ORIGINAL ARTICLE

A great diversity of Amerindian mitochondrial DNA ancestry is present in the Mexican mestizo population

Mariano Guardado-Estrada¹, Eligia Juarez-Torres², Ingrid Medina-Martinez¹, Ana Wegier³, Antonio Macías², Guillermo Gomez⁴, Fernando Cruz-Talonia⁴, Edgar Roman-Bassaure⁴, Daniel Piñero³, Susana Kofman-Alfaro¹ and Jaime Berumen^{1,2}

There are limited data on mitochondrial DNA (mtDNA) variation in the Mexican mestizo population. To examine the genetic diversity and matrilineal ancestry, the full mtDNA hypervariable regions I and II were sequenced in 270 unrelated mestizos from different regions of Mexico. A total of 202 different haplotypes were identified and the haplotype diversity was 0.9945. Amerindian haplotypes predominated in the sample with a proportion of 93.3%, followed by European (6.0%) and African haplotypes (0.7%). The frequency of the Amerindian haplogroups A2, B2, C1 and D1 was 51.1, 17.8, 18.5 and 5.9%, respectively. The frequency of Amerindian haplogroups was higher in the central region than in Mexico City, whereas it was the contrary for European haplogroups. This difference was accounted principally by the high frequency of B2 haplotypes in the central region. The minimum spanning network, the mismatch distribution and Tajima's *D* neutrality test suggest a population expansion for each Amerindian haplogroup, which could be initiated more recently for haplogroups A2 and D1. The present knowledge combined with other nuclear genetic markers will be essential in future association studies to correct for genetic substructure in mestizo populations.

Journal of Human Genetics (2009) 54, 695–705; doi:10.1038/jhg.2009.98; published online 16 October 2009

Keywords: Amerindian; control region; haplogroup; haplotype; mestizo; mtDNA; network

INTRODUCTION

A majority of the Mexican population is considered to be mestizo, which could be defined as a person who was born in Mexico, has a Spanish-derived last name and has a family of Mexican ancestors belonging to the third generation.¹ Mestizos originated from the admixture among the Spaniards, native Amerindians and Africans about 500 years ago.^{2,3} Mexico City and its metropolitan regions are located in the valley of Mexico, which is situated in central Mexico. This geographical area was the heart of the Aztec empire and the most populated region of Mesoamerica before the arrival of the Spaniards.⁴ In addition, a large amount of indigenous populations lived in the southern region of Mexico in well-established and populated Mesoamerican cities.^{5–7} In contrast, indigenous populations settled in the northern region of Mexico comprised few individuals who resided in small communities close to the rivers, and most of them were nomads and hunters.^{4,8} After the conquest was concluded, most Spaniards established in Mexico City and a large amount of the indigenous populations was displaced to small villages around the city in the valley of Mexico.⁵ During the following 300 years, colonial Mexico received a continuous flow of immigrants coming from Europe, in particular Spain, and to a lesser extent black slaves from Africa in the

sixteenth century, who were established essentially in the coastal region.⁹ In this period, the Spaniards expanded to the northern and southern regions of New Spain and founded the colonial cities, which were particularly located in areas that were rich in mineral reserves.¹⁰ According to historical data, estimation of admixture using classical nuclear genetic markers has shown that European, Amerindian and African ancestries vary across the geographical regions of the country.¹¹ Amerindian ancestry predominates in all regions, whereas the proportion of European and African ancestry is variable. There is a higher proportion of individuals with Caucasoid ancestry in Mexico City (35–71%) and the northern regions (35–55%) than in the central regions (21%) around Mexico City and the southern regions (4.2–43%) of Mexico, whereas the African ancestry predominates in the coastal regions (22–41%).¹² On the other hand, European ancestry, explored using mitochondrial DNA (mtDNA), was absent in the indigenous population^{13–16} and was very low in the mestizos either from the north (5.3%)¹⁷ or from Mexico City (3.2%).^{12,18} Although those Mexican mestizos come from modern populations, the fact that mtDNA is maternally inherited¹⁹ without recombination indicates that the mestizo mitochondrial genome has initially been inherited from indigenous women. This agrees with historical data because

¹Facultad de Medicina, UNAM, Mexico City, México; ²Unidad de Medicina Genómica, Hospital General de México, Mexico City, México; ³Departamento de Ecología Evolutiva, Instituto de Ecología, UNAM, Mexico City, México and ⁴Oncología, Servicio de Colposcopia, Hospital General de México, Mexico City, México
Correspondence: Dr J Berumen, Facultad de Medicina, Universidad Nacional Autónoma de México, Dr Balmis 148, Colonia Doctores, Delegación Cuauhtémoc, Mexico City, D.F. 06726, México.

E-mail: jaimeberumen@hotmail.com

Received 8 July 2009; accepted 15 September 2009; published online 16 October 2009

mating has been described essentially between Spaniard men and indigenous women.^{20,21} As mtDNA has a higher mutational rate than the nuclear genome,²² the former has been used as a tool in population genetics for making inferences and assumptions in ancestry and the migration patterns of human populations.²³ Furthermore, the Amerindian mtDNA has been well characterized and the haplogroups found in indigenous populations in the Americas were the first identified and named A, B, C, D and X.^{15,24,25} Those Amerindian haplogroups were recently termed A2, C1, B2 and D1, once there were more available data of the complete mtDNA genome sequence.^{26–28} As most studies that analyzed mtDNA polymorphism in the Mexican population used mainly restriction fragment length polymorphisms (RFLPs),^{13,17,18,29} and there are no available data from full control region sequences, the knowledge of mtDNA diversity is limited and it has not been extensively analyzed.

To describe the genetic variability in the mtDNA control region present in the Mestizo population, the hypervariable region I (HVR1) and hypervariable region II (HVR2) were analyzed in 270 unrelated Mexican mestizos. In addition, the mestizo population from Mexico City was compared with Mestizo populations from the central and coastal regions of Mexico. Besides the utility of the present knowledge for forensic purposes and population genetics, mtDNA in conjunction with other genetic markers could be helpful to estimate the individual ancestry proportion in admixed populations. These ancestry estimates will be useful to match adequately the cases and controls in genetic association studies and to avoid the bias caused by population stratification.³⁰

MATERIALS AND METHODS

Samples

Blood samples were collected from 270 unrelated healthy women attending the colposcopy screening program at the General Hospital of Mexico. All women were born in Mexico and had a Mexican ancestry for two generations. From the 270 individuals, 100 were born in Mexico City, 115 in the central, 47 in the

coastal (8 from the Gulf coastal and 39 from the Pacific coastal) and 8 in the northern regions of the country (Figure 1). All the women accepted to participate in the study and signed a written informative consent. The protocol and procedures in this study were reviewed and approved by the local ethics and scientific committee of the General Hospital of Mexico.

DNA extraction and sequencing

Whole genomic DNA was extracted using a commercial DNA extraction purification kit (Qiagen, Valencia, CA, USA) according to the manufacturer's protocol. The full control region (1261 bp) was amplified by PCR from 500 ng of whole genomic DNA in a 25 µl reaction containing 2 mM MgCl₂, 1× PCR buffer (50 mM KCl, 10 mM Tris-HCl and pH 8.3), 200 µM of each of dATP, dCTP, dGTP and dTTP, 1 µM of each primer and 1 U of Taq Platinum DNA polymerase (Invitrogen, Carlsbad, CA, USA). Reactions were carried out in a GeneAmp PCR System 2720 (Applied Biosystems, Carlsbad, CA, USA). After denaturing for 2 min at 94 °C, 32 cycles were performed by setting station 1 at 94 °C for 45 s, station 2 at 60 °C for 1 min and station 3 at 72 °C for 2 min. Finally, the PCR reaction was kept at 72 °C for another 7 min. DNA strands of both HVR1 (position 16024–16365) and HVR2 (position 73–340) were sequenced using the fluorescent cycle-sequencing method (BigDye Terminator Ready Reaction Kit, Applied Biosystems). Sequence analysis was performed using an ABI PRISM 3130 genetic analyzer system (Applied Biosystems). The sequence and position of the forward (F) and reverse (R) primers for PCR and sequencing are as follows:

L-15926 (5'-ATACACCAGTCTTGTAAC-3', position 15 907–15 926),
H-00580 (5'-TTGAGGAGGTAAGCTACATA-3', position 599–580),
H-16410 (5'-CACGGAGGATGGTGGTCAAG-3', position 16 414–16 395),
L-00015 (5'-CACCTATTAACCACTCACG-3', position 15–34) and
H-00383 (5'-AGGCTGGTGTAGGGTTCTT-3', position 383–364).

Data analysis

The mtDNA HVR1 and HVR2 sequences of each individual were compared with the revisited Cambridge Reference Sequence (rCSR)³¹ using the Applied Biosystems Seq Scape ver. 2.5 Software. The polymorphisms identified in each sample were stored in a Microsoft Office 2007 Access database to construct and identify the haplotypes, and calculate the allelic and haplotype frequencies.

