Original Paper

Human Heredity

Hum Hered 2001;51:97–106

Received: January 22, 1999
Revision received: September 14, 1999
Accepted: September 24, 1999

Genetic Diversity in an Andean Population from Peru and Regional Migration Patterns of Amerindians in South America: Data from Y Chromosome and Mitochondrial DNA

Luis A. Rodriguez-Delfin Verónica E. Rubin-de-Celis Marco A. Zago

Department of Clinical Medicine, School of Medicine, Ribeirão Preto, Brazil

Key Words
Mitochondrial DNA · Y chromosome · Haplotypes · Admixture · Quechua · Peruvian · Amerindian

Abstract
The genetic variability of a Quechua-speaking Andean population from Peru was examined on the basis of four Y chromosome markers and restriction sites that define the Amerindian mitochondrial DNA (mtDNA) haplogroups. Forty-nine out of 52 (90.4%) individuals had mtDNA which belonged to one of the four common Amerindian haplogroups, with 54% of the samples belonging to haplogroup B. Among 25 males, 12 had an Amerindian Y chromosome, which exists as four haplotypes defined on the basis of the DYS287, DYS199, DYS392 and DYS19 markers, three of which are shared by Amazonian Amerindians. Thus, there is a clear directionality of marriages, with an estimated genetic admixture with non-Amerindians that is 9 times lower for mtDNA than for Y chromosome DNA. The comparison of mtDNA of Andean Amerindians with that of people from other regions of South America in a total of 1,086 individuals demonstrates a geographical pattern, with a decreasing frequency of A and C haplotypes and increasing frequency of the D haplotype from the north of the Amazon River to the south of the Amazon River, reaching the lowest and the highest frequencies, respectively, in the more southern populations of Chile and Argentina. Conversely, the highest and lowest frequencies of the haplogroup B are found, respectively, in the Andean and the North Amazon regions, and it is absent from some southern populations, suggesting that haplotypes A, C and D, and haplotype B may have been dispersed by two different migratory routes within the continent.

Introduction

The autochthonous Andean populations are divided into two linguistic groups, Quechua and Aymara speakers. The Quechua are the main Andean population and is located in Peru, Bolivia and Ecuador. After the conquest of the Inca Empire by the Spanish in 1532, several other ethnic groups, especially of European and African origin, arrived in the region, and a variable degree of admixture occurred with the original Quechua inhabitants. The genetic heterogeneity of this population has been analyzed by classic protein markers, such as blood groups, serum proteins and erythrocyte enzymes. On the basis of the frequency of 10 loci, the Caucasian and African admixtures in the Quechua population have been estimated at 25 and 19%, respectively. Considering the frequency of the Rh alleles, an average admixture of 38% for the Quechua has been estimated [1], whereas for the Peru-

KARGER
Fax + 41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

0001–5652/01/0512–0097$17.50/0

Accessible online at:
www.karger.com/journals/hhe

Marco A. Zago
Department of Clinical Medicine, School of Medicine
Ribeirão Preto 14049-900 (Brazil)
Fax +55 16633 4009, E-Mail marazago@usp.br

vian population, the degree of genetic admixture has been calculated at 15% by the analysis of three loci [2].

Studies of mitochondrial DNA (mtDNA) variation between the Amerindians from North, Central and South America show that most of the mtDNA lineages belong to one of the four founding Amerindian haplogroups [3–9], and the distribution of the haplogroups follows a geographic variation pattern [10]. Additional Amerindian haplotypes have been proposed, some of which were found in Asian populations. The presence or absence of the *Hae*III 16517 site subdivides the Amerindian haplogroups into eight subtypes (A1, A2, B1, B2, C1, C2, D1, D2) [7, 11, 12]. Two haplotypes are defined by the lack of the markers identifying the four haplogroups A, B, C and D, the presence of the restriction sites *Dde*I 10394 and *Alu*I at 10397, and the absence or the presence, respectively, of *Hae*III at 16517 [13, 14]. However, there is still discordance concerning the acceptance and interpretation of these haplotype variations [15–18]. Reanalysis of sequence and restriction fragment length polymorphism (RFLP) data led Forster et al. [16] to propose another Amerindian founding haplogroup that is defined by lack of the *Dde*I site at nucleotide (nt) 1715. Similar joint analysis by Santos et al. [9] demonstrated that there is a high degree of haplotypic diversity within the different Amerindian haplogroups [19, 20].

In contrast to mtDNA studies, a limited number of native American populations have been studied with regard to the genetic diversity of the Y chromosome. The comparison of Y-specific DNA markers in several populations has suggested a reduced genetic variability of the Y chromosome in Amerindian populations, and that most Amerindian Y chromosomes probably derive from a single or two paternal lineages [12, 21–28]. With regard to Y-specific polymorphic markers, the most frequent alleles are the 186-bp allele at locus DYS19 [29], the 254-bp allele at locus DYS392, the 211-bp allele at locus DYS390, the 124-bp allele at locus DYS393, the T allele of the C→T mutation at locus DYS199 [30], and αII of the alphoid system [22].

A concomitant study of Y chromosome and mtDNA variations is suitable for examining genetic background, population structure in terms of the maternal and paternal DNA components, directionality of marriage patterns and gene flow. This approach has been employed by different authors: in native Mexican populations [8], in the Sinai Peninsula [31], in the Ojibwa [30], in Hispanic and Anglo groups in the San Luis Valley in the USA [32], in Uruguayans [33], and in Ethiopians [34].

With the purpose of better understanding the genetic diversity of the Andean Amerindian population in relation to genes of male and female origins, we analyzed the mtDNA (by RFLP) and Y-chromosome-specific markers of a sample of Peruvian individuals of Quechua origin. Our results show that the frequency of the mitochondrial haplogroups A–D in Peruvians is more similar to the Andean Amerindians than to the Amazon Amerindians, and the geographic distribution of A–D haplogroups may be interpreted as the consequence of at least two migratory routes within the continent. The Y chromosome haplotypes were similar to those observed in the Amazon region, although there was one haplotype that was found only among Peruvians. In addition, there is a clear directionality of marriage that governs the gene flow between Amerindian and other populations.

