mamitupo, L: Isla Ustupu, M: Isla Mulatupo Sasardi, N: Coetupu, O: Bayano, P: Nurra, Q: Wala.

The main subsistence practices of the Kuna are agriculture and fishing while hunting and gathering play a secondary role. Tourism has become increasingly important with the production of the mola, an applique-type handicraft worn by the women and sold worldwide. The Kuna maintain a strong political association with a clear hierarchical structure which, combined with a closed social organization, has led some to suggest that they were more resistant than other Amerind groups to the devastating effects of the European conquest (6,10). This is supported by the report that the Kuna show one of the lowest levels of Caucasian and Negroid admixture of all Amerind groups with mean admixture rates of 2.4 and 2.45%, respectively (11). Like other Chibchan groups (with the exception of the Ngöbé), the Kuna practice uxori-local marriage customs meaning that, upon marriage, a husband moves to his wife's home (11). Since moving to the Caribbean coast and adopting an insular life-style, Kuna families have tended to settle a single island and do not migrate far from that site during their lifetimes.

As an extension to our previous work on the linguistically related, but culturally distinct Ngöbé Amerinds of Panamá (12), we now present mtDNA sequence data on the Kuna Amerinds. We compare Kuna mtDNA sequence data with similarly sized samples obtained from five other New World indigenous groups representing Amerind and Na-Dene populations: 46 Chibcha-speaking Ngöbé from western Panamá (12), 27 Chibcha-speaking Huetar from central Costa Rica (13), 63 Nuu-Chah-Nulth of the Pacific Northwest (14), 38 Chilean Mapuche from the Chilean coast (15) and 40 Na-Dene Haida

of the Pacific Northwest (16). Although Amerind and Na-Dene groups are linguistically distinct, their geographic ranges overlap in the Pacific Northwest and there is controversy whether their entry to the New World was coincident, implying shared genetic ancestry, or not (12,17).

Recent tribal history, cultural mixing and the dramatic demographic consequences of European conquest caution against overzealous cross-cultural comparisons between sampled New World indigenous groups. However, the considerable polymorphism observed in mtDNA may facilitate statistical analyses of anthropological hypotheses concerning social formative processes and specific cultural histories. At a time of increasing tribal breakdown through admixture with non-indigenous peoples, mtDNA may be especially suited to tribal studies due to the probable asymmetry in the introduction of mtDNA genes relative to nuclear genes to indigenous populations (18).

## RESULTS

### DNA sequence results

Seven mtDNA haplotypes were revealed among the 63 Kuna individuals sequenced for nucleotide pairs 16 030–16 400 (numbered according to ref. 19) of the mitochondrial control region I (20) (Table 1). Four of the haplotypes have been previously reported: K3 was originally described in the Mapuche (15), K4 was reported for the Ngöbé (12), Nuu-Chah-Nulth (14) and Haida (16), K5 was detected in the Ngöbé (12) and Huetar (13) and K7 has been described in the Ngöbé (12), Huetar (13), Nuu-Chah-Nulth (14), Haida and Bella Coola (16). The most common Kuna haplotype, K1, with a frequency of 60%, was described for the first time.

Among the Kuna haplotypes, two distinct haplotype classes were observed (haplogroup A,  $n = 45$  and haplogroup B,  $n = 18$ ; haplogroup designations according to ref. 36), each dominated by a single haplotype with frequencies of 84 and 72%, respectively. Of the 370 nucleotide positions determined, 10 were polymorphic and distinguished the two mtDNA groups by a minimum of six mutations. Within groups, mtDNA haplotypes differed, on average, by one mutation. Of the 10 polymorphic sites determined, one was a C-to-A transversion and the remaining nine sites were transitions (four C-to-T, four T-to-C and one G-to-A). Nucleotide diversity,  $\pi$ , was determined to be 0.0109.

Three haplogroup A individuals were also sequenced for nucleotide pairs 39–380 of the mitochondrial control region II (Table 2). Of the 341 nucleotide positions determined, 17 were found to be polymorphic including seven transitions (one C-to-T, one T-to-C and five A-to-G), four insertions (the insertions at 303 and 304 were non-independent [data not shown] and were considered to be a single mutational event although the molecular mechanism of mutation is likely to be different than that of a single insertion mutation) and one 6 bp deletion (bp 106–111, counted as a single mutational event). Two individuals exhibited the 6 bp deletion that has been previously reported in the Ngöbé (12), Huetar (21) and Aymara of Chile (22). It has been suggested that this deletion may represent a Chibcha-specific mutation when found in haplogroup A individuals since all occurrences of the 6 bp deletion in the Aymara have been in haplogroup D backgrounds (23). Although preliminary, our results on control region II would

**Table 1.** Polymorphic nucleotides, restriction sites and frequency distributions of the seven mtDNA haplotypes reported in this study

ID	N	Collection sites											HaeIII	9bp	AluI	AluI	
			1	1	1	1	1	1	1	1	1	1					1
			6	6	6	6	6	6	6	6	6	6	6	6			
			1	1	2	2	2	2	3	3	3	3	3	3	663	8272/	13262
			1	8	1	2	7	9	1	1	6	6	6	6	8289		5176
			1	9	7	3	3	0	1	9	0	2	2				
Reference			C	T	T	C	G	C	T	G	C	T					
K1	38	A-F,H-L,N-Q	.	.	.	T	.	T	.	A	T	C	1	2	0	1	
K3	1	I	.	.	.	T	.	T	.	A	.	C	1	2	0	1	
K4	5	H,I,L	T	.	.	T	.	T	.	A	.	C	1	2	0	1	
K5	1	O	T	.	.	T	.	T	.	A	T	C	1	2	0	1	
K7	13	G,H,I	.	C	C	.	.	.	.	.	.	.	0	1	0	1	
K21	1	M	.	C	C	.	A	.	.	.	.	.	0	1	0	1	
K28	4	H	.	C	C	.	.	.	C	.	.	.	0	1	0	1	

See Figure 1 for location of collection sites. Numbering according to Anderson *et al.* (19).