Figure 1 Geographical location of individual birthplaces. Samples were grouped into those individuals who were born in Mexico City (black arrow; $n=100$), in the coastal (Veracruz, Michoacán, Guerrero, Oaxaca, Chiapas; |||| $n=47$), northern (Jalisco, San Luis Potosí, Zacatecas, Durango, Sinaloa; ● $n=8$) and central regions excluding Mexico City (Estado de México, Hidalgo, Puebla, Tlaxcala, Morelos, Querétaro, Guanajuato; ≡ $n=115$).

Calculations of the mean number of pairwise differences, nucleotide diversity, haplotype diversity indexes, mismatch distribution, Θ (τ) and Tajima's D neutrality test were performed using the population genetics software package Arlequin ver. 3.11.³² The motifs present in both regions that characterize the Amerindian,^{27,28,33} European^{34,35} and African³⁶ mitochondrial haplogroups were used to classify each sample by their haplogroup. The statistical parsimony algorithm was applied to construct a minimum spanning network with control region sequences using the TCS ver. 1.21 software.^{37,38} Positions not adjusting to an infinite sites model because of the presence of mutation heterogeneity in both regions (16 051, 16 093, 16 126, 16 129, 16 172, 16 187, 16 189, 16 223, 146, 150, 152, 182, 189, 195) were removed from the analysis to reduce the degree of homoplasy.³⁹ Although those sites were removed, not all the reticulations were eliminated in the network, and the remaining loops were solved using the algorithms based on the coalescent theory.⁴⁰ To evaluate the significance of differences in the proportion of mitochondrial haplogroups ancestry among the groups, a χ^2 test was performed. The significance of comparison of the mean pairwise differences and the nucleotide diversity was evaluated using the Student's t -test. All statistical tests were carried out using the Minitab statistical software.

RESULTS AND DISCUSSION

From 610 nucleotides analyzed, a total of 162 substitutions were found in 154 polymorphic positions (102 in HVR1 and 52 in HVR2). A total of 202 different haplotypes were identified in the whole sample ($n=270$). The number of polymorphic positions present in the haplotypes ranged from 1 to 16, with a mean of 8.9. In all, 168 haplotypes (83.1%) were identified only once and the most frequent haplotype (haplotype Mx7) was present in 15 (5.6%) individuals (see Table 1). It is noteworthy that most haplotypes (192; 95%) have not been reported earlier (Table 1).^{27,39,40–42} The haplotype diversity of the whole control region was 0.9945, with a mean number of pairwise differences of 11.14 and a nucleotide diversity of 0.0655 (see Table 2). Unlike HVR1, HVR2 had 15 sites that showed indels either as a single-point insertion (insertion at 180G) or as a part of polymorphic microsatellite (p248–249, p286–291, p303–309 and p311–315) and non-microsatellite (p106–111) tracts. In the poly-cytosine tract p303–309, only 26% of individuals showed seven cytosines like the rCSR, whereas most individuals (54.8%) had eight cytosines and 18.8% had nine cytosines (see Table 3). Interestingly, in the other microsatellite in 311–315, all but one individual had six cytosines, one more cytosine than in the rCSR. Both homopolymeric cytosine stretches are known to be highly variable;⁴³ however, in our sample only the microsatellite p303–309 was polymorphic. The frequency of alleles in the track p303–309 in our sample was similar to the frequencies reported in other populations.^{39,44–48} The deletion of two adenines in the poly-A track at positions p286–291 was always linked with the deletion of one adenine at position p248–249 and this combination was found in 17.1% of the individuals examined, whereas the remaining individuals (82.9%) had six adenines in the first and two adenines in the last positions similar to the rCSR. A deletion of six nucleotides from position 106 to 111, which has been described earlier in the Chibchan Amerindian population,⁴⁹ was found in four individuals. The insertion of guanine nucleotide in position 180 was found only in one individual.

Mitochondrial haplogroups

In the whole sample, 252 individuals (93.3%) were classified to one of the Amerindian haplogroups. The Amerindian haplogroups found were A2, B2, C1 and D1, with a frequency of 51.1, 17.8, 18.5 and 5.9%, respectively (Figure 2b). The frequency of the non-Amerindian haplogroups found in the sample was 6.7%, of which European haplogroups represented 5.9% (16 individuals) and the African

haplogroups 0.7% (2 individuals). The Amerindian ancestry proportion found in this sample is slightly higher but not significantly different from that earlier reported in a study carried out in the mestizo population from Mexico City (90.5%), where haplogroups were identified by RFLPs.¹⁸ In comparison with that study, which included persons with social security, individuals enrolled in this study included women who attended the Pap screening program at the General Hospital of Mexico, which regularly attends people with lower income lacking social security,⁵⁰ and it has been suggested earlier that the proportion of Amerindian ancestry could be higher in people coming from low socioeconomic strata than in those with a higher-level income in the Mexican mestizo population.^{11,51}

Amerindian mitochondrial haplogroups

Haplotype A2. The polymorphisms that characterize this haplogroup are the transitions at positions C16111T, T16223C, C16290T, G16319A, T16362C, T146C, A153G and A235G.^{27,28} In addition to these characteristic polymorphisms, 50 haplotypes, including the haplotype Mx7, showed the transitions A73G and A263G (Table 1). As most reported studies analyzed only the HVR1, the positions of this region were compared with reported sequences. HVR1 polymorphisms found in haplotype Mx7 (11.1% of all haplotypes when only HVR1 were considered) have been reported earlier in other studies but in lower proportions (5–8%).^{39–42} From 93 haplotypes that composed the haplogroup A2, 77 had a thymine at position 16111 and only 16 had the cytosine similar to the rCSR. Before the classification suggested by Bandelt,²⁶ haplogroup A was split into A1 and A2 by the transition at this position.³³ However, further studies reported that the C allele at position 16111 in the A2 haplotypes was instead a back mutation to the rCSR from T back to the C allele, because of the high mutation rate present in the mtDNA control region.²⁸ Transition at position C16111T is characteristic of the Amerindian A2 haplogroup because it is rarely found in Asian haplogroup A sequences.^{52,53} Furthermore, this C16111T polymorphism has rarely been seen in other haplogroups.^{54,55} Substitutions A153G and A235G were highly frequent in the haplogroup A2, with a frequency of 88.4 and 94.2%, respectively. Although these positions characterize the A2 haplogroup, they were rarely found in other Amerindian (C and B) and non-Amerindian haplogroups (H) in this sample. Interestingly, the transition T152C was found only in 17.8% of the A2 haplogroup sequences, whereas in Asian haplogroup A sequences it was found in 16 out of 18 sequences.⁵⁶

Haplotype B2. This haplogroup was present in 48 of the 270 individuals analyzed and it was the third most common of the Amerindian haplogroups found in this sample. The HVR1 motifs that are characteristic of haplogroup B2 are the transitions T16189C and T16217C.^{27,52} However, other positions associated with B2 haplotypes are the transversions A16182C and A16183C.^{28,39} The haplotype Mx101 was the most frequent from haplogroup B2 (four individuals) and included also the polymorphisms A16182C, A16183C, T16189C, T16217C, A73G and A263G (Table 1). Considering only the positions at HVR1, this haplotype has been reported earlier once in Latin America.^{40–42} Interestingly, the polymorphism A16183C was identified in all haplogroup B2 sequences. The haplogroup B2 in this study was further divided in two subgroups, based on transversion A16182C, which was found in 54.1% of the B2 haplotypes. These data highly contrast with an earlier study in the Mexican mestizo population,¹⁷ in which transversions A16182C and A16183C were not found in 33 individuals that had the HVR1 region analyzed. On the other hand, transversion 16183 has already been

