Mitochondrial DNA Analysis Reveals Substantial Native American Ancestry in Puerto Rico

J.C. MARTÍNEZ-CRUZADO,[1] G. TORO-LABRADOR,[1] V. HO-FUNG,[1,2] M.A. ESTÉVEZ- MONTERO,[1,3] A. LOBAINA-MANZANET,[1,4] D.A. PADOVANI-CLAUDIO,[1,5] H. SÁNCHEZ- CRUZ,[1,6] P. ORTIZ-BERMÚDEZ,[1,7] AND A. SÁNCHEZ-CRESPO[1]

Abstract To estimate the maternal contribution of Native Americans to the human gene pool of Puerto Ricans—a population of mixed African, European, and Amerindian ancestry—the mtDNAs of two sample sets were screened for restriction fragment length polymorphisms (RFLPs) defining the four major Native American haplogroups. The sample set collected from people who claimed to have a maternal ancestor with Native American physiognomic traits had a statistically significant higher frequency of Native American mtDNAs (69.6%) than did the unbiased sample set (52.6%). This higher frequency suggests that, despite the fact that the native Taíno culture has been extinct for centuries, the Taíno contribution to the current population is considerable and some of the Taíno physiognomic traits are still present. Native American haplogroup frequency analysis shows a highly structured distribution, suggesting that the contribution of Native Americans foreign to Puerto Rico is minimal. Haplogroups *A* and *C* cover 56.0% and 35.6% of the Native American mtDNAs, respectively. No haplogroup *D* mtDNAs were found. Most of the linguistic, biological, and cultural evidence suggests that the Ceramic culture of the Taínos originated in or close to the Yanomama territory in the Amazon. However, the absence of haplogroup *A* in the Yanomami suggests that the Yanomami are not the only Taíno ancestors.

This study seeks to contribute to our knowledge about the maternal contribution of Native Americans to the gene pool of Puerto Ricans, a people of mixed African, European, and Amerindian ancestry, and to investigate the migration history

[1]Department of Biology, University of Puerto Rico at Mayagüez, Mayagüez, Puerto Rico 00681-9012.
[2]School of Medicine, Medical Sciences Campus, University of Puerto Rico, San Juan, Puerto Rico.
[3]Department of Cellular and Molecular Medicine, John Hopkins University, Baltimore, Maryland.
[4]Graduate School, Loma Linda University, Loma Linda, California.
[5]Department of Neuroscience, Case Western Reserve University, Cleveland, Ohio.
[6]Department of Environmental Engineering, Polytechnic University of Puerto Rico, San Juan, Puerto Rico.
[7]Department of Bacteriology, Medical Microbiology and Immunology, University of Wisconsin at Madison, Madison, Wisconsin.

Human Biology, August 2001, v. 73, no. 4, pp. 491–511.

KEY WORDS: NATIVE AMERICAN MTDNA HAPLOGROUPS, TAÍNOS, HISTORICAL MIGRATION

of the extinct Taínos, the native people of Puerto Rico that were subjugated by the Spaniards.

Archaeological evidence shows that the Greater Antilles were already inhabited 8000 years ago by the Archaic culture, nomads who relied on shellfish, fish, and game for their subsistence (Rouse 1992). It is not known whether they migrated to the Antilles from Venezuela or from the peninsulas of Yucatán or Florida. However, from most of the linguistic, biological, and cultural evidence, it is apparent that the Ceramic culture that arrived in Puerto Rico 2200 years before the present (YBP) originated in South America, in or close to the territory occupied today by the Yanomami. The Taíno language belongs to the Arawakan family of languages. From the distribution of the Arawakan family of languages, Arawakan speakers have been inferred to have originated in the South American lowlands of the Amazon Basin (Noble 1965). More recently, linguists have proposed that the Proto-Arawakan speakers moved along north of the Negro River, passed through the Casiquiare Canal, and descended the Orinoco River, along the way producing the Proto-Maipuran language, which gave rise to the Proto-Maipuran subfamily of languages to which the Arawak (Lokono) and Taíno languages belong (Rouse 1992).

Similarly, using biological traits, Imbelloni (1938) traced the origin of the Taíno people from the Guyana lowlands and the Orinoco Valley into Amazonia. In addition, the cultural ancestry of the Taíno Indians may be traced back at least to the Ronquinan Saladoid people along the Orinoco River up to its junction with the Apure River. From there, however, they could have come from the west via the Apure and Meta Rivers or from the Amazon by the Negro River (Rouse 1992). Hence, the region in Venezuela occupied by the Yanomami, bordered by the Orinoco to the north and the Casiquiare to the west, is located within the path probably taken by the Taíno ancestors in their migration to the Orinoco lowlands.

Little is known about the genetic contribution to Puerto Ricans of the Taínos or Native Americans in general. The native population of Puerto Rico was decimated in the 16th century as a consequence of the Spanish colonization. However, Indian towns existed on the coast two decades after the start of the Spanish colonization in 1508 (Tapia 1854). Upon their emancipation in 1542, many of the remaining Taíno Indians stayed in the mountains in the central region of the island and were slowly assimilated through the following decades or centuries by the settler population.

The genetic contribution of the Taíno Indians is a matter of speculation. By 1530, gold was no longer found in quantities adequate for mining, and the gold rush that developed in other colonies of Spain like Mexico and Perú had severely reduced the mostly European settler population. The settler population remained low for centuries, thus probably contributing significantly to a higher relative abundance of Native American genotypes in the people that developed into Puerto Ricans than most historical accounts would imply. The 1777 census reported a total of only 70,210 inhabitants in Puerto Rico. These were classified into 31,951 whites, 24,164 free "pardos" (light-skinned mixes mostly of Indian and white an-

cestry), 4747 free blacks, 4249 black slaves, 3343 "mulato" (mixes of black and white ancestry) slaves, and 1756 pure Indians.

The composition of the Puerto Rico population changed dramatically during the 19th century. Upon the American Declaration of Independence in 1776, a strong commercial relationship between the United States and Puerto Rico commenced. With it, the sugar industry, which had collapsed by 1640, was resurrected, as was the slave trade (Fernández-Méndez 1970). In addition, because the Spanish Empire was now crumbling in Mexico and South America, wealthy white refugees emigrated to Puerto Rico in great numbers and stimulated the economy by developing the sugar and coffee industrues, and, to a lesser degree, the tobacco industry.

Through the Treaty of Paris that ended the Spanish-American War, Puerto Rico became a US colony in 1898. The first census carried out by the United States in 1899 reported a population of 953,243 inhabitants, 61.8% of them classified as white, 31.9% as mixed, and 6.2% as black (Sanger et al. 1900). Thus, an immigration wave with strong European and African components helped increase the population of Puerto Rico over thirteenfold during the 19th century, no doubt significantly diluting the Native American portion of the Puerto Rican gene pool. No major immigration wave occurred during the 20th century.

To gain a better understanding of the Native American gene contribution to Puerto Ricans, of the dynamics of the extensive admixture process that has occurred during the past five centuries, and of the origins of the Taínos, we are undertaking an extensive population genetics study. As a first step, the continental origins of the maternally inherited mitochondrial DNAs (mtDNAs) are determined here by means of haplogroup identification.

Mitochondrial DNA has long been a powerful tool to study the evolutionary and migration history of humans because of its rapid mutation rate (Brown et al. 1979, 1982; Miyata et al. 1982; Kocher and Wilson 1991; Torroni et al. 1994a) and its maternal, nonrecombinant mode of inheritance (Hutchison et al. 1974; Giles et al. 1980). Genetic bottlenecks associated with historical migrations followed by population expansions were often accompanied by mutations that became unique to the derived expanded populations. Thus, mtDNAs carrying such mutations or restriction fragment length polymorphisms (RFLPs), which can be detected by RFLP analysis, can be traced to particular population expansion events. Mitochondrial DNAs sharing such RFLPs belong to particular haplogroups (Ballinger et al. 1992; Torroni et al. 1993a, 1993b, 1994b, 1994c, 1996; Chen et al. 1995; Wallace 1995), which are thus defined by these RFLPs that can now be called haplogroup-specific markers. Because haplogroups are usually continent-specific, the continental origin of mtDNAs in a population of mixed ancestry can be traced by means of haplogroup identification.

Mutations can sometimes arise independently more than once. Thus, when identifying haplogroups from a population of mixed ancestry, it is wise to test not only for haplogroup-specific markers, but for RFLPs that are known to have been present in the mtDNAs in which the haplogroup-specific marker originally arose.

