Autosomal, mtDNA, and Y-Chromosome Diversity in Amerinds: Pre- and Post-Columbian Patterns of Gene Flow in South America

Natalia R. Mesa,[1] María C. Mondragón,[1] Iván D. Soto,[1] María V. Parra,[1] Constanza Duque,[1] Daniel Ortíz-Barrientos,[1] Luis F. García,[2] Iván D. Velez,[3] María L. Bravo,[4] Juan G. Múnera,[5] Gabriel Bedoya,[1] Maria-Cátira Bortolini,[6,7] and Andrés Ruiz-Linares[1,7]

[1]Laboratorio de Genética Molecular, Departamento de Fisiología y Bioquímica, [2]Laboratorio Central de Investigaciones, and [3]PECET, Facultad de Medicina, [4]Departamento de Biología, and [5]Departamento de Antropología, Universidad de Antioquia, Medellín, Colombia; [6]Departamento de Genética, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil; and [7]Department of Medical and Community Genetics, Imperial College School of Medicine, Northwick Park Hospital, Harrow, United Kingdom

To evaluate sex-specific differences in gene flow between Native American populations from South America and between those populations and recent immigrants to the New World, we examined the genetic diversity at uni- and biparental genetic markers of five Native American populations from Colombia and in published surveys from native South Americans. The Colombian populations were typed for five polymorphisms in mtDNA, five restriction sites in the β-globin gene cluster, the DQA1 gene, and nine autosomal microsatellites. Elsewhere, we published results for seven Y-chromosome microsatellites in the same populations. Autosomal polymorphisms showed a mean G_{ST} of 6.8%, in agreement with extensive classical marker studies of South American populations. MtDNA and Y-chromosome markers resulted in G_{ST} values of 0.18 and 0.165, respectively. When only Y chromosomes of confirmed Amerind origin were used in the calculations (as defined by the presence of allele T at locus DYS199), G_{ST} increased to 0.22. G_{ST} values calculated from published data for other South American natives were 0.3 and 0.29 for mtDNA and Amerind Y chromosomes, respectively. The concordance of these estimates does not support an important difference in migration rates between the sexes throughout the history of South Amerinds. Admixture analysis of the Colombian populations suggests an asymmetric pattern of mating involving mostly immigrant men and native women.

Introduction

Genetic data have been used for several decades in the study of human evolutionary history (Cavalli-Sforza 1998; Jorde et al. 1998; Ruiz-Linares 1999). Most often, population analyses have examined allele frequencies at autosomal (biparental) genetic markers (Cavalli-Sforza et al. 1994). The incorporation of mtDNA during the 1980s added a powerful tool to the geneticist's tool kit, since mtDNA does not recombine and is transmitted only by women (Stoneking and Soodyall 1996). More recently, the increasing number of polymorphic markers identified on the nonrecombining region of the Y chromosome is allowing analyses of male lineages along the lines of those performed using mtDNA (Hammer and Zegura 1997). Furthermore, the comparative analysis of markers on autosomes, mtDNA, and the Y chromosome now enables the scrutiny of male and female demographic patterns.

Using this approach, Seielstad et al. (1998) proposed that women had a higher migration rate than men throughout human evolution. Those authors observed that the differentiation of allele frequencies between populations is often much greater for Y-chromosome than for mtDNA polymorphisms. On a worldwide scale, single nucleotide polymorphisms showed a level of population structure that was 3.5 times higher for Y-linked than for mtDNA markers. A greater differentiation of allele frequencies at Y-linked microsatellites (STRs) than at autosomal STRs was also reported for a set of African populations. Finally, it was observed that, in Europe, the greater genetic differentiation of populations at increasing geographic distance was more pronounced for Y chromosome than for nuclear or mitochondrial markers. Although a smaller effective population size for men than for women, due to polygyny, could be involved in producing these patterns of population structure, the magnitude of the difference across markers was deemed to be incompatible with polygyny being solely responsible. A preferred explanation is that these differences are due mostly to a lower migration rate of men. This observation might relate to the well-

known phenomenon of patrilocality (the tendency for women to move into the locality inhabited by their male partners), which has been documented in about two-thirds of human populations (Murdock 1981).

Comparison of uni- and biparental genetic markers has also allowed the study of differences in male/female gene flow during more recent times. For example, analyses of mtDNA and autosomal markers in Mexican and Mexican American populations has revealed a frequency of Amerind mtDNA lineages in them of ~90% (Merriwether et al. 1997; Green et al. 2000). This is considerably higher than was expected on the basis of autosomal estimates of Amerind admixture for these populations of ~30%. The difference is indicative of directional mating involving preferentially immigrant men and Amerind women. Similarly, the analysis of sex-specific and autosomal markers in Brazilian individuals of African ancestry evidenced a substantial European admixture that was mediated mostly through men (Bortolini et al. 1999).

Thus far, no comparison of uni- and biparental genetic diversity has been made in the same Native American populations. Here we examine five Colombian native populations using mtDNA, autosomal, and Y-chromosome markers. We use these and published data to compare male and female migration rates during the evolution of native South American populations. Most of the Colombian Amerind populations examined show some degree of admixture, and these data are also informative in relation to the pattern of gene flow between natives and nonnative immigrants having occurred during the past few centuries.