Materials and Methods

This study included 52 Peruvians who live in the towns of Pasco (28 individuals) and Lima (24 individuals). All individuals studied were unrelated and apparently of nonmixed origin. Only those individuals who had at least one Quechua surname were included in the analysis. Additional information about the family background and place of birth was obtained by interviewing each participant.

DNA was obtained from whole blood by phenol-chloroform extraction and ethanol precipitation. The mtDNA was examined for the RFLPs at three sites and for the region V 9-bp deletion (8272–8289) that identify the four Amerindian haplogroups: haplotype A mtDNA is defined by the gain of a *Hae*III restriction site at nt position 663; haplotype B mtDNA by the presence of the 9-bp deletion; haplotype C mtDNA by the absence of a *Hin*cII 13259 site, and haplotype D mtDNA by the absence of an *Alu*I 5176 site. Two additional restriction sites were studied: *Alu*I at nt 10397, and *Dde*I at nt 10394. A positive *Hpa*I site at nt 3592 was used to identify a mtDNA lineage of African origin. The methods have been previously reported. Briefly, the segment for each site was amplified by PCR, digested with the appropriate restriction enzyme and analyzed by electrophoresis on polyacrylamide gel. To identify the 9-bp deletion, the amplified segment was directly analyzed by electrophoresis.

The 25 males were studied for four polymorphic Y-chromosome-specific loci: an Alu insertion (YAP) at DYS287, a point mutation of DYS199, and tetranucleotide and trinucleotide repeat loci DYS19 and DYS392. The primer sequences and PCR conditions have been published elsewhere [25, 35, 36].

For the comparison of mtDNA data, the Amerindian tribes were divided into four groups, on the basis of their geographical location: (1) Amazon tribes at the north of the Amazon River (North Amazon); (2) Amazon tribes at the south of the Amazon River (South Amazon); (3) tribes of the Andean Region (Andes), and (4) tribes in the southern regions of South America (Southern). The distribution of four Amerindian haplogroups among the four geographic regions was tested for homogeneity by the χ^2 test. The proportion of genetic variance attributable to geographic subdivision (structure) was estimated using the G_{ST} statistics [37, 38]. The relative contribution of

Table 1. Frequency and diversity of mtDNA haplogroups in Quechua speakers from the Peruvian Andes (present study) and other populations

Population	Frequency of Amerindian haplogroup, %							$h^{A/B}$
	n	A	B	C	D	χ^2	others	
Peruvians	52	3.9	53.8	17.3	19.2	–	5.8	0.651/0.609
Quechua[a]	19	26.3	36.8	5.3	31.6	–	–	0.731
Aymara[a]	172	6.4	67.4	12.2	14.0	–	–	0.510
Mexicans[b]	60	65.0	21.7	13.3	–	–	–	0.521
Yanomami[a]	83	–	6.0	72.4	9.6	12.0	–	0.455/0.312

$h^{A/B}$ = Diversity considering all haplogroups (A) and only A–D haplogroups (B). $\chi^2 = 4.58$; $p < 0.20$, considering haplogroups A–D for Quechua and Aymara populations.

[a] Data obtained from previous reports [10, 13, 14].

[b] The Mexican samples included 11 Mixtec, 6 Zapotec and 14 Mixed Mexican natives [8].

Table 2. Frequencies (%) of four Amerindian haplogroups in four regions of South America

Population	n	**A**	A1	A2	**B**	B1	B2	**C**	C1	C2	**D**	D1	D2	Other[2]	$h^{A/B}$
Amazon	415	**17.1**	7.2	9.5	**19.5**	14.3	0.3	**39.3**	19.7	23.8	**18.5**	7.5	12.9	6.3	0.743/0.715
North Amazon	262	**19.9**	8.2	9.5	**9.5**	13.8	0.4	**47.7**	22.4	24.1	**18.3**	5.2	11.6	4.6	0.691/0.662
South Amazon	153	**12.4**	3.2	9.2	**36.6**	16.1	0.0	**24.8**	9.7	22.6	**18.9**	16.1	17.7	7.2	0.753/0.718
Andean	244	**7.0**	0.7	3.7	**62.3**	56.9	7.3	**13.1**	3.6	8.0	**16.0**	4.4	12.4	1.6	0.566/0.552
Southern	427	**4.7**	1.7	2.6	**25.1**	19.4	4.7	**26.0**	3.0	20.7	**42.1**	8.2	39.6	2.1	0.691/0.678
Total	1,086	**9.9**	3.9	5.9	**31.3**	24.8	3.2	**28.2**	10.6	19.5	**27.3**	7.1	22.2	3.3	0.738/0.721

$h^{A/B}$ = Diversity considering all haplogroups (A) and only A–D haplogroups (B). Sources: previous reports [4–10, 13, 48, 49, 67, 76] and present study. 121 Amazon, 107 Andean and 195 Southern natives were not examined for the presence or absence of the *Hae*III 16517 site.

$\chi^2 = 222.6$; $p < 0.0001$, considering only haplogroups A, B, C and D for Amazon, Andean and Southern regions.

[1] North Amazon: 11 tribes; South Amazon: 9 tribes; Andean: Aymara and Quechua; Southern: 10 (4 ancient) native groups.

[2] Forty-one haplotypes include both additional Amerindian haplotypes and haplotypes shared by non-Amerindian populations.

each region to the total genetic differentiation (G_{ST}) was examined by removing from the analysis the subpopulation of one region each time and reestimating the G_{ST} values. The index of gene diversity was computed using the DISPAN program. Haplotype diversity (h) was calculated according to Nei [38].

Indirect methods have been developed for estimating the average level of gene flow among equilibrium and nonequilibrium populations with different geographic locations, from both DNA RFLP and sequence data and classic protein markers [39–44]. The gene flow as measured by the product Nm can be estimated from F_{ST} (G_{ST}) values on the basis of the island model of population structure [43, 45, 46], where N is the number of individuals in each subpopulation and m is the migration rate among subpopulations in each generation. An estimate of Nm can be obtained by the formula: $Nm = [(ds - 1)/2ds][1/G_{ST} - 1]$, where ds is the number of subpopulations sampled [43].