In this work, we use the *Dde*I and *Alu*I restriction motifs at the 10394 and 10397 nucleotide positions (np), according to the Cambridge reference sequence (Anderson et al. 1981).

Four major Native American haplogroups constitute 98% of the mtDNAs of the New World. Their specific markers and their restriction motifs are shown in Table 1. Haplogroups *A* and *B* lack the 10394 *Dde*I and 10397 *Alu*I sites, and under that context are defined by an *Hae*III site gain at np 663 and a 9-base-pair (bp) deletion within region V of Cann and Wilson (1983), between the cytochrome oxidase second subunit and the $tRNA^{Lys}$ genes (COII/$tRNA^{Lys}$ site), respectively. Haplogroups *C* and *D* possess both sites, and under that context can be defined by an *Alu*I site gain at np 13262 and an *Alu*I site loss at np 5176, respectively.

Because of the continent specificity of haplogroups, the maternal contribution of any continent to the gene pool of a mixed population can be estimated. In this work, we focus on the Native American contribution to Puerto Ricans. We find that the maternal contribution of Native Americans to the Puerto Rican gene pool is 52.6%. Furthermore, we find some geographic structuring of the Native American mtDNA diversity within Puerto Rico and considerable structuring in its haplogroup frequencies. The implications of this observation for the continental origins of the Taínos is discussed.

Materials and Methods

Samples. Two different sample sets were obtained following informed consent. The first sample set was biased for Native American ancestry. It included 56 Puerto Ricans living in communities known historically for their strong Indian component, or those not living in these communities but having mothers or maternal grandmothers with phenotypic traits identified by themselves as "Indian-like," such as straight, dark hair; pronounced cheekbones; almond-shaped, dark eyes; and bronzed skin color. Of these persons, 10 came from Sector el Treinta,

Table 1. RFLP Motifs Defining the Four Major Native American Haplogroups

10394 DdeI/ *10397* AluI *Motif*	*Haplogroup*	*Haplogroup-Specific Marker*	*Subhaplogroups and Their 16517* HaeIII *Motifs*
(–/–)	*A*	+663 *Hae*III	*A1* (+16517 *Hae*III)
			A2 (–16517 *Hae*III)
	B	9 bp del reg V	*B1* (+16517 *Hae*III)
			B2 (–16517 *Hae*III)
(+/+)	*C*	+13262 *Alu*I	*C1* (+16517 *Hae*III)
			C2 (–16517 *Hae*III)
	D	–5176 *Alu*I	*D1* (+16517 *Hae*III)
			D2 (–16517 *Hae*III)

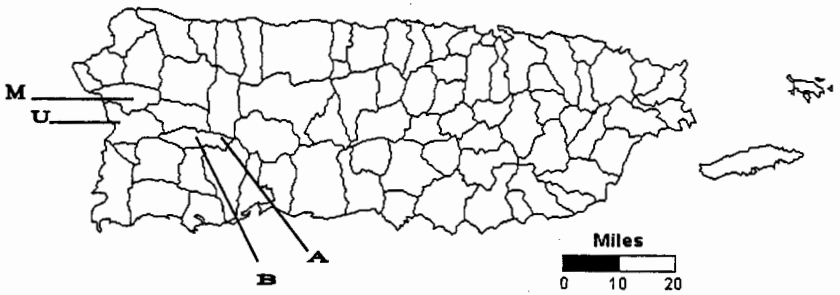

Figure 1. Geographic map of Puerto Rico. The localities from which samples were collected are indicated as follows: **A** = Indiera Alta, Maricao; **B** = Indiera Baja, Maricao; **M** = Miraflores, Añasco; **U** = University of Puerto Rico at Mayagüez.

Indiera Alta, Maricao; 8 from Indiera Baja, Maricao; and 5 from Miraflores, Añasco (Figure 1). Another 33 persons had mothers or maternal grandmothers with Indian-like phenotypic traits or who had originated from central Puerto Rico, an area that is considered to have the largest degree of Indian ancestry. This last group was composed of people associated with the University of Puerto Rico at Mayagüez or living in western Puerto Rico. The second, unbiased sample set was obtained independently from the first. It was taken at random from 38 Puerto Ricans associated with the University of Puerto Rico at Mayagüez or living in Mayagüez. In discussing comparisons between them, the last group of the biased sample set and the unbiased sample set will be referred to as the Native American–biased university group and the unbiased sample group, respectively.

Mitochondrial DNA Extraction. Volunteers provided three hair roots that were immersed in 0.5 mL of 5% Chelex (Sigma Chemical Co.) in a labeled 1.5 mL microcentrifuge tube. The tubes were stored in Styrofoam boxes cooled with cool packs until arrival at the lab on the same day. The samples were treated as in Walsh et al. (1991), with small modifications. They were placed at 58°C until the following day, when they were thoroughly vortexed, boiled for 8 min, vortexed again, and centrifuged at 13,000 *g* for 3 min to liberate the DNA. They were then stored at –20°C.

Polymerase Chain Reaction Amplification. Not more than eight fragments were amplified by the polymerase chain reaction (PCR) from each DNA sample for restriction or deletion analyses. The primers used for such amplifications and the restriction sites tested for their presence or absence in each fragment are shown in Table 2. The amplification reaction conditions were 1X PCR Buffer (Roche Molecular Biochemicals), 2.5 m*M* $MgCl_2$, 400 μ*M* dNTP, 1 μ*M* each primer, 5 μL of sample DNA, and 2.5 units of *Taq* DNA polymerase in a total

Table 2. Primers Used To Amplify Tested Sites

	Primers		Fragment
Sites Tested	*Name*[a]	*Sequence and Length*	*Size*[b]
10394 *Dde*I and	L10261	ACCTTCTTATTATTTGATCTAG (22 bp)	
10397 *Alu*I	H10425	TTAATGAGTCGAAATCATTCG (21 bp)	208 bp
663 *Hae*III	L612	ACCTCCTCAAAGCAATACACTG (22 bp)	
	H742	GTGCTTGATGCTTGTCCCTTTTG (23 bp)	176 bp
9 bp del Reg V	L8216	ACAGTTTCATGCCCATCGTC (20 bp)	
	H8296	ATGCTAAGTTAGCTTTACAG (20 bp)	121 bp[c]
13262 *Alu*I	L13233	CGCCCTTACACAAAATGACATCAA (24 bp)	
	H13392	TCCTATTTTTCGAATATCTTGTTC (24 bp)	208 bp
5176 *Alu*I	L5121	TAACTACTACCGCATTCCTA (20bp)	
	H5229	AAAGCCGGTTAGCGGGGGCA (20 bp)	149 bp
4577 *Nla*III	L4500	GGCCCAACCCGTCATCTAC (19 bp)	
	H4659	GAAGGATTATGGATGCGGTTG (21 bp)	200 bp
16517 *Hae*III	L16210	CCATGCTTACAAGCAAGT (18 bp)	
	H407	CTGTTAAAAGTGCATACCGCCA (22 bp)	807 bp
3592 *Hpa*I	L1636	AACACAAAGCACCCAACTTAC (21 bp)	
	H4083	GAAGTAGGGTCTTGGTGAC (19 bp)	2488 bp

a. Primers are named according to the strand extended from them (H for heavy, L for light) and th first base of the extension.
b. Primers included.
c. When no deletion or insertion occurs.

volume of 25 μL. The amplification reaction mixture was heated to 94°C for min before being subjected to 30 cycles of 1 min at 94°C, 1 min at 55°C, and 1. min at 72°C. One extension cycle of 10 min at 72°C completed the amplificatio

Restriction Digestion. The restriction digests were performed according t the manufacturer's conditions, using 10 to 15 μL of the amplification reactio and the DNA fractionated side-by-side to an undigested simile in a 2% or 3 agarose gel. To identify the 9-bp deletion at region V that defines haplogroup *B* i a (–/–) context (Table 1), the fragments amplified using primers L8216 an H8296 (Table 2) were fractionated in 4% NuSieve-1% standard agarose gels. Th agarose gels were stained with ethidium bromide and photodocumented und UV light.

DNA Sequencing. To determine the sequence of four PCR fragments ampl fied with primers L8216 (Table 2) and H8416 (5′ GGTGATGAGGAATAGTG TAAG 3′), the fragments were purified using the High Pure PCR Product Purif cation Kit (Roche Molecular Biochemicals) as instructed by the manufacture One hundred nanograms of each PCR product were sent together with 40 pic

moles of L8216 as the sequencing primer to the New Jersey Medical School Molecular Resource Facility for automated sequencing.