Subjects and Methods

Populations Studied

Five Native American populations living in "Resguardos" (native territories) from Colombia were examined in this study. These populations were selected

Table 1

Census Size and Linguistic Affiliation of the Colombian Amerind Populations Examined

Population	No. at Census	Linguistic Family	Linguistic Subfamily
Embera	71,000[a]	Chibchan-Paezan	Paezan
Ingano	18,000[a]	Andean	Andean
Ticuna	21,000[b]	Equatorial-Tucano	Macro-Tucano
Wayuu	305,000[b]	Equatorial-Tucano	Equatorial
Zenu	34,000[a]	...[c]	...[c]

[a] From Arango and Sánchez (1998).
[b] Includes individuals living in neighboring countries (Grimes and Grimes 1996).
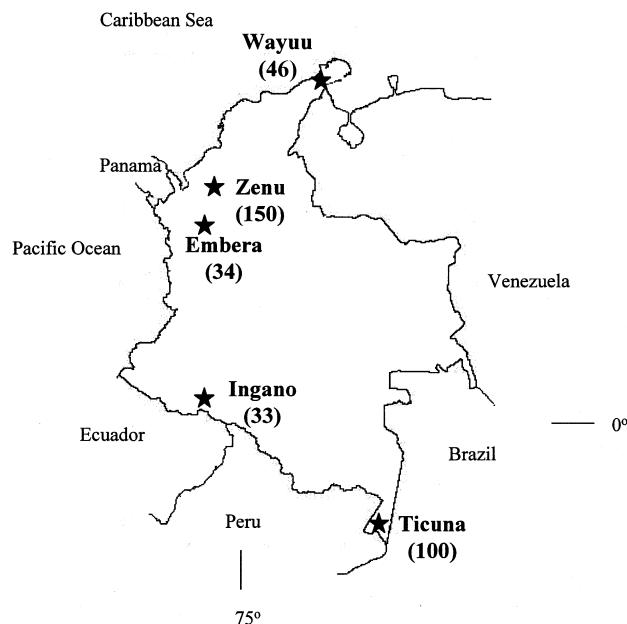[c] They have lost their native language (see text).

Figure 1 Map of Colombia showing the approximate geographic location of the five Amerind populations examined. The number of samples initially available for study is indicated in parentheses.

on the basis of their geographic dispersion within Colombia, their linguistic diversity, and their relatively large population size (table 1 and fig. 1). They include representatives of three of the four major linguistic subfamilies present in South America except Ge-Pano-Carib (Ruhlen 1991). The Zenu (also Senu or Sinu) currently speak only Spanish, and it is unclear what would be the linguistic affiliation of the language once spoken by them. According to Loukotka and Wilbert (1968), no record exists of the Zenu language. Nevertheless, Voegelin and Voegelin (1977) classify Zenu as a Carib language and indicate that Nutabe is a dialect of Zenu. However, this is in disagreement with Greenberg's classification, which includes Nutabe as a Chibchan language (Ruhlen 1991).

In every case, blood samples were collected from unrelated individuals. Each individual provided informed consent (this research received ethical approval of the Bioethics committee of the Universidad de Antioquia). Ticuna (also Tikuna, Tukuna, or Tucuna) samples were collected from the settlements of Arara ($n = 50$) and Cotuhé ($n = 50$), Departmento del Amazonas. Wayuu (also Guajiro, Goajiro, or Guajira) samples were collected in three localities: Guareripa ($n = 18$), La Loma ($n = 20$), and Ankalabary ($n = 8$), in the Departamento de la Guajira. Ingano (also Inga) samples were collected in San Miguel ($n = 10$) and Sibundoy ($n = 23$), Departamento del Putumayo. Zenu samples ($n = 150$)

were collected in San Andrés de Sotavento, Departamento de Córdoba. Embera (also Catíos, Katío, Chamí, or Chocó) samples ($n = 34$) were collected in Dabeiba, Departamento de Antioquia. DNA was extracted using standard protocols. Because of limited DNA availability, not all samples could be tested for every marker.

Autosomal Microsatellite Markers

The loci examined were LIPC, D15S100, D15S148, D15S102, D15S117, D15S125, D13S270, D13S126, and GABR3. These nine loci were included in the population survey of Bowcock et al. (1994), which included samples from Africa (Lisongo, Biaka, and Mbuti pygmies), America (Maya, Karitiana, and Surui), Asia (Cambodian, Chinese, and Japanese), Europe (Italian and northern European), and Oceania (Australian, Melanesian, and New Guinean). Microsatellites were typed radioactively or by fluorescent methods, using the primers and PCR conditions described in the Genome Database.

β-globin RFLPs

Five restriction-site polymorphisms were typed in the β-globin gene cluster by enzyme digestion of PCR products. The sites examined were *Hinc*II (5′ of ϵ), *Xmn*I (5′ of $^G\gamma$), *Hin*dIII (in IVS-II of the $^A\gamma$ and $^G\gamma$ genes), and *Hinc*II (3′ of $\psi\beta$). These polymorphic sites enable the identification of most of the 16 β-globin cluster haplotypes described in surveys of world populations (Flint et al. 1993). PCR primers, amplification conditions, and expected band sizes were as reported by Sutton et al. (1989). After amplification, PCR products were digested with the appropriate enzyme, run on agarose gels, and visualized by ethidium bromide UV fluorescence. Haplotype frequencies were calculated using the approach of Clark (1990). In ~20% of chromosomes, a restriction pattern was consistent with several haplotypes. In those cases, haplotype designations gave preference to those previously identified in Native Americans.

DQA1

Allelic variants at the DQA1 locus were typed by dot-blot oligonucleotide hybridization, using the PCR primers, probes, and hybridization conditions recommended by the 12th International Histocompatibility Workshop.

mtDNA Polymorphisms

Four restriction sites and the 9-bp COII/tRNA^{Lys} intergenic deletion variant were typed by PCR and agarose electrophoresis. The enzymes used and the location of the restriction sites examined (following Anderson's coordinates) were *Hae*III (663 bp), *Hinc*II (3,259 bp), *Alu*I (5,176 bp), and *Hae*III (6,517 bp). Experimental conditions were as described by Bailliet et al. (1994).

