Genetic Structure and Gene Flow in Gran Chaco Populations of Argentina: Evidence from Y-Chromosome Markers

DARÍO A. DEMARCHI[1] AND R. JOHN MITCHELL[2]

Abstract   The Gran Chaco region of central South America has been set-
tled by humans for only the last 4,000–5,000 years. To investigate population
structure and variation in this region's indigenous population, we scored
males from tribes of the Argentinean part of the Gran Chaco (Pilagá, Wichí,
and Toba, representing two major language groups, the Mataco and Guayc-
urú) for a number of Y-chromosome polymorphisms. The markers included
eight microsatellites (DYS19, DYS390, DYS391, DYS392, DYS393,
DYS437, DYS438, and DYS439) and the unique native American single nu-
cleotide polymorphism, DYS199. Sixty males (77%) from the total sample
carried the DYS199T chromosome, and these were the focus of the present
analysis. Unlike most other native Americans, Gran Chaco males show a
moderate level of diversity at the DYS19 locus but still less than that seen in
non–native Americans. The F_{ST} value for Y-chromosome markers in Gran
Chaco was 0.107, a value that is more than double that found for mtDNA
haplogroups in the same tribes but is not particularly high compared with
other Y-chromosome studies. Phylogenetic trees based on all eight microsa-
tellites showed relatively poor correlation of the tribes with either geography
or language, and this is possibly explained by their ecology. They are sea-
sonal hunters living in small bands, and under such circumstances drift will
be a powerful evolutionary force. An UPGMA tree based on five microsatel-
lites (DYS19, DYS390, DYS391, DYS392, and DYS393), however, showed
a more positive relationship, suggesting that DYS437, DYS438, and DYS439
may behave differently from the other microsatellites. No association was
found between maternal and paternal lineage distributions. The time to the
most recent common ancestor of the DYS199T chromosome is calculated to
lie between 13,000 and 26,000 years. This range is consistent with estimates
from other Y-chromosome studies as well as with estimates from mtDNA
and the archeology of the colonization of South America. We conclude that
the male lineages present in the contemporary Gran Chaco population reflect

[1]Departamento de Bioquímica Clínica, Facultad de Ciencias Químicas, and Museo de Antropología, Facul-
tad de Filosofía y Humanidades, Universidad Nacional de Córdoba, Córdoba, Argentina.
[2]Department of Genetics and Human Variation, School of Biochemistry and Genetics, La Trobe University,
Bundoora, Victoria, Australia.

KEY WORDS: Y CHROMOSOME, MICROSATELLITES, DYS199, DYS19, DYS390, DYS391,
DYS392, DYS393, DYS437, DYS438, DYS439, GRAN CHACO, AMERINDIANS, ARGENTINA, POPULA-
TION STRUCTURE, GENE FLOW, NETWORK ANALYSIS.

the level of diversity found in South America and that the region's male founders did not carry a restricted gene pool.

The Gran Chaco lies in the center of the South American continent, covering about 650,000 km^2 and encompassing parts of Argentina, Paraguay, and Bolivia. The region is a low-lying plain that is drained southward from Mato Grosso by the Paraguay River and southeastward by the Pilcomayo, Bermejo, and Salado Rivers, which originate near the Andean foothills. The southern Chaco becomes grassier and merges with the pampas, whereas northward along the Paraguay River in the east it becomes warmer and wetter and merges into tropical forest. Until the early 20th century, the Gran Chaco, especially the eastern and southern grasslands, had considerable amounts of game, such as guanaco, rhea, and many smaller animals. This accounts for the Quechua-given name *Gran Chaco,* or "great hunting" area (Steward and Faron 1959). The native inhabitants of the region generally lived in small bands of a few families and followed a simple nomadic life of hunter-fisher-gatherers. These populations have been grouped into several ethnolinguistic groups, named Mataco, Guaycurú, Tupí-Guaraní, Maskoy, Zamuco, and Lule-Vilela (Loukotka 1968).

Little is known about the origin of the Chaco tribes. The archeological data are scarce, and they mostly relate to sites of peripheral location, not to the core area. It is thought that the region has been inhabited for at least the last 4,000 or 5,000 years, and before this the entire area was an enormous swamp. Therefore the Gran Chaco could be one of the last regions in the world to be colonized by humans. Culturally as well as ecologically, the Gran Chaco is a transitional zone between the tropical plains of the Amazon basin and the pampas of Argentina. Along its western border it was wide open to influences from the Andean cultures, and in the east it abutted a subtropical region inhabited by Tupí-Guaraní tribes. With the Conquest, the Europeans gradually advanced into the region, initiating new movements of people and adding a new factor to the region's diversity.

