Am. J. Hum. Genet. 63:1862–1871, 1998

Characterization of Ancestral and Derived Y-Chromosome Haplotypes of New World Native Populations

Néstor O. Bianchi,[1] Cecilia I. Catanesi,[1] Graciela Bailliet,[1] Verónica L. Martinez-Marignac,[1] Claudio M. Bravi,[1] Lidia B. Vidal-Rioja,[1] René J. Herrera,[2] and Jorge S. López-Camelo[1]

[1]Instituto Multidisciplinario de Biología Celular, La Plata, Argentina; and [2]Department of Biological Sciences, College of Arts and Sciences, Florida International University, Miami

Summary

We analyze the allelic polymorphisms in seven Y-specific microsatellite loci and a Y-specific alphoid system with 27 variants (αh I–XXVII), in a total of 89 Y chromosomes carrying the DYS199T allele and belonging to populations representing Amerindian and Na-Dene linguistic groups. Since there are no indications of recurrence for the DYS199C→T transition, it is assumed that all DYS199T haplotypes derive from a single individual in whom the C→T mutation occurred for the first time. We identified both the ancestral founder haplotype, 0A, of the DYS199T lineage and seven derived haplogroups diverging from the ancestral one by one to seven mutational steps. The 0A haplotype (5.7% of Native American chromosomes) had the following constitution: DYS199T, αh II, DYS19/13, DYS389a/10, DYS389b/27, DYS390/24, DYS391/10, DYS392/14, and DYS393/13 (microsatellite alleles are indicated as number of repeats). We analyzed the Y-specific microsatellite mutation rate in 1,743 father-son transmissions, and we pooled our data with data in the literature, to obtain an average mutation rate of .0012. We estimated that the 0A haplotype has an average age of 22,770 years (minimum 13,500 years, maximum 58,700 years). Since the DYS199T allele is found with high frequency in Native American chromosomes, we propose that 0A is one of the most prevalent founder paternal lineages of New World aborigines.

Received March 31, 1998; accepted for publication October 14, 1998; electronically published November 17, 1998.

Address for correspondence and reprints: Dr. Néstor O. Bianchi, IMBICE, Calle 526 e/10 y 111900, La Plata, Argentina. E-mail: imbice@satlink.com

0002-9297/98/6306-0032$02.00

Introduction

Anthropological, archaeological, linguistic, odontological, and genetic tools have been used to reconstruct the history of the peopling of America. As a result of this multidisciplinary approach, it is generally accepted that the first colonization of America came from Asia during the last glaciation, through a Bering land bridge connecting both continents.

mtDNA is a molecule well suited to evolutionary studies, because of its maternal mode of inheritance, minimal recombination, and abundance of polymorphisms. Yet some of the hypotheses based on the interpretation of mitochondrial haplotypes are conflicting. Mitochondrial analysis has been invoked to support a multiwave-founder colonization of America (Torroni et al. 1993), whereas, on the other hand, mtDNA markers have also been interpreted as supporting a monophyletic colonization from Asia (Forster et al. 1996; Bonatto and Salzano 1997; Stone and Stoneking 1998). Archaeological studies seem to indicate an antiquity in the range of 11,000–33,000 years before the present (YBP) for the first settlements in Beringia and the New World (Hoeffecker et al. 1993; Szathmary 1993), whereas different laboratories working with mitochondrial polymorphisms have proposed times in the range of 14,000–55,000 YBP for this event (Horai et al. 1993; Torroni et al. 1994; Bonatto and Salzano 1997; Forster et al. 1997). Founder maternal Amerindian lineages initially had been estimated as being four (Torroni et al. 1993). Now it is assumed that there are ⩾10–13 such lineages, although there is no agreement on the molecular typification of some of these founder haplogroups (Bailliet et al. 1994; Merriwether et al. 1995; Forster et al. 1996, 1997; Bianchi et al. 1997). Some of those controversial hypotheses could perhaps be reinterpreted by resorting to the use of additional and complementary polymorphic DNA systems. In this regard, Y chromosome–specific regions are among the most promising.

The male-specific segment of the Y chromosome in mammals has no homologous counterpart, does not recombine, has all its genes in linkage disequilibrium, and

is paternally transmitted. In this regard, it is the male equivalent of mtDNA. Conversely, Y-specific genes are haploid, whereas mitochondrial genes are polyploid, and the mutation rate of Y DNA is much lower than that of mtDNA.

The association of two or more DNA markers defines a haplotype. Mitochondrial and Y-chromosome haplotypes are known to correlate with the ethnic origin of the population (Torroni et al. 1993; Hammer 1995; Merriwether et al. 1995; Pena et al. 1995; Underhill et al. 1996, 1997; Bianchi et al. 1997; Bravi et al. 1997; Karafet et al. 1997). We have reported elsewhere that the αh II form of the Y-specific alphoid satellite is associated with the A allele of the DYS19 microsatellite, giving rise to a Native American–specific haplotype that, thus far, has not been detected in any other geographic population (Pena et al. 1995). In 1996, Underhill et al. found that chromosomes sampled from speakers of Amerindian, Na-Dene, and Eskimo-Aleut languages exhibited the association of the DSY19A allele with a C→T transition at base position 181 of the DYS199 locus. Recently, Bianchi et al. (1997), by showing linkage disequilibrium of αh II, DYS199T, and DYS19A markers, defined the Native American aborigine Y chromosome with higher accuracy. The aim of the present report is to use additional polyallelic Y-specific markers to develop a better understanding of the origin and evolution of Native American Y chromosomes.