Data Analysis

Genetic diversity was estimated as $n/(n - 1)(1 - \Sigma p_i^2)$, where n is the sample size and p_i is the estimated allele frequency of allele i. In all cases (including linked markers), diversity was calculated for each individual polymorphism. Population structure was estimated by G_{ST}, defined as $(H_T - H_S)/H_T$, where H_T and H_S represent the expected heterozygosity under Hardy-Weinberg equilibrium for the total sample and the mean expected heterozygosity across populations, respectively (Nei 1987). Effective number of migrants (N_m) was calculated on the basis of Wright's island-migration model (Nei 1987). Ignoring mutation, at migration-drift equilibrium, the expectation is that $G_{ST} = 1/(4N_m + 1)$ for autosomal markers and $G_{ST} = 1/(N_m + 1)$ for uniparental markers. Confidence intervals around estimates of population structure were obtained by bootstrapping across loci. Differences between estimates were not considered significant if 95% confidence intervals overlapped (Weir 1996). For autosomal and Y-chromosome data, admixture estimation was performed by a least-squares approximation (Long 1991), using allele frequencies at each marker (i.e., haplotypic information was not used). For admixture calculations, classical marker data for 11 systems (ABO, Fy, Di, Tf, Rh, Hb, P, Gc, Kidd, Kell, and MN) were obtained by standard immunological and electrophoretic methods for the three populations sampled by N.R.M. and A.R.-L. (Embera, Inga, and Ticuna). Calculations were made using programs in the software packages ARLEQUIN (Schneider et al. 2000), GDA (Lewis and Zaykin 2000), GENETIX (Laboratoire Génome et Populations 2000), DISPAN (Ota 1993), RST CALC (Goodman 1997), and ADMIX (Long 1991).

Results

Autosomal Microsatellites

Figure 2 depicts the genetic relatedness of the 5 Colombian populations examined and of the 14 populations examined by Bowcock et al. (1994), on the basis of allele frequencies at nine STRs. This tree is consistent with that obtained by Bowcock et al. (1994), in that all populations cluster according to their continent of origin, and African populations, as a whole, appear to be more distantly related to the other populations. The separation of non-African continental populations is mostly undefined. Of interest, South American populations appear in a discrete cluster separate from the Mayan population from Central America. As expected, the boot-
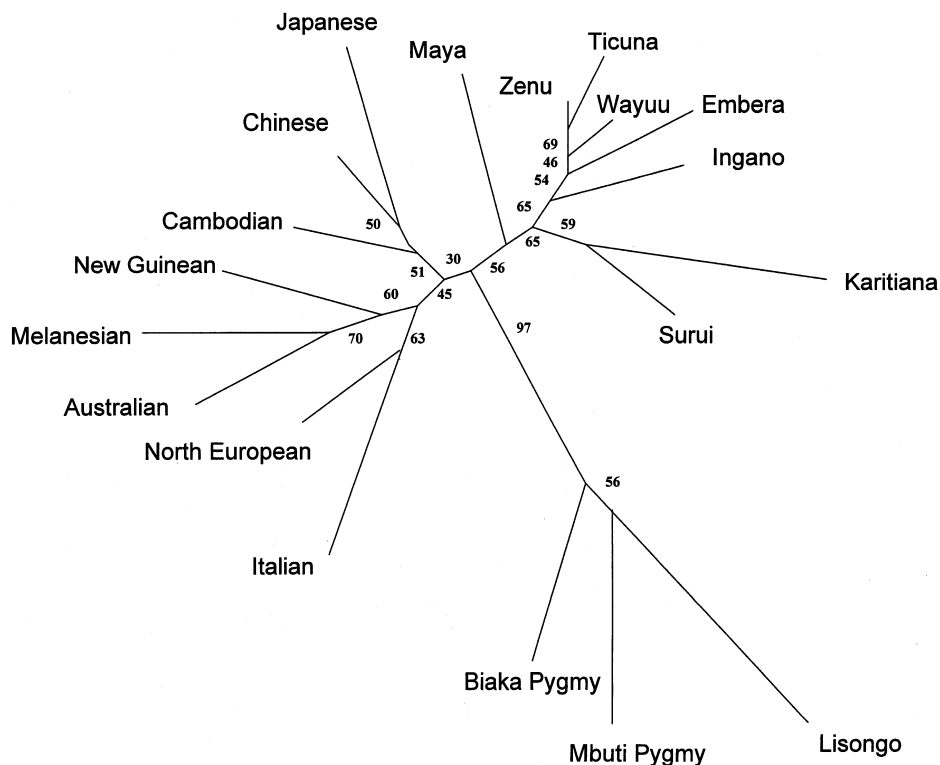
Figure 2   Neighbor-joining tree, based on nine autosomal microsatellites, relating 19 native populations from around the world. The tree was constructed from a pair-wise distance matrix, calculated using the DA genetic distance (Nei 1987). Numbers indicate percent bootstrap values for 2,000 replicates.

strap support for this tree is generally lower than that obtained by Bowcock et al. (1994), using 30 STRs.

β-Globin RFLPs

Table 2 compares haplotype frequencies at the β-globin gene cluster in the five Colombian populations with those reported for other Amerinds and for representative European and African populations. The haplotype frequencies observed in the Colombian populations agree with those seen in other Amerinds in showing a predominance of haplotypes 2 and 6. However, apart from the Ticuna, the populations we typed show less-extreme frequencies of these two haplotypes, together with the presence of additional haplotypes. Some degree of admixture is suggested by the presence in the Colombian populations (except in the Ticuna) of haplotypes 3 or 5, which are generally rare in Amerinds but have relatively high frequencies in African or European populations.

mtDNA Haplogroups

Table 3 shows the frequency in the Amerinds tested of the four major founder mtDNA lineages that have been identified in Native Americans (Torroni et al. 1993; Merriwether and Ferrell 1996). All the Ticuna individ-

uals examined carry haplotypes belonging to these haplogroups. In the other populations, ≤5% of individuals carry haplotypes lacking mutations diagnostic of founder lineages. These could represent novel founder haplogroups, back mutations at one of the diagnostic sites, or non-Amerind admixture. Preliminary mtDNA D-loop sequence analysis in 17 Colombian Amerind populations, including those examined here, suggests that haplotypes lacking characteristic RFLP changes mostly represent back mutations of C lineages (A.R.-L., unpublished data).