Recently, we reported the pattern and extent of variation in founding mtDNA haplogroups in five native populations from the Argentinean part of the Gran Chaco. The Chaco tribes exhibit the highest intragroup variation and the lowest intergroup diversity compared to other indigenous populations of South America (Demarchi et al. 2001). This pattern suggests substantial female gene flow inside the region as well as with neighboring populations. Interestingly, this finding contrasts with results obtained from the analysis of morphological traits (dermatoglyphics and anthropometrics) for the same populations, which indicate substantial differentiation among Chaco populations (Demarchi et al. 1998; Demarchi 2000).

In this paper we report the nature and extent of variation of paternal lineages among the aboriginal populations of the Gran Chaco using nine Y-chromosome markers: one single nucleotide polymorphism (SNP), DYS199, and eight

microsatellites or short tandem repeats (STRs). We also test the hypothesis of differential gene flow of males and females by comparing the results obtained from Y-chromosome markers with those from mtDNA haplogroups.

Subjects and Methods

Populations. The Wichí (or Mataco), the Toba, and the Pilagá historically have inhabited the Argentinean part of the Gran Chaco. Their traditional way of life was based on fishing, hunting, and the gathering of different wild products, although today many migrate for employment to the sugar plantations in the Salta and Jujuy provinces and for other occasional jobs. The Wichí belong to the Mataco linguistic stock, whereas the Toba and Pilagá are Guaycurú-speaking tribes (Loukotka 1968). The Wichí today have a total population of about 24,000 distributed across the provinces of Salta, Chaco, and Formosa. The Pilagá inhabit the province of Formosa (most of them in Departamento Patiño) and number approximately 27,000, whereas the Toba today number about 39,000 and live in settlements in the provinces of Chaco, Formosa, Salta, and Santa Fé (Hernández 1992).

DNA samples were obtained from 78 unrelated males (with informed consent) from different settlements in the Gran Chaco region of northeastern Argentina: 19 Pilagás, 16 Wichís, and 15 Tobas from the province of Formosa, and 15 Wichís and 13 Tobas from the province of Chaco. The sampling locations are shown in Figure 1. The Wichís of Formosa and Chaco and the Tobas of Formosa and Chaco were not pooled for analysis because each population constitutes an independent mating system, isolated from the others by hundreds of kilometers of thorny forest.

Y-Chromosome Variation in South Amerindians. In addition to an investigation of the listed five populations, we performed an extended analysis, including multiple dimensional scaling and correlation matrix analysis, by including another three population samples from the Argentinean part of the Gran Chaco scored for some of the Y-chromosome polymorphisms typed in the present study. These were the Wichí, Chorote, and Toba from Salta province reported by Bianchi et al. (1998). Finally, we computed molecular variation at three levels: intrapopulation, within geographic region, and for South America, including data for South Amerindian populations reported by Bianchi et al. (1998) and Tarrazona-Santos et al. (2001), based on a reduced set of Y-chromosome markers.

Systems Analyzed. All individuals were scored for six tetranucleotide (DYS19, DYS390, DYS391, DYS393, DYS437, and DYS439), one trinucleotide (DYS392), and one pentanucleotide (DYS438) repeat Y-chromosome polymorphisms. Five of the classic set of Y-chromosome short tandem repeats (STRs) or microsatellites (DYS19, DYS390, DYS391, DYS392, and DYS393) were typed

Figure 1. Approximate area of the Gran Chaco region and the sampling locations of the five tribal groups.

with a series of multiplex PCRs using primer sequences, PCR protocols, and allele nomenclature as described by Kayser et al. (1997). In addition, three Y-chromosome STRs (DYS437, DYS438, and DYS439), reported by Ayub et al. (2000), were scored. Although DYS438 has a simple array structure, DYS437 and DYS439 are complex, and therefore the relationship between variability and modal number of repeats is not straightforward. These three Y-chromosome STRs are referred to here as novel STRs, and DNA of known allele size for each locus was donated by C. Tyler-Smith, University of Oxford. Primer sequences, PCR protocols, and allele size nomenclature used in this study are described by Ayub et al. (2000). The only modification was that DYS437, DYS438, and DYS439 were scored in a triplex PCR.