Table 1 Haplotypes found in the 270 individuals

Hap Id ^a	HG ^b	n ^c	Hypervariable region I	Hypervariable region II
HG A2 motifs			16111T, 16223C, 16290T, 16319A, 16362C	146C, 153G, 235G
Mz39	A2	1	16051G, 16111T, 16223T, 16290T, 16319A, 16362C	73G, 146C, 152C, 153G, 235G, 263G
Mz15	A2	1	16069T, 16111T, 16223T, 16290T, 16319A, 16362C	73G, 146C, 153G, 235G, 263G
Mz85	A2	1	16069T, 16111T, 16223T, 16290T, 16319A, 16362C	73G, 263G
Mz75	A2	1	16075C, 16111T, 16172C, 16223T, 16239T, 16286T, 16290T, 16319A, 16362C	73G, 146C, 152C, 153G, 235G, 263G
Mz44	A2	1	16086C, 16111T, 16209C, 16223T, 16290T, 16319A, 16362C	73G, 146C, 153G, 235G, 263G
Mz18	A2	1	16086C, 16111T, 16223T, 16290T, 16319A, 16362C	73G, 146C, 153G, 235G, 263G
Mz76	A2	2	16092C, 16111T, 16136C, 16223T, 16257T, 16290T, 16311C, 16319A, 16344T, 16362C	73G, 146C, 153G, 235G, 263G
Mz1	A2	1	16092C, 16111T, 16223T, 16290T, 16319A, 16362C	73G, 153G, 235G, 263G
Mz45	A2	1	16092C, 16153A, 16223T, 16240G, 16290T, 16319A, 16362C	73G, 146C, 153G, 235G, 263G
Mz46	A2	1	16092C, 16168T, 16223T, 16290T, 16296T, 16319A, 16362C	73G, 146C, 153G, 235G, 263G
Mz47	A2	1	16093C, 16111T, 16129A, 16223T, 16290T, 16319A, 16362C	73G, 146C, 153G, 235G, 263G
Mz70	A2	1	16093C, 16111T, 16136C, 16223T, 16290T, 16311C, 16319A, 16362C	73G, 146C, 152C, 153G, 235G, 263G
Mz48	A2	1	16093C, 16111T, 16136C, 16290T, 16311C, 16319A, 16362C	73G, 146C, 153G, 235G, 263G
Mz49	A2	1	16093C, 16136C, 16223T, 16290T, 16311C, 16319A, 16362C	73G, 146C, 153G, 235G, 263G
Mz19	A2	1	16095T, 16111T, 16223T, 16290T, 16319A, 16362C	73G, 146C, 153G, 235G, 263G
Mz66	A2	1	16104A, 16111T, 16189C, 16192T, 16223T, 16290T, 16319A, 16362C	73G, 146C, 153G, 235G, 263G
Mz20	A2	1	16111T, 16124C, 16175G, 16223T, 16290T, 16319A, 16362C	73G, 146C, 235G, 263G
Mz50	A2	1	16111T, 16124C, 16223T, 16290T, 16319A, 16362C	73G, 146C, 151T, 153G, 235G, 263G
Mz71	A2	1	16111T, 16129A, 16189C, 16223T, 16290T, 16311C, 16319A, 16362C	73G, 146C, 153G, 215G, 235G, 263G
Mz51	A2	1	16111T, 16129A, 16213A, 16223T, 16290T, 16319A, 16362C	73G, 146C, 153G, 235G, 263G
Mz52	A2	1	16111T, 16129A, 16223T, 16234T, 16290T, 16319A, 16362C	73G, 146C, 153G, 235G, 263G
Mz67	A2	1	16111T, 16129A, 16223T, 16260T, 16290T, 16319A, 16362C	73G, 146C, 151T, 153G, 235G, 263G
Mz21	A2	3	16111T, 16129A, 16223T, 16290T, 16319A, 16362C	73G, 146C, 153G, 235G, 263G
Mz43	A2	1	16111T, 16129A, 16223T, 16290T, 16319A, 16362C	73G, 146C, 153G, 200G, 235G, 263G
Mz53	A2	1	16111T, 16129A, 16223T, 16290T, 16319A, 16362C	73G, 146C, 152C, 153G, 235G, 263G
Mz68	A2	1	16111T, 16131C, 16134T, 16223T, 16290T, 16319A, 16362C	73G, 75A, 146C, 153G, 235G, 263G
Mz77	A2	1	16111T, 16136C, 16172C, 16223T, 16247G, 16257T, 16274A, 16290T, 16319A, 16344T	73G, 146C, 153G, 235G, 263G
Mz78	A2	1	16111T, 16136C, 16223T, 16247G, 16257T, 16274A, 16290T, 16319A, 16344T, 16362C	73G, 146C, 153G, 235G, 263G
Mz79	A2	1	16111T, 16136C, 16223T, 16290T, 16299G, 16311C, 16319A, 16362C	73G, 146C, 153G, 198T, 215G, 235G, 263G
Mz54	A2	1	16111T, 16136C, 16223T, 16290T, 16319A, 16324C, 16362C	73G, 146C, 153G, 235G, 263G
Mz22	A2	2	16111T, 16136C, 16223T, 16290T, 16319A, 16362C	73G, 146C, 153G, 235G, 263G
Mz23	A2	2	16111T, 16172C, 16223T, 16290T, 16319A, 16362C	73G, 146C, 153G, 235G, 263G
Mz72	A2	1	16111T, 16176T, 16223T, 16239T, 16290T, 16298C, 16319A, 16362C	73G, 146C, 152C, 153G, 235G, 263G
Mz82	A2	1	16111T, 16183C, 16189C, 16217C, 16223T, 16239T, 16290T, 16319A, 16362C	73G, 146C, 153G, 226C, 228A, 235G, 263G
Mz80	A2	1	16111T, 16183C, 16189C, 16223T, 16239T, 16290T, 16319A, 16362C	73G, 146C, 153G, 226C, 228A, 235G, 263G
Mz55	A2	1	16111T, 16187T, 16223T, 16290T, 16319A, 16362C	73G, 146C, 152C, 153G, 235G, 263G
Mz73	A2	1	16111T, 16189C, 16223T, 16239T, 16290T, 16319A, 16325C, 16362C	73G, 146C, 152C, 153G, 235G, 263G
Mz69	A2	1	16111T, 16223T, 16234T, 16239T, 16290T, 16319A, 16362C	73G, 146C, 152C, 153G, 238G, 263G
Mz56	A2	1	16111T, 16223T, 16234T, 16290T, 16319A, 16360T, 16362C	97A, 146C, 153G, 235G, 263G, del 106-111
Mz24	A2	2	16111T, 16223T, 16234T, 16290T, 16319A, 16362C	73G, 146C, 153G, 235G, 263G
Mz86	A2	1	16111T, 16223T, 16234T, 16290T, 16319A, 16362C	73G, 249D, 263G, 290-291D
Mz25	A2	1	16111T, 16223T, 16239T, 16290T, 16319A, 16362C	73G, 146C, 153G, 235G, 263G
Mz81	A2	1	16111T, 16223T, 16239T, 16290T, 16319A, 16362C	73G, 146C, 151T, 152C, 153G, 215G, 235G, 263G, 335G
Mz26	A2	1	16111T, 16223T, 16239T, 16319A, 16362C	73G, 146C, 152C, 153G, 235G, 263G
Mz57	A2	1	16111T, 16223T, 16258G, 16290T, 16319A, 16362C	73G, 103A, 146C, 153G, 235G, 263G
Mz58	A2	1	16111T, 16223T, 16265C, 16290T, 16319A, 16330C	73G, 103A, 146C, 153G, 235G, 263G
Mz83	A2	1	16111T, 16223T, 16265G, 16290T, 16293G, 16311C, 16319A, 16362C	73G, 90C, 98T, 146C, 153G, 195C, 235G, 263G
Mz59	A2	2	16111T, 16223T, 16274A, 16290T, 16311C, 16319A, 16362C	73G, 146C, 153G, 235G, 263G
Mz2	A2	1	16111T, 16223T, 16274A, 16290T, 16319A, 16362C	73G, 146C, 153G, 263G
Mz27	A2	1	16111T, 16223T, 16274A, 16290T, 16319A, 16362C	73G, 146C, 153G, 235G, 263G
Mz60	A2	1	16111T, 16223T, 16274A, 16290T, 16319A, 16362C	73G, 146C, 153G, 182T, 235G, 263G
Mz61	A2	3	16111T, 16223T, 16284G, 16290T, 16319A, 16362C	73G, 146C, 152C, 153G, 235G, 263G
Mz87	A2	1	16111T, 16223T, 16284G, 16290T, 16319A, 16362C	73G, 249D, 263G, 290-291D
Mz28	A2	1	16111T, 16223T, 16290T, 16311C, 16319A, 16362C	73G, 146C, 153G, 235G, 263G
Mz62	A2	1	16111T, 16223T, 16290T, 16311C, 16319A, 16362C	73G, 146C, 153G, 199C, 235G, 263G