Results

All samples were tested for the 10394 *Dde*I / 10397 *Alu*I restriction motif and the four major Native American haplogroup-specific markers. All samples but the non–Native American samples of the unbiased sample set were also tested for the 16517 *Hae*III hypermutable site. In addition, all samples from the biased sample set were tested for the 4577 *Nla*III site. Finally, eight samples of the biased sample set were tested for the +3592 *Hpa*I marker that, in a (+/–) context, defines African haplogroups *L1* and *L2*, together known as superhaplogroup *L**. With one exception in each sample set, all Native American haplogroup-specific markers were found within the correct 10394 *Dde*I / 10397 *Alu*I context. In the biased sample set, the exception was sample B5 from Indiera Baja, which exhibited the 9-bp deletion specific for haplogroup *B* in a (+/–) context. In the unbiased sample set, the exception was U59, which exhibited the only haplogroup *D*–specific marker found in the study. It was found in a (+/–) context. These were not regarded as Native American mtDNAs.

In addition, only one mtDNA showed two haplogroup-specific markers: U22 in the biased sample set exhibited the specific markers for both haplogroups *A* and *B* within their correct (–/–) context. Because all four (–/–) mtDNAs with both of these haplogroup-specific markers whose control region have been sequenced have been shown to belong to haplogroup *A* (see Discussion), it was considered appropriate to regard it as belonging to haplogroup *A* and to include it in the Native American and haplogroup *A* totals.

The haplogroup distribution of the unbiased sample set (Table 3) shows that 20 of 38 (52.6%) mtDNAs were unambiguously identified as belonging to a Native American haplogroup. More than half of these (11) belonged to haplogroup *A*, whereas none in the sample belonged to haplogroup *D*. The haplogroup distribution of the biased sample set (Table 3) strongly suggests that not only haplogroup *A* but also haplogroup *C* prevails as a major Native American haplogroup in Puerto Rico. Of the 39 Native American mtDNAs in this sample set, 37 belong to one of these two haplogroups. Haplogroup *D* is again not represented in the sample set.

The haplogroup distribution of the Puerto Rican Native American mtDNA was further analyzed by testing for the hypermutable *Hae*III restriction sequence at position 16517. Each haplogroup may be subdivided into two subhaplogroups according to its 16517 *Hae*III motif and assigned the suffix "1" for the presence or "2" for the absence of this site (Table 1). Table 3 shows that there is a strong predominance of one or the other of the two motifs for each haplogroup. Specifically, for all samples combined, all haplogroup *B* and 20 of 21 haplogroup *C* mtDNAs possess the site, but 24 of 33 haplogroup *A* mtDNAs lack it.

Table 3. Haplogroup and Subhaplogroup Distribution in Native American Samples

	Distribution					
	Unbiased Sample Set		*Biased Sample Set*		*Sample Sets Combined*	
Native American Haplogroups	11 *A* (28.9)[a]	3 *A1* 8 *A2*	22 *A* (39.3)	6 *A1* 16 *A2*	33 *A* (35.1)	9 *A1* 24 *A2*
	3 *B* (7.9)	3 *B1* 0 *B2*	2 *B* (3.6)	2 *B1* 0 *B2*	5 *B* (5.3)	5 *B1* 0 *B2*
	6 *C* (15.8)	6 *C1* 0 *C2*	15 *C* (26.8)	14 *C1* 1 *C2*	21 *C* (22.3)	20 *C1* 1 *C2*
	0 *D* (0)	0 *D1* 0 *D2*	0 *D* (0)	0 *D1* 0 *D2*	0 *D* (0)	0 *D1* 0 *D2*
Total Native American mtDNAs	20 (52.6)		39 (69.6)		59 (62.8)	
Total Non–Native American mtDNAs	18 (47.4)		17 (30.4)		35 (37.2)	
Total	38 (100%)		56 (100%)		94 (100%)	

a. Percentages from total.

Table 4. Native American Subhaplogroup Distribution According to Collection Site i the Biased Sample Set

Subhaplogroup	*Indiera Alta*	*Indiera Baja*	*Miraflores*	*University at Mayagüez*
A1	4	0	0	2
A2	2	2	1	11
B1	0	1	0	1
B2	0	0	0	0
C1	0	1	3	10
C2	0	0	0	1
D1	0	0	0	0
D2	0	0	0	0
Total Native American mtDNAs	6	4	4	25
Total number of mtDNAs	10	8	5	33

Table 4 shows the Native American subhaplogroup distribution accordin to collection site in the biased sample set. Statistical analysis shows that the N tive American–biased university group has a significantly higher percentage Native American mtDNAs (25/33) than its unbiased counterpart [20/38; *G*-te with Williams' correction, $G(\text{adj}) = 4.06$, degrees of freedom (df) = 1, $p < 0.05$ This analysis suggests that, among the subjects composing the university grou of the biased sample set, either there exists a generally correct notion of the sen

blance of Indian phenotypic traits and an ability to identify them, or the central region actually has a larger Indian ancestry than elsewhere in western Puerto Rico. The Native American mtDNA frequency in the biased sample university group (75.8%) is even higher, although not significantly higher, than that found in the combination of the three communities known historically for their strong Indian component (60.9%).

Interestingly, the Indiera Alta distribution within haplogroup *A* is statistically significantly different from all other groups combined [G(adj) = 4.62, $df = 1$, $p < 0.05$]. Four of the six haplogroup *A* mtDNAs in Indiera Alta belong to subhaplogroup *A1*, whereas only 5 out of 27 do so in all other groups. Furthermore, when tested for the 9-bp deletion in region V, it was found that three of the four *A1* mtDNAs in Indiera Alta possessed a 9-bp insertion instead. These were the only three mtDNAs in the biased sample set with such an insertion. Only one other mtDNA in the unbiased sample set did (U65), and it belonged to subhaplogroup *A1* as well. The presence of the 9-bp insertion was confirmed in all four mtDNAs by DNA sequencing. It may be that these mtDNAs are identical by descent, and that the observed subhaplogroup distribution difference in Indiera Alta has historical meaning. The Indiera Alta population could have been founded by a population with few or maternally related women, distinct from the rest of the Native American population in Puerto Rico.

All samples from the Native American–biased sample set tested positive for the *Nla*III site at position 4577. The lack of such a restriction site in a (–/–) context is a specific marker for the Caucasian haplogroup V. This site is otherwise present in all other Caucasian and African mtDNAs tested to date. To our knowledge, however, no test for this site in Asian or Native American mtDNAs had been published. Our data suggest that the 4577 *Nla*III restriction site is also present in virtually all Puerto Rican Native American mtDNAs, and that the frequency of haplogroup V mtDNAs in Puerto Rico is very low.

Interestingly, the frequency of (+/-) mtDNAs among the non–Native American mtDNAs seemed higher for the biased sample set than for the unbiased one (Table 5). However, the difference was not significant at the 95% statistical level ($G = 3.53$; $df = 1, 0.1$; $p > 0.05$).

Table 5. Distribution in Native American Biased and Unbiased Sample Sets of Non–Native American mtDNAs Classified in Their Respective 10394 *Dde*I / 10397 *Alu*I Motifs

Restriction Motif	*Biased Sample Set*	*Unbiased Sample Set*	*Total*
(+/–)	15	11	26
(–/–)	2	7	9
Total	17	18	35

Discussion

The data from both sample sets suggest a strong Native American contribution to the mitochondrial gene pool of Puerto Ricans. Of the 56 samples in the biased sample set, 39 were identified as Native American in origin; of the 38 mtDNAs in the unbiased sample set, 20 were so identified. It could be argued that those five mtDNAs identified here as belonging to haplogroup *B* were not of Native American origin because the same 9-bp deletion that is specific to this haplogroup has arisen independently in Southeast Asia (Ballinger et al. 1992), Africa (Chen et al. 1995), Australia and the Pacific (Redd et al. 1995), and India (Watkins et al. 1999), as well as in Europe, although in this last case it has been found only in one mtDNA belonging to haplogroup *I* (Torroni et al. 1997). However, only the Asian-derived 9-bp deletion has been found within the –10394 *Dde*I / –10397 *Alu*I context, as those found here. The African and European deletions have been associated with a (+/–) context, and the Indian version with a (+/+) context. The Australian and Pacific deletions were not tested for their 10394 *Dde*I / 10397 *Alu*I motifs. Furthermore, of 21 mtDNAs with the deletion found by Alves-Silva et al. (1999) in Brazil, all 16 that were identified as Native American in origin according to their HV-1 sequence were (–/–), whereas all 4 that were identified as African were (+/–), as was the only European-derived one, which belonged to haplogroup *I*. Thus, it is reasonable to conclude that those (–/–) mtDNAs found here with the 9-bp deletion were Native American and belonged to haplogroup *B*.