Material and Methods

Samples Analyzed

We analyzed a total of 280 Native American Y chromosomes; 88 (31.4%) were DYS199C, and 192 (68.6%) were DYS199T (see below). DYS199T chromosomes were the ones selected for further studies; in 89 of them, we could analyze the eight polyallelic markers indicated below, whereas in the remaining cases, because of the limited amount of DNA, we could test only some of these markers. Table 1 provides details on the populations, either analyzed by us or tested by other groups, in which the DYS199T allele was found.

We also analyzed the paternal lineages in 40 certified families provided by CEPH, which comprised a total of 249 father-son events. The identification of the CEPH families included in the present report is given in table 2. All these families are of Caucasian ancestry; their geographic origins were the United States (28 families), France (10 families), and Venezuela (2 families) (Dausset et al. 1990).

Y-Specific Markers

We studied one biallelic system and seven polyallelic systems. The biallelic marker is a C→T transition in the 181-bp position of the DYS199 locus (Underhill et al. 1996). The polyallelic markers are the αh alphoid system with 27 different forms (αh I–XXVII) (Santos et al. 1996*a;* Bianchi et al. 1997) and the microsatellites DYS19 (tetranucleotide; 10 alleles), DYS389a (tetranucleotide; 7 alleles), DYS389b (tetranucleotide; 9 alleles), DYS390 (tetranucleotide; 10 alleles), DYS391 (tetranucleotide; 6 alleles), DYS392 (tetranucleotide; 8 alleles), and DYS393 (trinucleotide; 6 alleles) (de Knijff et al. 1997; Kayser et al. 1997). The methods used for testing these polymorphic markers have been described elsewhere (Santos et al. 1993, 1996*a;* Underhill et al. 1996; Kayser et al. 1997).

In the present report, the two DYS199 alleles are identified as C or T, the αh alleles are designated by means of roman numbers, and microsatellites are identified by the number of repeats, as used in the reference tables of de Knijff et al. (1997) and Kayser et al. (1997).

Results

Allelic Frequencies

Table 3 lists the allelic frequencies for each of the polymorphic systems analyzed in the 40 CEPH paternal lineages and in Native American Y chromosomes. For the estimation of allelic frequencies in Native Americans, we used the 89 DYS199T chromosomes in which we could test all Y-specific markers, plus the 47 DYS199T cases in which only part of these markers were analyzed (table 1). All Native American samples were αh II, whereas CEPH chromosomes exhibited seven different forms of αh, with αh II being the most frequent (table 3). CEPH and Native American Y chromosomes showed the same predominant allele in the DYS389a, DYS390, DYS391, and DYS393 loci and showed different predominant alleles in the DYS19, DYS389b, and DYS392 microsatellites (table 2; allelic distributions for each Y chromosome are shown in tables 4 and 5).

The allelic frequencies found in the 40 CEPH lineages are similar to those reported by other authors, for European and other human geographic populations (Deka et al. 1996; Roewer et al. 1996; de Knijff et al. 1997; Kayser et al. 1997). The predominance of the DYS19/13, DYS389b/27, and DYS392/14 alleles in our series of American aborigines confirms previous reports by Pena et al. (1995), Santos et al. (1996*b,* 1996*c*), Bianchi et al. (1997), de Knijff et al. (1997), and Kayser et al. (1997).

Mutation Rates

The fidelity of father-son allele transmission was assessed in the 1,743 generation events comprised by the 40 CEPH paternal lineages (249 Y chromosomes × 7 microsatellite loci). In two families (21 and 66), we