Table 1 Continued

Hap Id ^a	HG ^b	n ^c	Hypervariable region I	Hypervariable region II
Mz35	A2	1	16111T, 16223T, 16290T, 16319A	73G, 146C, 152C, 153G, 195C, 235G, 263G
Mz9	A2	2	16111T, 16223T, 16290T, 16319A	73G, 146C, 152C, 153G, 235G, 263G
Mz63	A2	2	16111T, 16223T, 16290T, 16319A, 16325C, 16362C	73G, 146C, 152C, 153G, 235G, 263G
Mz3	A2	1	16111T, 16223T, 16290T, 16319A, 16335G	73G, 146C, 153G, 235G, 263G
Mz30	A2	1	16111T, 16223T, 16290T, 16319A, 16335G	73G, 146C, 150T, 153G, 235G, 263G
Mz31	A2	1	16111T, 16223T, 16290T, 16319A, 16335G	73G, 146C, 152C, 153G, 235G, 263G
Mz29	A2	2	16111T, 16223T, 16290T, 16319A, 16335G, 16362C	73G, 146C, 153G, 263G, 310C
Mz4	A2	1	16111T, 16223T, 16290T, 16319A, 16344T, 16362C	73G, 146C, 235G, 263G
Mz64	A2	1	16111T, 16223T, 16290T, 16319A, 16357C, 16362C	73G, 146C, 153G, 235G, 263G, 295T
Mz5	A2	2	16111T, 16223T, 16290T, 16319A, 16360T, 16362C	146C, 153G, 235G, 263G, del 106–111
Mz14	A2	1	16111T, 16223T, 16290T, 16319A, 16362C	73G, 93G, 146C, 235G, 263G
Mz32	A2	6	16111T, 16223T, 16290T, 16319A, 16362C	73G, 103A, 146C, 153G, 235G, 263G
Mz33	A2	1	16111T, 16223T, 16290T, 16319A, 16362C	73G, 143A, 152C, 153G, 185A, 263G
Mz34	A2	1	16111T, 16223T, 16290T, 16319A, 16362C	73G, 146C, 152C, 153G, 235G, 263G
Mz6	A2	1	16111T, 16223T, 16290T, 16319A, 16362C	73G, 103A, 146C, 235G, 263G
Mz65	A2	1	16111T, 16223T, 16290T, 16319A, 16362C	73G, 146C, 152C, 153G, 195C, 235G, 263G
Mz7	A2	15	16111T, 16223T, 16290T, 16319A, 16362C	73G, 146C, 153G, 235G, 263G
Mz8	A2	1	16111T, 16223T, 16290T, 16319A, 16362C	97A, 146C, 153G, 235G, 263G, del 106–111
Mz90	A2	3	16111T, 16223T, 16290T, 16319A, 16362C	73G, 146C, 235G, 263G
Mz10	A2	1	16111T, 16223T, 16290T, 16362C	73G, 146C, 153G, 193G, 235G, 263G
Mz88	A2	1	16111T, 16223T, 16290T, 16362C	73G, 146C, 235G, 263G
Mz91	A2	2	16111T, 16223T, 16290T, 16362C	73G, 146C, 153G, 235G, 263G
Mz92	A2	1	16111T, 16241G, 16290T, 16362C	73G, 146C, 153G, 235G, 263G
Mz84	A2	1	16111T, 16290T, 16319A	73G, 146C, 235G, 263G
Mz89	A2	4	16111T, 16290T, 16319A, 16362C	73G, 146C, 235G, 263G
Mz36	A2	1	16129A, 16220C, 16223T, 16290T, 16319A, 16362C	73G, 146C, 153G, 235G, 263G
Mz11	A2	1	16136C, 16223T, 16290T, 16319A, 16362C	73G, 146C, 153G, 235G, 263G
Mz12	A2	1	16137G, 16223T, 16290T, 16319A, 16362C	73G, 146C, 153G, 235G, 263G
Mz74	A2	1	16153A, 16223T, 16240G, 16263C, 16266T, 16290T, 16319A, 16362C	73G, 146C, 153G, 225T, 235G, 263G
Mz37	A2	2	16153A, 16223T, 16240G, 16290T, 16319A, 16362C	73G, 146C, 153G, 235G, 263G
Mz40	A2	1	16167T, 16223T, 16290T, 16311C, 16319A, 16352C, 16362C	73G, 146C, 153G, 235G, 263G
Mz41	A2	1	16187T, 16223T, 16232T, 16290T, 16319A, 16362C	73G, 146C, 152C, 153G, 235G, 263G
Mz42	A2	1	16187T, 16223T, 16290T, 16319A, 16362C	73G, 146C, 152C, 153G, 235G, 263G, 320T
Mz38	A2	1	16213A, 16223T, 16290T, 16319A, 16325C, 16362C	73G, 146C, 153G, 235G, 263G
Mz16	A2	1	16223T, 16240G, 16290T, 16319A, 16362C	73G, 146C, 153G, 195C, 235G, 263G
Mz17	A2	1	16223T, 16270T, 16290T, 16319A, 16362C	73G, 146C, 153G, 214G, 235G, 263G
Mz13	A2	1	16223T, 16290T, 16319A, 16362C	73G, 146C, 153G, 200G, 235G, 263G
Mz93	A2	7	16223T, 16290T, 16319A, 16362C	73G, 146C, 153G, 235G, 263G
HG B2 motifs			16182C, 16183C, 16189C, 16217C	73G, 263G
Mz122	B2	1	16092C, 16111T, 16183C, 16189C, 16217C, 16325C	73G, 228A, 263G
Mz98	B2	1	16092C, 16182C, 16183C, 16189C, 16217C, 16335G, 16362C	73G, 94A, 263G, 324G
Mz111	B2	1	16092C, 16182C, 16183C, 16189C, 16217C, 16344T	73G, 263G
Mz112	B2	1	16093C, 16148T, 16183C, 16185T, 16217C	73G, 146C, 263G
Mz105	B2	1	16104T, 16183C, 16189C, 16217C, 16362C	73G, 263G
Mz94	B2	2	16111T, 16182C, 16183C, 16189C, 16217C, 16294T	73G, 93G, 228A 263G
Mz95	B2	1	16129A, 16145A, 16183C, 16187T, 16189C, 16217C, 16323C	73G, 195C, 263G
Mz96	B2	2	16140C, 16183C, 16189C, 16217C, 16270G, 16278T	73G, 146C, 248G, 263G
Mz106	B2	1	16144C, 16182C, 16183C, 16189C, 16217C	73G, 263G
Mz100	B2	2	16164G, 16183C, 16189C, 16217C	73G, 263G
Mz113	B2	1	16172C, 16182C, 16183C, 16184T, 16189C, 16217C	73G, 263G
Mz114	B2	1	16173T, 16183C, 16189C, 16217C, 16311C, 16352C	73G, 263G
Mz107	B2	1	16179A, 16183C, 16189C, 16217C, 16278T	73G, 263G
Mz97	B2	1	16181C, 16182C, 16183C, 16189C, 16217C, 16362C	73G, 146C, 182T, 263G
Mz101	B2	4	16182C, 16183C, 16189C, 16217C	73G, 263G
Mz108	B2	1	16182C, 16183C, 16189C, 16217C	73G, 263G, 320T
Mz119	B2	1	16182C, 16183C, 16189C, 16217C	73G, 150T, 263G, 308T
Mz128	B2	1	16182C, 16183C, 16189C, 16217C	73G, 146C, 153G, 235G, 263G
Mz115	B2	1	16182C, 16183C, 16189C, 16217C, 16250G	73G, 182T, 263G
Mz123	B2	1	16182C, 16183C, 16189C, 16217C, 16259T, 16357C, 16362C	73G, 263G