In addition, it could be argued that many of the mtDNAs found here that belong to any of the four major Native American haplogroups were due to modern Asian immigration, since all four Native American haplogroups have their origin in Asia (Schurr et al. 1990; Torroni et al. 1993b). However, this is very unlikely. The 1899 census reported only 75 Asians, 66 of whom were men, in Puerto Rico (Sanger et al. 1900). Furthermore, Asian immigration to Puerto Rico through the 20th century was scant to the point that censuses taken every 10 years throughout the century never included any Asian country or region among those reported as origins for foreign residents, despite typically naming over 10 such countries or regions.

The only mtDNA found with the 9-bp deletion in a (+/–) context, *B5*, was probably African in origin. Alternatively, it could be a Native American mtDNA that gained the *Dde*I site at np 10394 by mutation. Although it failed to produce the +3592 *Hpa*I motif (data not shown) characteristic of haplogroups *L1* and *L2* (Chen et al. 1995), which together encompass approximately 70% of sub-Saharan mtDNAs, it could belong to the African haplogroup *L3*. A (+/–) mtDNA with the 9-bp deletion that belongs to African haplogroup *L3* was recently reported in Brazil (Alves-Silva et al. 1999). Because this deletion has been shown to have arisen independently multiple times in Africa (Soodyall et al. 1996) and in Asia (Betty et al. 1996; Watkins et al. 1999), an African origin for *B5* is more likely than a Native American origin.

blance of Indian phenotypic traits and an ability to identify them, or the central region actually has a larger Indian ancestry than elsewhere in western Puerto Rico. The Native American mtDNA frequency in the biased sample university group (75.8%) is even higher, although not significantly higher, than that found in the combination of the three communities known historically for their strong Indian component (60.9%).

Interestingly, the Indiera Alta distribution within haplogroup *A* is statistically significantly different from all other groups combined [G(adj) = 4.62, $df = 1$, $p < 0.05$]. Four of the six haplogroup *A* mtDNAs in Indiera Alta belong to subhaplogroup *A1*, whereas only 5 out of 27 do so in all other groups. Furthermore, when tested for the 9-bp deletion in region V, it was found that three of the four *A1* mtDNAs in Indiera Alta possessed a 9-bp insertion instead. These were the only three mtDNAs in the biased sample set with such an insertion. Only one other mtDNA in the unbiased sample set did (U65), and it belonged to subhaplogroup *A1* as well. The presence of the 9-bp insertion was confirmed in all four mtDNAs by DNA sequencing. It may be that these mtDNAs are identical by descent, and that the observed subhaplogroup distribution difference in Indiera Alta has historical meaning. The Indiera Alta population could have been founded by a population with few or maternally related women, distinct from the rest of the Native American population in Puerto Rico.

All samples from the Native American–biased sample set tested positive for the *Nla*III site at position 4577. The lack of such a restriction site in a (–/–) context is a specific marker for the Caucasian haplogroup V. This site is otherwise present in all other Caucasian and African mtDNAs tested to date. To our knowledge, however, no test for this site in Asian or Native American mtDNAs had been published. Our data suggest that the 4577 *Nla*III restriction site is also present in virtually all Puerto Rican Native American mtDNAs, and that the frequency of haplogroup V mtDNAs in Puerto Rico is very low.

Interestingly, the frequency of (+/-) mtDNAs among the non–Native American mtDNAs seemed higher for the biased sample set than for the unbiased one (Table 5). However, the difference was not significant at the 95% statistical level ($G = 3.53$; $df = 1, 0.1$; $p > 0.05$).

Table 5. Distribution in Native American Biased and Unbiased Sample Sets of Non–Native American mtDNAs Classified in Their Respective 10394 *Dde*I / 10397 *Alu*I Motifs

Restriction Motif	*Biased Sample Set*	*Unbiased Sample Set*	*Total*
(+/–)	15	11	26
(–/–)	2	7	9
Total	17	18	35

Only one other mtDNA had a haplogroup-specific marker out of its 10394 *Dde*I / 10397 *Alu*I context. U59 had the 5176 *Alu*I site loss specific for haplogroup *D* mtDNAs but within a (+/–) rather than a (+/+) context. An extensive literature search (Ballinger et al. 1992; Torroni et al. 1993a, 1993b, 1994b, 1994c, 1996, 1997; Chen et al. 1995; Starikovskaya et al. 1998) did not produce any other mtDNA like it, although two Tibetans, two Taiwanese Han, two Sabah aborigines, one Siberian, and one Mayan were found with the 5176 *Alu*I site loss in a (–/–) context. Further RFLP and HV-1 sequence analysis will be needed to identify the biological origin of this mtDNA.

The data presented here are consistent with the HV-1 sequences that Sánchez-Crespo (1999) obtained from four human skeletal remains unearthed in Arecibo, Puerto Rico. These remains were radiocarbon dated at 1420 YBP (calibrated to A.D. 590 to 675, 2 sigma, 95% probability, Beta Analytic). Interestingly, all of the Arecibo samples were identical, exhibiting the 16298C and 16327C mutations characteristic of Native American haplogroup *C* mtDNAs (Torroni et al. 1993a).

Our data are also consistent with the results of Abujoub (1994), who sequenced the HV-1 region from 50 mtDNAs collected at random in the southern city of Ponce. Although Abujoub estimated the Native American contribution at 68%, our reanalysis of his data using more recent published sequences of Caucasian (Calafell et al. 1996; Torroni et al. 1996; Richards et al. 1996) and African (Graven et al. 1995; Soodyall et al. 1996) origin showed that some of the mtDNAs initially identified as Native American were really African or Caucasian. Only 23 of the 50 mtDNAs studied could be identified as belonging to a Native American haplogroup; in this case, they belonged to either haplogroup *A* (11) or *C* (12). Of the remainder, 12 were identified as African and 3 as Caucasian in origin. Of the 12 sequences that could not be identified, none could belong to haplogroup *B*, and only 2 may have belonged to Native American haplogroup *D*. Because haplogroup *D* does not have base substitutions in the HV-1 region as specific as those of the remaining major Native American haplogroups (Torroni et al. 1993a), the evidence provided by the HV-1 sequence is too weak to classify these mtDNAs as belonging to haplogroup *D*. Furthermore, the absence of haplogroup *D* in both of our sample sets strongly suggests that haplogroup *D* is very rare in Puerto Rico. However, these two mtDNAs exhibited transitions at two of the three sites where most haplogroup *D* mtDNAs do (Easton et al. 1996; Stone and Stoneking 1998; Alves-Silva et al. 2000). Namely, they exhibited transitions at np 16223 and 16362 but not at 16325. Thus, RFLP analyses would be needed to clarify the identity of these mtDNAs.

In conclusion, if the Abujoub sample was representative of the Puerto Rican population, the contribution of the four major Native American lineages would be estimated between 46% and 50%, depending on the identities of the two potential haplogroup *D* mtDNAs. This range is very close to the 52.6% estimated from our unbiased sample set. It must be emphasized, however, that neither our sample sets nor Abujoub's sample set are necessarily representative of Puerto

Rico as a whole. His consists mainly of residents from the south side of the island while ours are mainly from residents of the west side of the island.

Interestingly, among the 12 Abujoub sequences whose haplogroup could not be identified, 3 possessed the transitions at np 16223 and 16278 that characterize haplogroup *X* mtDNAs, although they lacked the transition at np 16189 that usually accompanies them (Torroni et al. 1996; Brown et al. 1998; Smith et al. 1999; Alves-Silva et al. 2000; Richards et al. 2000). Haplogroup *X* was first described as a minor Caucasian haplogroup, accounting for 4.5% of a combined sample of Finnish, Swedish, and Tuscan (Torroni et al. 1996) mtDNAs, and later shown also to represent a minor Native American founding lineage (Forster et al. 1996). It is widespread in Native North Americans. It is found in at least seven unrelated language families, accounting for approximately 3% of their mtDNAs (Smith et al. 1999). Although a literature search did not produce mention of modern haplogroup *X* mtDNAs in South America, they were found in 3 of the 18 skeletal remains from the Brazilian Amazon with datings estimated at 500 to 4000 YBP, which were studied by Ribeiro-Dos-Santos et al. (1996). It is noteworthy that these ancient haplogroup *X* mtDNAs showed transitions at np 16223 and 16278 but not at np 16189, just as the aforementioned Abujoub sequences. Thus, it is possible that haplogroup *X* was present in the Taíno population.