**Table 2.** Polymorphic nucleotides of individuals sequenced for mtDNA control regions I and II

ID	N	Region II										Region I																						
		1	1	1	1	1	1	1	1	2	2	2	3	3	3	1	1	1	1	1	1	1	1	1	1									
		6	7	0	0	0	0	1	1	4	5	9	0	3	6	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
		4	3	6	7	8	9	0	1	6	3	0	0	5	3	3	4	3	1	8	1	2	7	9	1	1	6	6	6	6	6	6	6	
																			1	9	7	3	0	0	1	9	0	2	2	3	3	3	3	3
Reference		C	A	G	G	A	G	C	A	T	A	X	A	A	A	X	X	X	C	T	T	C	C	C	C	G	C	T						
K1	1	.	.	X	X	X	X	X	X	C	G	.	G	G	G	C	C	C	.	.	.	T	.	T	.	A	T	C						
K2	1	.	.	X	X	X	X	X	X	C	G	.	G	G	G	.	C	C	.	.	.	T	.	T	.	A	T	C						
K3	1	T	G	.	.	.	.	.	.	C	G	A	.	G	G	.	C	.	.	.	T	.	T	.	A	.	C							

'I' indicates an insertion mutation following the specified position; 'X' represents a deletion relative to the sequence(s) being compared

support this conclusion. Due to the non-independence of the insertions at bp 303 and 304, individuals K1 and K2 could not be considered to differ in their region I and region II sequences. Individual K3, however, differed from K1 and K2 at a single position in region I, but differed at six positions in region II (counting the 6 bp deletion and 303/304 insertions as single mutational events each) suggesting the possibility that region II carries more nucleotide diversity than previously reported (20). Owing to the limited number of control region II sequences presented, no region II data were used in any of the statistical, phylogenetic or pairwise difference analyses presented below.

### Restriction fragment length polymorphism (RFLP) results

Two of the four Amerind mtDNA haplogroups defined by Schurr *et al.* (24) and Torroni *et al.* (25) were observed among the Kuna surveyed: 45 individuals represented haplogroup A and 18 individuals represented haplogroup B (Table 1). There was a one-to-one correspondence between the two haplogroups identified by nucleotide sequence analysis and the two haplogroups determined by RFLP analysis. A previous study revealed only A haplotypes in the Kuna based on RFLP analysis of 16 individuals (17). However, the previous study collected samples from a single island, Río Azucar, whereas

the present study analyzed samples from 14 islands (including Río Azucar) and three inland sites.

### Statistical and phylogenetic analyses of the Kuna

Phylogenetic analysis of the Kuna resolved the two mtDNA haplogroups described above. Haplogroup B individuals derived from just four of the 14 San Blas island populations and from none of the mainland populations; therefore these four island populations clustered in the phylogenetic analysis (PAUP, data not shown). This relationship was investigated further using AMOVA to test for population subdivision. Several different population subdivisions were tested (Table 3). Significant interpopulation variances were found for all population subdivisions in which populations carrying haplogroup B were tested as part of the same group (e.g. 1–6). However, when these haplotypes were split or excluded, no significant interpopulation variance was detected (e.g. 7, 8).

### Statistical and phylogenetic analyses of Amerind and Na-Dene groups

Haplotype and nucleotide diversity values were calculated for the combined dataset of Kuna, Ngöbé, Huetar, Nuu-Chah-Nulth, Mapuche and Haida and are summarized in Table 4. The Chibchan and Na-Dene groups exhibited the lowest

Table 3. AMOVA test of Kuna populations

Populations	Regional description	Interpopulation variance	Significance
(1) A-N vs. O-Q	Caribbean vs. Pacific	24%	$p = 0.004$
(2) A-F,O vs. G-N,P,Q	West vs. East	31%	$p < 0.001$
(3) A-F vs. G-N	W San Blas vs. E. San Blas	27%	$p = 0.008$
(4) A-N vs. O	San Blas vs. Bayano	22%	$p = 0.004$
(5) A-N,P,Q vs. O	San Blas + Darién vs. Bayano	21%	$p = 0.01$
(6) G,H,I,M vs. others	Populations with hap. B vs. populations without hap. B	46%	$p < 0.001$
(7) A-H vs. I-N	W. San Blas vs. E. San Blas (splits hap. B)	-2.0%	$p = 0.40$
(8) A-F vs. O	W. San Blas vs. Bayano	-5.1%	$p = 0.36$

haplotype and nucleotide diversities as predicted by previous work on the Chibcha-speaking Ngöbé (12).

The MST network pictured in Figure 2 provided a graphical summary of the relationships between the 59 mtDNA haplotypes that have been assayed in the six groups. Although only one of many possible MSTs, it was consistent with the principal features of the parsimony and neighbor-joining analyses (data not shown). A principal result derived from the phylogenetic analyses was the lack of phylogeographic structure in the mtDNA haplotype data; there were examples of identical haplotypes shared by more than one group (KNHC, KNCA, KNH, CM, KM, CA1 and CA2) as well as genetically dissimilar haplotypes observed within the same group (K1, K28:Kuna; N1, N15: Ngöbé; C7, C26: Nuu-Chah-Nulth; A2, A6: Haida).

MST-based AMOVA results revealed significant levels of interpopulation variance among the six groups tested here. Approximately 20% of the variance was partitioned among populations and 80% was partitioned within populations (tested as six populations within one region,  $p < 0.001$ ). When the three Chibchan groups and the remaining two Amerind groups were divided into two regions, 7.4% between-region and 13% interpopulation variance was detected ( $p < 0.001$ ). However, when the five Amerind groups and the single Na-Dene group were divided into two regions, no significant regional subdivision was revealed ( $p = 0.34$ ).

Interpopulation variances were also tested for pairs of populations with significant levels determined in all cases (Table 5;  $p = 0.002$  for Kuna/Ngöbé,  $p = 0.012$  for Ngöbé/Huetar analyses and  $p < 0.001$  for the remaining 13 paired comparisons). Interpopulation variances for Ngöbé, Nuu-Chah-Nulth, Mapuche and Haida populations had been calculated previously (12). Interpopulation variances ranged from 10% for the Ngöbé/Huetar comparison to 38% for the Mapuche/Haida comparison. Chibcha/Chibcha comparisons showed slightly lower interpopulations variances than other population combinations. Although the population pair separated by the greatest geographic distance revealed the highest level of interpopulation variance (Mapuche/Haida—38%), increased geographic distance did not always correlate with increased interpopulation variance. For example, the second-most geographically distant pair, Mapuche/Nuu-Chah-Nulth, had an interpopulation variance of only 13%.