Table 1 Continued

Hap Id ^a	HG ^b	n ^c	Hypervariable region I	Hypervariable region II
Mz124	B2	1	16182C, 16183C, 16189C, 16217C, 16270G, 16278T	73G, 146C, 263G
Mz116	B2	2	16182C, 16183C, 16189C, 16217C, 16278T	73G, 146C, 263G
Mz117	B2	1	16182C, 16183C, 16189C, 16217C, 16291T	73G, 195C, 263G
Mz125	B2	2	16182C, 16183C, 16189C, 16217C, 16295T	73G, 146C, 152C, 263G
Mz126	B2	1	16182C, 16183C, 16189C, 16217C, 16295T	73G, 152C, 263G, 310C
Mz118	B2	2	16182C, 16183C, 16189C, 16217C, 16319A	73G, 146C, 263G
Mz127	B2	1	16182C, 16183C, 16189C, 16217C, 16342C	73G, 146C, 152C, 263G
Mz120	B2	1	16183C, 16189C, 16200G, 16217C, 16219C	73G, 146C, 263G
Mz104	B2	1	16183C, 16189C, 16217C	73G, 114G, 263G
Mz99	B2	1	16183C, 16189C, 16217C	73G, 263G
Mz121	B2	1	16183C, 16189C, 16217C, 16232A, 16278T, 16362C	73G, 263G
Mz129	B2	1	16183C, 16189C, 16217C, 16259T, 16357C	73G, 146C, 152C, 263G
Mz130	B2	1	16183C, 16189C, 16217C, 16270G, 16278T	73G, 143A, 146C, 263G
Mz102	B2	2	16183C, 16189C, 16217C, 16278T	73G, 263G
Mz109	B2	1	16183C, 16189C, 16217C, 16278T	73G, 234G, 263G
Mz103	B2	1	16183C, 16189C, 16217C, 16316G	73G, 263G
Mz110	B2	2	16183C, 16189C, 16217C, 16323C	73G, 263G, 267C
HG C1 motifs			16223T, 16298C, 16325C, 16327T	73G, 263G
Mz131	C1	1	16051G, 16093C, 16188T, 16204A, 16223T, 16325C, 16327T, 16362C	73G, 249D, 263G, 290–291D
Mz162	C1	1	16051G, 16094C, 16223T, 16298C, 16311C, 16325C, 16327T	73G, 263G, 290–291D
Mz132	C1	1	16051G, 16188T, 16204A, 16223T, 16325C, 16327T, 16362C	73G, 200G, 263G, 290–291D
Mz161	C1	1	16051G, 16188T, 16204A, 16223T, 16325C, 16327T, 16362C	73G, 263G
Mz164	C1	2	16051G, 16188T, 16204A, 16223T, 16325C, 16327T, 16362C	73G, 249D, 263G, 290–291D
Mz142	C1	1	16051G, 16188T, 16223T, 16325C, 16327T, 16362C	73G, 146C, 152C, 153G, 235G, 263G
Mz154	C1	1	16051G, 16188T, 16223T, 16325C, 16327T, 16362C	73G, 249D, 263G, 290–291D
Mz133	C1	1	16051G, 16223T, 16298C, 16311C, 16325C, 16327T	73G, 146C, 152C, 263G
Mz163	C1	1	16051G, 16223T, 16298C, 16311C, 16325C, 16327T	73G, 195C, 249D, 263G, 290–291D
Mz146	C1	1	16051G, 16223T, 16298C, 16325C, 16327T	73G, 249D, 263G, 290–291D
Mz155	C1	1	16051G, 16223T, 16298C, 16325C, 16327T	73G, 249D, 263G, 310C, 290–291D
Mz165	C1	1	16051G, 16223T, 16298C, 16325C, 16327T	73G, 150T, 194T, 249D, 263G, 290–291D
Mz134	C1	1	16075C, 16223T, 16298C, 16325C, 16327T	73G, 150T, 199C, 204C, 249D, 263G, 290–291D
Mz166	C1	1	16086C, 16172C, 16181G, 16223T, 16298C, 16325C, 16327T	73G, 249D, 263G, 290–291D
Mz135	C1	1	16086C, 16223T, 16295T, 16298C, 16319A, 16325C, 16327T	73G, 194T, 249D, 263G, 290–291D
Mz139	C1	1	16086C, 16223T, 16295T, 16298C, 16319A, 16325C, 16327T, 16362C	73G, 194T, 249D, 263G, 290–291D
Mz167	C1	1	16126C, 16259T, 16298C, 16311C, 16325C, 16327T	73G, 75A, 249D, 263G, 290–291D
Mz140	C1	3	16129A, 16172C, 16223T, 16298C, 16311C, 16325C, 16327T	73G, 146C, 152C, 249D, 263G, 290–291D
Mz168	C1	1	16129A, 16223T, 16224C, 16234T, 16298C, 16325C, 16327T	73G 249D 263G 290–291D
Mz147	C1	1	16153A, 16223T, 16298C, 16325C, 16327T	73G, 249D, 263G, 290–291D
Mz138	C1	1	16172C, 16223T, 16298C, 16325C, 16327T	73G, 143A, 195C, 249D, 260A, 263G, 290–291D
Mz148	C1	1	16179T, 16223T, 16298C, 16325C, 16327T	73G 249D 263G 290–291D
Mz136	C1	1	16181G, 16223T, 16261T, 16292T, 16298C, 16325C, 16327T	73G, 234G, 249D, 263G, 290–291D
Mz169	C1	1	16181G, 16223T, 16292T, 16298C, 16325C, 16327T	73G, 234G, 249D, 263G, 290–291D
Mz156	C1	1	16181G, 16223T, 16298C, 16325C, 16327T	73G, 151T, 249D, 263G, 290–291D
Mz141	C1	1	16186T, 16223T, 16293G, 16298C, 16325C, 16327T	73G, 146C, 174T, 199C, 249D, 263G, 290–291D
Mz170	C1	1	16223T, 16294T, 16298C, 16325C, 16327T	73G, 203A, 204C, 249D, 263G, 290–291D
Mz157	C1	1	16223T, 16298C, 16301T, 16325C, 16327T	73G, 215G, 249D, 263G, 290–291D
Mz149	C1	1	16223T, 16298C, 16311C, 16325C, 16327T	73G, 249D, 263G, 290–291D
Mz137	C1	1	16223T, 16298C, 16325C, 16327T	73G, 103A, 146C, 153G, 235G, 263G
Mz144	C1	5	16223T, 16298C, 16325C, 16327T	73G, 249D, 263G, 290–291D
Mz150	C1	1	16223T, 16298C, 16325C, 16327T	150T, 239C, 263G
Mz152	C1	1	16223T, 16298C, 16325C, 16327T	73G, 185A, 249D, 263G, 290–291D
Mz153	C1	2	16223T, 16298C, 16325C, 16327T	73G, 215G, 249D, 263G, 290–291D
Mz159	C1	1	16223T, 16298C, 16325C, 16327T	73G, 146C, 153G, 263G
Mz160	C1	1	16223T, 16298C, 16325C, 16327T	73G, 203A, 204C, 249D, 263G, 290–291D
Mz171	C1	1	16223T, 16298C, 16325C, 16327T	73G, 94A, 150T, 249D, 263G, 309T, 290–291D
Mz151	C1	1	16223T, 16298C, 16325C, 16327T, 16354T	73G, 249D, 263G, 290–291D
Mz158	C1	1	16223T, 16298C, 16325C, 16327T, 16362C	73G, 215G, 249D, 263G, 290–291D
Mz143	C1	2	16223T, 16325C, 16327T	73G 249D 263G 290–291D
Mz145	C1	1	16298C, 16325C, 16327T	73G, 215G, 249D, 263G, 290–291D

Table 1 Continued

Hap Id ^a	HG ^b	n ^c	Hypervariable region I	Hypervariable region II
HG D1 motifs			16223T, 16325C, 16362C	73G, 263G
Mz175	D1	1	16093C, 16223T, 16325C, 16362C	73G, 263G
Mz181	D1	1	16093C, 16223T, 16325C, 16362C	73G, 146C, 263G
Mz172	D1	1	16129A, 16142T, 16192T, 16223T, 16264T, 16265G, 16325C, 16362C	73G, 146C, 263G, 324A
Mz183	D1	1	16142T, 16223T, 16274A, 16325C, 16362C	73G, 146C, 263G
Mz182	D1	1	16189C, 16223T, 16325C, 16362C	73G, 263G, 324G
Mz173	D1	1	16223T, 16274A, 16325C	73G, 263G
Mz176	D1	1	16223T, 16274A, 16325C, 16362C	73G, 263G
Mz177	D1	1	16223T, 16292T, 16325C, 16362C	73G, 263G
Mz180	D1	1	16223T, 16325C	73G, 150T, 185A, 263G
Mz174	D1	3	16223T, 16325C, 16362C	73G, 263G
Mz178	D1	1	16223T, 16325C, 16362C	73G, 200G, 263G
Mz179	D1	2	16223T, 16325C, 16362C	73G, 211G, 263G
Mz184	D1	1	16223T, 16325C, 16362C	73G, 147A, 188G, 234G, 263G
European haplogroups				
Mz185	H	1		263G
Mz188	H	1	16051G, 16129C, 16189C, 16256T	73G, 146C, 263G
Mz186	H	1	16240G	152C, 263G
Mz187	H	1	16362C	73G, 263G
Mz189	H	1	16362C	73G, 103A, 146C, 153G, 235G, 263G
Mz190	J	1	16069T, 16126C, 16215T	73G, 185A, 188G, 228A, 263G, 295T
Mz191	J	1	16069T, 16126C, 16294T	73G, 152C, 185A, 228A, 263G, 295T
Mz193	K	1	16209C, 16223T, 16311C	73G, 150T, 189G, 200G, 263G
Mz192	K	1	16213A, 16224C, 16311C	73G, 146C, 263G
Mz197	T	1	16126C, 16292T, 16294T	73G, 146C, 152C, 252C, 263G, 279C
Mz196	T	1	16126C, 16294T	73G, 263G
Mz198	U	1	16189C, 16192T, 16223T, 16245T, 16278T, 16294T, 16309G	73G, 143A, 146C, 152C, 195C, 263G
Mz201	U	1	16189C, 16192T, 16270T, 16319A	73G, 150T, 263G
Mz199	U	1	16192T	93G, 95C, 263G
Mz200	U	1	16192T, 16270T, 16319A	73G, 263G
Mz202	V	1	16093C, 16298C	263G
African haplogroups				
Mz195	L	1	16051G, 16223T, 16278T, 16318G, 16362C	73G, 263G
Mz194	L	1	16124C, 16223T, 16278T, 16362C	73G, 263G

^aHaplotype Id.^bHaplogroup.^cNumber of samples.Characteristic motifs of each Amerindian haplogroup taken from Tamm *et al.*²⁷ are shown in bold.

reported in all B haplotypes analyzed in Latin America and almost all from Asia, whereas transversion A16182C was absent in the Argentinian population⁴² and it was found in low frequency in Colombia and Brazil (25%).^{40–42,53} The polymorphism A16182C is widely distributed among Asian B haplogroups, mainly in the subhaplogroups B4a and B4b.^{53,55,56} In fact, it has been suggested that the Amerindian haplogroup B2 descended from the haplogroup B4b.^{26,57} According to the updated phylogeny, B haplotypes having the transversion A16182C belong to subhaplogroup B2c.²⁸

Haplogroup C1. Haplogroup C1 was present in 50 (18.5%) individuals of the total sample and was the second most common haplogroup. Earlier reports have shown that the frequency of haplogroup C1 varies across distinct native American ethnic groups in Mexico, having a frequency similar to haplogroup B2.^{16,29} C1 haplogroup is defined by polymorphisms C16223T, T16298C, T16325C, C16327T, A73G, A263G, and the adenine deletions at positions 249,

290 and 291 of HVR2.^{26–28} The haplotype Mx144 had the characteristic polymorphisms of C1 haplogroup and it was the most commonly detected (five individuals; Table 1). Considering only the HVR1, this haplotype has been reported earlier in populations from Mexico and Latin America.^{17,40} Although the adenine deletions in the poly-A tracts of HVR2 are characteristic of haplogroup C, they had been rarely reported in other haplogroups.^{40,46} For instance, in Asian populations these deletions have been found in the haplogroups A, F and R9 instead.^{53,57} In contrast, in the haplogroups C4c and L2b the deletion of one adenine at position 248–249 is not accompanied by deletion of two adenines at positions 290–291.^{27,40} Interestingly, the transition A16051G, which has been associated with haplogroup C1d, was found in 32% (16 individuals) of the C1 haplotypes in this sample.