In that case, because our study did not screen for haplogroup *X* mtDNAs, the Native American ancestry of Puerto Ricans may be higher than estimated here. Haplogroup *X* is defined by the lack of a *Dde*I site at np 1715 in a (–/–) context. In our unbiased sample set, 7 out of 38 mtDNAs had the (–/–) motif and did not belong to any of the four major Native American haplogroups. If all of them belonged to the Native American haplogroup *X*, the estimate would increase from 52.6% to 71.1%. However, we are not aware of a reliable method to distinguish Native American from Caucasian haplogroup *X* mtDNAs in a population of mixed ancestry such as that in Puerto Rico. Thus, finding haplogroup *X* mtDNAs in modern Puerto Ricans would not necessarily imply a higher Native American ancestry.

Another group of Native American mtDNAs was found in 10 of 83 Yanomami mtDNAs (Easton et al. 1996). These lacked any haplogroup-specific marker but had the +10394 *Dde*I/+10397 *Alu*I motif found in mtDNAs belonging to haplogroups *C* or *D*. A neighbor-joining tree of Native American and Mongolian lineages based on control region sequences failed to find a common origin for them (Stone and Stoneking 1998), although they shared a transition at np 16325 and generally tended to associate with haplogroup *D* sequences in the tree. These mtDNAs were not found in our study, since all mtDNAs that exhibited the (+/+) motif were shown to belong to haplogroup *C*.

The percentage contribution of Native Americans to the Puerto Rican mitochondrial gene pool estimated here could come as a surprise to some. The number of Taíno Indians in Puerto Rico was estimated by some historians at only 20,000 at the start of the Spanish colonization in 1508. The conditions imposed by colonization presumably led to the massive emigration to the Lesser Antilles of most

of those Taínos not killed by war, the abuses of slavery, or diseases brought by the Europeans (Anderson-Córdova 1995). The 1530 census reported only 1148 Indians in Puerto Rico (Brau 1904), 675 slaves and 473 servants, and by many accounts the Taínos were exterminated before the end of the 16th century.

It is apparent, however, that the 1530 census ignored what could have been thousands of Taínos living in the inhospitable mountains of the central region of Puerto Rico. Indian emancipation shortly thereafter (1542) may have propitiated the slow and relatively peaceful assimilation of this population into the colonial society. In addition, the size of the Taíno population at the time of Spanish colonization is in dispute, some authors estimating it at 600,000 (Abbad 1788). Furthermore, Indian slave immigration to Puerto Rico was common before Indian emancipation. These were mostly brought from the islands of Margarita, Trinidad, Aruba, and Bonaire, and from the coast of Venezuela, but some were brought from the Yucatán peninsula (Zavala 1948) or Brazil (Tapia 1854). Moreover, up until the American Revolution, European immigration to Puerto Rico was scant, and European emigration quite common (Fernández-Méndez 1970). Even African slaves were not introduced in very large numbers until the beginning of the 19th century, when they were brought in to work in the sugar industry that was then stimulated by wealthy refugees from the Spanish colonies in revolution. Thus, before the large immigration wave of the 19th century, the 1777 census put the total population of Puerto Rico at only 70,210 inhabitants, and the Native American genetic component may have been the largest one at the time.

The Native American–biased university group had a statistically significant higher frequency of Native American mtDNAs than that of the unbiased group. It is possible that, because of the large Indian ancestry found, Indian physiognomic traits persist in the population and are remembered by the people even though the Taíno culture has long been extinct. However, it should be kept in mind that only the female ancestry can be tracked through mtDNA studies. Because the majority of Europeans who arrived in Puerto Rico, especially before the 19th century, were men, and because the social stigma attached to European women bearing children of non-European men was much more severe than that attached to European men fathering children of non-European women, it is reasonable to expect that mtDNA studies will underestimate the Caucasian ancestry and consequently overestimate all other ancestries in Puerto Rico.

For example, recent studies using mtDNA and Y-chromosome Amerindian markers have found strikingly disparate contributions of Amerindian men and women to the population of Brazil. Whereas 59% of the mtDNAs in Belém were identified as Amerindians, less than 5% of the Y chromosomes were (Batista dos Santos et al. 1999). In another study on Brazil, virtually all Y chromosomes examined by Carvalho-Silva et al. (2001) were of European origin, while more than 60% of the mtDNAs were Amerind. Similarly, whereas 71.0% of the nuclear DNA of Canary Islanders is European in origin, Pinto et al. (1996) found that only 35.6% of the mtDNA is. The combination of these results suggests that the vast majority of the paternal contribution to the Canarian gene pool must be of

European origin. An even more drastic asymmetry was found in Antioquia, Colombia. In this province, approximately 94% of the Y chromosomes were of European origin and only 1% was Amerind, but close to 90% of the mtDNAs were of Native American origin (Carvajal-Carmona et al. 2000).

Ancient DNA studies have shown that haplogroup frequencies in native populations of North America have remained stable over thousands of years (O'Rourke et al. 2000). Thus, haplogroup frequencies may be useful to trace the origin of the Taínos to their continental ancestors. The haplogroup distribution in the Puerto Rican Indian mtDNAs is highly structured, with haplogroups *A* and *C* accounting for 91.5% of all mtDNAs. A highly structured population with reduced mtDNA diversity has been observed frequently in Native American tribes in North, Central, and South America (Torroni et al. 1993a: Santos et al. 1994; Batista et al. 1995; Kolman et al. 1995; Merriwether and Ferrell 1996; Ward et al. 1996; Kolman and Bermingham 1997), suggesting that they were founded by groups of women with low mtDNA diversity and that they developed mostly in isolation. This observation suggests that most of the Native American mtDNAs found in Puerto Ricans have a common origin in the Taínos and that the number of women with Native American mtDNAs migrating to Puerto Rico after the Spanish arrival has been small compared with the number of female descendants that the Taíno women were able to leave in the colony. Otherwise, because the documented immigration from other Spanish colonies of Indians and mixed women that could have had Native American mtDNAs stemmed from various localities (see above), a high immigration rate would have been expected to reduce the frequency of haplogroups *A* and *C* and the haplogroup distribution structuring. The highly structured haplogroup distribution also suggests that the Taínos had reduced genetic diversity. This scenario is consistent with the repeated water crossings over the Lesser Antilles that a colonization originating in the Orinoco delta would entail.

To further characterize the genetic variability of the Native American mtDNAs of Puerto Rico and to analyze their possible pre-Columbian origin in more detail, the haplogroups were divided into subhaplogroups according to the presence (1) or absence (2) of the 16517 *Hae*III site. Like other sites in the mtDNA control region (Wakeley 1993), the 16517 *Hae*III site has been determined to be hypermutable, since it has been found in both states in high frequencies in all world haplogroups described to date. It is, thus, of low informative value for human population migration and expansion studies (Forster et al. 1996), which generally use restriction sites that change, on average, every 24,420 years (Torroni et al. 1998). However, this site might be useful for studies of very recent migration and expansion events, such as the arrival of the Ceramic culture to Puerto Rico 2200 years ago (Rouse 1992).