#### Pairwise genetic difference distributions

Mismatch distributions were calculated for the Kuna, Ngöbé and Huetar (Fig. 3). Mismatch distribution refers to the pairwise

Table 4. Haplotype and nucleotide diversities for Kuna, Ngöbé, Huetar, Mapuche, Nuu-Chah-Nulth and Haida populations

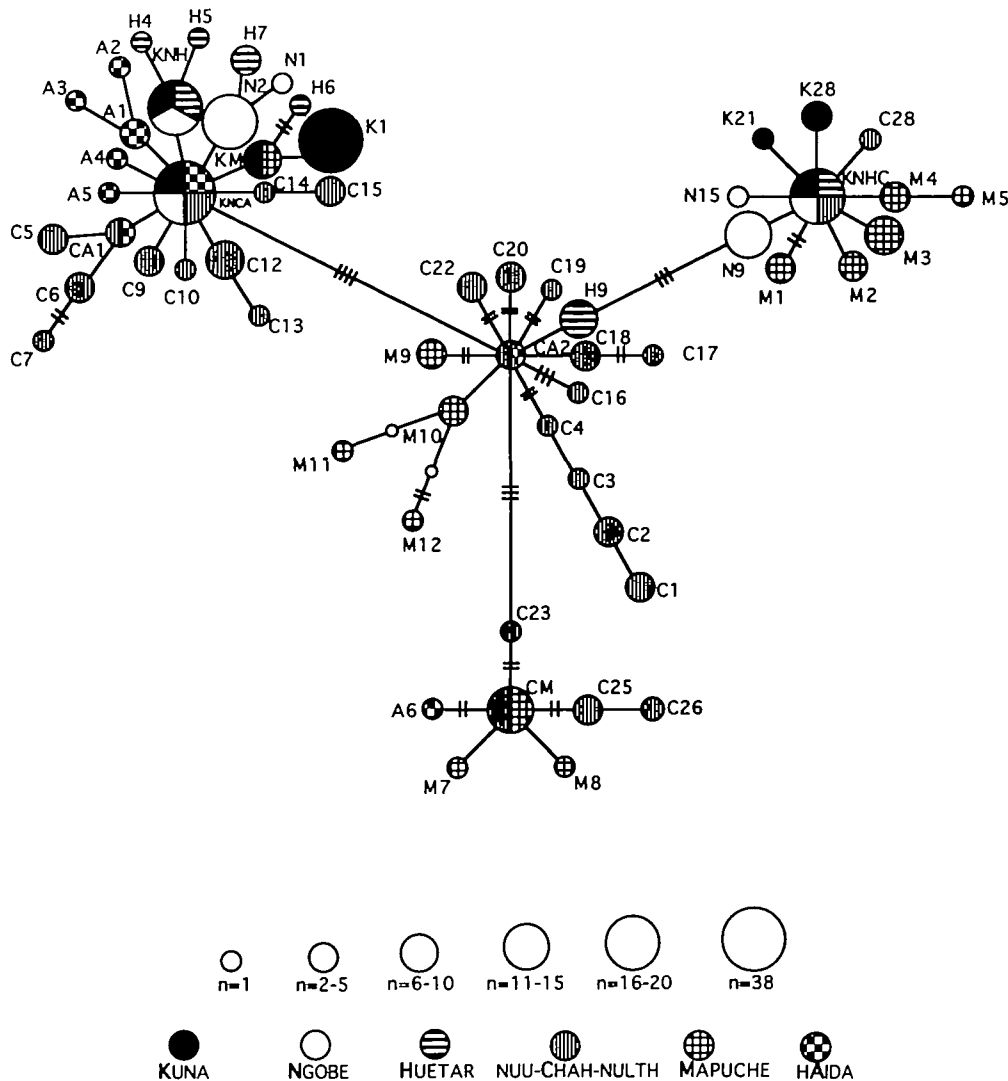
Group	Haplotype diversity $h$	Nucleotide diversities	
		$\pi$	E(v)
Kuna	0.59	0.011	2.1
Ngöbé	0.76	0.013	2.7
Huetar	0.71	0.011	3.1
Mapuche	0.91	0.016	5.0
Nuu-Chah-Nulth	0.95	0.016	5.5
Haida	0.69	0.007	3.8

count of nucleotide differences within a population (26,27). All groups shared a peak in their mismatch distributions at six to nine units of mutational time. A second major peak was revealed in the mismatch distributions of all three Chibchan populations suggesting a second, more recent expansion. This expansion occurred between zero and one unit of mutational time. While the Kuna and Ngöbé distributions revealed only two peaks separated by a flat area on the graph, the Huetar displayed a third peak in this region. Although the significance of this feature is not completely understood, its explanation may lie in the fact that the Huetar carry three, not two, of the four mtDNA haplogroups defined by Schurr *et al.* (24) and Torroni *et al.* (25).

When the three Chibchan groups were pooled and analyzed concurrently (Fig. 4), the shared second peak occurred at one unit of mutational time suggesting a Chibchan expansion at approximately 10 000 years BP (one mutational unit is equivalent to 9300 years as determined in ref. 27 for control region I data). The uniqueness of this feature is better appreciated when compared with mismatch distributions of non-Chibchan groups (Fig. 4). Neither the Nuu-Chah-Nulth nor the Mapuche displayed a second major peak, although they shared a single peak at approximately seven mutational units similar to that exhibited by the Chibchan groups.