Intrapopulation Diversity

Table 4 compares the genetic diversity of the Colombian populations for the different sets of markers examined, including values calculated from Y-chromosome STR (Y-STR) data reported elsewhere (Ruiz-Linares et al. 1999). The Embera show some reduction of diversity for all systems examined, except for the β-globin gene cluster RFLPs, thus suggesting that this settlement is likely to have undergone a recent bottleneck. At these same markers, the Ticuna show a marked reduction in diversity. Only one Ticuna settlement was typed for the β-globin gene cluster RFLPs, and this settlement

Table 2

Percentage Frequency of β-Globin Gene Cluster Haplotypes in Native Populations from South America, Europe, and Africa

		HAPLOTYPE																
POPULATION	n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	REFERENCE
South America:																		
Embera	24	12.5	45.9	16.7			8.3	8.3	8.3									Present study
Ingano	36	13.9	49.9			16.7	16.7	2.8										Present study
Ticuna	46		86.9					2.2		4.4			6.5					Present study
Wayuu	58		70.7	1.7		15.5	5.2					6.4						Present study
Zenu	64	6.2	77.4	2.2		1.6	7.8	1.6					1.6			1.6		Present study
Mapuche	86	5.8	57.0	2.2	2.3	3.5	26.7	1.2								1.2		Kaufman et al. (1998)
Xavante	60		60.0			15.0	18.3	5.0									1.7	Bevilaqua et al. (1995)
Zoró	60		93.0				7.0											Bevilaqua et al. (1995)
Gavião	58	1.7	88.0	1.7			3.4	5.2										Bevilaqua et al. (1995)
Suruí	44		81.8			4.5	11.4					2.3						Bevilaqua et al. (1995)
Wai-Wai	56		87.5				12.5											Bevilaqua et al. (1995)
Yanomami	34	2.9	91.2				5.9											Guerreiro et al. (1994)
Kayapó	44		79.5				20.5											Guerreiro et al. (1994)
Wayana-Ap.	34		82.4				14.7			2.9								Guerreiro et al. (1994)
Wayampi	30		86.7			3.3	6.7	3.3										Guerreiro et al. (1994)
Arara	30		83.3		3.3		10.0			3.3								Guerreiro et al. (1994)
Europe:																		
England	37		43.8			40.0	13.5	2.7										Guerreiro et al. (1994)
Germany	16		43.8	6.2		31.3	6.2	12.5										Guerreiro et al. (1994)
Greece	64		46.7	1.6	1.6	14.1	25.0	6.3			1.6						3.1	Guerreiro et al. (1994)
Italy	169		66.9		0.6	21.3	11.2											Guerreiro et al. (1994)
Cyprus	120		70.0		2.5	20.0	7.5											Guerreiro et al. (1994)
Africa:																		
Bantu-speaking	47		12.8	40.4	10.6	10.6	14.9							2.1	4.3		4.3	Guerreiro et al. (1994)
Benin	30		3.3	56.7	13.3		6.7							10.0	3.3	6.7		Guerreiro et al. (1994)
Nigeria	35		8.6	60.0	17.1	11.4											2.9	Guerreiro et al. (1994)
Senegal	14		7.1	28.6	14.3	50.0												Long et al. (1990)

NOTE.—Populations examined in this article are underlined.

also shows some reduction in Y-STR diversity (data not shown), suggesting another instance of a recent population constriction.

The mean diversity values obtained for South America with the Y-chromosome and autosomal STR loci were 0.542 and 0.697, respectively. These values are similar to those reported for the same markers in other continents except for Africa, which shows a higher level of diversity (Bowcock et al. 1994; Jorde et al. 2000).

Interpopulation Diversity

Table 5 shows the level of genetic structure detected with the different sets of markers typed in the Colombian populations. The mean G_{ST} value observed for Y STRs is very similar to that obtained with mtDNA (0.165 vs. 0.18, respectively). These values are considerably higher than the mean value obtained from nuclear markers in the same populations (0.068). The increased variability of allele frequencies at uniparental markers is expected, given their smaller effective population size relative to autosomes. Noticeably, the estimated number of migrants obtained with the different genetic systems is very similar ($N_m \approx 4$). A bootstrap analysis indicates that the

difference between estimates of population structure based on mtDNA and Y markers are not statistically significant.

Table 5 also shows G_{ST} values estimated from data of previous surveys in other native South American populations for mtDNA, Y-chromosome STRs, and two of the three autosomal systems examined in the Colombian populations. The level of genetic structure observed in South America is generally somewhat higher than that for the Colombian populations, possibly because of the larger number of populations included and their wider geographic distribution. However, the estimated effective number of migrants is again similar for the different genetic systems ($N_m \approx 1$–3). Specifically, for mtDNA and Y STRs, the value of N_m is very close (2.5 and 2.3, respectively).

