

---

## ***Genetic Structure Analysis of Three Hispanic Populations from Costa Rica, Mexico, and the Southwestern United States Using Y-Chromosome STR Markers and mtDNA Sequences***

REBECA CAMPOS-SÁNCHEZ,<sup>1</sup> RAMIRO BARRANTES,<sup>2</sup> SANDRA SILVA,<sup>1</sup> MICHAEL ESCAMILLA,<sup>3</sup> ALFONSO ONTIVEROS,<sup>4</sup> HUMBERTO NICOLINI,<sup>5</sup> RICARDO MENDOZA,<sup>6</sup> RODRIGO MUNOZ,<sup>7</sup> AND HENRIETTE RAVENTOS<sup>1</sup>

**Abstract** Two hundred seventeen male subjects from Costa Rica, Mexico, and the Hispanic population of the southwestern United States were studied. Twelve Y-chromosome STRs and the HVSI sequence of the mtDNA were analyzed to describe their genetic structure and to compare maternal and paternal lineages. All subjects are part of two NIMH-funded studies to localize schizophrenia susceptibility genes in Hispanic populations of Mexican and Central American ancestry. We showed that these three populations are similar in their internal genetic characteristics, as revealed by analyses of mtDNA and Y-chromosome STR diversity. These populations are related through their maternal lineage in a stronger way than through their paternal lineage, because a higher number of shared haplotypes and polymorphisms are seen in the mtDNA (compared to Y-chromosome STRs). These results provide evidence of previous contact between the three populations and shared histories. An analysis of molecular variance revealed no genetic differentiation for the mtDNA for the three populations, but differentiation was detected in the Y-chromosome STRs. Genetic distance analysis showed that the three populations are closely related, probably as a result of migration between close neighbors, as indicated by shared haplotypes and their demographic histories. This relationship could be an important common feature for genetic studies in Latin American and Hispanic populations.

Population genetic studies can be used to compare relationships and differences between populations and to determine the genetic structures of naturally

<sup>1</sup>Centro de Investigación en Biología Celular y Molecular, Universidad de Costa Rica, San Pedro, SJ, Costa Rica.

<sup>2</sup>Escuela de Biología, Universidad de Costa Rica, San Pedro, SJ, Costa Rica.

<sup>3</sup>Department of Psychiatry, University of Texas Health Science Center, San Antonio, TX.

<sup>4</sup>INFOSAME, Monterrey, Mexico.

<sup>5</sup>Grupo de Estudios Médicos y Familiares Carracci S.C., Mexico City, Mexico.

<sup>6</sup>David Geffen School of Medicine, University of California, Los Angeles, CA.

<sup>7</sup>Family Health Centers of San Diego, San Diego, CA.

*Human Biology*, October 2006, v. 78, no. 5, pp. 551–563.

Copyright © 2006 Wayne State University Press, Detroit, Michigan 48201-1309

KEY WORDS: GENETIC STRUCTURE, Y-CHROMOSOME STR, mtDNA, HVI, COSTA RICA, MEXICO, SOUTHWESTERN UNITED STATES, U.S. HISPANIC POPULATION, SCHIZOPHRENIA, *DYS19*, *DYS385*, *DYS389*, *DYS390*, *DYS391*, *DYS392*, *DYS437*, *DYS438*, *DYS439*.

occurring populations (Hey and Machado 2003). Molecular analyses of Y-chromosome STR markers and mtDNA haplotypes are ideal for determining the geographic structure of human populations, assuming that the markers act as neutral markers (Behar et al. 2004; Harpending and Rogers 2000; Pakendorf and Stoneking 2005; Walsh 2001). The fast evolutionary rate of STRs and mtDNA and small effective population sizes make these markers more sensitive to genetic differences among groups than autosomal markers. Moreover, comparing the patterns of variation from the Y chromosome and the mtDNA allows one to understand the paternal and maternal history of the populations. This is especially useful for admixed populations (Kayser et al. 2003).

The origins of many South, Central, and North American populations, including Costa Rica, Mexico, and the United States, can be traced to ancestral migrations from Asia through the Bering Strait (the Amerindian heritage) (Collins-Schramm et al. 2004) and posterior migrations and admixture during the colonization period of the 15th–16th centuries (Collins-Schramm et al. 2004; Green et al. 2000). To characterize the genetic relationships among the Hispanic populations of Mexico, the southwestern United States, and Costa Rica, we studied the genetic structure of these populations using microsatellite markers from the nonrecombining region of the Y chromosome and by sequencing the HVSI region of the mtDNA. The sample was drawn from two studies, funded by the National Institute of Mental Health (NIMH), to map schizophrenia susceptibility genes (one study mapped schizophrenia susceptibility genes in the Hispanic population and the other study is a population-based study of schizophrenia in Costa Rica).