#### DISCUSSION

The Kuna, and neighboring Chibchan groups, exhibited the lowest levels of mitochondrial diversity yet reported for Amerind populations. Figure 2 provides a striking visual depiction of the low diversity of the Chibcha. The Kuna, Ngöbé and Huetar were each represented by only seven of the



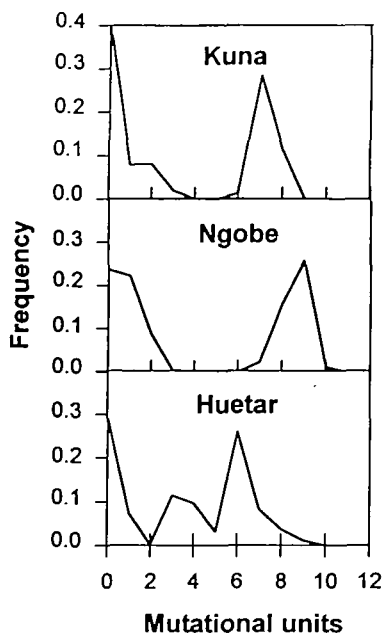
**Figure 2.** Minimum spanning tree illustrating the relationships between haplotypes of the five Amerind groups and one Na-Dene group. The size of each circle reflects the frequency of that haplotype and different circle fill patterns represent each of the six groups. Circles with more than one fill pattern represent haplotypes that were found in more than one population and are labeled by a combination of letters representing the groups that share that haplotype. Each line connecting the circles represents a single mutation except where cross-hatches enumerate mutations greater than one. All Kuna haplotypes are included. The four unique Ngöbé sequences are labeled according to Kolman *et al.* (12). The five unique Huetar sequences are labeled according to Santos *et al.* (13). The 28 Nuu-Chah-Nulth sequences from Ward *et al.* (14) are labeled C1–C28. The 13 Mapuche sequences from Ginther *et al.* (15) are labeled M1–M13. The nine Haida sequences from Ward *et al.* (16) are labeled A1–A9. The four clusters centered on haplotypes KNCA, KNHC, CM and CA2 correspond to lineages A, B, C and D, respectively (36). The two unlabeled open circles represent haplotypes that were observed in the Horai *et al.* (43) collection of Chilean Amerind mtDNA sequence.

59 haplotypes pictured in the MST, for a total of 16 haplotypes, as compared with 28 haplotypes for the Nuu-Chah-Nulth, 13 for the Mapuche and nine for the Haida. Furthermore, the Kuna (and Ngöbé) carried only two of the four mtDNA haplogroups observed in the non-Chibchan Amerind groups.

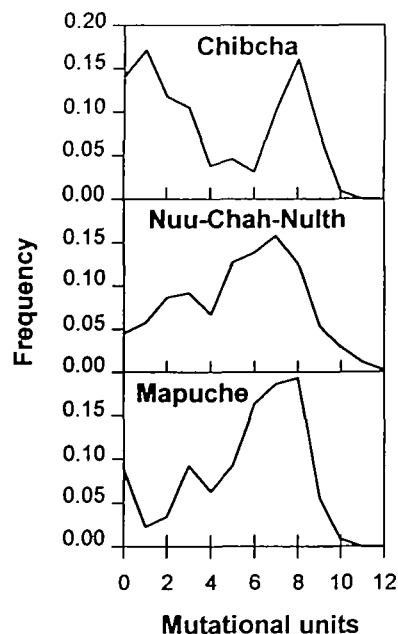
A previous study of mtDNA diversity in 16 Kuna revealed only group A haplotypes but was based on a population sample taken from a single island (18). Our results based on 63 Kuna indicate that the absence of group B haplotypes in the previous study was due to inadequate sampling of the Kuna populations and not loss through genetic drift as argued by Torroni and

coworkers (18). We do not intend criticism of the earlier Kuna results as they were reported in a study that broadly characterized mtDNA variation in six Chibcha-speaking populations in addition to the Kuna. Although our analysis of 63 Kuna may also overlook the true extent of Kuna mtDNA diversity, comparison across the two Kuna studies makes clear the need for extensive sampling, both numerical and geographical, when attempting to genetically characterize human populations.

The reduced mtDNA diversity noted for the Kuna was also apparent at the nucleotide level whether calculated as a current



**Figure 3.** Mismatch distributions for Kuna, Ngöbé and Huetar populations. The X-axis represents the number of pairwise differences in units of mutational difference (1 unit is equivalent to 9300 years; 27). The Y-axis represents the frequency of each pairwise difference and the scale is 0.0–0.4 for all graphs.



**Figure 4.** Mismatch distributions for the combined dataset of three Chibchan populations (Kuna, Ngöbé and Huetar) and for the Nuu-Chah-Nulth and Mapuche. The X-axis represents the number of pairwise differences in units of mutational difference. The Y-axis represents the frequency of each pairwise difference and the scale of the Y-axis is 0.0–0.2 for all graphs.

generation,  $\pi$ , or a long term,  $E(v)$ , value (Table 4). The Kuna had a  $\pi$  value that was 69% that of the Nuu-Chah-Nulth and the Mapuche and a  $E(v)$  value that was 38–42% that of the Nuu-Chah-Nulth and the Mapuche, respectively. The diversity values for the Ngöbé and Huetar were similarly reduced relative to the other Amerind groups. Excoffier and Langaney (28) have suggested that, in order to speculate on the evolutionary history of a group, one should use an estimator of nucleotide diversity such as  $E(v)$  that is independent of haplotype frequencies and thus measures a long-term average rather than a current generation value. Therefore, the low levels of genetic diversity determined for the Chibchan groups do not simply reflect the current situation, but imply a history of reduced diversity. It must be noted, however, that it is not possible to trace Chibchan history prior to their migration to the New World at which time a reduction in genetic diversity of all migrating groups is thought to have occurred.

The reduced mtDNA diversity common to the Kuna and other lower Central American Chibchan groups may reflect cultural differences among Amerind groups. For example, Barrantes *et al.* (29) found increased genetic diversity at the nuclear level for several South American groups, which was attributed to their practice of village propagation by the fissioning of pre-existing villages along family lines. Although the three Chibchan groups studied here share an agriculture/fishing-based economy, they practice significantly different marriage traditions. Specifically, the Kuna and Huetar practice uxorilocal customs (husband moves to the wife's home) while the Ngöbé practice virilocal customs (wife moves to the husband's home) (11). This difference suggests that the Kuna and Huetar may have lower levels of mtDNA flow between

communities relative to the Ngöbé leading to elevated levels of genetic diversity by virtue of the increased isolation of Kuna and Huetar communities. As the three groups share similarly reduced mtDNA diversities, our data provide no indication that measures of mtDNA diversity are sensitive to such cultural differences as marriage traditions. On the other hand, the virtual restriction of haplogroup B to three neighboring islands (a single group B haplotype was detected on one non-adjacent island) may indicate that geographic proximity plays a more important part in determining patterns of mtDNA variation than marriage traditions.