Haplogroup D1. This haplogroup was the least frequent Amerindian haplogroup and it was found in only 16 (5.9%) individuals. The HVR1 polymorphisms characteristic of this haplogroup are transitions

Table 2 Genetic diversity indexes calculated for the whole sample and Amerindian and European haplogroups

Summary statistics	Total samples	Amerindians				European
		A2	B2	C1	D1	
<i>n</i>	270	138	48	50	16	16
No. substitutions	162	94	59	55	20	43
Polymorphic sites	154	93	59	55	19	43
Transitions	143	88	50	55	18	40
Transversions	19	6	9	0	2	3
Haplotypes	202	93	37	41	13	16
Mean number of pairwise differences	11.1 ± 0.13	*4.7 ± 0.12	5.8 ± 0.24	5.9 ± 0.23	3.7 ± 0.35	8.7 ± 0.51
Nucleotide diversity	0.065 ± 0.11	*0.027 ± 0.01	0.034 ± 0.01	0.034 ± 0.01	0.021 ± 0.02	0.051 ± 0.04
Tajima's D	−1.82**	−2.37**	−1.96**	−1.95**	−1.42**	−1.37**
Haplotype diversity	0.9945	0.9823	0.9876	0.9875	0.9667	1

P* < 0.01, Student's *t*-test; comparing A2 vs B2, A2 vs C1, A2 vs B2.Tajima *D* test, *P* < 0.01.**Table 3** Allele frequencies of microsatellites loci in the tracks 303–309 and 311–315 from hypervariable region II

	% (n)
<i>p</i> 303–309	
(C) ₆	0.3 (1)
(C) ₇	26.0 (70)
(C) ₈	54.8 (148)
(C) ₉	18.8 (51)
<i>p</i> 311–315	
(C) ₆	99.7 (269)
(C) ₇	0.3 (1)

C16223T, T16325C and T16362C.²⁷ Mx174 was the most frequent D1 haplotype found (three individuals) in the sample and had the characteristic polymorphisms plus the transitions A73G and A263G (Table 1). Considering only HVR1 polymorphisms, this haplotype has been found earlier in South America.⁴¹ The haplogroup D1 belongs to the branch of Asian haplogroup D4; however, the transition at position 16325 is not found in D haplogroups from Asia.^{53,57} The frequencies of haplogroup D1 in Amerindian populations are lower in Central America and higher in North and South America.^{15,16,58}

Genetic diversity of mitochondrial Amerindian haplogroups

Molecular diversity indexes were calculated for each Amerindian haplogroup (see Table 2). Interestingly, although haplogroup A2 was the most frequent in the whole sample, it had a significantly lower mean number of pairwise differences (4.7) and nucleotide diversity (0.027) compared with the B2 (5.8; 0.034) and C1 (5.9; 0.034) haplogroups (*P* < 0.01, Student's *t*-test). The genetic diversity indexes of haplogroup A2 were higher than those of the haplogroup D1, but they were not significant (Table 2). The mismatch distribution was plotted for each Amerindian haplogroup and all haplogroups showed a unimodal distribution (data not shown). Moreover, the Tajima's *D* values calculated for haplogroups A2, B2, C1 and D1 were significantly negative (see Table 2). A significantly negative Tajima's *D* value with a unimodal mismatch distribution also suggests a demographic expansion.^{59,60} τ calculated for haplogroups A2, B2, C1 and D1 was 2.7, 6.1, 5.1 and 1.8, respectively. Using τ and the mutation rate estimated for mtDNA control region (HV1R: 1.64×10^{-7} and HV2R: 2.29×10^{-7}),⁶¹

the time to demographic expansion calculated for the haplogroups A2, B2, C1 and D1 was 11.6, 25.9, 21.6 and 7.9 kya before present, respectively. Therefore, the lower genetic diversity present in the haplogroups A2 and D1 could be explained by a recent expansion of these haplogroups in the Mexican mestizo population.⁶² It is important to emphasize that these calculations in the present study were performed using data from the control region, which has a higher mutational rate compared with the coding region.⁶³

Non-Amerindian haplogroups

European haplogroups present in the sample (16 individuals) were H (1.9%), J (0.7%), K (0.7%), T (0.7%), U (1.9%) and V (0.4%). Haplogroup H frequency is similar to that earlier reported for Mexico City.¹⁸ Interestingly, this haplogroup is the most frequently found in Spain.^{34,35} The lower frequency of European mitochondrial haplogroups in Mexico is related to the rare mating between European women and Amerindian or mestizo men through Mexican history.² The two individuals with African haplotypes belong to haplogroup L3. The frequency of African haplogroups found in the sample was lower than that found earlier (4.4%).¹⁷

Minimum spanning network topology of mitochondrial haplogroups

The individual sequences that were earlier classified to one of the mitochondrial haplogroups by the control region motifs were represented in the minimum spanning network (Figure 2a). It is clearly seen that individuals with the same haplogroup are grouped. Therefore, the network grouped Amerindian haplotypes into the four main haplogroups (A2, B2, C1 and D), whereas the non-Amerindian haplotypes were separated into two different groups, one for the Europeans and the other for the Africans. All four Amerindian haplogroups represented in the network had a star-like pattern with a main internal node connected to other less frequent haplotypes (Figure 2a). This observed star-like pattern has been described earlier in Amerindian haplogroups^{28,59} and also suggests population expansion.⁶⁴ The A2 haplotypes were grouped into three main internal nodes. Two nodes have the transition C16111T but are separated by transition A153G. The third node includes those haplotypes with the back-transition T16111C. Haplogroup B2 is represented in the minimum spanning network by two main internal nodes with multiple connections to other haplotypes in star-like topology. Both nodes were separated by transversion A16182C. It is noteworthy that B2 haplo-

Figure 2 Mitochondrial haplogroups in the Mexican mestizo population. (a) Minimum spanning network constructed with mitochondrial DNA (mtDNA) control region sequences from the Mexican mestizo population. Colored circles in the network show the haplotypes within each haplogroups and their relationships; the circle sizes are proportional to the haplotypes frequencies. Black circles represent inferred haplotypes and the black arrow shows the rCSR localization in the network. The main mutational steps are depicted. (b) Mitochondrial haplogroup frequency found in the total sample; the colors of the haplogroups are the same for both figures.

types are closely related to the European H haplogroup because both are derived from the ancestral R haplogroup, which belongs to the macrohaplogroup N.⁶⁵ The C1 and D1 haplotypes had only one main internal node connected to other singletons.

Distribution of mitochondrial haplogroups by geographical region of birthplace

There was no significant difference in Amerindian ancestry among the individuals grouped by birthplace region (see Table 4). However, the sample of Mexico City (91.0%) and coastal region (91.6%) showed the lower frequency of Amerindian ancestry, whereas the sample of the central region had the higher percentage of Amerindian ancestry (95.7%). This agrees with the fact that most of the extant indigenous communities, such as the Nahuatl, Mazahua and Otomi, are concentrated in the central region of Mexico.^{66,67} The frequency of haplogroup A2 was similar in all regions; however, the transition C16111T was significantly more frequent in A2 haplotypes of Mexico City (94.2%, 49 out of 52) than in the central region (75.4%, 43 out of 57; $P < 0.01$, χ^2). It is noteworthy that the frequency of the haplogroup B2 was significantly higher in the sample of the central region compared

with Mexico City and the coastal region ($P < 0.01$, χ^2). This agrees with the higher frequency of B haplotypes in some indigenous populations from the central regions, such as Nahuatl (27–40%), or from the south, such as Mixe (33.3%).^{13,15,16} The data also suggest that the higher frequency of Amerindian ancestry in the central region is accounted essentially by the increased frequency of B haplotypes. The frequency of the transversion A16182C in B2 haplotypes was more frequent in the central region (55.5%, 15 out of 27 individuals) than in Mexico City (41.6%, 5 out of 12 individuals) B2 haplotypes, but this difference was not significant. In contrast, the haplogroup C1 was more frequent in individuals from Mexico City (21%) than from the central (15.6%) and coastal regions (17%), but this difference was not significant. The lower frequency of C haplotypes in the central region agrees with the low frequency of this haplogroup in indigenous populations from the valley of Mexico.¹³ In contrast, half of the eight individuals from the north region had this haplotype, which agrees with the high frequency of this haplogroup in indigenous populations from the north, such as the Huichol (41%) and Tarahumara (38%) populations.^{13,15,16} The frequency of haplogroup D1 was similarly low in all geographical regions analyzed. More than a half of