In this study we report the maternal and paternal relationships among the three groups, based on the analysis of 12 Y-chromosome STRs and HVSI mtDNA sequences. Similarities between these populations have not been reported previously. Our results demonstrate that these three populations are genetically related to each other.

## Materials and Methods

**Population Samples.** All subjects were recruited as part of the NIMH-funded studies to map schizophrenia susceptibility genes; informed consent was obtained from all participants. Two hundred seventeen DNA samples from unrelated males from Costa Rica (120 samples), Mexico (55 samples from Mexico City and Monterrey), and the southwestern United States (42 samples from the Hispanic populations of Los Angeles, San Diego, and San Antonio) were studied. Subjects recruited from the United States had to have some family ancestry from Mexico or Central America. Some were probands (i.e., they had a previous diagnosis of schizophrenia) and others were family members, but none of the 217 subjects were related to each other. The geographic coordinates of the regions of study are the following: Central Valley of Costa Rica, 9.45–10.15° N, 83.45–84.30° W;

San Antonio, Texas, 29.25° N, 98.29° W; Los Angeles, California, 34.03° N, 118.14° W; San Diego, California, 32.42° N, 117.09° W; Monterrey, Mexico, 25.40° N, 100.19° W; and Mexico City, Mexico, 19.24° N, 99.09° W.

**Y-Chromosome STR Typing.** Twelve microsatellites from the Y chromosome were amplified using the PowerPlex Y-System kit (Promega, Madison, Wisconsin) in a 12.5- $\mu$ l reaction volume, as specified by the manufacturer. The Y-chromosome STRs studied were *DYS19*, *DYS385 I* and *II*, *DYS389 I* and *II*, *DYS390*, *DYS391*, *DYS392*, *DYS437*, *DYS438*, and *DYS439*. Detection of the amplified fragments was done using the ABI Prism 377 (Applied Biosystems, Foster City, California). PowerTyper Y Macro (Promega) was used to assign the alleles.

**Sequencing of the mtDNA.** Hypervariable segment I (HVSI) of the mitochondrial DNA was amplified in a PCR reaction volume of 50- $\mu$ l using primers L15997 (5'-CAC CAT TAG CAC CCA AAG CT) and H16401 (5'-TGA TTT CAC GGA GGA TGG TG). The first amplification was done in a volume of 50  $\mu$ l with 4  $\mu$ l DNA (2 ng/ $\mu$ l), 2  $\mu$ l of each primer (5  $\mu$ M), 5  $\mu$ l of STR Buffer 10x (Promega), 3  $\mu$ l Taq DNA polymerase (5 U/ $\mu$ l) (Promega), and water. The amplification protocol used was the following: 95°C for 5 min; 35 cycles of 94°C for 45 s, 65°C for 45 s, and 72°C for 90 s; followed by 72°C for 7 min and a final hold at 4°C. To confirm amplification, the products were separated on a 2% agarose gel with TBE 1  $\times$  buffer. PCR products were purified with the Wizard PCR Preps kit (Promega), according to the manufacturer's instructions, and quantified through spectrophotometry. Double-stranded sequencing was carried out using the BigDye Terminator v3.1 Cycle Sequencing RR-100 (Applied Biosystems), as specified in the manual. The products were precipitated with EDTA and ethanol and resuspended in formamide and blue dextran before the separation on the ABI Prism 377 sequencer. The software Sequencing Analysis 3.3 (Applied Biosystems) was used to register the data and obtain the electropherograms. The quality of the sequences was evaluated with BioEdit (Hall 1999) and then was aligned with the Cambridge Reference Sequence (Anderson et al. 1981) using Clustal. Three hundred nineteen nucleotides from 16047 to 16365 were used for all the statistical analyses.

**Data Analysis.** For the Y-chromosome STR data we estimated allele and haplotype frequencies. Gene and haplotype diversities were estimated according to the method of Nei (1987). Polymorphism frequency and nucleotide diversities were estimated for the mtDNA sequences. In addition, we studied mean pairwise differences for the Y-chromosome STRs (based on the number of alleles) and for mtDNA (Tamura and Nei 1993). Shared haplotypes were used to determine relation between populations.

**Table 1.** Gene Diversity for Each of the 12 Y-Chromosome STRs Analyzed in the Hispanic Populations Related to Schizophrenia from Costa Rica, the Southwestern United States, and Mexico

Marker	Costa Rica	Southwestern United States	Mexico
<i>DYS19</i>	0.697	0.584	0.601
<i>DYS385I</i>	0.798	0.680	0.758
<i>DYS385II</i>	0.793	0.714	0.804
<i>DYS389I</i>	0.609	0.544	0.432
<i>DYS389II</i>	0.769	0.681	0.704
<i>DYS390</i>	0.675	0.573	0.573
<i>DYS391</i>	0.507	0.580	0.567
<i>DYS392</i>	0.682	0.608	0.719
<i>DYS393</i>	0.502	0.387	0.286
<i>DYS437</i>	0.628	0.557	0.508
<i>DYS438</i>	0.702	0.627	0.653
<i>DYS439</i>	0.671	0.658	0.563
Mean gene diversity <sup>a</sup>	0.675 (0.349)	0.614 (0.324)	0.608 (0.321)

a. Standard deviations are given in parentheses.