Genetic distances between the populations of the four regions were estimated from the frequencies of the four mtDNA haplogroups using the method of Reynolds et al. [47] (F_{ST}), and a phylogenetic tree was constructed with the UPGMA algorithm. Each restriction site was considered as one locus with two alleles. For comparative purposes, eleven Amerindian tribes of Central America were included in this analysis [5, 8, 48–50]. The reliability of the clusters obtained by UPGMA was tested by bootstrap with 1,000 replications, computing a majority-rule consensus tree. All analyses were carried out using the Phylip package [51].

Results

Mitochondrial DNA

Tables 1 and 2 compare the frequencies of the four mtDNA haplogroups A–D in Peruvians with other relevant native American populations, in a total of 1,086 individuals. All four Amerindian haplogroups were found in the Peruvians, and the frequencies of these haplotypes are similar to those of the Quechua and Aymara Andean populations. The haplogroups A and B display the lowest and highest frequencies, with 4 and 50%, respectively.

Table 3. Comparison of the frequencies of positivity for *Dde*I at 10394 and *Alu*I at 10397 in haplotypes of the haplogroup B from relevant populations

Haplotype types	Population	N	Haplogroup B			Reference
			n	–/–	+/–	
	Peruvian	52	30	28	2	present study
AM118	Zapotec	15	5	4	1	8
17.25; 8.12	Mongolian	103	10	7	3	50
AS36,55	Vietnam	28	5	3	2	6, 53
AS60	Taiwan	20	8	7	1	6, 53
AS100	Korean	13	2	1	1	6, 53
1[Tharus b]	Tharus	107	8	0	5	77
Total		338	68	50	15	

N = Total number of individuals studied; n = number of haplotypes belonging to haplogroup B.

Table 4. Frequency of six Y chromosome lineages defined on the basis of four polymorphic markers in Peruvians

YAP	DYS199	DYS392	DYS19	Frequency, %
–	T	257	186[1]	20
–	T	254	186[1]	20
–	T	254	190[2]	4
–	C	254	186[1]	4
–	C	245	190	8
–	C	245	194	8
–	C	245	198	4
–	C	251	190	20
–	C	251	194	8
+	C	245	190	4

[1] Haplotypes shared with Amazon tribes.
[2] Haplotypes found only in Peruvians. Other haplotypes were shared with nonnative populations.

With regard to the presence or absence of *Hae*III at nt 16517, each haplogroup can be classified into subtypes 1 and 2 [7]. This polymorphic site has been found in all Amerindian haplogroups of South America [9, 13, 14]. The frequencies obtained were 1.9% for A1 and A2; 42 and 8% for B1 and B2; 9.6 and 7.7% for C1 and C2, and 1.9 and 17.3% for D1 and D2 haplotypes.

Three out of 52 (5.8%) samples analyzed did not have all the specific Amerindian mtDNA restriction sites and were classified as non-Amerindian haplogroups. These other haplotypes are probably the product of admixture. One mtDNA that had the *Hpa*I site at nt 3592 was probably of African origin, since this marker identifies the major African haplogroup, L. [52]. Another haplotype defined by *Hinc*II at nt 13259, *Alu*I at nt 5176, *Hae*III at nt 16517, *Dde*I at nt 10394, the loss of *Hae*III at nt 663 and *Alu*I at nt 10397, and the lack of the 9-bp deletion, was found in two Peruvian mtDNA. The same haplotype has been observed in Anglo (12.6%) and Hispanic (1.1%) populations by Merriwether et al. [32], and corresponds to haplotype X4.

Two mtDNA haplotypes, one with and the other without the *Hae*III site at nt 16517, displayed the 9-bp deletion, but possessed a *Dde*I restriction site at nt 10394 and lacked *Alu*I at nt 10397. The Amerindian haplogroup B has always been associated with the loss of the *Dde*I 10394 and *Alu*I 10397 sites.

The distribution of the haplotype with the 9-bp deletion associated with the *Dde*I site at 10394 and absence of an *Alu*I site at 10397 in Asian and Amerindian populations indicates that this haplotype is of Asian origin and represents a founder haplotype for the native populations of America (table 3), with low frequency both in Asia and America (4.4%), and represents 22% of the haplogroup B (36.4% of haplogroup B in Asians and 8.6% in America). Five Asians and 2 Amerindians from Central and South America presented this haplotype. The phylogenetic tree shows that this haplotype forms an independent cluster within the lineage B [5, 8, 50, 53]. This haplotype was not found among the Amazon Amerindians.

Y Chromosome

Twenty-five Peruvian males were examined for four Y-chromosome-specific polymorphic markers (table 4). The insertion of the *Alu* element at locus DYS287 (YAP) was found in 1 Y chromosome. The allele C of DYS199 was observed in 14 subjects (56%). The four alleles of

Table 5. G_{ST} statistics by haplogroup and South America region

Condition		H_T	H_S	G_{ST}	G'_{ST}	Dif[1]
Haplogroup	A	0.204	0.196	0.038	5.0	
	B	0.452	0.376	0.170	21.0	
	C	0.413	0.378	0.084	10.9	
	D	0.372	0.349	0.061	8.0	
	All loci	0.360	0.325	0.099	12.7	
Region						
North and South Amazon, Andean		0.357	0.320	0.103	14.8	–2.1
Andean, South Amazon, Southern		0.345	0.323	0.063	9.2	+3.5
Andean, North Amazon, Southern		0.360	0.314	0.128	18.1	–5.4
North and South Amazon, Southern		0.363	0.342	0.059	8.6	+4.1

$G'_{ST} = Dm/H'_T$.
[1] Calculated as $(12.7–G'_{ST})$.

DYS392 (245, 251, 254 and 257 bp) were found with similar frequencies. For DYS19, the most frequent alleles were the 186-bp (44%) and 190-bp (36%) alleles.