Admixture Analyses

Bi- and uniparental markers were also used to estimate levels of admixture, and results are shown in table 6. In the case of the mtDNA polymorphisms, the Amerind contribution was simply taken as the aggregate frequency of the major founding lineages, with allowance

Table 3

Percentage Frequency of mtDNA RFLP Haplogroups

		HAPLOGROUP				
POPULATION	n	A	B	C	D	Other
Embera	22	73	23	0	0	5
Ingano	27	15	44	37	0	4
Ticuna	54	13	15	39	33	0
Wayuu	40	25	35	38	0	3
Zenu	37	19	41	30	5	5

for the possibility that the other lineages observed could have an Amerind, African, or European origin. Although the admixture estimates carry large error margins, they broadly indicate that all populations, except the Ticuna, have considerable non-Amerind admixture (~20%–30%, on the basis of the autosomal results). This is in agreement with the greater geographical isolation of the Ticuna, who live in a remote Amazonian location (fig. 1). In the admixed populations, the Amerind component detected with nuclear markers is always smaller than that observed with mtDNA. This suggests that the non-Amerind component of the Native American gene pool was probably introduced through men. This is in agreement with the Y-STR–based estimates of admixture which, apart from the Ingano, show the smallest Amerind component of the three sets of markers. Estimates for the Ingano have the poorest fit to the admixture model used, as indicated by their relatively high mean square error (table 6). Taken together, these results are consistent with an asymmetric pattern of gene flow involving mainly immigrant men and Native American women.

Discussion

The comparison of genetic structure parameters based on uni- and biparental genetic markers, such as that made by Seielstad et al. (1998), has been criticized on the grounds that they did not include uniform sets of markers and populations (Stoneking 1998). In this article, we examined population structure for native South Americans on the basis of published data, as well as by typing a uniform set of markers in the same five Amerind populations. The level of genetic structure detected with autosomal markers in Colombian Amerinds is highly consistent with the one estimated for South America based on a large number of classical markers and populations (~6%) (Cavalli-Sforza et al. 1994). This figure is about twice the value observed in other continents, confirming that native populations from South America show the greatest differentiation in allele frequencies of any continent. This important variation in allele frequencies most likely reflects strong drift, probably because of the small effective size of tribal groups and the

common occurrence of founder effects throughout much of the evolutionary history of Amerind populations.

Estimates of population structure based on mtDNA and Y markers were found to be similar to each other (G_{ST} of 0.18 vs. 0.165, respectively), the difference not being statistically significant. This observation does not support an important difference in the long-term migration rate of men and women between the populations examined. However, the analysis of table 6 indicates that admixture could particularly affect estimates of population subdivision based on Y chromosome markers. We thus repeated the population-structure estimation, using only Y chromosomes of definite Amerind ancestry—those carrying allele T at locus DYS199 (Ruiz-Linares et al. 1999). This reduced the Y-chromosome data set by ~50%, and we also excluded the Ingano from the analysis, because, in that population, only one chromosome with the DYS199T allele was detected. In this restricted data set, G_{ST} increases slightly relative to that obtained when all the data are used (0.226 vs. 0.165, respectively), but the level of population structure is, again, not significantly different from that detected with mtDNA. As shown in table 5, using data from a survey of DYS199T chromosomes in 12 native South American populations, we obtained a G_{ST} of 0.30, which is almost identical to the value obtained from extensive surveys of mtDNA variation in South American natives (~0.29). Because the Ticuna show no evidence of admixture, we also examined levels of population structure between the two settlements sampled for this population. We obtained a G_{ST} for autosomal microsatellites of 0.017. This value is consistent with a G_{ST} of 0.02 that we calculated, using data for 15 classical markers reported in six Ticuna settlements from Brazil (Neel et al. 1980). G_{ST} estimates based on mtDNA and Y-chromosome markers in the Colombian Ticuna were, respectively, 0.238 and 0.205, again indicating no significant difference in population-structure estimates across markers.

Table 4

Genetic Diversity, Estimated with Uniparental and Biparental Genetic Markers

Population	mtDNA	Y STR[a]	β-globin	DQA1	STR
Embera	.242	.225	.402	.504	.655
Ingano	.350	.521	.482	.596	.687
Ticuna	.359	.403	.101[b]	.581	.682
Wayuu	.366	.583	.393	.566	.680
Zenu	.289	.653	.274	.601	.666
South America[c]	.380	.542	.353	.636	.697

[a] Estimated from data obtained by Ruiz-Linares et al. (1999).
[b] Only samples from one settlement were available for typing the β-globin gene cluster restriction sites.
[c] Estimated from the data obtained for the five Colombian populations.

Table 5

G_{ST} and Effective N_m Estimated with Uni- and Biparental Genetic Markers for Native Populations from Colombia And South America

	mtDNA	Y STRs	β-globin	DQA1	STRs
Colombia:					
G_{ST}	.180	.165	.087	.116	.061
N_m	4.5	5.1	2.6	1.9	3.8
Mean G_{ST}	.180	.165		.068	
Mean N_m	4.5	5.1		3.4	
South America:					
G_{ST}	.287[a]	.299[b]	.076[c]	.172[d]	Not done
N_m	2.5	2.3	3.0	1.2	Not done

[a] Data compiled by Bortolini and Salzano (1996) on 14 RFLPs typed in 20 native populations.

[b] Six STR loci in 12 populations, considering only chromosomes carrying the T allele at locus DYS199 (Bianchi et al. 1998).

[c] Three restriction sites in the β-globin gene cluster typed in 11 populations (Guerreiro et al. 1994; Bevilaqua et al. 1995).

[d] Allele frequencies in 21 populations (Cerna et al. 1993; Guedez et al. 1994; Yunis et al. 1994; Trachtenberg et al. 1996; Weg-Remers et al. 1997).