Table 1

Populations Analyzed

Population	No. of Samples			Linguistic Phylum[b]	Geographic Region	Reference(s)
	Total	DYS199C	DYS199T (Full/Partial[a])			
Even	65	64	1	Altaic	Russia	Karafet et al. (1997)
Chukchi	30	25	5	Chukchi-Kamchatkan	Russia	Karafet et al. (1997), Lell et al. (1997)
Siberian Eskimo	56	46	10	Eskimo-Aleut	Russia	Karafet et al. (1997), Lell et al. (1997)
North American Eskimo	10	3	7	Eskimo-Aleut	United States	Karafet et al. (1997), Underhill et al. (1996)
Tanana	12	7	5	Na-Dene	United States	Karafet et al. (1997)
Navajo	80/10	39/4	41/6 (2/3)	Na-Dene	New Mexico	Karafet et al. (1997), Underhill et al. (1996), present study
Ojibwa	16	16	0	Amerindian	United States	Scozzari et al. (1997)
Cheyenne	45	37	8	Amerindian	United States	Karafet et al. (1997)
Havasupai	10	5	5	Amerindian	United States	Karafet et al. (1997)
Pima	24	14	10	Amerindian	United States	Karafet et al. (1997)
Pueblo	18	9	9	Amerindian	United States	Karafet et al. (1997)
Sioux	17	13	4 (2/1)	Amerindian	Minnesota	Present study
Zuni	16	7	9 (4/5)	Amerindian	United States	Present study
Seminole	25	13	12	Amerindian	Florida	Lell et al. (1997)
Maya	10	2	8 (6/1)	Amerindian	Yucatan, Mexico	Present study
Mixe	14	2	12	Amerindian	Mexico	Lell et al. (1997)
Mixteca	10	3	7	Amerindian	Mexico	Lell et al. (1997)
Zapoteca	21	9	12	Amerindian	Mexico	Karafet et al. (1997), Lell et al. (1997)
Chimila	11	4	7 (4/2)	Amerindian	Colombia	Present study
Wayuu	23	15	8	Amerindian	Venezuela	Karafet et al. (1997)
Guahibo	31	4	27	Amerindian	Venezuela	Scozzari et al. (1997)
Ticuna	1	0	1 (1/0)	Amerindian	Brazil	Present study
Arara	8	0	8	Amerindian	Amazonia, Brazil	Rodriguez-Delfin et al. (1997)
Kayapo	10	0	10	Amerindian	Amazonia, Brazil	Rodriguez-Delfin et al. (1997)
Wayam Apal	9	1	8	Amerindian	Amazonia, Brazil	Rodriguez-Delfin et al. (1997)
Wayampi	10	0	10	Amerindian	Amazonia, Brazil	Rodriguez-Delfin et al. (1997)
Yanomana	9	2	7	Amerindian	Amazonia, Brazil	Rodriguez-Delfin et al. (1997)
Karitiana	15	0	15	Amerindian	Brazil	Underhill et al. (1996)
Surui	17	1	16	Amerindian	Brazil	Underhill et al. (1996)
Lengua	24	8	16 (5/2)	Amerindian	Southern Paraguay	Present study
Ayoreo	9	4	5 (5/0)	Amerindian	Southern Paraguay	Present study
Wichí	26	2	24 (12/10)	Amerindian	Salta, Argentina	Present study
Toba	5	1	4 (4/0)	Amerindian	Salta, Argentina	Present study
Chorote	9	1	8 (5/3)	Amerindian	Salta, Argentina	Present study
Mocoví	20	7	13 (0/12)	Amerindian	Santa Fe, Argentina	Present study
Atacameño	7	2	5	Amerindian	Atacama, Chile	Present study
Susque	19	1	18 (18/0)	Amerindian	Jujuy, Argentina	Present study
Humahuaqueño	10	0	10 (10/0)	Amerindian	Jujuy, Argentina	Present study
Huilliche	25	13	12 (0/0)	Amerindian	Southern Chile	Present study
Mapuche	24	10	14 (3/5)	Amerindian	Rio Negro, Argentina	Present study
Pehuenche	21	6	15 (0/0)	Amerindian	Southern Chile	Present study
Tehuelche	16	3	13	Amerindian	Chubut, Argentina	Present study
Total	838/280	399/88	439/192 (89/47)			

Note.—Data from the present study are underlined.

[a] Full = full testing, which included all DYS199T chromosomes in which the seven microsatellites and αh systems could be analyzed; Partial = partial testing, which included only DYS199T chromosomes in which only one to six polyallelic systems could be analyzed.

[b] According to Ruhlen (1997).

Table 2

CEPH Haplotypes

	Haplotype								
Lineage	DYS199	α	DYS19	DYS389a	DYS389b	DYS390	DYS391	DYS392	DYS393
37	C	I	14	11	27	22	10	11	12
1375	C	I	15	10	30	24	10	11	13
1345	C	II	13	10	27	24	11	13	13
1346	C	II	14	9	25	24	10	13	13
1331	C	II	14	9	26	25	10	13	14
1377	C	II	14	9	27	25	11	13	14
1344	C	II	14	10	26	24	10	13	13
1333	C	II	14	10	26	24	11	11	13
02	C	II	14	10	26	24	11	13	13
1408	C	II	14	10	26	24	13	14	14
21	C	II	14	10	28	24	10	13	13
104	C	II	14	10	27	24	11	13	13
17	C	II	14	10	28	24	11	14	13
1418	C	II	14	10	27	25	11	13	13
35	C	II	14	11	27	23	11	13	13
884	C	II	14	11	27	24	10	11	13
1341	C	II	14	11	28	23	10	13	13
1416	C	II	14	11	28	25	11	11	13
13293	C	II	15	9	27	21	10	11	14
1334	C	II	15	10	26	24	11	13	13
1332	C	II	16	11	27	23	10	11	13
1349	C	II	16	11	29	25	11	11	13
1413	C	III	14	9	26	22	11	11	13
1350 and 13294	C	III	14	9	26	23	10	11	13
1362	C	III	14	10	28	22	10	11	13
1340	C	III	15	9	26	22	10	11	13
1424	C	III	15	9	26	24	10	11	12
12	C	III	17	9	26	25	11	11	14
1421	C	III	15	10	29	24	10	11	12
1347	C	III	15	11	30	23	10	12	15
1420	C	III	15	11	29	23	10	15	14
23	C	III	16	10	26	22	10	12	14
28	C	III	16	10	28	23	10	13	15
13291 and 1329	C	IV	15	9	27	22	10	11	14
45	C	IV	15	11	29	22	10	11	14
66	C	V	14	9	27	24	10	11	12
102	C	IX	15	10	27	21	10	12	15
1423	C	XII	14	11	30	25	10	11	14

found a mutation in an F_2 male. However, all the male offspring from these two apparently mutated males had the same allele as was seen in the F_1 ancestor and in all the other males in the lineage. Therefore, it is likely that the two mutations observed were somatic, probably produced by the lymphoblast transformation used to immortalize CEPH lymphocytes (Weber and Wong 1993). Thus, the rate of mutation was 0, with a 95% upper confidence-interval limit of .0025. Two other direct estimations of mutation rates for Y-specific microsatellites have been reported in the literature. Heyer et al. (1997) found three mutations in nine Y-specific microsatellite loci (which include the seven loci analyzed in the present report) and in 213 independent meiotic events; this combination of loci and meioses represents a total of 1,917 generations. In addition, Kayser et al. (1997) found two DYS19 slippage mutation events in 626 father-son pairs. If we pool the data from the present report with the data from the reports by Heyer et al. (1997) and Kayser et al. (1997), the average mutation rate (μ) is .0012, with 95% Poisson confidence-interval limits of .00046–.0028.