The simultaneous analysis of four loci (/YAP/DYS199/DYS392/DYS19/) resulted in the identification of ten haplotypes, three of them with the T allele of DYS199 (11 chromosomes). Six of the seven haplotypes with the C allele of DYS199 were considered of nonnative origin, since these haplotypes were shared by Brazilian Blacks, Whites, Japanese and Africans: /+/C/245/, /–/C/245/ and /–/C/251/ lineages (13 chromosomes). The /–/C/254/ (4%), /–/T/254/ (24%) and /–/T/257/ (20%) lineages were observed in Amerindian populations (12 chromosomes). One haplotype (/–/T/254/190/) was found only in Peruvians (4%) and three others were shared with other native Amazon groups.

Admixture

Genetic admixture was estimated as the proportion of the population with non-Amerindian Y chromosomes or mtDNA haplotypes. For mtDNA, haplotypes that did not show all specific markers of the Amerindian haplogroups [5–7, 10] were considered to be non-Amerindian. The genetic admixture calculated on the basis of mtDNA haplotypes for the Peruvian population was 5.8%. The Y haplotypes specific for Amerindians have been defined elsewhere [28]. The genetic admixture of the same Peruvian population calculated on the basis of Y chromosome markers varies between 52 and 56%, depending on whether one considers as Amerindian only the 11 chromosomes with a T allele at DYS199 (56% admixture), or whether one includes as Amerindian the chromosome with the /–/C/254/186/ haplotype (52%).

Genetic Diversity

The diversity of Y chromosomes in the Peruvian sample (0.871) is high, due to the European/African admixture, and near to the value estimated previously for five Amazon tribes (0.899) [28]. When the diversity is estimated only for Y haplotypes which are specific for Amerindians, this value diminishes to 0.697. Comparison with the Amazon tribes shows that three of four haplotypes are shared with the Peruvian Andeans, indicating that there is lower degree of Y-specific genetic differentiation among the Andean than among the Amazon Amerindians. Similar results are obtained for mtDNA diversity (tables 1, 2). The haplotypic diversity of the Peruvian sample (0.609) and the Andean natives (0.552) is lower than for the Amazon (0.715) and the Southern (0.678) populations.

Geographic Structure and Gene Flow

The analysis of mtDNA variation in South America showed that 13% of genetic differentiation can be attributed to the geographic locations of Amerindian populations and that 87% of the variation occurs within each region. The B haplogroup is responsible for most of the geographic diversity (table 5), followed by haplogroup C. When we investigated the degree of contribution of each region to the total interregional variation (G'_{ST} = 12.7%), we observed that the North Amazon and Andean regions increase the differentiation level, while the South Amazon and Southern regions diminish it (table 5). This result is due to the high frequencies of haplogroups B and C in the former regions; in contrast, the frequency of haplogroup B in the South Amazon and Southern regions is half that of the Andean (table 2).

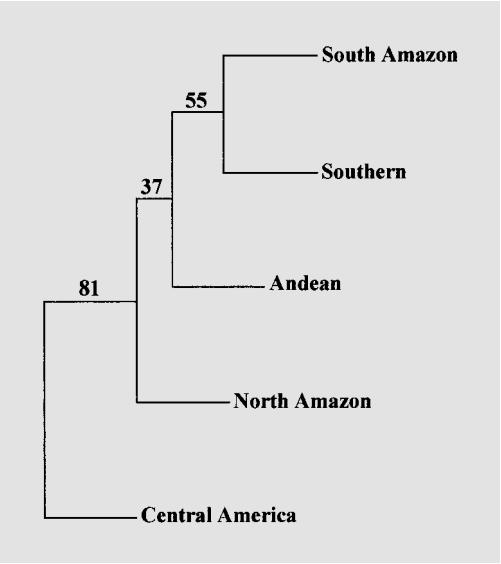

Fig. 1. A consensus tree calculated by the UPGMA algorithm from F_{ST} distances based on the mtDNA haplogroups obtained by RFLP analysis. The numbers indicate the percentage of times that the branch recovered in 1,000 replications.

Table 6. Genetic diversity index and Nm values for each region pair

Region pair	H_T	H_S	G_{ST}	Nm
North Amazon, South Amazon	0.316	0.343	0.051	4.7
North Amazon, Andean	0.357	0.302	0.153	1.4
Andean, Southern	0.335	0.306	0.086	2.6
Andean, South Amazon	0.326	0.316	0.031	7.8
South Amazon, Southern	0.357	0.347	0.027	8.9

$Nm = [(ds-1)/2ds][1/G_{ST}-1]$.

The degree of interpopulational differentiation ($G_{ST} \sim 10\%$) for the four South American regions is similar to the values obtained by analyses of nuclear genes, variable number of tandem repeats and RFLPs in Asian, European and American populations [29, 54, 55–57].

To test whether the low level of genetic differentiation ($G_{ST} = 9.2\%$) and the total diversity ($H_T = 0.345$) for the Andean, South Amazon and Southern regions, as compared with the other regions (table 5), was caused by gene flow, we analyzed the gene flow between each pair of regions. The gene flow measured as Nm (using G_{ST} values) probably does not represent the real flow and is not the best estimation of this demographic parameter [58], but this measure permitted us to observe the overall trend. Our analysis suggests that the gene flow between the Andean and South Amazon and the South Amazon and Southern groups (7.8 and 8.9) is higher than between the other regions (1.4–4.7) (table 6). Also, it suggests that geographic proximity is probably not the only factor that facilitates the gene flow among Amerindian tribes; for instance, the flow between the Andean and Southern regions was 2.6.

Phylogenetic Analysis

Since genetic drift is the major factor of variation of the A–D Amerindian haplogroup frequencies in native Americans, and assuming that recurrent mutations did not occur either in the ancestral (Paleo-Indians) or derived populations (Amerindians), the F_{ST} distance seems the most appropriate to represent the evolutionary relationship of the populations of the four geographic regions (table 2). The UPGMA tree obtained shows that the Andean population clusters in a different group in relation to the Amazon tribes, whereas the fact that the Southern and South Amazon Amerindians cluster into the same group indicates a closer relationship between the Southern natives and some Amazon tribes (fig. 1), especially those located at the south of the Amazon River. The separation of Andean, South Amazon and Southern tribes is less consistent (lower bootstrap values), probably due to genetic flow between the tribes from those regions.