It has been suggested that biallelic markers on the Y chromosome show, on a worldwide scale, a higher level of population structure than Y STRs (Jorde et al. 2000). This difference might relate to the specific mutational properties of the markers, particularly the high mutation rate at STR loci and their propensity to recurrent mutation. To assess the effect of mutation on our estimates, we evaluated population structure on the basis of the Colombian Y-STR data, using the R_{ST} statistic of Slatkin (1995), which incorporates stepwise mutations. We obtained R_{ST} values of 0.2 and 0.23, using, respectively, all Y chromosomes or only those carrying the DYS199 T allele. The increase seen for the full data set relative to a G_{ST} estimate of 0.165 is probably due to the inclusion of some highly divergent haplotypes resulting from admixture. The almost identical G_{ST} and R_{ST} values obtained with the restricted data set (0.226 vs. 0.23, respectively) suggests that mutation has had little effect on the level of population structure detected. Furthermore, there is no significant correlation between estimates of population structure and locus variability (data not shown)—thus, there seems to be no evidence that our findings are strongly influenced by the mutational characteristics of the Y-STR loci examined.

In conclusion, our results suggest that there has been no important difference in the migration rate of males and females throughout the history of native South American populations. This could be indicating that neither matrilocality nor patrilocality has predominated. In fact, of the 34 South Amerind populations examined in the ethnographic study of Burton et al. (1996), half were considered predominantly patrilocal

and half predominantly matrilocal. There would thus seem to be a concordance between the type of social organization and the level of genetic structure detected with different types of genetic markers, as reported by Seielstad et al. (1998). However, ethnographic studies mostly document interactions occurring between groups within populations, and it is not clear what type of interaction predominates between populations, which is the level we have examined, or the effect that local interactions could have on a wider geographic scale (Stoneking 1998).

Admixture analysis suggests a pattern of directional mating that preferentially involved immigrant men and native women. Although it would be desirable to expand sample sizes, in order to have more accurate

Table 6

Admixture Estimates Based on Bi- and Uniparental Genetic Markers

POPULATION	PARENTAL CONTRIBUTION[a] (%)			MSE[b] (%)
	Native American	African	European	
Embera:				
Autosomal[c]	79 ± 11	21 ± 11	0	9
Y-STR	75 ± 12	11 ± 8	14 ± 11	13
mtDNA	95–100	0–5	0–5	
Ingano:				
Autosomal[d]	69 ± 11	31 ± 10	0	14
Y-STR	80 ± 21	0	20 ± 21	42
mtDNA	96–100	0–4	0–4	
Ticuna:				
Autosomal[d]	100 ± 2	0	0	8
Y-STR	97 ± 7	0	3 ± 7	12
mtDNA	100	0	0	
Wayuu:				
Autosomal[d]	82 ± 8	18 ± 8	0	10
Y-STR	48 ± 12	6 ± 6	46 ± 13	12
mtDNA	97–100	0–3	0–3	
Zenu:				
Autosomal[c]	75 ± 21	7 ± 20	18 ± 34	8
Y-STR	35 ± 34	3 ± 12	62 ± 36	29
mtDNA	95–100	0–5	0–5	

[a] Calculated on the basis of published data. In the case of classical markers, we used the compilations of Mourant et al. (1976), Roychoudhury and Nei (1988), Salzano and Callegari-Jacques (1988), and Cavalli-Sforza et al. (1994). For the other markers, the sources were as indicated in table 5, plus additional data for DQA1 (Schnittger et al. 1997; Luo et al. 1999; Migot-Nabias et al. 1999), β-globin (Hewitt et al. 1996), and Y STRs (Ruiz-Linares et al. 1996; Zago et al. 1996; Kayser et al. 1997; Perez-Lezaun et al. 1997; Rodriguez-Delfin et al. 1997; Bergen et al. 1999; Jorde et al. 2000). Autosomal STRs were not included in the calculations because of the scarcity of data for putative parental populations.

[b] Mean square error, which represents the proportion of allele frequency variation unexplained by the admixture model (Long 1991).

[c] Based on allele frequencies at the HLA DQA1 gene and at three restriction sites in the β-globin gene cluster.

[d] Based on allele frequencies at DQA1, three restriction sites in the β-globin gene cluster and 11 classical markers (see Subjects and Methods section).

estimates of admixture, the same pattern has been detected in non-Amerind populations from Mexico (Merriwether et al. 1997; Green et al. 2000) and Colombia (Carvajal-Carmona et al. 2000). This observation agrees with historical records showing that during the colonial period (late 15th to early 19th centuries), most of the immigrants to Spanish America were male (Sanchez-Albornoz 1977). This trend is also likely to have been reinforced by the preferential incorporation of immigrant women and their descendants into *criollo* society rather than into native American populations. The analysis of uniparental genetic markers in a non-Amerind population from Colombia has shown that Amerind male lineages have almost entirely disappeared from that population but that, at the mtDNA level, Amerind lineages are by far predominant (Carvajal-Carmona et al. 2000). This suggests a marked pattern of directional mating in which native women and immigrant men established neo-American populations. Concomitant with the demographic collapse of Amerinds during the first century of contact, this process seems to have led to the virtual "absorption" (through women) of native populations into the *criollo* population. Perhaps not surprisingly, the results presented here suggest that South Amerinds who still maintain their identity have frequently been exposed to similar demographic pressures.