To analyze the hierarchical genetic structure of the three populations, we performed an analysis of molecular variance (AMOVA) among and within populations. The AMOVA compares gene frequencies among populations and differences (mutations) between alleles (Excoffier 2001). For the Y-chromosome microsatellite data  $R_{ST}$  was estimated, and for mtDNA  $F_{ST}$  was calculated. Also, we estimated genetic distances for the three populations according to the method of Tamura and Nei (1993) for mtDNA and according to the method of Nei (1972) for Y-chromosome STRs. These methods analyze the average genetic differences between local, regional, and continental populations by computing distances between pairs of populations and then analyzing them to determine how past evolutionary events could explain the patterns of genetic distance observed in present-day populations (Relethford 2001).

All mtDNA analyses were performed using a gamma A value (heterogeneity of mutation rates) of 0.26 and a transition-transversion ratio of 15. The statistical packages Arlequin, version 2.000 (Schneider et al. 2000), and Arlequin, version 1.1 (Schneider et al. 1997), were used for all calculations.

## Results

**Y-Chromosome STR Variability.** Table 1 shows the gene diversities for all 12 Y-chromosome STRs in Costa Rica, the southwestern United States, and Mexico. For 9 of the 12 markers Costa Rica shows a slightly higher variability inside the population. When allele frequencies were compared between these populations, all three populations shared the most common allele for 9 of the

**Table 2.** Diversity Measures Obtained by the Study of 12 Y-Chromosome STRs in the Hispanic Populations of Costa Rica, the Southwestern United States, and Mexico

	Costa Rica	Southwestern United States	Mexico
<i>N</i> <sup>a</sup>	120	42	55
Number of haplotypes	103	40	51
Shared haplotypes	14	2	2
Haplotype diversity <sup>b</sup>	0.997 (0.002)	0.998 (0.006)	0.995 (0.005)
Mean number of pairwise differences <sup>b</sup>	8,099 (3,786)	7,368 (3,516)	7,300 (3,470)

a. Number of individuals.

b. Standard deviations are given in parentheses.

12 markers (*DYS385I\*11*, *DYS385II\*14*, *DYS389I\*13*, *DYS390\*24*, *DYS393\*13*, *DYS437\*14*, *DYS438\*12*, *DYS439\*12*, and *DYS19\*14*; data not shown).

On average, gene diversities were 0.675, 0.614, and 0.608 for Costa Rica, the U.S. Southwest, and Mexico, respectively (Table 2). The most informative marker, based on gene diversity, for all three populations was *DYS385III*, and the least informative marker was *DYS393*.

For the 12 Y-chromosome STRs, 103 haplotypes were obtained with a diversity of 99.7% for Costa Rica, 40 haplotypes with a diversity of 99.8% for the U.S. Southwest, and 51 haplotypes with a diversity of 99.5% for Mexico (see Table 2). Although the populations showed high diversities, some shared haplotypes were found within populations. Fourteen haplotypes were shared inside Costa Rica, eleven by two individuals and three by three individuals. Two haplotypes were shared by two individuals in the U.S. Southwest, and in Mexico one haplotype was shared by four individuals and a second haplotype was shared by two individuals (data not shown).

**HVSI Variability.** The analysis of the mtDNA (np 16047 to 16365) in the three populations showed 52, 39, and 56 polymorphic sites in Costa Rica, the U.S. Southwest, and Mexico, respectively (Table 3). Transitions were more common (48 in Costa Rica, 36 in the U.S. Southwest, and 53 in Mexico) than transversions (4 in Costa Rica and Mexico, 3 in the U.S. Southwest). Eighty-eight polymorphisms were present in the three populations together.

There were 18 polymorphisms shared among the three populations, and 4 more were shared between Costa Rica and the U.S. Southwest, 12 between Costa Rica and Mexico, and 7 between the U.S. Southwest and Mexico. There were also length variation polymorphisms at nucleotide positions 16188.1C, 16192.1C, and 16193.1C that were not introduced in the sequence analysis (Cordaux et al. 2003; Salas et al. 2004). Nucleotide diversity reveals heterozygosity at the nucleotide level and was approximately 0.02 for all three populations, showing another similarity between them (see Table 3).