Since no αh mutation was found in the 40 CEPH lineages, the mutation rate of this system should be lower than 1/249 generation events.

Y Haplotypes

Tables 2 and 4 detail the combination of markers giving rise to Y-specific haplotypes. The 40 CEPH lineages

Table 3

Allelic Frequencies

Allele	Frequency in[a]	
	CEPH	Native American Sample (No. of Y ChromosomesTested)
DYS199:		
C	1	... (280)
T	...	1
αh:		
I	.050	...
II	.500	1
III	.300	... (97)
IV	.075	...
V	.025	...
IX	.025	...
XII	.025	...
DYS19:		
12		.008
13	.025	.817
14	.550	.150 (120)
15	.300	.025
16	.100	...
17	.025	...
DYS389a:		
9	.325	.081
10	.400	.676 (111)
11	.275	.243
DYS389b:		
24	...	.010
25	.025	.048
26	.325	.136
27	.325	.320 (103)
28	.150	.272
29	.100	.204
30	.075	.010
DYS390:		
21	.005	.008
22	.200	.024
23	.200	.248 (125)
24	.375	.568
25	.175	.152
DYS391:		
9	...	.069
10	.650	.857 (120)
11	.325	.074
13	.025	...
DYS392:		
11	525	...
12	.075	.019
13	.325	.179
14	.050	.490 (106)
15	.025	.198
16	...	.104
17	...	.009
DYS393:		
12	.100	.023
13	.550	.870 (131)
14	.275	.107
15	.075	...

Note.—The allele composition in each individual is given in tables 2 and 4.

[a] The most frequent allele in each group is underlined.

represent a sample of European Y chromosomes randomly selected. In this group we found 38 haplotypes, with two pairs of lineages (lineages 1350 /13294 and 13291/13292; table 2) sharing the same haplotype. Native American lineages, on the other hand, have Y chromosomes derived from a common ancestor (see below) and are not phylogenetically related to the CEPH sample. The aborigine group exhibited 61 haplotypes in 89 individuals, as a result of several Y chromosomes having the same haplotype (haplotype 0A and 2f, five cases each; haplotype 1c, six cases; haplotypes 1d, 2g, and 4a, three cases each; haplotypes 1a, 1e, 2d, 2h, 3b, 3e, 3f, 3i, and 3j, two cases each; see table 5). Table 5 shows the genetic diversity for each locus. The estimation of average genetic diversity for all loci (not shown in table 5) (Nei 1986, 1987) for CEPH and Native American samples is .565 ± .231 and .367 ± .279, respectively (Fisher 4.43; $P = .037$), significantly lower for the latter group of Y chromosomes.

By measuring the number of mutations separating the two most distant haplotypes, we can also generate an additional measurement of intragroup haplotype variability. For microsatellites, we accept the stepwise model of mutations (Ota and Kimura 1973); accordingly, a change in two repeats in the same locus is counted as two mutations. For the αh system, we follow the pathway of mutations detailed by Santos et al. (1996*a*). A pairwise comparison of CEPH haplotypes shows that 102 and 1349 are the pair of most distant haplotypes. For microsatellites, the distance between these two lineages is 12 mutational steps. Moreover, haplotype 102 has the αh IX form, and haplotype 1349 has the αh II variant; since there is no direct conversion between II and IX, both alleles have probably evolved independently from a V form, via two deletions for the V→II conversion and two duplications for the V→IX conversion (Santos et al. 1996*a*). Thus, the total number of changes between the 102 and 1349 CEPH lineages is 16. On the other hand, the two most distant Native American lineages (0A and 7a–7c), are separated by only seven microsatellite allelic shiftings (table 4).

Discussion

More than 2,500 Y chromosomes of wide geographic origin have been tested for the presence of the DYS199T allele (Santos et al. 1996*c*; Underhill et al. 1996, 1997; Bianchi et al. 1997; Karafet et al. 1997; Lell et al. 1997; Scozzari et al. 1997). Thus far, this allele has been found only in populations belonging to Amerindian (8 North American, 4 Central American, and 24 South American tribes), Na-Dene (2 tribes tested), Eskimo-Aleut (2 populations), Altaic (1 population), and Chukchi-Kamchatkan (1 population) linguistic phyla (table 1). The population frequency of the DYS199T allele is .35–.95, with

South American tribes and tribes with low genetic admixture showing the highest rates of the T allele (Santos et al. 1996*c*; Underhill et al. 1996,1997; Bianchi et al. 1997; Karafet et al. 1997). In North American tribes, the frequency of the T allele ranges from .0 (in the Ojibwa; table 1) to .56 (in the Zuni; table 1), with an average of .376; this relatively low prevalence of the marker could be due to non–Native American gene admixture, to a relevant incidence of DYS199C Native American founder haplotypes, or to a combination of both causes.

Since, thus far, no indication of recurrence for the DYS199T allele has been found, it seems reasonable to assume that all Y chromosomes exhibiting this allele derive from a single ancestor who carried the mutation for the first time. The 89 Y chromosomes listed in table 5 derive from this ancestor. In this regard, the DYS199T allele is equivalent to the Y-specific *Alu* insert, which is known to have occurred in a single individual in Africa and from which all extant YAP^+ Y chromosomes found worldwide derive (Hammer 1995). The phylogenetic association of Native American haplotypes listed in table 5 is further supported by the high frequency of shared haplotypes (43 of 89 cases; table 4) and by the low H for this group ($H = .367 \pm .279$), compared with the haplotype sharing (4 of 40 cases; table 2) and higher H for the CEPH group ($H = .565 \pm .231$).