Acknowledgments

Electronic-Database Information

The URL for data in this article is as follows:

The Genome Database, http://www.gdb.org/

References

Arango R, Sánchez E (1998) Los pueblos indígenas de Colombia. Tercer Mundo, Bogotá

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

Bergen AW, Wang CY, Tsai J, Jefferson K, Dey C, Smith KD, Park SC, Tsai SJ, Goldman D (1999) An Asian-Native American paternal lineage identified by RPS4Y resequencing and by microsatellite haplotyping. Ann Hum Genet 63:63–80

Bevilaqua LR, Mattevi VS, Ewald GM, Salzano FM, Coimbra CJ, Santos RV, Hutz MH (1995) β-globin gene cluster haplotype distribution in five Brazilian Indian tribes. Am J Phys Anthropol 98:395–401

Bianchi NO, Catanesi CI, Bailliet G, Martinez-Marignac VL, Bravi CM, Vidal-Rioja LB, Herrera RJ, Lopez-Camelo JS (1998) Characterization of ancestral and derived Y-chromosome haplotypes of New World native populations. Am J Hum Genet 63:1862–1871

Bortolini MC, Salzano FM (1996) mtDNA diversity in Amerindians and other human populations—how different are they? Braz J Genet 19:527–534

Bortolini MC, Silva WA, Castro-de-Guerra D, Remonatto G, Mirandola RM, Hutz MH, Weimer TA, Zago MA, Salzano FM (1999) African-derived South American populations: a history of symmetrical and asymmetrical matings according to sex revealed by bi and uniparental genetic markers. Am J Hum Biol 11:551–563

Bowcock AM, Ruiz-Linares A, Tomfohrde J, Minch E, Kidd JR, Cavalli-Sforza LL (1994) High resolution of human evolutionary trees with polymorphic microsatellites. Nature 368:455–457

Burton ML, Moore CC, Whiting JWM, Romney AK (1996) Regions based on social structure. Curr Anthropol 37:87–123

Carvajal-Carmonn LG, Soto ID, Pineda N, Ortíz-Barrientos D, Duque C, Ospina-Duque J, McCarthy M, Montoya P, Alvarez VM, Bedoya G, Ruiz-Linares A (2000) Strong Amerind/white sex bias and a possible Sephardic contribution among the founders of a population in northwest Colombia Am J Hum Genet 67:1287–1295 (in this issue)

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

Cavalli-Sforza LL, Menozzi P, Piazza A (1994) The history and geography of human genes. Princeton University Press, Princeton

Cerna M, Falco M, Friedman H, Raimondi E, Maccagno A, Fernández-Viña M (1993) Differences in HLA Class II Alleles of Isolated South American Indian Populations from Brazil and Argentina. Hum Immunol 37:213–220

Clark AG (1990) Inference of haplotypes from PCR-amplified samples of diploid populations. Mol Biol Evol 7:111–122

Flint J, Harding RM, Clegg JB, Boyce AJ (1993) Why are some genetic diseases common: distinguishing selection from other processes by molecular analysis of globin gene variants. Hum Genet 91:91–117

Goodman SJ (1997) RST CALC: a collection of computer programs for calculating unbiased estimates of genetic differentiation and determining their significance for microsatellite data. Mol Ecol 6:881–885

Green LD, Derr JN, Knight A (2000) mtDNA affinities of the peoples of North-Central Mexico. Am J Hum Genet 66:989–998

Grimes JE, Grimes BF (1996). Ethnologue: language family index to the thirteenth edition of the Ethnologue. Summer Institute of Linguistics, Dallas

Guedez YB, Layrisse Z, Dominguez E, Rodriguez-Larralde A, Scorza J (1994) Molecular analysis of MHC class II alleles

and haplotypes (DRB1, DQA1 and DQB1) in the Bari Amerindians. Tissue Antigens 44:125–128

Guerreiro JF, Figueiredo MS, Zago MA (1994) β-globin gene cluster haplotypes of Amerindian populations from the Brazilian Amazon region. Hum Hered 44:142–149

Hammer NF, Zegura SL (1997) The role of the Y chromosome in human evolutionary studies. Evol Anthropol 4:116–133

Hewitt R, Krause A, Goldman A, Campbell G, Jenkins T (1996) β-globin haplotype analysis suggests that a major source of Malagasy ancestry is derived from Bantu-speaking Negroids. Am J Hum Genet 58:1303–1308

Jorde LB, Bamshad M, Rogers AR (1998) Using mitochondrial and nuclear DNA markers to reconstruct human evolution. Bioessays 20:126–136

Jorde LB, Watkins WS, Bamshad MJ, Dixon ME, Ricker CE, Seielstad MT, Batzer MA (2000) The distribution of human genetic diversity: a comparison of mitochondrial, autosomal, and Y-chromosome data. Am J Hum Genet 66:979–988

Kaufman L, Carnese FR, Goicoechea A, Dejean C, Salzano FM, Hutz MH (1998) β-globin gene cluster haplotypes in the Mapuche Indians of Argentina. Genet Mol Biol 21: 435–437

Kayser M, Caglià A, Corach D, Fretwell N, Gehrig C, Graziosi G, Heidorn F, et al (1997) Evaluation of Y-chromosomal STRs: a multicenter study. Int J Legal Med 110:125–133

Laboratoire Génome et Populations (2000) GENETIX, logiciel sous WindowsTM pour la génétique des populations. Université de Montpellier II, France

Lewis P, Zaykin D (2000) Genetic data analysis: computer program for the analysis of allelic data. Department of Statistics, North Carolina State University, Raleigh

Long JC (1991) The genetic structure of admixed populations. Genetics 127:417–428

Long JC, Chakravarti A, Boehm CD, Antonarakis S, Kazazian HH (1990) Phylogeny of human β-globin haplotypes and its implications for recent human evolution. Am J Phys Anthropol 81:113–130

Loukotka C, Wilbert J (1968). Classification of South American Indian languages. Vol 7. Latin American Center, University of California, Los Angeles, Los Angeles

Luo M, Blanchard J, Pan Y, Brunham K, Brunham RC (1999) High-resolution sequence typing of HLA-DQA1 and -DQB1 exon 2 DNA with taxonomy-based sequence analysis (TBSA) allele assignment. Tissue Antigens 54:69–82

Merriwether DA, Ferrell RE (1996) The four founding lineage hypothesis for the New World: a critical reevaluation. Mol Phylogenet Evol 5:241–246