Discussion

The present study of the genetic diversity of the Peruvian population provides: (1) the first information about Y chromosomes in the Quechua Amerindians; (2) additional data and diversity measurements on mtDNA RFLPs of this population; (3) admixture estimates for male and female non-Amerindians genes, and (4) a comparison of mtDNA data of the Amerindians from the Andean and the Amazon regions of South America that suggests an evolutionary relationship between these two Amerindian groups.

The comparison of markers of maternal and paternal origins permits us to distinguish the male and female contributions to the genetic admixture of the population. In native Amerindian populations, our results and data previously reported show that males are responsible for the major part of the European admixture, through a directional mating of European males and Amerindian females

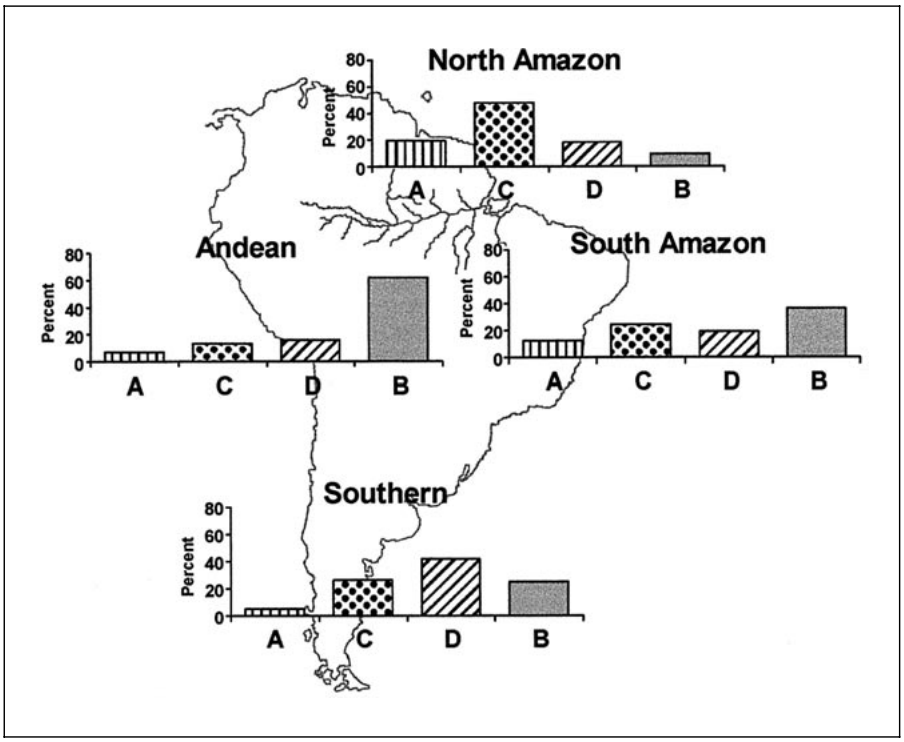

Fig. 2. Distribution of mtDNA haplogroups A, B, C and D in South American Amerindians.

[8, 30, 32]. A comparison of admixture indicates that the rate of nonindigenous genes of female origin in the Peruvian Quechua (5.8%) is times lower than that of nonindigenous genes of male origin (52%).

The average admixture calculated on the basis of our mitochondrial and Y chromosome data (30%) is twice the value obtained for Peruvians by Modiano et al. [2] with classic protein markers (15%), and near the 25–38% values obtained elsewhere for the Quechua [1]. The African admixture in Peruvians is lower; 1 of 5 males and 1 of 12 females.

This profound dissimilarity in terms of DNA inherited exclusively from the father (Y DNA) or from the mother (mtDNA) is not completely unexpected in view of the history of this population after the arrival of the Europeans. The first Spanish who settled in the region were soldiers and volunteer adventurers seeking wealth, rather than farmers with families. The predominance of European males in relation to females facilitated the male admixture with the natives. Furthermore, the social attitude during the first centuries of colonization encouraged or at least tolerated mating between European men and indigenous women, but condemned the opposite.

A comparison of our data with the literature on mtDNA RFLPs of native American populations of South America shows that the distribution of the four Amerindian haplogroups follows a defined variation pattern, with a high prevalence of each haplogroup confined to a specific geographic region (table 6, fig. 2), contrary to the idea proposed by Rothhammer and Bianchi [59]. This variation can be interpreted as a consequence of at least two population movements of the first Amerindians who entered the continent, followed by genetic drift for the populations in each region, which increased the genetic differences. The frequencies of the haplogroups A, C, and D suggest a migratory route of one group from the North Amazon region towards the South Amazon and the Southern (Argentina and Chile) regions, with a secondary flow to the Andes. The frequencies of the haplogroups A and C diminish from the north to the south of the Amazon River and to the Southern populations, while the haplogroup D frequencies increase in the same direction (north-south cline). The highest frequencies of the haplogroups C and D are found in the populations at the north of the Amazon River and in Southern natives. Conversely, the lowest and highest frequencies of haplogroup B occur in the North Amazon (9%) and Andean (61.5%) groups, respectively, suggesting a migratory route in a different direction to that of the haplogroups A, C, and D. The dispersion of the haplogroup B within South America

seems to have occurred from the Andean to the Amazon and Southern regions, but this movement would have occurred later, mainly after the expansion of the haplogroups C and D to the Southern region. This theory is reinforced by the low frequency or lack of the 9-bp deletion in some modern and ancient Southern groups, such as Yagham and Tierra del Fuego aborigines [59–61], and the finding of certain B haplotypes that are restricted to the Andes and have not been found among Amazon Amerindians (table 2). We must always consider that some genetic flow does occur between these populations, as we have demonstrated the occurrence of rare and complex gene rearrangements in individuals of tribes belonging to different linguistic groups from distant geographic regions [62].