It is usually accepted that, in a group of phylogenetically related individuals, the most frequent molecular markers represent shared ancestral characteristics that are due to the retention of traits found in the common ancestor (Stewart 1993). Thus, we attempted to reconstruct the Y haplotype of this ancestor, by combining the predominant alleles observed in each one of the seven loci analyzed (table 3). The haplotype obtained is 0A, which represents 5.7% of haplotypes listed in table 4. All other haplotypes listed in that table derive from 0A and can be sorted into seven haplogroups diverging from 0A by one to seven mutational steps. The letters following the numbers in the designation column of table 4identify the haplotypes within each haplogroup.

By using the Y microsatellite mutation rate (μ), .0012, we can calculate the age of 0A. The probability of observing one mutation in any of the seven loci analyzed is given by the binomial distribution $P_{(x)} = \binom{n}{k}\mu^k\ (1\ \mu)^{n-k}$, where n is the loci number and k is the number of mutations to occur; the figure obtained is .0083. The equation $M_d/P_{(x)}$, where $M_d = 7$ is the mutational distance between 0A and 7a, gives 843, which represents the number of generations expected to produce the most divergent haplogroup. If an average generation time of 27 years (Underhill et al. 1996) is assumed, the antiquity of 0A can be estimated to be 22,770 years, with minimum and maximum bounds of 13,500 and 58,700 years, respectively. These figures are approximately the same as estimates of the time of entry into America that are based on classic genetic markers (Cavalli-Sforza et al. 1994), mtDNA (Forster et al. 1997), and archaeological remains (Adovasio et al. 1990). table 4 shows that the frequency of nonancestral alleles varies for each microsatellite, suggesting a locus-specific mutation rate. In the future, the accurate estimation of these individual rates will serve to confirm or correct the estimation of the 0A age obtained by use of the average mutation rate. It is worth mentioning here that the finding of αh II in all Native American haplotypes indicates a mutation rate <1/22,770, or 4.4×10^{-4}, for this specific form.

Our conclusions in this section are based on the analysis of 89 Y chromosomes, representing a limited number of Native American populations. Therefore, we should raise the question of whether our assumptions can be extended to all DYS199T chromosomes. The characterization of the 0A ancestral haplotype is based on allele frequencies. The predominance of αh II and DYS19/13 alleles in Native American populations is well supported by data in the literature (Pena et al. 1995; Deka et al. 1996; Santos et al. 1996*b*, 1996*c*; Underhill et al. 1996; Bianchi et al. 1997). The most frequent alleles reported by us for microsatellites DYS389-393 have also been found by de Knijff et al. (1997) and Kayser et al. (1997), in independent analyzes of Mapuches, Tehuelches, Wichis, and Yanomami Amerindians. Moreover, Deka et al. (1996) have also detected the same predominant DYS390 allele in Bribri and Pehuenche Indians. It is interesting to mention here that Inuit (Canadian Eskimo) show the same predominant alleles as do Amerindians, for the DYS19, DYS389a, DYS389b, and DYS390 loci (see de Knijff et al. 1997; reference table to Kayser et al. 1997). The analysis in table 4shows that Y chromosomes at one and two mutational steps from 0A represent 40.4% of the total (36 cases); groups 3 and 4 comprise 38.2% of the lineages (34 cases); groups 5 and 6 have 12.3% of Y chromosomes (11 cases), and 3 cases (3.4%) show a genetic distance of seven mutational steps from 0A. This inverse correlation between haplotype frequency and genetic distance from 0A reflects the clustering of the most recently diverged haplotypes around 0A, a phenomenon expected to occur if the identification of the ancestral haplotype is correct. Furthermore, the decrease of haplotype frequencies at greater genetic distance from 0A indicates that the finding of Y chromosomes more distant than seven mutational steps is unlikely.

0A was found in two Wichi samples, two Toba samples, and one Susque sample. Although Wichis and Tobas belong to the same linguistic group (table 1), they are separated by a distance of 800 km, and no documented family relationship among them is known to exist. Moreover, the Susque sample belongs to a Central