**Table 3.** Diversity Measures in the mtDNA Sequences from the Hispanic Populations of Costa Rica, the Southwestern United States, and Mexico

	<i>Costa Rica</i>	<i>Southwestern United States</i>	<i>Mexico</i>
$N^a$	120	42	55
Number of polymorphic sites	52	39	56
Number of unique polymorphisms	18	10	19
Nucleotide diversity <sup>b</sup>	0.024 (0.012)	0.022 (0.012)	0.026 (0.014)
Number of haplotypes	39	33	43
Shared haplotypes	17	4	8
Haplotype diversity <sup>b</sup>	0.930 (0.012)	0.981 (0.011)	0.989 (0.007)
Mean number of pairwise differences <sup>b</sup>	7.657 (3.595)	6.881 (3.303)	8.445 (3.967)

a. Number of individuals.

b. Standard deviations are given in parentheses.

In the Costa Rica sample 39 mtDNA haplotypes were found, with a haplotype diversity of 93%, and 17 of the haplotypes were shared. The presence of 6 haplotypes shared by 20, 19, 11, 7, and 6 individuals is interesting. This suggests homogeneity within the sample. In the U.S. Southwest sample four haplotypes were shared (two shared by four individuals, one shared by three individuals, and one shared by two individuals) and in the Mexico sample one haplotype was shared by four individuals, two haplotypes were shared by three individuals, and five haplotypes were shared by two individuals (see Table 3).

**Shared Haplotypes Between Populations.** We found that Mexico and the U.S. Southwest share Y-chromosome STR haplotypes, one of which is the most represented in Mexico (7.3%). Just one haplotype was shared between Costa Rica and the U.S. Southwest. None of the haplotypes were shared between Mexico and Costa Rica or between all three populations (Table 4) for the Y-chromosome STRs.

In contrast to the Y-chromosome STR data, two mtDNA haplotypes were shared by all three populations (Table 5). There were also additional haplotypes that were shared only between subjects from Costa Rica and Mexico and only between subjects from the U.S. Southwest and Mexico.

**AMOVA.** The analysis of the populations by AMOVA for the Y-chromosome STR data reveals that 96% of the differences are the result of variations within populations ( $R_{ST} = 0.04$ ,  $P < 0.001$ ). This is also significant when the variation is separated per locus (Table 6). For *DYS392* ( $R_{ST} = 0.051$ ,  $P < 0.001$ ) the difference between populations is only 5.149%, and for *DYS439* 97.8% of the variation is within populations with a probability of 0.038 (data not shown).

When the sample is subdivided into the six regions from Costa Rica, the U.S. Southwest, and Mexico, similar results are obtained (96.25% of the variation

**Table 4.** Shared Haplotypes and Their Description Among the Hispanic Populations from Costa Rica, the Southwestern United States, and Mexico for the 12 Y-Chromosome STRs<sup>a</sup>

Haplotype	Southwestern			Description <sup>b</sup>
	Costa Rica	United States	Mexico	
HY25	3 (0.025)	2 (0.048)		11/14, 14, 30, 24, 11, 13, 13, 15, 12, 11, 14
HY106		1 (0.024)	1 (0.018)	11/14, 13, 29, 25, 11, 13, 13, 14, 12, 12, 14
HY119		1 (0.024)	1 (0.018)	13/14, 14, 30, 24, 9, 11, 13, 14, 10, 10, 13
HY125		1 (0.024)	4 (0.073)	11/15, 13, 29, 24, 11, 13, 13, 15, 12, 12, 14
HY135		1 (0.024)	1 (0.018)	11/14, 14, 30, 24, 10, 13, 13, 14, 12, 12, 14

a. Haplotype frequencies are given in parentheses for each population.  
 b. Order of the haplotypes: *DYS385III*, *DYS389I*, *DYS389II*, *DYS390*, *DYS391*, *DYS392*, *DYS393*, *DYS437*, *DYS438*, *DYS439*, *DYS19*.

is within the samples,  $R_{ST} = 0.038$ ,  $P < 0.01$ ; see Table 6). This result is also significant when we separate the variation by locus ( $R_{ST} = 0.019$ ,  $P < 0.001$ ), for *DYS392* ( $R_{ST} = 0.044$ ,  $P < 0.01$ ), *DYS439* ( $R_{ST} = 0.043$ ,  $P < 0.01$ ), and *DYS389II* ( $R_{ST} = 0.037$ ,  $P < 0.01$ ).

For the mtDNA the AMOVA did not show differences between the populations ( $F_{ST} = 0.006$ ,  $P > 0.05$ ) or between the six regions ( $F_{ST} = 0.005$ ,  $P > 0.304$ ).