The Y chromosome and mtDNA variation patterns do not permit a clear conclusion as to whether the Andean Quechua and Aymara populations are derived from the Amazon population [63, 64], or if they represent two independently differentiated Amerindian populations, one located in the Andes region and the other in the tropical forest [65]. Previous studies, including multivariate analyses comparing the gene frequency of several blood proteins and the HLA system, and dermatoglyphic, craniometrical and morphological features with linguistic and geographic/climatic variables, have produced controversial results [1, 57, 61, 63–72]. In contrast, our DNA marker data demonstrate that there exist significant genetic differences between the Amazon, the Andean and Southern natives for RFLP-mtDNA, whereas the differences in the Y-specific markers between Andean and Amazon natives are low, although the number of males studied is still too small. The differences in mtDNA haplogroup frequencies can be interpreted on the basis of two ancient migratory routes: one along the Pacific coast (mainly Amerindians with the B haplogroup), and the other within the Amazon region (mainly Amerindians with the A and C haplogroups). Although it is not necessarily so, these migratory movements could, finally, be related to two migration waves that could have entered America as proposed in the literature [48, 72–75], and that may be associated with two morphological patterns found in ancient and recent Amerindians [61]. The higher diversity of the B haplogroup as compared to the A, C and D haplogroups [9] in South American Amerindians would thus be explained by the accumulation of new haplotypes of the B haplogroup during the expansion of the first Amerindians in North, Central and South America.

Two additional alternatives must be considered. Selective advantage has not been described for any mtDNA haplogroup. By contrast, a rapid population expansion might conceivably produce the observed picture of the predominance of haplogroup B among the Andeans. The territorial expansion of the Inca Empire into a subcontinental power was accompanied by the submission and extinction of numerous small populations, probably followed by the rapid population growth of the Incas. If haplogroup B was the sole or predominant haplogroup in the small initial Inca population, genetic drift would have caused it to become the predominant haplogroup of the most important and organized pre-Columbian population of South America, which would also explain the present-day data on haplogroup distribution in South America.

The present distribution of mtDNA and Y chromosome DNA, as well as other genetic nuclear markers, is subject to the influence of both local selective (warfare and famine) and stochastic (genetic drift and founder effect) forces [1, 59]. Further investigations of ancient and contemporary Amerindian samples for Y chromosome, mtDNA and other nuclear markers may allow an interpretation of the present genetic variability of Amerindians and identification of the major ancient migratory routes during the expansion of Paleo-Indians in South America.

Acknowledgments

We especially thank Drs R. Cardenas and S. Olvizuri from Nacional Guillermo Almenara Hospital, Lima, Peru, for helping us collect the samples. This research was partly supported by CNPq and Fundação de Apoio à Pesquisa do Estado de S. Paulo (FAPESP), Brazil. The authors are indebted to Ms. Julia Y. Komoto and Ms. Marli H. Tavela for their expert technical help.

References

1 Salzano F, Callegari-Jacques SM: South American Indians: A Case Study in Evolution. Oxford, Clarendon Press, 1988.

2 Modiano G, Bernini L, Carter ND, Santachiara-Benerecetti AS, Detter JC, Baur EW, Paolucci AM, Gigliani F, Morpurgo G, Santolamazza C, Scozzari R, Terrenato L, Meera Khan P, Nijenhuis LE, Kanashiro VK: A survey of several red cell and serum genetic markers in a Peruvian population. Am J Hum Genet 1972; 24:111–123.

3 Schurr TG, Ballinger SW, Gan Y-Y, Hodge JA, Merriwether DA, Lawrence DN, Knowler WC, Weiss KW, Wallace DC: Amerindian mitochondrial DNAs have rare Asian mutations at high frequencies, suggesting they derived from four primary maternal lineages. Am J Hum Genet 1990;46:613–622.

4 Horai S, Kondo R, Nakagawa-Hattori Y, Hayashi S, Sonoda S, Tajima K: Peopling of the Americas, founded by four major lineages of mitochondrial DNA. Mol Biol Evol 1993;10: 23–47.

5 Torroni A, Schurr TG, Cabell MF, Brown MD, Neel JV, Larsen M, Smith DG, Vullo CM, Wallace DC: Asian affinities and continental radiation of the four founding Native American mtDNAs. Am J Hum Genet 1993;53:563–590.

6 Torroni A, Sukernik RI, Schurr TG, Starikovskaya YB, Cabell MF, Crawford MH, Comuzzie AG, Wallace DC: mtDNA variation of aboriginal Siberians reveals distinct, genetic affinities with Native Americans. Am J Hum Genet 1993;53:591–608.

7 Bailliet A, Rothhammer F, Carnese FR, Bravi CM, Bianchi NO: Founder mitochondrial haplotypes in Amerindian populations. Am J Hum Genet 1994;54:27–33.

8 Torroni A, Chen Y-C, Semino O, Santachiara-Beneceretti AS, Scott CR, Lott MT, Winter M, Wallace DC: mtDNA and Y-chromosome polymorphisms in four Native American populations from Southern Mexico. Am J Hum Genet 1994;54:303–318.

9 Santos SEB, Ribeiro-Dos-Santos AKC, Meyer D, Zago MA: Multiple founder haplotypes of mitochondrial DNA in Amerindians revealed by RFLP and sequencing. Ann Hum Genet 1996;60:305–319.

10 Merriwether DA, Rothhammer F, Ferrell RE: Distribution of the four founding lineage haplotypes in Native Americans suggests a single wave of migration for the New World. Am J Phys Anthropol 1995;98:411–430.

11 Bianchi NO, Bailliet A, Bravi CM: Peopling of the Americas as inferred through the analysis of mitochondrial DNA. Braz J Genet 1995;18: 661–668.

12 Bianchi NO, Bailliet A: Further comments on the characterization of founder Amerindian mitochondrial haplotypes. Am J Hum Genet 1997;61:244–246.

13 Easton RD, Merriwether DA, Crews DE, Ferrell RE: mtDNA variation in the Yanomami: Evidence for additional New World founding lineages. Am J Hum Genet 1996;59:213–225.

14 Merriwether DA, Hall WW, Vahlne A, Ferrell RE: mtDNA variation indicates Mongolia may have been the source for the founding population for the New World. Am J Hum Genet 1996;59:204–212.

15 Torroni A, Wallace DC: mtDNA haplogroups in Native Americans. Am J Hum Genet 1995; 56:1234–1236.

16 Forster P, Harding R, Torroni A, Bandelt H-J: Origin and evolution of Native American mtDNA variation: A reappraisal. Am J Hum Genet 1996;59:935–945.