Table 4

Native American Haplotypes

Haplotype Designation	Haplotype									No. of Cases	Population(s)[a]
	DYS199	αh	DYS19	DYS389a	DYS389b	DYS390	DYS391	DYS392	DYS393		
0A	T	II	13	10	27	24	10	14	13	5	Wichi, T, Susque
1a	T	II	14	10	27	24	10	14	13	2	Wichi, Tehuelche
1b	T	II	13	10	27	24	10	15	13	1	Wichi
1c	T	II	13	10	28	24	10	14	13	6	Toba, Ayoreo, Susque
1d	T	II	13	10	27	23	10	14	13	3	Lengua
1e	T	II	13	10	26	24	10	14	13	2	Maya, Susque
1f	T	II	13	10	27	24	10	13	13	1	Wichi
1g	T	II	13	10	27	25	10	14	13	1	Susque
1h	T	II	13	9	27	24	10	14	13	1	Navajo
2a	T	II	14	10	26	24	10	14	13	1	Mapuche
2b	T	II	13	10	26	24	10	15	13	1	Wichi
2c	T	II	13	10	27	25	10	15	13	1	Wichi
2d	T	II	13	10	28	24	10	15	13	2	Tehuelche, Chorote
2e	T	II	13	10	28	23	9	14	13	1	Sioux
2f	T	II	13	10	29	24	10	14	13	5	Chorote, Ayoreo, Susque
2g	T	II	13	10	27	25	9	14	13	3	Toba, Chimila
2h	T	II	13	11	27	25	10	14	13	2	Susque, Humahuaqueño
2i	T	II	13	10	25	24	10	14	13	1	Susque
2j	T	II	13	10	27	23	10	13	13	1	Mapuche
2k	T	II	13	11	28	24	10	14	13	1	Chorote
3a	T	II	13	10	29	24	10	15	13	1	Wichi
3b	T	II	13	10	28	23	10	15	13	2	Chorote, Toba
3c	T	II	13	10	28	25	10	15	13	1	Wichi
3d	T	II	14	10	28	24	10	15	13	1	Tehuelche
3e	T	II	13	11	28	25	10	14	13	2	Wichi, Humahuaqueño
3f	T	II	13	10	28	25	10	13	13	2	Chimila
3g	T	II	13	11	28	23	10	14	13	1	Sioux
3h	T	II	15	11	27	24	10	14	13	1	Toba
3i	T	II	13	11	29	24	10	14	13	2	Susque
3j	T	II	13	10	25	24	10	13	13	2	Susque, Mapuche
3k	T	II	13	10	26	23	10	13	13	1	Lengua
3l	T	II	13	10	27	23	11	15	13	1	Zuni
3m	T	II	13	10	27	23	10	13	14	1	Maya
3n	T	II	13	10	27	24	10	17	13	1	Zuni
4a	T	II	13	10	26	23	10	13	14	3	Susque, Wichi
4b	T	II	13	11	29	24	11	14	13	1	Maya
4c	T	II	14	10	29	24	10	13	13	1	Wichi
4d	T	II	13	10	26	22	10	13	13	1	Lengua
4e	T	II	14	10	26	23	10	15	13	1	Tehuelche
4f	T	II	14	10	26	24	10	15	12	1	Tehuelche
4g	T	II	13	11	27	23	10	15	14	1	Humahuaqueño
4h	T	II	13	10	29	23	10	14	14	1	Humahuaqueño
4i	T	II	13	9	26	24	10	16	13	1	Navajo
4j	T	II	13	10	29	23	10	15	13	1	Chorote
4k	T	II	13	11	29	24	10	15	13	1	Ticuna
4l	T	II	14	10	28	23	11	14	13	1	Susque
4m	T	II	13	10	26	23	10	13	14	1	Maya
5a	T	II	13	9	28	23	9	13	13	1	Maya
5b	T	II	13	11	28	23	10	15	14	1	Humahuaqueño
5c	T	II	13	10	28	23	10	16	14	1	Humahuaqueño
5d	T	II	14	9	26	24	11	15	13	1	Zuni
5e	T	II	14	10	29	24	10	13	14	1	Susque
6a	T	II	14	9	25	24	10	16	13	1	Zuni
6b	T	II	14	10	29	22	10	13	13	1	Ayoreo
6c	T	II	13	11	28	23	11	15	14	1	Humahuaqueño
6d	T	II	14	11	28	23	10	16	14	1	Humahuaqueño
6e	T	II	13	11	27	22	10	16	14	1	Humahuaqueño
6f	T	II	13	11	29	23	11	13	13	1	Maya

(*continued*)

Table 4 continued

Haplotype Designation	Haplotype									No. of Cases	Population(s)[a]
	DYS199	αh	DYS19	DYS389a	DYS389b	DYS390	DYS391	DYS392	DYS393		
7a	T	II	13	11	29	23	10	16	14	1	Humahuaqueño
7b	T	II	14	10	27	21	10	12	14	1	Tehuelche
7c	T	II	14	11	28	23	10	16	14	1	Susque

[a] Mutations of 0A are underlined.

Andean population. Thus, it seems very unlikely that the presence of 0A in these five males originated via a recent common ancestor. Moreover, if we group 0A with the most recently diverged haplotypes (haplogroups 1 and 2), we can observe that this cluster contains 13 populations belonging to different geographic regions and linguistic groups.

Rodriguez-Delfin et al. (1997) tested the DYS199 locus and the DYS19, DYS390, DYS392, and DYS393 microsatellite loci in the Y chromosomes of 45 Amerindians belonging to five Amazonian tribes. Determination of the most frequent microsatellite alleles in the 39 DYS199T chromosomes showed coincidence with our data, for the DYS19 and DYS393 loci, but not for the DYS390 and DYS392 loci. However, a critical analysis of the report by Rodriguez-Delfin et al. (1997) shows the same haplotype both in the entire Arara sample (eight cases) and in the eight Kayapo chromosomes carrying the DYS199T form. This clearly indicates that individuals from these two tribes are the offspring of two recent male ancestors. When Arara and Kayapo Y chromosomes are considered to represent 2, and not 16, different lineages, the microsatellite alleles that Rodriguez-Delfin et al. (1997) have reported as being most frequent are the same as those that are most frequent in our own study. Furthermore, 8% of Rodriguez-Delfin et al.'s DYS199T chromosomes are potential candidates to have the 0A haplotype.

On the basis of the aforementioned considerations, we think it reasonable to predict that an increase in the number of Native American Y chromosomes analyzed will allow us to identify new haplotypes in each haplogroup, will probably allow us to detect the presence of 0A in other Native American populations, and will perhaps allow us to demonstrate a correlation between a given allele frequency and the geographic origin of Native American populations. In this regard, it is interesting to mention that Humahuaqueño chromosomes (Dipierri et al. 1998) have a clear predominance of allele 14 in the DYS393 locus (table 4). On the other hand, no substantial changes in the characterization and dating of 0A are expected to result from an increase in the number of individuals studied. In fact, one additional mutational step in the distance from 0A will produce an increase of only 14% (3,253 years) in the estimated average age of 0A. Further analyzes of Y haplotypes in other Native American populations, mainly those from North and Central America, will be required for validation of our assumptions.