Table 6 shows the subhaplogroup distribution in the Native American mtDNAs of Puerto Rico, together with those of the Yanomami, non-Yanomami Amazonians, and Mexican or Central American Amerinds found in the literature (Merriwether and Ferrell 1996; Kolman and Bermingham 1997). No such infor-

Table 6. Native American mtDNA Subhaplogroup Percentage Distribution in Puerto Rico, the Amazon, and Mexico and Central America[a]

Subhaplogroup	*Puerto Rico (n = 59)*	*Yanomama (n = 107)*	*Non-Yanomama Amazon (n = 91)*[b]	*Mexico and Central America*[c] *(n = 267)*
A1	15.3	0	7.7	11.6
A2	40.7	0	8.8	43.4
B1	8.5	7.5	14.3	30.3
B2	0	0.9	0	0
C1	33.9	35.5	7.7	6.0
C2	1.7	33.6	23.1	7.5
D1	0	2.8	15.4	0.4
D2	0	10.3	22.0	0.7
X6	0	6.5	1.1	0
X7	0	2.8	0	0

a. MtDNAs classified as "other" are not included. These are 8 mtDNAs from Mexico and Central America and 3 from the non-Yanomama Amazon.

b. Includes mtDNAs from the following tribes: 10 Piaroa, 10 Makiritare, 12 Wapishana, 10 Macushi, 25 Ticuna, 10 Marubo, and 14 Kraho.

c. Includes mtDNAs from the following tribes: 15 Mixtec-Alta, 14 Mixtec-Baja, 15 Zapotec, 16 Mixe, 23 Maya, 24 Bribri-Cabecar, 13 Boruca, 20 Guatuso, 20 Teribe, 16 Guaymi, 16 Kuna, 44 Emberá, and 31 Wounan.

mation was found for Native Floridians, although it is well known that the mtDNA of the 7000-year-old remains found in southwestern Florida (Pääbo et al. 1988) belonged to the rare Native American haplogroup *X* (Hauswirth et al. 1994). Two important observations come to light. First, the proportion of *C1* to *C2* mtDNAs is very much higher in Puerto Rico than in the Yanomami, the rest of the Amazon, or Mexico and Central America. This observation combined with the hypermutability of the 16517 *Hae*III site suggests that subhaplogroup *C1* may represent a very recent human population expansion in pre-Columbian Puerto Rico. Many Native American RFLP haplotypes have been found to be identical between themselves except for the 16517 *Hae*III site, including tribal-specific haplotypes (Torroni et al. 1992, 1993a, 1994a). Even in populations less than 10,000 years old, pairs of haplotypes like these have been found (Starikovskaya et al. 1998). Thus, subhaplogroup *C1* in Puerto Rico may be much younger than 10,000 years old, and could represent the arrival of the Ceramic culture 2200 years ago.

Second, subhaplogroup *C1* is not common in Mexico–Central America or in the non-Yanomami Amazonians, but is common in the Yanomami. This finding is consistent with a close genealogical relationship between the subhaplogroup *C1* mtDNAs in Puerto Rico and the Yanomami, as may be expected according to cultural, linguistic, and biological evidence that suggests a South America origin for the Ceramic culture of Puerto Rico, close to the territory occupied today by

the Yanomami. However, it is noteworthy that the Yanomami have other subhaplogroups at relatively high frequencies that are absent in Puerto Rico (Table 6).

In addition, haplogroup *A*, which encompasses the majority of the Puerto Rico Indian mtDNAs, is absent in the Yanomami and uncommon in the rest of the Amazon. Hence, a strong non-Amazonian contribution to the Taíno gene pool cannot be discarded from the data.

Some archaeological evidence points towards a strong influence from extinct cultures along the northern coast of South America in Puerto Rican cultures that preceded the Taínos (Chanlatte and Narganes 2001), but mtDNA sampling in this region has not been performed. In addition, the very high frequency of haplogroup *A* in Mexico and Central America (Table 6) and the fact that the Archaics that preceded the Ceramic culture may have originated from the Yucatán peninsula should not be ignored. The finding of haplotype *U22*, a haplogroup *A* mtDNA with the 9-bp deletion in region V, supports a Mexican–Central American origin for haplogroup *A* mtDNAs. Only five other mtDNAs having the 663 *Hae*III site and the region V 9-bp deletion have been identified elsewhere, and all have been from Mexico or Central America: one from the Maya (Schurr et al. 1990), one from the Boruca (Torroni et al. 1993a), and three from the Mixtecs from the Baja (Torroni et al. 1994d).

Thus, the simplest scenario, which is often not the true one when it comes to historical human population migration events, would be that haplogroup *A* represents the Archaic culture that migrated from the Yucatán peninsula through Cuba and Hispaniola, and that haplogroup *C* represents the Ceramic culture that migrated later from the Amazon through the Lesser Antilles. The less frequent Puerto Rican haplogroup *B* could represent minor founding lineages or modern migrations. More thorough characterization of the Native American mtDNAs of Puerto Rico and reduced median network analysis (Bandelt et al. 1995) may better distinguish the native, more closely related and frequent mtDNA haplotypes from recent immigrants, and shed some light on the continental origin of the Taínos. Genetic diversity studies in native mtDNAs should lead to an estimation of the time at which the last genetic bottleneck in the population that gave rise to the Taínos occurred.

Regarding subhaplogroup *A1*, it is interesting that four of the nine mtDNAs identified as belonging to this subhaplogroup exhibited a 9-bp insertion in region V, thus further subdividing haplogroup *A*. The fact that three of the mtDNAs with the insertion were found in the same locality (Indiera Alta) suggests a common origin for these mtDNAs. Insertions in region V have been reported only sporadically and in small numbers. They have been found to be produced by a 3- to 4-bp expansion of the C-run that constitutes 5 bp of the 9-bp repeated unit (Ballinger et al. 1992; Torroni et al. 1993b; Watkins et al., 1999), by the addition of a third 9-bp unit (Shields et al. 1992; Passarino et al. 1993; Torroni et al. 1997; Lum and Cann 1998; Alves-Silva et al. 1999), or by an apparent combination of both. Thus, a 12-bp insertion has been reported for this region in three Pygmies from

the Central African Republic (Chen et al. 1995). DNA sequencing showed that the insertion corresponds to an addition of a third 9-bp unit in the three Indiera Alta mtDNAs and in U65. Furthermore, no nucleotide differences were observed among any of the four sequences, even when the sequenced region extended for 176 bp (nucleotide positions 8241 to 8416). Thus, the four mtDNAs may have a common origin, distinct from that of most of the Native American mtDNAs in Puerto Rico.

The Indiera Alta Indian population was first reported in the census of 1777 as a group of "pure Indians." Brau (1904) maintains that this population was constituted by a group of Indians from southwestern Puerto Rico that in 1570 moved deep into the mountains to remain isolated from the colonial society for centuries. Our data, although scant, suggest that this was a very small group.

Although our data are consistent with those of Abujoub (1994), neither study used a sample representative of the Puerto Rican population. We have started a study, the results of which will be presented elsewhere, that will identify the biological origin of Native American, sub-Saharan African, and Caucasian mtDNAs based on the continent-specificity of haplogroups. This study entails testing samples from a representative group of Puerto Ricans for all known haplogroup-specific markers. However, to our knowledge, the –4577 *Nla*III specific marker for Caucasian haplogroup V (Torroni et al. 1996) had not been tested in Asians or Native Americans. The presence of this motif in Native American mtDNAs could lead to the wrong identification of those mtDNAs that had a mutation eliminating their Native American haplogroup-specific marker. This could mislead the study towards a conclusion significantly deviated from the true Native American and Caucasian mtDNA ancestries of Puerto Ricans if such mtDNAs were to represent a founding lineage of the island. Thus, it was considered a prudent preliminary step to test for the presence of the –4577 *Nla*III motif on all samples of the biased sample set. The lack of such marker in all mtDNAs strongly suggested that it is indeed found only in Caucasian mtDNAs belonging to haplogroup V, and that haplogroup V mtDNAs are very infrequent in Puerto Rico.

The information obtained from a study that identifies the biological origin of mtDNAs in a sample set representative of Puerto Ricans should include their geographic distribution within Puerto Rico. The apparent association between (+/-) non–Native American mtDNAs and the biased sample set, the set with a higher percentage of Native American mtDNAs (Table 5), would suggest that the sub–Saharan African component predominates in the non–Native American ancestry of groups of strong Native American ancestry. Approximately 91% of sub-Saharan African mtDNAs exhibit the (+/–) motif (Chen et al. 1995), whereas only 25% of Tuscans and 19% of Scandinavians do (Torroni et al. 1996). Thus, it should be reasonable to expect that Puerto Rican groups with a higher (+/–) frequency may have a stronger African ancestry. The apparent association between (+/–) non–Native American mtDNAs and Native American mtDNAs could reflect the historical socioeconomic segregation of ethnic groups. However, the

the Central African Republic (Chen et al. 1995). DNA sequencing showed that the insertion corresponds to an addition of a third 9-bp unit in the three Indiera Alta mtDNAs and in U65. Furthermore, no nucleotide differences were observed among any of the four sequences, even when the sequenced region extended for 176 bp (nucleotide positions 8241 to 8416). Thus, the four mtDNAs may have a common origin, distinct from that of most of the Native American mtDNAs in Puerto Rico.

The Indiera Alta Indian population was first reported in the census of 1777 as a group of "pure Indians." Brau (1904) maintains that this population was constituted by a group of Indians from southwestern Puerto Rico that in 1570 moved deep into the mountains to remain isolated from the colonial society for centuries. Our data, although scant, suggest that this was a very small group.