Table 4 Mitochondrial haplogroups frequencies according to the geographical region of birth place

		Amerindian haplogroups						
Region	n	Amer % (n)	Eur % (n)	Afr % (n)	A2 % (n)	B2 % (n)	C1 % (n)	D1 % (n)
Mexico City	100	91.0 (91)	9.0 (9)	0.0 (0)	52.0 (52)	12.0 (12)	21.0 (21)	6.0 (6)
Center ^a	115	95.7 (110)	3.5 (4)	0.8 (1)	49.5 (57)	23.5 (27)*	15.6 (18)	7.0 (8)
Coast ^b	47	91.5 (43)	6.4 (3)	2.1 (1)	55.3 (26)	17.0 (8)	14.9 (7)	4.2 (2)
North ^c	8	100.0 (8)	0.0 (0)	0.0 (0)	37.5 (3)	12.5 (1)	50.0 (4)	0.0 (0)
All	270	93.3 (252)	5.9 (16)	0.7 (2)	51.1 (138)	17.8 (48)	18.5 (50)	5.9 (16)

* $P < 0.01$, χ^2 test; Center vs Mexico City, Center vs South, Center vs North.

^aEstado de México, Hidalgo, Puebla, Tlaxcala, Morelos, Queretaro and Guanajuato.

^bOaxaca, Chiapas, Guerrero, Veracruz and Michoacan.

^cJalisco, San Luis Potosi, Zacatecas, Durango and Sinaloa.

the individuals with European haplogroups were concentrated in Mexico City, which is consistent with differences in the matrilineal ancestry in the country owing to the historical fact that European women established mainly in Mexico City and larger colonial cities.²¹

CONCLUSIONS

The present work is the first survey to analyze the full mitochondrial control region in the Mexican mestizo population. Although the haplotype diversity was similarly high in other populations, most of the haplotypes identified have not been reported earlier. A higher frequency of Amerindian ancestry was found in individuals from the central region of Mexico than in those from Mexico City and the coastal areas. These differences could be because of a heterogeneous settlement of indigenous communities in the country, with a higher concentration of these communities in the central region. Mexico City has higher frequency of European haplogroups than other regions, in agreement with the historical fact that the majority of European women established in this city. The Tajimas's *D* neutrality test performed in this sample and the topology of the minimum spanning tree could suggest a demographic expansion of all Amerindian mitochondrial haplogroups in the Mexican mestizo population. The τ values calculated for these Amerindian haplogroups showed that the time to expansion was recent for haplogroups A2 and C1. The present knowledge will contribute to future forensic, evolutionary and demographic studies in Mexican mestizos. These data, in addition to other genetic markers, will be useful to estimate the ancestry proportion of the subjects in association studies to better select cases and controls and avoid the bias caused by population stratification.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This work was supported by the National Council of Science and Technology (CONACYT), grant numbers 8135/A1 and 24341 and the National University of Mexico (UNAM), grant number SDI.PTID.05.2 (to JB).

- Gorodezky, C., Alaeza, C., Vázquez-García, M. N., de la Rosa, G., Infante, E., Balladares, S. *et al.* The genetic structure of Mexican Mestizos of different locations: tracking back their origins through MHC genes, blood group systems, and microsatellites. *Hum. Immunol.* **62**, 979–991 (2001).

- Ochoa, L. *Conquista, Transculturación y Meztizaje. Raíz y Origen de México* (UNAM, Ciudad de México, México, 1995).
- Serrano, C. *Meztizaje y Características Físicas de la Población Mexicana* (Spanish). *Arquel. Mexicana.* **11**, 64–67 (2004).
- Smith, M. E. in *The Oxford Encyclopedia of Mesoamerican Cultures. The Civilizations of Mexico and Central America* Vol. 3 (ed. Carrasco, D.) 290–294 (Oxford Univ. Press, USA, 2001).
- Humboldt, A. *Ensayo Político sobre el Reino de la Nueva-España*, (Editorial Porrúa, Ciudad de México, México, 1822).
- Wolf, E. *Pueblos y culturas de Mesoamérica*, (Era, Ciudad de México, México, 1967).
- Niederberger, C. in *Historia general de América Latina*, Vol. I (ed. UNESCO) 117–150 (Trotta, Spain, 1999).
- Fernandez, I. *Historia de Mexico I*, (Pearson, Ciudad de México, México, 2006).
- Beltran, A. The slave trade in Mexico. *Hisp. Am. Hist. Rev.* **24**, 412–431 (1944).
- Benitez, A. *Mexico Mestizo*, (Fondo de Cultura Económica, México, 2002).
- Lisker, R. *Estructura Genética de la Población Mexicana*, (Salvat, Mexico City, 1981).
- Bonilla, C., Gutierrez, G., Parra, E., Kline, C. & Shriver, M. Admixture analysis of a rural population of the state of Guerrero, Mexico. *Am. J. Phys. Anthropol.* **128**, 861–869 (2005).
- Peñalosa-Espinosa, R., Arenas-Aranda, D., Cerda-Flores, R. M., Buentello-Malo, L., González-Valencia, G., Torres, J. *et al.* Characterization of mtDNA haplogroups in 14 Mexican indigenous populations. *Hum. Biol.* **79**, 313–320 (2007).
- Torroni, A., Schurr, T. G., Yang, C. C., Szathmari, E. J., Williams, R. C., Schanfield, M. S. *et al.* Native American mitochondrial DNA analysis indicates that the Amerind and the Nadene populations were founded by two independent migrations. *Genetics* **130**, 153–162 (1992).
- Torroni, A., Schurr, T. G., Cabell, M. F., Brown, M. D., Neel, J. V., Larsen, M. *et al.* Asian affinities and continental radiation of the four founding native American mtDNAs. *Am. J. Hum. Genet.* **53**, 563–590 (1993).
- Torroni, A., Chen, Y. S., Semino, O., Santachiara-Benecere, A. S., Scott, C. R., Lott, M. T. *et al.* mtDNA and Y-chromosome polymorphisms in four native American populations from southern Mexico. *Am. J. Hum. Genet.* **54**, 303–318 (1994).
- Green, L., Derr, J. & Knight, A. mtDNA affinities of the peoples of North-Central Mexico. *Am. J. Hum. Genet.* **66**, 989–998 (2000).
- Martinez-Marignac, V., Valladares, A., Cameron, E., Chan, A., Perera, A., Globus-Goldberg, R. *et al.* Admixture in Mexico City: implications for admixture mapping of type 2 diabetes genetic risk factors. *Hum. Genet.* **120**, 807–819 (2007).
- Giles, R., Blanc, H., Cann, H. & Wallace, D. Maternal inheritance of human mitochondrial DNA. *Proc. Natl Acad. Sci. USA* **77**, 6715–6719 (1980).
- Castillo, B. D. *Del Historia Verdadera de la Conquista de la Nueva España*, (El Colegio de Mexico AC, México, 2005).
- Serrano, C. *Meztizaje e Historia de la Población en México* (Con un esbozo antropológico de los lacandones de Chiapas) in *Polimorfismo Génico (HLA) en Poblaciones Hispanoamericanas* (eds Martin, A. & Garcia, P.) (Real Academia de Ciencias Exactas, Físicas y Naturales, Madrid, 1996).
- Miller, F. J., Rosenfeldt, F. L., Zhang, C., Linnane, A. W. & Nagley, P. Precise determination of mitochondrial DNA copy number in human skeletal and cardiac muscle by a PCR-based assay: lack of change of copy number with age. *Nucl. Acids Res.* **31**, e61 (2003).
- Brown, W. Polymorphism in mitochondrial DNA of humans as revealed by restriction endonuclease analysis. *Proc. Natl Acad. Sci. USA* **77**, 3605–3609 (1980).
- Schurr, T., Ballinger, S. W., Gan, Y. Y., Hodge, J. A., Merriwether, D. A., Lawrence, D. N. *et al.* 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 (1990).
- Horai, S., Kondo, R., Nakagawa-Hattori, Y., Hayashi, S., Sonoda, S. & Tajima, F. Peopling of the Americas, founded by four major lineages of mitochondrial DNA. *Mol. Biol. Evol.* **10**, 23–47 (1993).