Several authors have proposed a multiwave colonization of the America. Neves and Pucciarelli (1991) and Powell et al. (1998) compared the cranial morphology of early South American remains with worldwide human remains from the Late Pleistocene and Holocene and reached the conclusion that the Americas were occupied by undifferentiated premongoloid populations before the migration of differentiated mongoloid colonizers to America. Roosvelt et al. (1996) analyzed Paleo-Indian campsites in the Brazilian Amazon and found evidence of a cultural tradition contemporary to but different from the Clovis Paleo-Indian culture of North America. The conclusion drawn from these results is that big-game hunters were probably not the only migrants into America. Linguistic, dental, and genetic data have been interpreted as supporting a three-wave migration into America, giving rise to Amerindians, Na-Denes, and Aleut-Eskimos (Greenberg et al. 1986). Moreover, Neel et al. (1994) found that many Amerindian tribes are endemically infected with the human T-cell lymphotrophic virus type II (HTLV-II) and that this viral infection is also present in native populations from Mongolia but not in those from Siberia. On the basis of these findings, Neel et al. (1994) have proposed that the ancestors of the first migrants to the New World entered the continent ~30,000 YBP and were not derived from north or central Siberia but from populations inhabiting Mongolia, Manchuria, or the extreme southeastern border of Siberia. Later migrations into America, on the other hand, would have originated in Siberia and would have been free of HTLV-II virus infection.

Recent studies using mtDNA markers to resolve the timing and number of prehistoric migrations into America propose founding times of 20,000–25,000 YBP (Forster et al. 1996) or 22,000–55,000 YBP (Bonatto and Salzano 1997). In the out-of-Asia hypothesis, Siberia is considered to be the geographic region of origin of Native American populations, and Beringia is given the role of a corridor (Forster et al. 1996). In the out-of-Beringia

Table 5

***H* for CEPH and Native American Y Chromosomes**

	H FOR			
			Subpopulation	
LOCUS	Total	Intrapopulation	CEPH	Native American
DYS199	.500	.000	.000	.000
αh	.414	.329	.665 ± .042	.000 ± .000
DYS19	.671	.453	.603 ± .040	.301 ± .034
DYS389a	.602	.568	.667 ± .014	.479 ± .030
DYS389b	.772	.755	.759 ± .024	.764 ± .011
DYS390	.722	.713	.746 ± .022	.694 ± .020
DYS391	.391	.363	.477 ± .039	.256 ± .035
DYS392	.777	.644	.617 ± .037	.680 ± .023
DYS393	.454	.419	.613 ± .041	.232 ± .032

NOTE.—Data are estimated according to the method of Nei (1986, 1987)

proposal, Beringia is assumed to be the place where Native American ancestors differentiated before migrating into the New World (Bonatto and Salzano 1997). In spite of the discrepancies, both groups of investigators coincide in suggesting that, after an early colonization event, the passage from Beringia to North America became obliterated by the coalescence of glaciers (18,000–12,000 YBP), isolating—and hence producing the genetic and linguistic differentiation of—the populations north and south of the glaciers. The discovery of a high frequency of DYS199T Y chromosomes in tribes belonging to Amerindian, Na-Dene, and Eskimo-Aleut linguistic groups (table 1) seems to favor the hypothesis of a common origin for all Native American populations (Szathmary 1984). Moreover, if future studies confirm our assumption, then the 0A haplotype could have been one of the predominant founder Native American lineages.

Acknowledgments

This work was supported by grants from National Council of Research, Argentina (CONICET); Research Commission of the Province of Buenos Aires, Argentina; National Agency for the Development of Science and Technology, Argentina (ANPCyt); and Argentine-Brazilian Center for Biotechnology (CABBIO). We wish to thank C. Horgan (IMBICE, Argentina) for his technical assistance, J. Dipierri (University of Jujuy, Argentina) for providing Susque and Atacameño samples, P. Underhill (Stanford University) for the Ticuna sample, and J. Ferrer (University of Pennsylvania) for providing Lengua and Ayoreo samples.

References

Adovasio JM, Donahue J, Stuckenrath R (1990) The Meadcroft rock shelter: radiocarbon chronology 1975–1990. Am Antiquity 55:348–353

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

Bianchi NO, Bailliet G, Bravi CM, Carnese FR, Rothhammer F, Martinez-Marignac VL, Pena SDJ (1997) Origin of Amerindian Y-chromosomes as inferred by the analysis of six polymorphic markers. Am J Phys Anthropol 102:79–89

Bonatto SL, Salzano FM (1997) A single and early migration for the peopling of the Americas supported by mitochondrial DNA sequence data. Proc Natl Acad Sci USA 94:1866–1971

Bravi CM, Sans M, Bailliet G, Martínez Marignac VL, Portas M, Barreto I, Bonilla C, et al (1997) Characterization of mitochondrial DNA and Y-chromosome haplotypes in a Uruguayan population of African ancestry. Hum Biol 69: 641–652

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

Dausset J, Cann H, Cohen D, Lathrop M, Lalouel JM, White R (1990) Centre d'Étude du Polymorphisme Humain (CEPH): collaborative genetic mapping of the human genome. Genomics 6:575–577