The reduced mtDNA diversity of the Chibchan groups may indicate passage through post-conquest population bottlenecks. It is known that the Amerind populations of Central America crashed following the Spanish conquest and evidence suggests that the Panamanian groups were particularly decimated (30, 31). However, the first unequivocal documentation of the Kuna in Panamá occurred in the 1600s (6), a time of decreased exploration by the Spanish, leading to speculation that the Kuna had no or limited direct contact with the Spanish and, therefore, may have escaped the most devastating effects of European contact.

An alternative, but not mutually exclusive, interpretation posits that the reduction of mtDNA diversity in the Chibcha pre-dates the Spanish conquest of the Americas. Analysis of eight Chibchan groups (including the Kuna results presented here) has revealed a preponderance of A and B haplogroups in the Chibcha with a low frequency of group D haplotypes documented in the Huetar and Boruca (13,32). The fact that all Chibchan groups studied thus far exhibit reduced levels of

**Table 5.** Interpopulation variances for Kuna, Ngöbé, Huetar, Mapuche, Nuu-Chah-Nulth and Haida populations

	Ngöbé	Huetar	Mapuche	Nuu-Chah-Nulth	Haida
Kuna	16%	20%	26%	21%	25%
Ngöbé		10%	21%	14%	18%
Huetar			27%	12%	16%
Mapuchet				13%	38%
Nuu-Chah-Nulth					15%

mtDNA diversity argues strongly for an explanation that provides for a simultaneous origin or cause of the reduced diversity in the Chibcha. The most parsimonious explanation for the reduced mtDNA diversity exhibited by all eight Chibchan groups suggests their common ethnogenesis from a small founding population. The reduction in mtDNA diversity may reflect a smaller sampling of the Amerind migration wave and of the Amerind founding haplogroups carried in that population.

It is difficult to confirm or discredit either of the two scenarios presented above and, in fact, they may both have played a part in Chibchan history. However, Harpending *et al.*'s (27) pairwise genetic difference analysis offers a means of estimating the date of a Chibchan expansion following the proposed population bottleneck. Mismatch distributions of the Kuna, Ngöbé, Huetar, Nuu-Chah-Nulth and Mapuche (Figs 3 and 4) shared a major peak at six to nine differences indicating an ancient expansion (presumably in Asia) at 84 000–55 000 years BP, which corresponds well with the first major expansion of human populations estimated at 80 000–30 000 years BP by Harpending *et al.* (27).

Mismatch distributions of all three Chibchan groups revealed an additional major peak between zero and one unit of mutational time. The fact that similar pairwise difference distributions were detected using datasets from three separate Chibchan studies makes it unlikely that the distributions were due to sampling of a particular mtDNA haplotype, but instead suggests that the distributions reflect the unique population history of the Chibcha. Simulation analyses were performed to determine the set(s) of conditions which would reproduce the pairwise difference distributions revealed in the Chibchan groups. Using sample sizes similar to the Chibchan samples, the simulations demonstrated that only a mild population bottleneck ( $n = 1500\text{--}3000$ ) at one unit of mutational time produced a major peak between zero and one unit of mutational time while retaining a major peak at six to nine mutational units. A more severe bottleneck eliminated the peak at six to nine units presumably because ancestral lineages had been lost and a more recent bottleneck produced only a slight peak at zero to one mutational unit while retaining a basic Poisson distribution (similar to the Nuu-Chah-Nulth and Mapuche distributions) presumably because effects of the bottleneck were not yet manifest. The coincident timing of the zero to one peak in the three Chibchan analyses suggested that the hypothesized population bottleneck occurred simultaneously in these groups. Pooled analysis of the three Chibchan groups provided an independent estimate of 10 000 years BP for the Chibchan population expansion. Although this estimate carries a large but indeterminate error, it should be noted that the date of 10 000 years BP falls during the 10 000–7000 years BP time period proposed for ethnogenesis of the Chibchan lin-

guistic stock based on genetic, linguistic and archeologic evidence (2,4,33).

A second, major peak was also observed in the Na-Dene Haida at one to two mutational units (12) suggesting that the Haida may have undergone a recent bottleneck and expansion comparable with the Chibcha. It is interesting to note that, in addition to a second peak in their mismatch distributions, the Haida and the Chibchan groups also shared the lowest mtDNA diversities reported for New World indigenous groups.

### Kuna ancestry

The mitochondrial evidence strongly suggested that the Kuna were of Chibchan origin and not a mixture of linguistically diverse Amerind groups. Neighboring Choc-speaking groups have been found to exhibit increased levels of mtDNA diversity, including all four Amerind mtDNA haplogroups, as opposed to the Chibcha (Kolman and Bermingham, in preparation). The mtDNA evidence, therefore, would seem to disprove Howe's suggestion that the Kuna were formed from remnant Choc-speaking groups inhabiting the Darien and northern Colombia during the time period immediately following the Spanish conquest (9). This leaves open the possibility, however, that the Kuna represent an agglomeration of various Chibcha-speaking groups. Of course, the mitochondrial data address Kuna ancestry exclusively along matriline and the data presented here demonstrate only that Kuna women derive from Chibchan stock. Past and present co-habitation of the Kuna with Chocoan groups, who currently have smaller population numbers but greater mtDNA diversity than the Kuna, offers the possibility that non-Chibchan origins will be detected in the Kuna at the nuclear level. If non-Chibchan origins in the Kuna are indicated by the nuclear data, this would have fascinating implications for formation of the Kuna cultural group by suggesting Chibchan descent of Kuna women, but Chocoan, or mixed, descent of Kuna men.

### Isthmian migrations

Archeological and genetic evidence have been used to counter the previously held idea that lower Central America served as a continuous migration corridor between cultures to the north and south. Archeological data suggest that contemporary Chibchan groups descended from a single ancestral population which fragmented, *in situ*, into small, sedentary units during the Holocene (3,4). Barrantes *et al.* (33) and Thompson *et al.* (34) analyzed several Chibcha-specific protein polymorphisms and proposed that the Chibcha had maintained an isolated, long-term presence in lower Central America. The presence of the four Amerind mtDNA haplogroups north and south of Chibchan territory not only provides evidence that the Kuna samples were relatively free of maternal admixture, but also argues against a history of intermarriage between the Chibcha

and neighboring groups. In conjunction with the broader discontinuity between the Chibcha, Nuu-Chah-Nulth and Mapuche, our mtDNA results support the archeological and anthropological view of isolated, *in situ* development of contemporary Amerind populations in lower Central America.