**Genetic Distances.** Analysis of genetic distances showed that the three populations are not statistically significantly distinguishable by either Y-chromosome STR data or mtDNA sequences (Table 7), allowing us to infer a close relationship

**Table 5.** Shared Haplotypes and Their Description Among the Hispanic Populations from Costa Rica, the Southwestern United States, and Mexico for the mtDNA HVSI Region<sup>a</sup>

Haplotype	Southwestern			Description (Number Minus 16000) <sup>b</sup>
	Costa Rica	United States	Mexico	
HM4	19 (0.158)	4 (0.095)	4 (0.073)	111T, 223T, 290T, 319A, 362C
HM10	6 (0.050)	2 (0.048)	1 (0.018)	182C, 183C, 189C, 217C
HM20	1 (0.008)		1 (0.018)	111C, 209C, 223T, 290T, 319A, 360T, 362C
HM39	1 (0.008)		1 (0.018)	183C, 189C, 217C, 278T
HM43		1 (0.024)	1 (0.018)	111T, 223T, 290T, 362C
HM49		4 (0.095)	1 (0.018)	223T, 298C, 325C, 327T
HM53		1 (0.024)	1 (0.018)	111T, 183C, 189C, 217C
HM54		3 (0.071)	3 (0.055)	223T, 290T, 319A, 362C
HM57		1 (0.024)	1 (0.018)	182C, 183C, 189C, 224C, 311C
HM69		1 (0.024)	2 (0.036)	182C, 183C, 189C, 217C, 295T

a. The frequency of the haplotype in the population is given in parentheses.  
 b. The description of the haplotype includes the substitution and the position minus 16000.

**Table 6.** AMOVA Results in the Three Hispanic Populations or Regions for the 12 Y-Chromosome STRs and the HVSI Sequence

Population	12 Y-Chromosome STRs			mtDNA	
	Variance	$R_{ST}$	$R_{ST}$ per Locus	Variance	$F_{ST}$
Three populations (Costa Rica, U.S. Southwest, Mexico)					
Between populations	4.00	0.040 <sup>a</sup>	0.016 <sup>b</sup>	0.64	0.006
Within populations	96.00			99.36	
Six regions (Costa Rica, San Diego, San Antonio, Los Angeles, Monterrey, and Mexico City)					
Between populations	4.53	0.038 <sup>a</sup>	0.019 <sup>b</sup>	0.96	0.005
Between populations within regions	-0.77			-0.47	
Within populations	96.25			99.51	

a.  $P < 0.01$ .b.  $P < 0.001$ .

between them. By simple observation, Costa Rica is farther from the U.S. Southwest, and Mexico is closer to the American Southwest, with an intermediate value between Costa Rica and Mexico for both types of data.

## Discussion

The Y-chromosome STR results confirmed the assumption that the three populations are related because of shared history and migration (Melton et al. 2001), more so between Mexico and the U.S. Southwest, which are geographic neighbors with a common history (the southwestern United States was part of Mexico until the mid 19th century). Frequent allele sharing is evidence of genetic relatedness (Bianchi et al. 1998), so the comparison of these data with previous studies would allow us to answer how representative the present samples are to the natural populations they belong to. Two recent publications on Mexican mestizos show the same most frequent allele for six Y-chromosome STRs (Rangel-Villalobos et al. 2001a, 2001b) in the mestizo sample and the current sample from Mexico. Previous studies for the admixed populations from Costa Rica and the U.S. Southwest with Y-chromosome STRs are insufficient to demonstrate the representativeness of the present sample from Costa Rica and the U.S. Southwest, so more extensive research on this issue should be done.

The genetic similarity of these three populations could be the result of shared inheritance from Europe (Spain and Portugal) during the colonial period and the slave trade from Africa. This possibility was evaluated by a simple comparison of allele frequencies from Spain and Portugal (Carracedo et al. 2001;

**Table 7.** Genetic Distances Between the Hispanic Populations from Costa Rica, the Southwestern United States, and Mexico for the 12 Y-Chromosome STRs (Above Diagonal) and the HVSI Sequence of the mtDNA (Below Diagonal)

<i>mtDNA Y-Chromosome STR</i>	<i>Costa Rica</i>	<i>Southwestern United States</i>	<i>Mexico</i>
Costa Rica		0.075	0.073
Southwestern United States	0.015		0.072
Mexico	0.003	-0.002	

Gusmão et al. 2003), the main genetic European contributors to the gene pool of the Americas. For six Y-chromosome STRs these two European populations share the most frequent allele with the three current populations, showing an important relationship between these five populations in their paternal lineages. Using Y-chromosome STRs, Gusmão et al. (2003) grouped Costa Rican mestizos with the Iberian-rooted populations (including Spain and Portugal) in a neighbor-joining tree, showing evidence of a relationship between the populations. Also, the U.S. Southwest population shares four of the most frequent alleles with the Y-chromosome Reference Database for European and Hispanic Americans, which reports five Y-chromosome STR markers. Moreover, African inheritance is evident in the Costa Rica population, which shares four frequent alleles with Costa Rica Africans (Gusmão et al. 2003; Jiménez-Arce 2000). This information shows evidence of the admixing of these populations.