Although our data are consistent with those of Abujoub (1994), neither study used a sample representative of the Puerto Rican population. We have started a study, the results of which will be presented elsewhere, that will identify the biological origin of Native American, sub-Saharan African, and Caucasian mtDNAs based on the continent-specificity of haplogroups. This study entails testing samples from a representative group of Puerto Ricans for all known haplogroup-specific markers. However, to our knowledge, the –4577 *Nla*III specific marker for Caucasian haplogroup V (Torroni et al. 1996) had not been tested in Asians or Native Americans. The presence of this motif in Native American mtDNAs could lead to the wrong identification of those mtDNAs that had a mutation eliminating their Native American haplogroup-specific marker. This could mislead the study towards a conclusion significantly deviated from the true Native American and Caucasian mtDNA ancestries of Puerto Ricans if such mtDNAs were to represent a founding lineage of the island. Thus, it was considered a prudent preliminary step to test for the presence of the –4577 *Nla*III motif on all samples of the biased sample set. The lack of such marker in all mtDNAs strongly suggested that it is indeed found only in Caucasian mtDNAs belonging to haplogroup V, and that haplogroup V mtDNAs are very infrequent in Puerto Rico.

The information obtained from a study that identifies the biological origin of mtDNAs in a sample set representative of Puerto Ricans should include their geographic distribution within Puerto Rico. The apparent association between (+/-) non–Native American mtDNAs and the biased sample set, the set with a higher percentage of Native American mtDNAs (Table 5), would suggest that the sub–Saharan African component predominates in the non–Native American ancestry of groups of strong Native American ancestry. Approximately 91% of sub-Saharan African mtDNAs exhibit the (+/–) motif (Chen et al. 1995), whereas only 25% of Tuscans and 19% of Scandinavians do (Torroni et al. 1996). Thus, it should be reasonable to expect that Puerto Rican groups with a higher (+/–) frequency may have a stronger African ancestry. The apparent association between (+/–) non–Native American mtDNAs and Native American mtDNAs could reflect the historical socioeconomic segregation of ethnic groups. However, the

data did not show a significant difference in the frequency of (+/–) mtDNA among the non–Native American mtDNAs between the biased sample set and th unbiased one.

Acknowledgments We are grateful to the volunteers who provided samples for th study. We thank Rob Donnelly (New Jersey Medical School Molecular Resource Facility for the oligonucleotide primers used in this project and for his automated DNA sequencin services. We also thank Duane A. Kolterman (University of Puerto Rico at Mayagüez) an two excellent anonymous reviewers for their useful comments on the analyses and th manuscript. This study was supported by the MBRS/National Institute of Health and th RUI/National Science Foundation through grants to J.C. Martínez-Cruzado.

Received 7 February 2001; revision received 12 March 2001.

Literature Cited

Abbad, F.I. 1788. *Historia geográfica, civil y natural de la Isla de San Juan Bautista de Puerto Rico* San Juan, PR: Ediciones Porta Coeli.

Abujoub, A.A.R. 1994. Polymorphism of the mitochondrial DNA control region in the Puerto Rica population. M.Sc. thesis, Michigan State University, East Lansing, Michigan.

Alves-Silva, J., M. da Silva Santos, P.E.M. Guimaraes et al. 2000. The ancestry of Brazilian mtDN lineages. *Am. J. Hum. Genet.* 67:444–461.

Alves-Silva, J., M.S. Santos, S.D.J. Pena et al. 1999. Multiple geographic sources of region V 9-b deletion haplotypes in Brazilians. *Hum. Biol.* 71:245–259.

Anderson, S., A.T. Bankier, B.G. Barrell et al. 1981. Sequence and organization of the human mito chondrial genome. *Nature* 290:457–465.

Anderson-Córdova, K. 1995. Aspectos demográficos de los cacicazgos Taínos. In *Proceedings of th Fifteenth International Congress for Caribbean Archaeology*, R. Alegría and M. Rodríguez eds. San Juan, PR: Centro de Estudios Avanzados de Puerto Rico y el Caribe, 351–366.

Ballinger, S.W., T.G. Schurr, A. Torroni et al. 1992. Southeast Asian mitochondrial DNA analysis re veals genetic continuity of ancient Mongoloid migrations. *Genetics* 130:139–152.

Bandelt, H.J., P. Forster, B.C. Sykes et al. 1995. Mitochondrial portraits of human populations usin median networks. *Genetics* 141:743–753.

Batista, O., C. Kolman, and E. Bermingham. 1995. Mitochondrial DNA diversity in the Kun Amerinds of Panamá. *Hum. Mol. Genet.* 4:921–929.

Batista dos Santos, S.E., J.D. Rodrigues, A.K.C. Ribeiro dos Santos et al. 1999. Differential contribu tion of indigenous men and women to the formation of an urban population in the Amazon re gion as revealed by mtDNA and Y-DNA. *Am. J. Phys. Anthropol.* 109:175–180.

Betty, D.J., A.N. Chin-Atkins, L. Croft et al. 1996. Multiple independent origins of the COII/tRNA(Lys) intergenic 9-bp mtDNA deletion in aboriginal Australians. *Am. J. Hum Genet.* 58:428–433.

Brau, S. 1904. *Historia de Puerto Rico.* San Juan, PR: Editorial Coqui.

Brown, M.D., S.H. Hosseini, A. Torroni et al. 1998. MtDNA haplogroup X: An ancient link between Europe/Western Asia and North America? *Am. J. Hum. Genet.* 63:1852–1861.

Brown, W.M., M. George Jr., and A.C. Wilson. 1979. Rapid evolution of animal mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* 76:1967–1971.

Brown, W.M., E.M. Prager, A. Wang et al. 1982. Mitochondrial DNA sequences of primates: Tempo and mode of evolution. *J. Mol. Evol.* 18:225–239.

Calafell, F., P. Underhill, A. Tolun et al. 1996. From Asia to Europe: Mitochondrial DNA sequence variability in Bulgarians and Turks. *Ann. Hum. Genet.* 60:35–49.

Cann, R.L., and A.C. Wilson. 1983. Length mutations in human mitochondrial DNA. *Genetics* 104:699–711.

Carvajal-Carmona, L.G., I.D. Soto, N. Pineda et al. 2000. Strong Amerind/white sex bias and a possible Sephardic contribution among the founders of a population in northwest Colombia. *Am. J. Hum. Genet.* 67:1287–1295.

Carvalho-Silva, D.R., F.R. Santos, J. Roche et al. 2001. The phylogeography of Brazilian Y-chromosome lineages. *Am. J. Hum. Genet.* 68:281–286.

Chanlatte, L.A., and I.M. Narganes. 2001. La importancia del sitio La Hueca en la arqueología antillana. In *Proceedings of the Eighteenth International Congress for Caribbean Archaeology*, J. Haviser, ed. Curaçao, Netherland Antilles: The Jacob Gelt Dekker Institute for Advanced Cultural Studies (in press).

Chen, Y.S., A. Torroni, L. Excoffier et al. 1995. Analysis of mtDNA variation in African populations reveals the most ancient of all human continent-specific haplogroups. *Am. J. Hum. Genet.* 57:133–149.

Easton, R.D., D.A. Merriwether, D.E. Crews et al. 1996. MtDNA variation in the Yanomami: Evidence for additional New World founding lineages. *Am. J. Hum. Genet.* 59:213–225.

Fernández-Méndez, E. 1970. *Historia cultural de Puerto Rico.* San Juan, PR: Ediciones "El Cemí."

Forster, P., R. Harding, A. Torroni et al. 1996. Origin and evolution of Native American mtDNA variation: A reappraisal. *Am. J. Hum. Genet.* 59:935–945.

Giles, R.E., H. Blanc, H.M. Cann et al. 1980. Maternal inheritance of human mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* 77:6715–6719.

Graven, L., G. Passarino, O. Semino et al. 1995. Evolutionary correlation between control region sequence and restriction polymorphisms in the mitochondrial genome of a large Senegalese Mandenka sample. *Mol. Biol. Evol.* 12:334–345.

Hauswirth, W.W., C.D. Dickel, D.J. Rowold et al. 1994. Inter- and intrapopulation studies of ancient humans. *Experientia* 50:585–591.

Hutchison, C.A., J.E. Newbold, S.S. Potter et al. 1974. Maternal inheritance of mammalian mitochondrial DNA. *Nature* 251:536–538.

Imbelloni, J. 1938. Tabla clasificatoria de los indios: regiones biológicas y grupos raciales humanos de América. *Physis: Revista de la sociedad argentina de ciencias naturales* (Buenos Aires) 12:229– 249.