Deka R, Jin L, Shriver MD, Mei Yu L, Saha N, Barrantes R, Chakraborty R, et al (1996) Dispersion of human Y chromosome haplotypes based on five microsatellites in global populations. Genome Res 6:1177–1184

de Knijff P, Kayser M, Caglia A, Corach D, Fretwell N, Gehrig C, Graziosi G, et al (1997) Chromosome Y microsatellites: population genetic and evolutionary aspects. Int J Legal Med 110:134–140

Dipierri JE, Alfaro E, Martinez-Marignac VL, Bailliet G, Bravi CM, Cejas S, Bianchi NO (1998) Paternal directional mating in two Amerindian subpopulations located at different altitude in the northwest of Argentina. Hum Biol 70: 1001–1010

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

——— (1997) Reply to Bianchi and Bailliet. Am J Hum Genet 61:246–247

Greenberg JH, Turner CG, Zegura LZ (1986) The settlement of the Americas: a comparison of the linguistic, dental and genetic evidence. Curr Anthropol 27:477–497

Hammer MF (1995) A recent common ancestry for human Y chromosomes. Nature 378:376–378

Heyer E, Puymirat J, Dieltjes P, Bakker E, de Knijff P (1997) Estimating Y chromosome specific microsatellite mutation frequencies using deep rooting pedigrees. Hum Mol Genet 6:799–803

Hoffecker JF, Powerds RW, Goebel T (1993) The colonization of Beringia and the peopling of the new world. Science 259: 46–53

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

Karafet T, Zegura SL, Vuturo-Brady J, Posukh O, Osipova L, Wiebe V, Romero F, et al (1997) Y-chromosome markers and trans–Bering Strait dispersals. Am J Phys Anthropol 102:301–314

Kayser M, Caglia 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

Lell JT, Brown MD, Schurr TG, Sukernik IR, Starikovskaya YB, Torroni A, Moore LG, et al (1997) Y chromosome polymorphisms in Native American and Siberian populations: identification of Native American Y chromosome haplotypes. Hum Genet 100:536–543

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

Neel JV, Biggar RJ, Sukernik RI (1994) Virologic and genetic studies relate Amerindian origins to the indigenous people of the Mongolia/Manchuria/Southeastern Siberia region. Proc Natl Acad Sci USA 91:10737–10741

Nei M (1986) Definition and estimation of fixation indexes. Evolution 40:643–645

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

Neves WA, Puciarelli HM (1991) Morphological affinities of the first Americans: an exploratory analysis based on early South American human remains. J Hum Evol 21:261–273

Ota T, Kimura M (1973) A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. Genet Res 22:201–204

Pena SDJ, Santos FR, Bianchi NO, Bravi CM, Carnese FR, Rothhammer F, Gerelsaikhan T, et al (1995) A major founder Y-chromosome haplotype in Amerindians. Nat Genet 11:15–16

Powell J, Neves WA, Ozolins E, Pucciarelli HM. Afinidades biológicas extra-continentales de los dos esqueletos más antiguos de América: implicaciones para el poblamiento del Nuevo Mundo. Antropol Física Latinoam (in press)

Rodriguez-Delfin L, Santos SEB, 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

Roewer L, Kayser M, Dieltjes P, Nagy M, Bakker E, Krawczak M, de Knijff P (1996) Analysis of molecular variance (AMOVA) of Y-chromosome–specific microsatellites in two closely related human populations. Hum Mol Genet 5: 1029–1033

Roosvelt AC, Lima da Costa M, Lopes Machado C, Michab M, Mercier N, Valladas H, Feathers J, et al (1996) Paleoindian cave dwellers in the Amazon: the peopling of the Americas. Science 272:373–384

Ruhlen M (1997) A guide to the world's languages. Vol 1: Classification. Stanford University Press, Stanford

Santos FR, Bianchi NO, Pena SDJ (1996*a*) Worldwide distribution of human Y-chromosome haplotypes. Genome Res 6:601–611

Santos FR, Gerelsaikhan T, Munkhtuja B, Oyunsuren T, Epplen JE, Pena SDJ (1996*b*) Geographic differences in the allelic frequencies of the human Y-linked tetranucleotide polymorphism DYS19. Hum Genet 97:309–313

Santos FR, Pena SDJ, Epplen JT (1993) Genetic and population study of a Y-linked tetra-nucleotide repeat DNA polymorphism with a simple non-isotopic technique. Hum Genet 90:655–656

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

Scozzari R, Cruciani F, Santolamazza P, Sellitto D, Cole DEC, Rubin LA, Labuda D, et al (1997) mtDNA and Y chromosome–specific polymorphisms in modern Ojibwa: implications about the origin of their gene pool. Am J Hum Genet 60:241–244

Stewart C-B (1993) The powers and pitfalls of parsimony. Nature 361:603–607

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

Szathmary EJ (1984) Peopling of northern North America: clues from genetic studies. Acta Anthropogenet 8:79–109

——— (1993) Genetics of aboriginal North Americans. Evol Anthropol 1:202–220

Torroni A, Neel JV, Barrantes R, Schurr TG, Wallace DC (1994) Mitochondrial DNA "clock" for the Amerinds and its implications for timing their entry into North America. Proc Natl Acad Sci USA 91:1158–1162

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

Underhill PA, Jin L, Lin AA, Qasim Mehdi S, Jenkins T, Vollrath D, Davis RW, et al (1997) Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. Genome Res 7: 996–1005

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

Weber JL, Wong C (1993) Mutation of human short tandem repeats. Hum Mol Genet 2:1123–1128