- 26 Bandelt, H. J., Hernstadt, C., Yao, Y. G., Kong, Q. P., Kivisild, T., Rengo, C. *et al.* Identification of native American founder mtDNAs through the analysis of complete mtDNA sequences: some caveats. *Ann. Hum. Genet.* **67**, 512–524 (2003).
- 27 Tamm, E., Kivisild, T., Reid, M., Metspalu, M., Smith, D. G., Mulligan, C. J. *et al.* Beringian standstill and spread of native American founders. *PLoS ONE* **2**, e829 (2007).
- 28 Achilli, A., Perego, U. A., Bravi, C. M., Coble, M. D., Kong, Q. P., Woodward, S. R. *et al.* The phylogeny of the four Pan-American MtDNA haplogroups: implications for evolutionary and disease studies. *PLoS ONE* **3**, e1764 (2008).
- 29 González-Oliver, A., Márquez-Morfin, L., Jiménez, J. & Torre-Blanco, A. Founding Amerindian mitochondrial DNA lineages in ancient Maya from Xcaret, Quintana Roo. *Am. J. Phys. Anthropol.* **116**, 230–235 (2001).
- 30 Marchini, J., Cardon, L. R., Phillips, M. S. & Donnelly, P. The effects of human population structure on large genetic association studies. *Nat. Genet.* **36**, 512–517 (2004).
- 31 Andrews, R., Kubacka, I., Chinnery, P. F., Lightowlers, R. N., Turnbull, D. M. & Howell, N. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet.* **23**, 147 (1999).
- 32 Excoffier, V., Laval, L. & Schneider, S. Arlequin ver 3.11: an integrated software package for population genetics data analysis. *Evol. Bioinform. Online* **1**, 47–50 (2005).
- 33 Forster, P., Harding, R., Torroni, A. & Bandelt, H. Origin and evolution of native American mtDNA variation: a reappraisal. *Am. J. Hum. Genet.* **59**, 935–945 (1996).
- 34 Torroni, A., Huoponen, K., Francalacci, P., Petrozzi, M., Morelli, L., Scozzari, R. *et al.* Classification of European mtDNAs from an analysis of three European populations. *Genetics* **144**, 1835–1850 (1996).
- 35 Alvarez, J. C., Johnson, D. L., Lorente, J. A., Martinez-Espin, E., Martinez-Gonzalez, L. J., Allard, M. *et al.* Characterization of human control region sequences for Spanish individuals in a forensic mtDNA data set. *Leg. Med.* **9**, 293–304 (2007).
- 36 Salas, A., Richards, M., Lareu, M. V., Scorzari, R., Coppa, A., Torroni, A. *et al.* The African diaspora: mitochondrial DNA and the Atlantic slave trade. *Am. J. Hum. Genet.* **74**, 454–465 (2004).
- 37 Templeton, A. R., Crandall, K. A. & Sing, C. F. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* **132**, 619–633 (1992).
- 38 Clement, M., Posada, D. & Crandall, K. A. TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* **9**, 1657–1659 (2000).
- 39 Allard, M. W., Polansky, D., Wilson, M. R., Monson, K. L. & Budowle, B. Evaluation of variation in control region sequences for hispanic individuals in the SWGDAM mtDNA data set. *J. Forensic Sci.* **51**, 566–573 (2006).
- 40 Barbosa, A., Antonio, L., Dalmó, d. S., Valdir, A. & Luiz, B. Mitochondrial DNA control region polymorphism in the population of Alagoas state, North-Eastern Brazil. *J. Forensic Sci.* **53**, 142–146 (2008).
- 41 Salas, A., Acosta, V., Alvarez-Iglesias, M., Cerezo, C., Philips, M. V. & Carracedo, L. A. The mtDNA ancestry of admixed Colombian populations. *Am. J. Hum. Biol.* **20**, 584–591 (2008).
- 42 Salas, A., Jaime, J. C., Alvarez-Iglesias, V. & Carracedo, A. Gender bias in the multiethnic genetic composition of central Argentina. *J. Hum. Genet.* **53**, 662–674 (2008).
- 43 Stewart, J., Fisher, C. L., Aagaard, P. J., Wilson, M. R., Isenberg, A. R., Polansky, D. *et al.* Length variation in HV2 of the human mitochondrial DNA control region. *J. Forensic Sci.* **46**, 862–870 (2001).
- 44 Vigilant, L., Stoneking, M., Harpending, H., Hawkes, K. & Wilson, A. C. African populations and the evolution of human mitochondrial DNA. *Science* **253**, 1503–1507 (1991).
- 45 Yong Ji, Z., Qing Song, X., Zhe Jia, Z., Hai Yu, L. & Jung Bin, L. Haplotype diversity in mitochondrial DNA hypervariable region I, II and III in northeast China Han. *Forensic Sci. Int.* **149**, 267–269 (2005).
- 46 Crespiello, M., Luque, J. A., Paredes, M., Fernandez, R., Ramirez, E. & Valverde, J. L. Mitochondrial DNA sequences for 118 individuals from northeastern Spain. *Int. J. Legal Med.* **114**, 130–132 (2000).
- 47 Salas, A., Lareu, V., Calafell, F., Bertranpetit, J. & Carracedo, A. mtDNA hypervariable region II (HVII) sequences in human evolution studies. *Eur. J. Hum. Genet.* **8**, 964–974 (2000).
- 48 Turchi, C., Buscemi, L., Previdere, C., Grignani, P., Brandstätter, A., Achilli, A. *et al.* Italian mitochondrial DNA database: results of a collaborative exercise and proficiency testing. *Int. J. Legal Med.* **122**, 199–204 (2008).
- 49 Santos, M. & Barrantes, R. D-loop mtDNA deletion as a unique marker of Chibchan Amerindians. *Am. J. Hum. Genet.* **55**, 413–414 (1994).
- 50 Arredondo, A., Nájera, P. & Leyva, R. Atención médica ambulatoria en México: el costo para los usuarios. *Salud Públ. Méx.* **41**, 18–26 (1999).
- 51 Relethford, J., Stern, M., Gaskill, S. & Hazuda, H. Social class, admixture, and skin color variation in Mexican-Americans and Anglo-Americans living in San Antonio, Texas. *Am. J. Phys. Anthropol.* **61**, 97–102 (1983).
- 52 Easton, R. D., Merriwether, D. A., Crews, D. E. & Ferrell, R. E. mtDNA variation in the Yanomami: evidence for additional new world founding lineages. *Am. J. Hum. Genet.* **59**, 213–225 (1996).
- 53 Kong, Q. P., Bandelt, H. J., Sun, C., Yao, Y. G., Salas, A., Achilli, A. *et al.* Updating the East Asian mtDNA phylogeny: a prerequisite for the identification of pathogenic mutations. *Hum. Mol. Genet.* **15**, 2076–2086 (2006).
- 54 Tanaka, M., Cabrera, V. M., González, A. M., Larruga, J. M., Takeyasu, T., Fuku, N. *et al.* Mitochondrial genome variation in Eastern Asia and the peopling of Japan. *Genome Res.* **14**, 1832–1850 (2004).
- 55 Mabuchi, T., Susukida, R., Kido, A. & Oya, M. Typing the 1.1 kb control region of human mitochondrial DNA in Japanese individuals. *J. Forensic Sci.* **52**, 355–363 (2007).
- 56 Yao, Y. G., Kong, Q. P., Bandelt, H. J., Kivisild, T. & Zhang, Y. P. Phylogeographic differentiation of mitochondrial DNA in Han Chinese. *Am. J. Hum. Genet.* **70**, 635–651 (2002).
- 57 Kivisild, T., Tolk, H. V., Parik, J., Wang, Y., Papiha, S. S., Bandelt, H. J. *et al.* The emerging limbs and twigs of the East Asian mtDNA tree. *Mol. Biol. Evol.* **19**, 1737–1751 (2002).
- 58 Salzano, F. M. Molecular variability in Amerindians: widespread but uneven information. *An. Acad. Bras. Ciénc.* **74**, 223–263 (2002).
- 59 Harpending, H. C., Batzer, M. A., Gurven, M., Jorde, L. B., Rogers, A. R. & Sherry, S. T. Genetic traces of ancient demography. *Proc. Natl Acad. Sci. USA* **95**, 1961–1967 (1998).
- 60 Tajima, F. Statistical analysis of DNA polymorphism. *Jpn. J. Genet.* **68**, 567–595 (1993).
- 61 Soares, P., Ermini, L., Thomson, N., Mormina, M., Rito, T., Röhl, A. *et al.* Correcting for purifying selection: an improved human mitochondrial molecular clock. *Am. J. Hum. Genet.* **84**, 740–759 (2009).
- 62 Schneider, S. & Excoffier, L. Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics* **152**, 1079–1089 (1999).
- 63 Torroni, A., Neel, J. V., Barrantes, R., Schurr, T. G. & Wallace, D. C. Mitochondrial DNA “clock” for the Amerinds and its implications for timing their entry into North America. *Proc. Natl Acad. Sci. USA* **91**, 1158–1162 (1994).
- 64 Kaessmann, H. & Pääbo, S. The genetical history of humans and the great apes. *J. Int. Med.* **251**, 1–18 (2002).
- 65 Maca-Meyer, N., Gonzalez, A., Larruga, J., Flores, C. & Cabrera, V. Major genomic mitochondrial lineages delineate early human expansions. *BMC Genet.* **2**, 13 (2001).
- 66 Valdes de Montano, L. La población indígena. Hacia la expansión demográfica de los grupos étnicos. *Demos* **2**, 17–18 (1989).
- 67 INEGI. *Perfil Sociodemográfico de la Población Hablante de Náhuatl* (INEGI, México, 2005).