### Colonization of the New World

The novel finding of reduced mtDNA diversity in three linguistically related Amerind groups addresses the utility of genetic diversity measures for evaluating the number of migration event(s) to the New World. Based on linguistic, dental character and genetic evidence, three distinct 'waves' of migration have been proposed for the founding of contemporary Amerind, Na-Dene and Eskimo-Aleut groups in the New World (35). Recent support for the three wave hypothesis has come from mtDNA-based studies, which suggest that mtDNA diversity differences between Amerind and Na-Dene groups indicate two distinct migrations for these groups (17,25). However, conclusions based solely on mtDNA diversity measures may be misleading. For example, the AMOVA results presented here suggest that the most geographically distant Amerind groups, the Nuu-Chah-Nulth and Mapuche, are among the most closely related genetically. A consideration of the geographic location of the groups studied and the mtDNA haplotypes contributing to their diversity values suggests it is more likely that the missing haplogroups in the Chibchan populations create an artificially close relationship between the Nuu-Chah-Nulth and Mapuche.

We believe that a similar oversimplification of mtDNA diversity values has been used in support of the three wave hypothesis. In the past, the subset of the four founding Amerind haplogroups that a population carries has been used to identify the migration wave to which a group belongs. For example, the observation that the Na-Dene exhibited only mtDNA haplogroups A and B (36) or, more recently, only haplogroup A (17) was offered as evidence that the Na-Dene arrived in a migration distinct from the Amerinds who carried haplogroups A, B, C and D. However, the Kuna, Ngöbé, Bribri and Cabécar also carried only lineages A and B, yet no one questions their Amerind ancestry. As presented previously (12), the molecular evidence suggests that the Na-Dene and the Chibcha may simply be smaller samples of the migration from Asia and that the Na-Dene do not necessarily represent an independent migration event.

## MATERIALS AND METHODS

### Sampled populations

Since migrating to Panamá from Colombia, the Kuna Amerinds have inhabited three major territories in Panamá: the San Blas Comarca along the Caribbean coast, the Bayano Lake region in eastern Panamá and the Darien jungle near Colombia (Fig. 1). The San Blas Comarca is a vast territory encompassing 320 000 ha of forested coastline and hundreds of mangrove islands inhabited by a large percentage of the Kuna while the other two areas are currently inhabited by only a small number of Kuna. The Kuna population is composed of approximately 47 000 members, 67% of whom live in the island communities and coastal villages of the Comarca (1990 Panamanian government census). Roughly 2000 Kuna live along the Bayano, Chucunaque and Tuira rivers on the Pacific slope of Panamá and in the Darien Province. Approximately 14 000 Kuna have moved out of these territories into Panamá City, Colon and rural areas of Panamá (1990 Panamanian government census), while several hundred still live in Colombia.

Samples were collected from 53 individuals (45 blood samples and eight hair samples) representing the geographic range of the Panamanian Kuna

(Fig. 1). Extensive biographical information was recorded for each individual including the name, birthplace and language of his/her parents and grandparents whenever possible. Additionally, 10 serum samples were donated by the Gorgas Memorial Laboratory of Panamá City, Panamá which were collected between October 1983 and January 1985. Only unrelated individuals with no obvious, i.e. recent, admixture with other Amerind or non-Amerind groups were chosen for analysis. Despite our precautions, it is possible that some sampled individuals may be related due to the insular life-style of the Kuna.

### Isolation, amplification and sequencing of DNA

Ten to twenty milliliters of blood were collected into Vacutainer tubes (Becton Dickinson) containing ACD buffer (citric acid/dextrose). DNA was isolated from the blood samples following the procedure of Kolman *et al.* (12) which involved lysis of the erythrocytes, pelleting of the leucocytes and treatment with proteinase K and SDS, followed by salt precipitation of protein material and ethanol precipitation of the DNA. DNA was isolated from hair samples by incubation of the hair root(s) in a proteinase K/SDS buffer at 68°C overnight followed by four organic extractions (one phenol, two phenol/chloroform and one chloroform). The aqueous layer was further purified by filtration through a Centricon-100 filter. DNA isolation from serum samples required incubation of 100 µl serum with an equal volume of 2× extraction buffer (1× = 10 mM Tris-HCl, pH 8.0, 2 mM EDTA, 10 mM NaCl, 1% SDS, 8 mg/ml DTT and 0.4 mg/ml proteinase K) at 55°C overnight followed by organic extraction and Centricon-100 filtration as described for the hair samples. One microliter of DNA was used in all amplification reactions, although it was necessary to perform dilutions of four serum samples.

Primers L15997:H16401 (14) and L00029:H00408 (12) were used to amplify regions I and II of the mtDNA control region, respectively (20). (The number in the primer designation identifies the 3' ends according to the reference sequence [ref. 19] while L and H refer to the light and heavy strands, respectively.) Reaction components were as described in Kolman *et al.* (12). For region I sequence, PCR reactions were carried out for 27 cycles using the following temperature profile: 94°C for 45 seconds, 57°C for 30 s and 72°C for 30 s. Single-stranded DNA was generated by lambda exonuclease (Gibco BRL) digestion of one DNA strand initiated by a 3' phosphate. Two PCR reactions were carried out for each sample with only one of the primers containing a 3' phosphate in each case. The lambda exonuclease reaction was performed at 37°C for 30 min. After incubation at 96°C for 5 min to denature the enzyme, the samples were purified through Centricon-30 microconcentrators (Amicon Inc.). For region II sequence, single-stranded DNA was generated by a second, asymmetric PCR performed as previously described (12).