Five of the 18 shared polymorphisms in the mtDNA between Costa Rica, the U.S. Southwest, and Mexico have frequencies higher than 54.2% and are also present in Amerindian and mestizo populations from Mexico and Costa Rica (Meléndez 2001; Santos et al. 1994; Green et al. 2000). This may indicate part of the genetic admixed content of the studied populations and the similarity between their maternal lineages, also supported by their similar nucleotide diversity values. These data can be extrapolated to the U.S. Southwest Hispanic population because of shared history (Maza-Zavala 1992).

The Costa Rica sample shows an overrepresentation of some haplotypes (HM4, HM7, and HM13), which could be evidence of a more homogeneous maternal lineage in the Costa Rican samples. Nonetheless, further studies are needed to answer whether this result can be generalized to the Costa Rican population or whether it is an effect of the selection of the sample for the presence of schizophrenia in a family member.

Shared haplotypes could be the result of relationship between populations because of migrations and/or shared backgrounds or parallel mutations (identity by state). The sharing of more Y-chromosome STR haplotypes between Mexico and the Hispanic population from the southwestern United States is explained by known migration behaviors of these populations favored by reduced geographic distances and shared political history through the mid 19th century (CEPAL

1999; Melton et al. 2001). A possible explanation for the lack of shared haplotypes across all three populations is the reduction of particular Y chromosomes resulting from native population bottlenecks in the three populations. Another explanation is the distinctive waves of colonization that occurred during several periods throughout history, causing heterogeneity in different regions of the Americas. A third explanation could be the admixture among Amerindians, Europeans, and Africans within each population, increasing heterogeneity. Last, mutations make founding haplotypes differ in their descendents, thus increasing diversity.

When we analyzed shared haplotypes between the three studied Hispanic populations and samples of whites, Hispanics, and African Americans from the U.S. Southwest for the 12 Y-chromosome STRs, we found a relationship between the two groups of populations, suggesting that shared haplotypes are common in Latin American and Hispanic populations because of the documented historical admixture.

The shared haplotypes among the three study populations for mtDNA are relevant because they provides direct evidence of a shared ancestral relationship between these groups. This relationship was also observed when we analyzed nine more populations from Europe, Asia, the Americas, and Australia, in which many more haplotypes were shared between Latin American and Hispanic populations than between the Hispanic populations and the other non-Hispanic populations (data not shown). Again, this situation could be the result of migrations, parallel mutations, and ancestral female lineages, which can be tested by analyzing the presence and proportion of Amerindian haplogroups. In the Costa Rican population from the present study the proportion of Amerindian haplogroups is 84% (haplogroup A, 52.5%; haplogroup B, 27.5%; and haplogroup C, 4.1%), as was previously reported (Carvajal-Carmona et al. 2003). For the Mexican samples 59% of the mitochondrial haplogroups correspond to Amerindian haplogroups (haplogroup A, 40%; haplogroup B, 25.5%; and haplogroup C, 1.8%), which is much less than what has been seen in previous reports (Green et al. 2000). For the U.S. Southwest sample 67.2% of the sample corresponds to Amerindian lineages (haplogroup A, 33%; haplogroup B, 14.3%; and haplogroup C, 11.9%). In proportion these percentages agree with the genetic history of the Americas, in which haplogroup A is the most frequent among Amerindians. The current results, which show increased mtDNA contribution from Amerindians, show that directional mating, the asymmetric gene contribution of maternal or paternal lineages in cases of ethnic gene admixture, occurred in all three study populations. Specifically, the haplotypes shared by the three populations are part of haplogroup A in the case of HM4 and haplogroup B for HM10, assigned by polymorphisms in the HVSI region, which should be confirmed with polymorphisms in the coding region (Green et al. 2000). These two haplogroups are represented in Mexican mestizos (Green et al. 2000) and Costa Rican Amerindians (Santos et al. 1994) in similar proportions to the present data, indicating a

direct relationship between Costa Rica and Mexico through maternal Amerindian lineages.

Patrilocal behavior has been used to explain the reduction in mitochondrial differences between populations. This custom, in which the wife goes to live in the husband's location, causes a higher rate of migration in women than in men (Cavalli-Sforza 1998; Pérez-Lezaun et al. 1999) and the spread of maternal mitochondrial haplotypes. Recently, however, it has been shown that this rate is not so different from the rate for men (Wilder et al. 2004) and depends on the social and mating system of the human population under study (Hamilton et al. 2005).

The mutation rate of mitochondrial DNA is high, especially in the control region. Mutations sometimes can either make haplotypes identical not due to ancestry, or, conversely, lead to differences between individuals with the same ancestral mtDNA (Relethford 2001). This may influence the distribution of the haplotypes (Cavalli-Sforza 1998).