17 Forster P, Harding R, Torroni A, Bandelt H-J: Reply to Bianchi and Bailliet. Am J Hum Genet 1997;61:246–247.

18 Stone AC, Stoneking M: mtDNA analysis of a prehistoric Oneota population: Implications for the peopling of the New World. Am J Hum Genet 1998;62:1153–1170.

19 Ward RH, Frazier BI, Dew-Jager K, Paabo S: Extensive mitochondrial diversity within a single Amerindian tribe. Proc Natl Acad Sci USA 1991;88:8720–8724.

20 Ward RH, Redd A, Valencia D, Frazier B, Paabo S: Genetic and linguistic differentiation in the Americas. Proc Natl Acad Sci USA 1993; 90:10663–10667.

21 Pena SDJ, Santos FR, Bianchi NO, Bravi C, Carnese FR, Rothhammer F, Gerelsaikhan T, Munkhtuja B, Oyunsuren T: A major founder Y chromosome haplotype in Amerindians. Nat Genet 1995;11:15–16.

22 Santos FR, Hutz MH, Coimbra CEA, Santos RV, Salzano F, Pena SDJ: Further evidence for the existence of a major founder Y chromosome haplotype in Amerindians. Braz J Genet 1995;18:669–672.

23 Santos FR, Gerelsaikhat T, Monkhtuja B, Oyunsuren T, Epplen J, Pena SDJ: Geographic differences in the allele frequencies of the human Y-linked tetranucleotide polymorphism DYS19. Hum Genet 1996;97:309–313.

24 Santos FR, Rodriguez-Delfin LA, Pena SDJ, Moore J, Weiss KM: North and South Amerindians may have the same major founder Y chromosome haplotype. Am J Hum Genet 1996;58:1369–1370.

25 Underhill PA, Jin L, Zemans R, Oefner PJ, Cavalli-Sforza LL: A pre-Columbian Y chromosome-specific transition and its implications for human evolutionary history. Proc Natl Acad Sci USA 1996;93:196–200.

26 Lell JT, Brown MD, Schurr TG, Sukernik RI, Starikouskaya YB, Torroni A, Moore LG, Troup GM, Wallace DC: Y chromosome polymorphisms in native American and Siberian populations: Identification of Native American Y chromosome haplotypes. Hum Genet 1997; 100:536–543.

27 Karafet T, Zegura SL, Vuturo-Brady J, Posukh O, Osipoval L, Wiebe V, Romero F, Long JC, Harihara S, Jin F, Dashnyam B, Gerelsaikhan T, Omoto K, Hammer MF: Y chromosome markers and Trans-Bering Strait dispersals. Am J Phys Anthropol 1997;102:301–314.

28 Rodriguez-Delfin LA, Santos SEB, Zago MA: Diversity of the human Y chromosome of South American Amerindians: A comparison with Blacks, Whites and Japanese from Brazil. Ann Hum Genet 1997;61:439–448.

29 Zago MA, Silva WA Jr, Tavella MH, Santos SEB, Guerreiro JF, Figueiredo MS: Interpopulational and intrapopulational genetic diversity of Amerindians as revealed by six variable number of tandem repeats. Hum Hered 1996; 46:274–289.

30 Scozzari R, Cruciani F, Santolamazza P, Sellito D, Cole DEC, Rubin LA, Labuda D, Marini E, Succa V, Vona G, Torroni A: mtDNA and Y chromosome-specific polymorphisms in Modern Ojibwa: Implications about the origin of their gene pool. Am J Hum Genet 1997;60: 241–244.

31 Salem A-H, Badr FM, Gaballah MF, Paabo S: The genetics of traditional living: Y chromosomal and mitochondrial lineages in the Sinai Peninsula. Am J Hum Genet 1996;59:741–746.

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

33 Bravi CM, Sans M, Bailliet G, Martínez-Marignac VL, Portas M, Barreto I, Bonilla C, Bianchi NO: Characterization of mitochondrial DNA and Y-chromosome haplotypes in a Uruguayan population of African ancestry. Hum Biol 1997;69:641–652.

34 Passarino G, Semino O, Quintana-Murci L, Excoffier L, Hammer M, Santachiara-Benerecetti AS: Different genetic components in the Ethiopian population, identified by mtDNA and Y-chromosome polymorphisms. Am J Hum Genet 1998;62:420–434.

35 Hammer MF, Horai S: Y chromosomal DNA variation and the peopling of Japan. Am J Hum Genet 1995;56:951–962.

36 Jobling MA, Tyler-Smith C: Fathers and sons: The Y chromosome and human evolution. Trends Genet 1995;11:449–456.

37 Nei M: Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci USA 1973;70: 523–538.

38 Nei M: Molecular Evolutionary Genetics. New York, Columbia University, 1987.

39 Takahata N, Palumbi S: Extranuclear differentiation and gene flow in the finite island model. Genetics 1985;109:441–457.