Merriwether DA, Huston S, Iyengar S, Hamman R, Norris JM, Shetterly SM, Kamboh MI, Ferrell RE (1997) Mitochondrial versus nuclear admixture estimates demonstrate a past history of directional mating. Am J Phys Anthropol 102:153–159

Migot-Nabias F, Fajardy I, Danze PM, Everaere S, Mayombo J, Minh TN, Renaut A, Georges AJ (1999) HLA class II polymorphism in a Gabonese Banzabi population. Tissue Antigens 53:580–585

Mourant AE, Kopepc AC, Domaniewska-Sobczak K (1976) The distribution of the human blood groups, and other polymorphisms. 2d ed. Oxford University Press, London

Murdock GP (1981) Atlas of world cultures. University of Pittsburgh Press, Pittsburgh

Neel JV, Gershowitz H, Mohrenweiser HW, Amos B, Kostyu DD, Salzano FM, Mestriner MA, Lawrence D, Simoes AL, Smouse PE, Oliver WJ, Spielman RS, Neel JV, Jr (1980) Genetic studies on the Ticuna, an enigmatic tribe of Central Amazonas. Ann Hum Genet 44:37–54

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Ota T (1993) DISPAN: genetic and phylogeny analysis. Penn State University, University Park, PA

Perez-Lezaun A, Calafell F, Seielstad M, Mateu E, Comas D, Bosch E, Bertranpetit J (1997) Population genetics of Y-chromosome short tandem repeats in humans. J Mol Evol 45:265–270

Rodriguez-Delfin L, Santos SE, Zago MA (1997) Diversity of the human Y chromosome of South American Amerindians: a comparison with blacks, whites, and Japanese from Brazil. Ann Hum Genet 61:439–448

Roychoudhury AK, Nei M (1988) Human polymorphic genes: world distribution. Oxford University Press, New York

Ruhlen M (1991) A guide to the world's languages. Stanford University Press, Stanford, CA

Ruiz-Linares A (1999) Microsatellites and the reconstruction of the history of human populations. In: Goldstein DB, Schlotterer C (eds) Microsatellites: evolution and applications. Oxford University Press, Oxford

Ruiz-Linares A, Nayar K, Goldstein DB, Hebert JM, Seielstad MT, Underhill PA, Lin AA, Feldman MW, Cavalli-Sforza LL (1996) Geographic clustering of human Y-chromosome haplotypes. Ann Hum Genet 60:401–408

Ruiz-Linares A, Ortiz-Barrientos D, Figueroa M, Mesa N, Munera JG, Bedoya G, Velez ID, Garcia LF, Perez-Lezaun A, Bertranpetit J, Feldman MW, Goldstein DB (1999) Microsatellites provide evidence for Y chromosome diversity among the founders of the New World. Proc Natl Acad Sci USA 96:6312–6317

Salzano FM, Callegari-Jacques SM (1988) South American Indians: a case study in evolution. Vol 6. Clarendon, Oxford

Sánchez-Albornoz N (1977) La poblacion de America Latina: desde los tiempos precolombinos al año 2000. 2d ed. Alianza Editorial, Madrid

Schneider S, Roessli D, Excoffier L (2000) ARLEQUIN version 2000: a software for population genetic data analysis. Genetics and Biometry Laboratory, University of Geneva, Geneva

Schnittger L, May J, Loeliger CC, Gallin MY, Erttmann KD, Bienzle U, Kremsner PG, Meyer CG (1997) HLA DRB1-DQA1-DQB1 haplotype diversity in two African populations. Tissue Antigens 50:546–551

Seielstad MT, Minch E, Cavalli-Sforza LL (1998) Genetic evidence for a higher female migration rate in humans. Nat Genet 20:278–280

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. Genetics 139:457–462

Stoneking M (1998) Women on the move. Nat Genet 20:219–220

Stoneking M, Soodyall H (1996) Human evolution and the mitochondrial genome. Curr Opin Genet Dev 6:731–736

Sutton M, Bouhassira EE, Nagel RL (1989) Polymerase chain reaction amplification applied to the determination of β-like globin gene cluster haplotypes. Am J Hematol 32:66–69

Torroni A, Schurr TG, Cabell MF, Brown MD, Neel JV, Lar-

sen M, Smith DG, Vullo CM, Wallace DC (1993) Asian affinities and continental radiation of the four founding Native American mtDNAs. Am J Hum Genet 53:563–590

Trachtenberg EA, Keyeux G, Bernal JE, Rhodas MC, Erlich HA (1996) Results of Expedicion Humana. I. Analysis of HLA class II (DRB1-DQA1- DPB1) alleles and DR-DQ haplotypes in nine Amerindian populations from Colombia. Tissue Antigens 48:174–181

Voegelin CF, Voegelin FM (1977) Classification and index of the world's languages. Elsevier, New York

Weg-Remers S, Brenden M, Schwarz E, Witzel K, Schneider PM, Guerra LK, Rehfeldt IR, Lima MT, Hartmann D, Petzl-Erler ML, de Messias IJ, Mauff G (1997) Major histocompatibility complex (MHC) class III genetics in two Amerindian tribes from southern Brazil: the Kaingang and the Guarani. Hum Genet 100:548–556

Weir BS (1996) Genetic data analysis II: methods for discrete population genetic data. Sinauer, Sunderland, MA

Yunis JJ, Ossa H, Salazar M, Delgado MB, Deulofeut R, de la Hoz A (1994) Major histocompatibility complex class II alleles and haplotypes and blood groups of four amerindian tribes of Northern Colombia. Hum Immunol 41:248–258

Zago MA, Silva WA, Tavella MH, Santos SEB, Guerreiro JF, Figuereido MS (1996) Interpopulational and intapopulational genetic diversity of Amerindians as revealed by six variable number of tandem repeats. Hum Hered 46:274–289