Seven microliters of the Centricon-30 retentate was sequenced using Sequenase Version 2.0 (US Biochemical). For region I sequence, primers L15997 (14), L16191 (5'-CCCATGCTTACAAGCAAGTA-3') and L16106 (5'-GCCAGCCACCATGAATATTG-3') were used to sequence the light strand and primers H16401 (14) and H16226 (5'-GCAGTTGATGTGTGATAGTT-3') were used to sequence the heavy strand of each sample. For region II sequence, PCR primers L00029 and H00408 (12) were used to sequence the light and heavy strands, respectively. Reaction products were separated by electrophoresis through 6% polyacrylamide gels containing 7 M urea. Owing to a T-to-C mutation at position 16189 which caused the *Taq* polymerase to stutter, we could only confidently read sequence in each direction until the location of the transition for the 18 individuals carrying this mutation. An internal primer downstream of this mutation (L16191) allowed us to read sequence of the light strand from position 16 230 to 16 400. In the remaining 45 individuals, the sequences analyzed were completely verified through full overlap of heavy and light strands.

### RFLP analysis

Limited RFLP analysis was performed on the sequenced individuals to allow comparison of our data with the four major mtDNA restriction site classes defined by Torroni *et al.* (25). Four sets of primers were used in PCR reactions (run for 28 cycles) to screen for presence of the *HaeIII* site at bp 663 (haplogroup A), the COII/tRNA<sup>Lys</sup> intergenic 9 bp deletion (haplogroup B), the *AluI* site at bp 13262 (haplogroup C), and absence of the *AluI* site at bp 5176 (haplogroup D). The primer pairs used were L00577:H00743, L08215:L08297, L13232: H13393 (12) and L05099 (5'-CCTAACTACTA-CGCGATTCCTAC-3'):H05333 (5'-CCTCGATAATGGCCCATTTGGGC-3'). Thermal cycler and PCR conditions were as described in Kolman *et al.* (12) except that L00577:H00743 and L05099:H05333 used the following temperature profile: 94°C for 30 s, 62°C for 15 s and 72°C for 15 s. PCR products were electrophoresed through 15% polyacrylamide gels or 4% MetaPhor (FMC BioProducts) agarose gels.



## Diversity, phylogenetic and statistical analysis

Haplotype diversity,  $h$  (37) and two measures of nucleotide diversity,  $\pi$  (38) and  $E(v)$  (39) were calculated. Sequence differences between mtDNA haplotypes were measured using Kimura's two-parameter model. These genetic distance data were summarized using the neighbor-joining (NJ) algorithm in NTSYS (40). Phylogenetic trees were constructed with the Phylogenetic Analysis Using Parsimony package (PAUP; 41), although computational limitations prevented us from analyzing all possible trees. Minimum spanning trees (MSTs) were constructed by hand following the recommendations of Excoffier *et al.* (42) and used in the analysis of molecular variance (AMOVA) program provided by L. Excoffier. The MST used for AMOVA analysis of all six indigenous groups is shown in Figure 2. Significance of variance terms was tested by permuting the original data 1000 times.

Pairwise genetic difference analyses were computed using programs provided by S. Sherry. Population expansion estimates from these analyses were based on the divergence rate for mtDNA control region I calculated by Harpending *et al.* (27). Simulations of pairwise genetic difference data were performed using programs provided by H. Harpending.

## Data analyzed

In analyses specific to the Kuna, we compared 370 bases (bp 16030–16400) across the mtDNA control region I for all 63 individuals studied. In addition to the 63 Kuna mtDNA sequences presented here, mtDNA sequences from five additional New World indigenous groups were also analyzed: 46 Ngöbé sequences (12), 27 Huetar sequences (13), 38 Mapuche sequences (15); the fourth haplotype of pattern 3 was not used because of missing data at position 16362), 63 Nuu-Chah-Nulth sequences (14) and 40 Haida sequences (16; sequence 35 was eliminated because of suspected non-Native American origin). To compare haplotype and nucleotide diversity between the six groups, we used only the 334 nucleotide positions common to all studies (bp 16050–16383). The same bases were used for PAUP, NJ, AMOVA and pairwise genetic difference analyses of the six groups. Polymorphic positions 16182 and 16183 were excluded from the analyses because of their non-independent association with the T–C mutation at position 16189 (43).

## ACKNOWLEDGMENTS

We gratefully acknowledge the participation of the Kuna people in our study. Françoise Guionneau-Sinclair provided essential assistance in obtaining blood samples and in the collection of biographical data on each individual. We gratefully acknowledge the gift of serum samples from Gorgas Memorial Laboratories. The assistance of Marivi Walker, Digna Pérez, Lucía Jorge and Chimene Longwater in obtaining, processing and analyzing samples was invaluable. Our sincere appreciation goes to Shawn McCafferty for laboratory and analytical assistance. We owe a special debt of thanks to Henry Harpending and Steve Sherry for their assistance in the analysis and simulation of the pairwise genetic difference data. We extend our heartfelt appreciation to Richard Cooke for his archeological expertise in discussion of conclusions based on the genetic data and his critical review of the manuscript.