40 Slatkin M: Detecting small amounts of gene flow from phylogenies of alleles. Genetics 1989;121:609–612.

41 Slatkin M, Barton NH: A comparison of three indirect methods for estimating average levels of gene flow. Evolution 1989;43:1349–1368.
42 Slatkin M, Maddison WP: A cladistic measure of gene flow inferred from the phylogenies of alleles. Genetics 1989;123:603–613.
43 Hudson RR, Slatkin M, Maddison WP: Estimation of levels of gene flow from DNA sequence data. Genetics 1992;132:583–589.
44 Slatkin M: Isolation by distance in equilibrium and non-equilibrium populations. Evolution 1993;47:264–279.
45 Takahata F: Gene identity and genetic differentiation of population in the finite island model. Genetics 1983;104:497–512.
46 Crow JF, Aoki K: Group selection for a polygenic behavioral trait: Estimating the degree of population subdivision. Proc Natl Acad Sci USA 1984;81:6073–6077.
47 Reynolds J, Weir BS, Cockerhan CC: Estimation on the coancestry coefficient: Basis for a short-term genetic distance. Genetics 1983; 105:767–779.
48 Torroni A, Schur TG, Yang C-C, Szathmary EJ, Williams RC, Schanfield MS, Troup GA, Knowler WC, Lawrence DN, Weiss KM, Wallace DC: Native American mitochondrial DNA analysis indicates that the Amerind and the Nadene populations were founded by two independent migrations. Genetics 1992;130: 153–162.
49 Batista O, Kolman CJ, Bermingham E: Mitochondrial DNA diversity in the Kuna Amerinds of Panama. Hum Mol Genet 1995;4:921–929.
50 Kolman CJ, Sambuughin N, Bermingham E: Mitochondrial DNA analysis of Mongolian populations and implications for the origin of New World founders. Genetics 1996;142: 1321–1334.
51 Felsenstein J Phylip: Phylogeny Inference Package, Version 3.5c (distributed by the author). Seattle, Department of Genetics, University of Washington, 1993.
52 Chen Y-S, Torroni A, Excoffier L, Santachiara-Benerecetti AS, Wallace DC: Analysis of mtDNA variation in African populations reveals the most ancient of all human continent-specific haplogroups. Am J Hum Genet 1995; 57:133–149.
53 Ballinger SW, Schurr TG, Torroni A, Gan YY, Hodge JA, Hassan K, Chen K-H, Wallace DC: Southeast Asian mitochondrial DNA analysis reveals genetic continuity of ancient Mongolia migrations. Genetics 1992;130:139–152.
54 Livshits G, Nei M: Relationship between intrapopulational and interpopulational genetic diversity in man. Ann Hum Biol 1990;17:501–513.
55 Bortolini MC, Baptista C, Callegari-Jacques SM, Weimer TA, Salzano FM: Diversity in protein, nuclear DNA and mtDNA in South Amerinds – Agreement or discrepancy? Ann Hum Genet 1998;62:133–145.
56 Jorde LB, Bamshad MJ, Watkins WS, Zenger R, Fraley AE, Krakowiak PA, Carpenter KD, Soodyall H, Jenkins T, Rogers AR: Origins and affinities of modern humans: A comparison of mitochondrial and nuclear genetic data. Am J Hum Genet 1995;57:523–538.
57 Rothhammer F: Biodiversity of South American Indians. A collation of multifactorial and single gene marker systems between Andean and Tropical forest tribes. Braz J Genet 1995; 18:657–659.
58 Stoneking M, Jorde LB, Bhatia K, Wilson AC: Geographic variation in human mitochondrial DNA from Papua New Guinea. Genetics 1990; 124:717–733.
59 Rothhammer F, Bianchi NO: Origin and distribution of B mtDNA lineage in South America. Am J Hum Genet 1995;56:1247–1248.
60 Turbon D, Lalueza A, Perez-Perez A, Prats E, Moreno P, Pons P: Absence of the 9 bp mtDNA region V deletion in ancient remains of aborigines from Tierra del Fuego. Ancient DNA Newslett 1994;2:24–26.
61 Lahr MM: Patterns of modern human diversification: Implications for Amerindian origins. Yearb Phys Anthropol 1995;38:163–198.
62 Zago MA, Santos EJM, Clegg JB, Guerreiro JF, Martinson JJ, Norwich J, Figueiredo MS: α-Globin gene haplotype in South American indians. Hum Biol 1995;67:535–546.
63 Rothhammer F, Silva C: Peopling of Andean South America. Am J Phys Anthropol 1989;78: 403–410.
64 Rothhammer F, Silva C: Gene geography of South America: Testing models of population displacement based on archeological evidence. Am J Phys Anthropol 1992;89:441–446.
65 O'Rourke DH, Suarez BK: Patterns and correlates of genetic variation in South Amerindians. Ann Hum Biol 1985;13:13–31.
66 O'Rourke DH, Mobarry A, Suarez BK: Patterns of genetic variation in native America. Hum Biol 1992;64:417–434.
67 Fox CL: Mitochondrial DNA haplogroups in four tribes from Tierra del Fuego-Patagonia: Inferences about the peopling of the Americas. Hum Biol 1996;68:855–871.
68 Rothhammer F, Silva C, Callegari-Jacques SM, Llop E, Salzano F: Gradients of HLA diversity in South American Indians. Ann Hum Biol 1997;24:197–208.
69 Demarchi DA, Marcellino AJ: Dermatoglyphic relationships among South Amerindian populations. Hum Biol 1998;70:579–596.
70 Munford D, Zanini MC, Neves WA: Human cranial variation in South America: Implication for the settlement of the New World. Braz J Genet 1995;18:673–688.
71 Callegari-Jacques SM, Salzano FM, Constans J, Maurieres P: Gm haplotype distribution in Amerindians: Relationship with geography and language. Am J Phys Anthropol 1993;90:427–444.
72 Dugoujon J-M, Mourrieras B, Senegas M-T, Guitard E, Sevin A, Bois E, Hazout S: Human genetic diversity (immunoglobulin Gm allotypes) linguistic data and migrations of Amerindian tribes. Hum Biol 1995;67:231–249.
73 Torroni A, Neel JV, Barrantes R, Schurr TG, Wallace DC: Mitochondrial DNA 'clock' for the Amerinds and its implications for timing their entry into North America. Proc Natl Acad Sci USA 1994;91:1158–1162.
74 Wallace DC: Mitochondrial DNA variation in human evolution, degenerative disease and aging. Am J Hum Genet 1995;57:201–223.
75 Starikouskaya YB, Sukernik RI, Schurr TG, Kogelnik AM, Wallace DC: mtDNA diversity in Chukchi and Siberian Eskimos: Implications for the genetic history of ancient Beringia and the peopling of the New World. Am J Hum Genet 1998;63:1473–1491.
76 Kolman CJ, Bermingham E, Cooke R, Ward RH, Arias TD, Guionneau-Sinclair F: Reduced mtDNA diversity in the Ngobe Amerinds of Panama. Genetics 1995;140:275–283.
77 Passarino G, Semino O, Modiano G, Santachiara-Benerecetti AS: CO II/tRNALys intergenic 9-bp deletion and other mtDNA markers clearly reveal that the Tharus (Southern Nepal) have Oriental affinities. Am J Hum Genet 1993;53:609–618.