The three populations studied here descend predominantly from five mitochondrial native Amerindian haplogroups (A, B, C, D, and X; Bianchi et al. 1998), which maintain their identity to the present day. In the Mexican mestizo population 59% (current study) to 90% (Green et al. 2000) of the mtDNA is of Amerindian ancestry, compared to 83% (Carvajal-Carmona et al. 2003) to 84% (current study) in the Costa Rican population. These values confirm the relationship of the populations through their maternal lineages. The Y chromosome reflects the predominantly European background of these populations. The most common allele for six Y-chromosome STRs was shared between these three populations and Spain and Portugal (Carracedo et al. 2001; Gusmão et al. 2003). All three groups (Costa Rica, Mexico, and the U.S. Southwest) therefore show mtDNA inherited primarily from Amerindians and Y-chromosome DNA inherited primarily from Spanish and Portuguese ancestors.

Although AMOVA was significant for the Y-chromosome STRs in the three populations and six regions, the percentage of differentiation was low (4–4.5%). This significance was due primarily to the results of three markers (*DYS392*, *DYS439*, and *DYS389II*), in which the populations did not share the most common allele, the difference between the frequencies of the most representative allele was extreme, and the number of alleles present was different.

As expected from previous studies, the genetic distances based on Y-chromosome STR and mtDNA data show small differences. The geographic proximity, shared political history, and frequent migration could explain the similarity between these populations that is revealed by the gene frequencies (Cavalli-Sforza 1998; Cavalli-Sforza and Feldman 2003; CEPAL 1999).

In conclusion, in this study we describe in detail the genetic relationship between the Hispanic populations of Costa Rica, Mexico and the southwestern United States. As previously reported, the main contribution to the male lineage is of Spanish-Portuguese ancestry and the main contribution to the maternal lineage comes from a probable small effective group of Amerindian ancestors. A

higher variation was found for the Y chromosome than for the mtDNA, suggesting female movements during the history that reduced the differences between the populations.

*Acknowledgments* We are indebted to the patients and family members who participated in this study. This study was supported in part by the National Institute of Mental Health through grants R01-MH61884 and R01-MH60881 (awarded to Michael Escamilla) and the International Center for Genetic Engineering and Biotechnology, Project CRP/COS98-01 (grant awarded to Henriette Raventos). The data are available upon request from Henriette Raventos. For U.S. population data on Y-chromosome STRs, we used the *Y-Chromosome Haplotype Reference Database* (available at <http://www.ystr.org/usa>).

*Received 15 June 2005; revision received 21 February 2006.*

## Literature Cited

- Anderson, S., A. T. Bankier, B. G. Barrell et al. 1981. Sequence and organization of the human mitochondrial genome. *Nature* 290:457–465.
- Behar, D.M., D. Garrigan, M. E. Kaplan et al. 2004. Contrasting patterns of Y-chromosome variation in Ashkenazi Jewish and host non-Jewish European populations. *Hum. Genet.* 114:354–365.
- Bianchi, N. O., C. I. Catanesi, G. Bailliet et al. 1998. Characterization of ancestral and derived Y-chromosome haplotypes of New World native populations. *Am. J. Hum. Genet.* 63:1862–1871.
- Carracedo, A., A. Beckmann, A. Bengts et al. 2001. Results of a collaborative study of the EDNAP Group regarding the reproducibility and robustness of the Y-chromosome STRs *DYS19*, *DYS389 I and II*, *DYS390*, and *DYS393* in a PCR pentaplex format. *Forensic Sci. Int.* 119:28–41.
- Carvajal-Carmona, L. G., R. Ophoff, S. Service et al. 2003. Genetic demography of Antioquia (Colombia) and the Central Valley of Costa Rica. *Hum. Genet.* 112:534–541.
- Cavalli-Sforza, L. L. 1998. The DNA revolution in population genetics. *Tr. Genet.* 14:60–65.
- Cavalli-Sforza, L. L., and M. W. Feldman. 2003. The application of molecular genetic approaches to the study of human evolution. *Nat. Genet.* 33(suppl.):266–274.
- CEPAL (Comisión Económica para América Latina y el Caribe). 1999. *Migración y desarrollo en América del Norte y Centroamérica: Una visión sintética*. Santiago, Chile: Naciones Unidas.
- Collins-Schramm, H. E., B. Chima, T. Morii et al. 2004. Mexican American ancestry-informative markers: Examination of population structure and marker characteristics in European Americans, Mexican Americans, Amerindians, and Asians. *Hum. Genet.* 114:263–271.
- Cordaux, R., N. Saha, G. R. Bentley et al. 2003. Mitochondrial DNA analysis reveals diverse histories of tribal populations from India. *Eur. J. Hum. Genet.* 11:253–264.
- Excoffier, L. 2001. Analysis of population subdivision. In *Handbook of Statistical Genetics*, D. J. Balding, M. Bishop, and C. Cannings, eds. Chichester, England: Wiley, 271–307.
- Green, L. D., J. N. Derr, and A. Knight. 2000. mtDNA affinities of the peoples of north-central Mexico. *Am. J. Hum. Genet.* 66:989–998.
- Gusmão, L., P. Sánchez-Diz, C. Alves et al. 2003. Results of the GEP-ISFG collaborative study on the Y-chromosome STRs *GATA A10*, *GATA C4*, *GATA H4*, *DYS437*, *DYS438*, *DYS439*, *DYS460*, and *DYS461*: Population data. *Forensic Sci. Int.* 135:150–157.