## REFERENCES

- Dillehay, T.D. and Meltzer, D.J. (1991) *The First Americans: Search and Research*. Boca Raton, CRC Press, pp. 287–294.
- Constenla-Umaña, A. (1991) *Las Lenguas del Area Intermedia: Introduccion a su Estudio Areal*, Editorial de la Universidad de Costa Rica, San José, pp. 34–51.
- Linares, O.F. and Ranere, A.J. (1980) *Peabody Monographs* No. 5. Harvard University Press, Cambridge, MA, pp. 233–247.
- Cooke, G.R. and Ranere, A.J. (1992) In Lange, F.W. (ed.), *Wealth and Hierarchy in the Intermediate Area*. Dumbarton Oaks Research Library and Collection, Washington, DC, pp. 243–315.
- Romoli, K. (1987) *Los de la Lengua Cueva*. Ediciones Tercer Mundo, Colombia, p. 222.
- Howe, J. (1986) *The Kuna Gathering: Contemporary Village Politics in Panamá*. Latin American Monographs, No. 67. Univ. of Texas Press, Austin, p. 326.
- Costello R.W. (1983) In Bort J.R. and Helms M. (eds), *Local Reactions to Development Policies*. Department of Anthropology, University of Missouri, Columbia, pp. 91–100.
- Suer, F.R. (1979) *The Effect of Demographic Change on Agriculture in San Blas*. Ph.D. dissertation, Univ. of Arizona, University Microfilms, Ann Arbor.
- Howe, J. (1978) *Revista Panameña de Antropología* 2, 30–47.
- Howe, J. (1974) *Village Political Organization among the San Blas Kuna*. Ph.D. dissertation, University of Pennsylvania.
- Barrantes, R. (1993) *Evolucin en el Trpico: Los Amerindios de Costa Rica y Panamá*. Editorial de la Universidad de Costa Rica, San José, pp. 70, 122–123.
- Kolman, C.J., Bermingham, E., Cooke, R., Ward, R.H., Arias, T.D. and Guionneau-Sinclair, F. (1995) *Genetics*. In Press.
- Santos, M., Ward, R.H. and Barrantes, R. (1994) *Hum. Biol.*, 66, 963–977.
- Ward, R.H., Frazier, B.L., Dew-Jager, K. and Pääbo, S. (1991) *Proc. Natl Acad. Sci. USA*, 88, 8720–8724.
- Ginther, C., Corach, D., Penacino, G.A., Rey, J.A., Carnese, F.R., Hutz, M.H., Anderson, A., Just, J., Salzano, F.M. and King, M.C. (1993) In Pena, S.D.J., Chakraborty, R., Epplen, J.T. and Jeffreys, A.J. (eds), *DNA Fingerprinting: State of the Science*. Birkhäuser Verlag, Basel, Switzerland, pp. 211–219.
- Ward, R.H., Redd, A., Valencia, D., Frazier, B. and Pääbo, S. (1993) *Proc. Natl Acad. Sci. USA*, 90, 10663–10667.
- Torroni, A., Schurr, T.G., Cabell, M.F., Brown, M.D., Neel, J.V., Larsen, M., Smith, D.G., Vullo, C.M. and Wallace, D.C. (1993) *Am. J. Hum. Genet.*, 53, 563–590.
- Torroni, A., Chen, Y.-S., Semino, O., Santachlara-Beneceretti, A.S., Scott, C.R., Lott, M.T., Winter, M. and Wallace, D.C. (1994) *Am. J. Hum. Genet.* 54, 303–318.
- Anderson, S., Bankier, A.T., Barrell, B.G., deBruijn, M.H.L., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., Schreier, P.H., Smith, A.J.H., Staden, R. and Young, I.G. (1981) *Nature*, 290, 457–465.
- Vigilant, L., Pennington, R., Harpending, H., Koehler, T.D. and Wilson, A.C. (1989) *Proc. Natl Acad. Sci. USA*, 86, 9350–9354.
- Santos, M. and Barrantes, R. (1994) *Am. J. Hum. Genet.* 55, 413–414.
- Merriwether, D.A. (1993) *Mitochondrial DNA Variation in South American Indians*. Ph.D. dissertation, Univ. of Pittsburgh.
- Merriwether, D.A., Ferrell, R.E. and Rothhammer, F. (1995) *Am. J. Hum. Genet.* In Press.
- Schurr, T.G., Ballinger, S.W., Gan, Y.-Y., Hodge, J.A., Merriwether, D.A., Lawrence, D.N., Knowler, W.C., Weiss, K.M. and Wallace, D.C. (1990) *Am. J. Hum. Genet.*, 46, 613–623.
- Torroni, A., Schurr, T.G., Yang, C.C., Szathmary, E.J.E., Williams, R.C., Schanfield, M.S., Troup, G.A., Knowler, W.C., Lawrence, D.N., Weiss, K.M. and Wallace, D.C. (1992) *Genetics*, 130, 153–162.
- Rogers, A.R. and Harpending, H.C. (1992) *Mol. Biol. Evol.*, 9, 552–569.
- Harpending, H.C., Sherry, S.T., Rogers, A.L. and Stoneking, M. (1993) *Curr. Anthropol.*, 34, 483–496.
- Excoffier, L. and Langanay, A. (1989) *Am. J. Hum. Genet.*, 44, 73–85.
- Barrantes, R., Smouse, P.E., Neel, J.V., Mohrenweiser, H.W. and Gerhowitz, H. (1982) *Am. J. Phys. Anthropol.*, 58, 201–214.
- Newson, L. (1986) *The Cost of Conquest: Indian Decline in Honduras under Spanish Rule*. Westview Press, Boulder, pp. 95, 110, 128.
- Cook, N.D. (1992) In Verano, J.W. and Ubelaker, D.H. (eds), *Disease and Demography in the Americas*. Smithsonian Institution Press, Washington, pp. 209–210.
- Torroni, A., Neel, J.V., Smouse, P.E. and Barrantes, R. (1994b) *Proc. Natl Acad. Sci. USA*, 91, 1158–1162.
- Barrantes, R., Smouse, P.E., Mohrenweiser, H.W., Gershowitz, H., Azofeifa, J., Arias, T.D. and Neel, J.V. (1990) *Am. J. Hum. Genet.*, 46, 63–84.
- Thompson, E.A., Neel, J.V., Smouse, P.E. and Barrantes, R. (1992) *Am. J. Hum. Genet.*, 51, 609–626.
- Greenberg, J.H., Turner II, C.G. and Zegura, S.L. (1986) *Curr. Anthropol.*, 27, 477–497.
- Wallace, D.C. and Torroni, A. (1992) *Hum. Biol.*, 64, 403–416.
- Nei, M. and Roychoudhury, A.K. (1974) *Genetics*, 76, 379–390.
- Nei, M. and Tajima, F. (1981) *Genetics*, 76, 379–390.
- Watterson, G.A. (1975) *Theor. Popul. Biol.*, 7, 256–276.
- Rohlf, F.J. (1993) NTSYS-pc: Numerical Taxonomy and Multivariate Analysis System, Version 1.8, Exeter Software, Setauket, NY.
- Swofford, D.L. (1989) *PAUP, Phylogenetic Analysis Using Parsimony*, Version X. Illinois Natural History Survey, Champaign, IL.
- Excoffier, L., Smouse, P.E. and Quattro, J.M. (1992) *Genetics*, 131, 479–491.
- Horai, S., Kondo, R., Nakagawa-Hattori, Y., Hayashi, S., Sonoda, S. and Tajima, K. (1993) *Mol. Biol. Evol.* 10, 23–47.