Kocher, T.D., and A.C. Wilson. 1991. Sequence evolution of mitochondrial DNA in humans and chimpanzees: control region and a protein-coding region. In *Evolution of life: Fossils, molecules, and culture*, S. Osawa and T. Honjo, eds. Tokyo: Springer, 391–413.

Kolman, C.J., and E. Bermingham. 1997. Mitochondrial and nuclear DNA diversity in the Chocó and Chibcha Amerinds of Panamá. *Genetics* 147:1289–1302.

Kolman, C.J., E. Bermingham, R. Cooke et al. 1995. Reduced mtDNA diversity in the Ngöbé Amerinds of Panamá. *Genetics* 140:275–283.

Lum, J.K., and R.L. Cann. 1998. mtDNA and language support a common origin of Micronesians and Polynesians in Island Southeast Asia. *Am. J. Phys. Anthropol.* 105:109–119.

Merriwether, D.A., and R.E. Ferrell. 1996. The four founding lineage hypothesis for the New World: A critical reevaluation. *Mol. Phylogenet. Evol.* 51:241–246.

Miyata, T., H. Hayashida, R. Kikuno et al. 1982. Molecular clock of silent substitution: At least six-fold preponderance of silent changes in mitochondrial genes over those in nuclear genes. *J. Mol. Evol.* 19:28–35.

Noble, G.K. 1965. *Proto–Arawakan and Its Descendants.* Bloomington, IN: Indiana University Publications in Anthropology and Linguistics, no. 38.

O'Rourke, D.H., M.G. Hayes, and S.W. Carlyle. 2000. Spatial and temporal stability of mtDNA haplogroup frequencies in native North America. *Hum. Biol.* 72:15–34.

Pääbo, S., J.A. Gifford, and A.C. Wilson. 1988. Mitochondrial DNA sequences from a 7000-year-old brain. *Nucleic Acids Res.* 16:9775–9787.

Passarino, G., O. Semino, G. Modiano et al. 1993. COII/tRNA^{Lys} intergenic 9-bp deletion and other mtDNA markers clearly reveal that Tharus of southern Nepal have Oriental affinities. *Am. J. Hum. Genet.* 53:609–618.

Pinto, F., A.M. González, M. Hernández et al. 1996. Genetic relationship between Canary Islanders and their African and Spanish ancestors inferred from mitochondrial DNA sequences. *Ann. Hum. Genet.* 60:321–330.

Redd, A.J., N. Takezaki, S.T. Sherry et al. 1995. Evolutionary history of the COII/tRNA^{Lys} intergenic 9 base pair deletion in human mitochondrial DNAs from the Pacific. *Mol. Biol. Evol.* 12:604–615.

Ribeiro-Dos-Santos, A.K.C, S.E.B. Santos, A.L. Machado et al. 1996. Heterogeneity of mitochondrial DNA haplotypes in pre-Columbian natives of the Amazon region. *Am. J. Phys. Anthropol.* 101:29–37.

Richards, M., H. Côrte-Real, P. Forster et al. 1996. Paleolithic and Neolithic lineages in the European mitochondrial gene pool. *Am. J. Hum. Genet.* 59:185–203.

Richards, M., V. Macaulay, E. Hickey et al. 2000. Tracing European founder lineages in the Near Eastern mtDNA pool. *Am. J. Hum Genet.* 67:1251–1276.

Rouse, I. 1992. *The Taínos: Rise and Decline of the People Who Greeted Columbus.* New Haven, CT: Yale University Press.

Sánchez-Crespo, A. (1999). Variación genética en una población indígena en Puerto Rico. M.Sc. thesis, University of Puerto Rico at Mayagüez, Mayagüez, Puerto Rico.

Sanger, J.P., H. Gannett, and W.F. Willcox. 1900. *1899 Census Report of Puerto Rico.* Washington, DC: Government Printing Office.

Santos, M., R.H. Ward, and R. Barrantes. 1994. MtDNA variation in the Chibcha Amerindian Huétar from Costa Rica. *Hum. Biol.* 66:963–977.

Schurr, T.G., S.W. Ballinger, Y.Y. Gan et al. 1990. Amerindian mitochondrial DNAs have rare Asian mutations at high frequencies, suggesting they derived from four primary maternal lineages. *Am J. Hum. Genet.* 46:613–623.

Shields, G.F., K. Hecker, M.I. Voevoda et al. 1992. Absence of the Asian-specific region V mitochondrial marker in native Beringians. *Am. J. Hum. Genet.* 50:758–765.

Smith, D.G., R.S. Malhi, J. Eshleman et al. 1999. Distribution of mtDNA haplogroup X among Native North Americans. *Am J. Phys. Anthropol.* 110:271–284.

Soodyall, H., L. Vigilant, A.V. Hill et al. 1996. mtDNA control-region sequence variation suggests multiple independent origins of an "Asian-specific" 9-bp deletion in Sub-Saharan Africans. *Am. J. Hum. Genet.* 58:595–608.

Starikovskaya, Y.B., R.I. Sukernik, T.G. Schurr et al. 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, A.C., and M. Stoneking. 1998. mtDNA analysis of a prehistoric Oneota population: implications for the peopling of the New World. *Am. J. Hum. Genet.* 62:1153–1170.

Tapia, A. 1854. *Biblioteca histórica de Puerto Rico.* Barcelona, Spain: Ediciones RVMBOS.

Torroni, A., H.J. Bandelt, L. D'Urbano et al. 1998. MtDNA analysis reveals a major late Paleolithic population expansion from southwestern to northeastern Europe. *Am. J. Hum. Genet.* 62:1137–1152.

Torroni, A., Y.S. Chen, O. Semino et al. 1994d. mtDNA and Y-chromosome polymorphisms in four Native American populations from southern Mexico. *Am. J. Hum. Genet.* 54:303–318.

Torroni, A., K. Houponen, P. Francalacci et al. 1996. Classification of European mtDNAs from an analysis of three European populations. *Genetics* 144:1835–1850.

Torroni, A., M.T. Lott, M.F. Cabell et al. 1994b. mtDNA and the origin of Caucasians: Identification of ancient Caucasian-specific haplogroups, one of which is prone to a recurrent somatic duplication in the D-loop region. *Am. J. Hum. Genet.* 55:760–776.

Torroni, A., J.A. Miller, L.G. Moore et al. 1994c. Mitochondrial DNA analysis in Tibet: Implications

for the origin of the Tibetan population and its adaptation to high altitude. *Am. J. Phys. Anthropol.* 93:189–199.

Torroni, A., J.V. Neel, R. Barrantes et al. 1994a. Mitochondrial DNA "clock" for the Amerinds with its implications for timing their entry into North America. *Proc. Natl. Acad. Sci. USA* 91:1158–1162.

Torroni, A., M. Petrozzi, L. D'Urbano et al. 1997. Haplotype and phylogenetic analyzes suggest that one European-specific mtDNA background plays a role in the expression of Leber Hereditary Optic Neuropathy by increasing the penetrance of the primary mutations 11778 and 14484. *Am. J. Hum. Genet.* 60:1107–1121.

Torroni, A., T.G. Schurr, M.F. Cabell et al. 1993a. Asian affinities and continental radiation of the four founding Native American mtDNAs. *Am. J. Hum. Genet.* 53:563–590.

Torroni, A., R.I. Sukernik, T.G. Schurr et al. 1993b. mtDNA variation of aboriginal Siberians reveals distinct genetic affinities with Native Americans. *Am. J. Hum. Genet.* 53:591–608.

Wakeley, J. 1993. Substitution rate variation among sites in hypervariable region 1 of human mitochondrial DNA. *J. Mol. Evol.* 37:613–623.

Wallace, D.C. 1995. Mitochondrial DNA variation in human evolution, degenerative disease, and aging. *Am. J. Human. Genet.* 57:201–223.

Walsh, P.S., D.A. Metzger, and R. Higuchi. 1991. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques* 10:506–513.

Ward, R.H., F.M. Salzano, S.L. Bonatto et al. 1996. Mitochondrial DNA polymorphism in three Brazilian Indian tribes. *Am. J. Hum. Biol.* 8:317–323.

Watkins, W.S., M. Bamshad, M.E. Dixon et al. 1999. Multiple origins of the mtDNA 9-bp deletion in populations of South India. *Am. J. Phys. Anthropol.* 109:147–158.

Zavala, S. 1948. *Estudios indianos.* México, D.F., Mexico: Publicaciones del Colegio de México.