- Hall, T. A. 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41:95–98.
- Hamilton, G., M. Stoneking, and L. Excoffier. 2005. Molecular analysis reveals tighter social regulation of immigration in patrilocal populations than in matrilocal populations. *Proc. Natl. Acad. Sci. USA* 102:7476–7480.
- Harpending, H., and A. Rogers. 2000. Genetic perspectives on human origins and differentiation. *Annu. Rev. Genomics Hum. Genet.* 1:361–385.
- Hey, J., and C. A. Machado. 2003. The study of structured populations: New hope for a difficult and divided science. *Nat. Rev. Genet.* 4:535–543.
- Jiménez-Arce, G. 2000. *Variación del cromosoma Y en la población afrocostarricense de Limón, Costa Rica*. Ph.D. dissertation, Universidad de Costa Rica, San Pedro, San José.
- Kayser, M., S. Brauer, H. Schädlich et al. 2003. Y-chromosome STR haplotypes and the genetic structure of U.S. populations of African, European, and Hispanic ancestry. *Genome Res.* 13:624–634.
- Maza-Zavala, D. F. 1992. *Hispanoamérica-Angloamérica: Causas y factores de su diferente evolución*. Madrid, Spain: Editorial MAPFRE.
- Meléndez, E. L. 2001. *Polimorfismos del ADN mitocondrial en una muestra general de la población costarricense y su aplicabilidad en la identificación individual*. Ph.D. dissertation, Universidad de Costa Rica, San Pedro, San José.
- Melton, T., S. Clifford, M. Kayser et al. 2001. Diversity and heterogeneity in mitochondrial DNA of North American populations. *J. Forensic Sci.* 46:46–52.
- Nei, M. 1972. Genetic distance between populations. *Am. Natur.* 106:283–292.
- Nei, M. 1987. *Molecular Evolutionary Genetics*. New York: Columbia University Press.
- Pakendorf, B., and M. Stoneking. 2005. Mitochondrial DNA and human evolution. *Annu. Rev. Genomics Hum. Genet.* 6:165–183.
- Pérez-Lezaun, A., F. Calafell, D. Comas et al. 1999. Sex-specific migration patterns in Central Asian populations, revealed by analysis of Y-chromosome short tandem repeats and mtDNA. *Am. J. Hum. Genet.* 65:208–219.
- Rangel-Villalobos, H., A. R. Jaloma-Cruz, L. Cerda-Aguilar et al. 2001a. La huella genética del DNA en varones: Haplotipos del cromosoma Y en una población mexicana analizando cinco STRs. *Rev. Invest. Clin.* 53:401–406.
- Rangel-Villalobos, H., A. R. Jaloma-Cruz, L. Sandoval-Ramírez et al. 2001b. Y-chromosome haplotypes for six short tandem repeats (STRs) in a Mexican population. *Arch. Med. Res.* 32:232–237.
- Relethford, J. H. 2001. Genetic history of the human species. In *Handbook of Statistical Genetics*, D. J. Balding, M. Bishop, and C. Cannings, eds. Chichester, England: Wiley, 813–846.
- Salas, A., M. Richards, M. V. Lareu et al. 2004. The African Diaspora: Mitochondrial DNA and the Atlantic slave trade. *Am. J. Hum. Genet.* 74:454–465.
- Santos, M., R. H. Ward, and R. Barrantes. 1994. mtDNA variation in the Chibcha Amerindian Huetar from Costa Rica. *Hum. Biol.* 66:963–977.
- Schneider, S., J. M. Kueffer, D. Roessli et al. 1997. *Arlequin, Version 1.1*. Geneva: University of Geneva.
- Schneider, S., D. Roessli, and L. Excoffier. 2000. *Arlequin, Version 2.000: A Software for Population Genetics Data Analysis*. Geneva: University of Geneva.
- Tamura, K., and M. Nei. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* 10:512–526.
- Walsh, B. 2001. Estimating the time to the most recent common ancestor for the Y chromosome or mitochondrial DNA for a pair of individuals. *Genetics* 158:897–912.
- Wilder, J. A., S. B. Kingan, A. Mobasher et al. 2004. Global patterns of human mitochondrial DNA and Y-chromosome structure are not influenced by higher migration rates of females versus males. *Nat. Genet.* 36:1122–1124.

Copyright of Human Biology is the property of Wayne State University Press and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.