The DYS199 C-T transition is both unique to native Americans and extremely widespread among them (Pena et al. 1995; Underhill et al. 1996; Santos et al. 1999). Its universality could be due to its presence in the founder gene pool

of all native Americans or to a later geographic expansion of this haplogroup from its source population across the continents. The former idea is the most likely explanation. Previous Y-chromosome studies (Santos et al. 1999; Underhill et al. 2001) as well as mtDNA (Forster et al. 1996) and archeological data (Morell 1990) indicate that native American ancestors are about 20,000 years old. The DYS199 locus was typed using the protocol of Santos et al. (1999).

Data Analysis.   Allele and haplotype frequencies were estimated by direct counting. Population genetic structure was investigated using an analysis of molecular variance (AMOVA) (Excoffier et al. 1992), as implemented in the program Arlequin, version 1.1 (Schneider et al. 1997). This program was also used to calculate pairwise F_{ST} genetic distances, using the estimator proposed by Reynolds et al. (1983), with nonzero significance evaluated by a randomization test. Arlequin was also used to estimate gene diversity (Nei 1987) on the basis of Y-chromosome haplotypes and the average gene diversity across loci. Estimates of gene flow were calculated from the formula given by Slatkin and Barton (1989):

$$N_m = \frac{0.5(1 - G_{ST})}{G_{ST}}.$$ (1)

From distance matrices, we constructed trees and genetic maps using the unweighted pair-group method with arithmetic mean (UPGMA) and the neighbor-joining method using multiple dimensional scaling (Kruskal 1964). Association between genetic distance matrices derived from mtDNA haplogroups and Y-chromosome markers and associations with geography (computed as linear distances on the map) and language (ordinal language distances between groups computed based on hierarchical classification) in the Argentinean Chaco were examined using the Mantel test (Mantel 1967).

The program Network, version 3.0 (Bandelt et al. 1995; http://www.fluxusengineering.com), was used to construct median-joining networks of haplotypes and assumes the single-step model of mutation. The time to the most recent common ancestor (TMRCA) of the DYS199T lineage is estimated by the method described by Forster et al. (1996), which measures age based on mutational steps in the network from the specified root node (ρ statistic). The ρ statistic is inherently unbiased: Past demography will have influenced the shape of the evolutionary tree, but this affects only the error of the time estimate rather than systematically increasing or decreasing the time estimate itself.

In a worldwide study of Y-chromosome microsatellites Forster et al. (2000) found that their mutation rate is a function of both the mean number of repeats and the motif length, and they also noted differences between the trinucleotide DYS392 and the tetranucleotides. DYS392 is evolving much faster than the tetranucleotides and has violated the single repeat mutation model at least twice in

Europeans, but this seems not to have occurred in native Americans. No detailed data on mutation rates are available for the DYS437, DYS438, and DYS439 loci. A time estimate for the DYS199T mutation is performed using all eight microsatellites and the six microsatellites with a tetranucleotide motif. The mutation rate was set at 2.8×10^{-3} based on the findings of Kayser et al. (2000). Generation times of 30 years and 35 years were used, the latter being suggested specifically for Y-chromosome loci (Tremblay and Vezina 2000).

Results and Discussion

Sixty of the 78 individuals (76.9%) carry the native American–specific C to T transition at the DYS199 locus, and these males are the subject of this paper, because the DYS199C chromosomes are indicative of nonaboriginality. Most DYS199C individuals are from the Pilagá (10/19), although this chromosome was present in the other four tribes (two Tobas from Formosa, two Tobas from Chaco, three Wichí from Chaco, and one Wichí from Formosa). Allele frequencies at the eight microsatellite loci in DYS199T males are presented in Table 1 and the distribution of haplotypes is given in the Appendix.

Estimates of gene diversity and gene flow by locus are presented in Table 2. The trinucleotide repeat DYS392 exhibits the highest diversity, as reported before in native Americans (Bianchi et al. 1998). The G_{ST} values for each STR are quite variable, with DYS437, DYS438, and DYS439 being the systems showing the largest intergroup variation. This finding among the relatively novel microsatellites suggests that these loci, not yet extensively used in population studies, may be useful estimators of local population differences. N_m values for each locus are also quite variable because they are calculated using G_{ST} values, but overall they indicate a moderate level of gene flow in the Gran Chaco.

The DYS19 locus shows a moderate level of diversity in relation to the other loci. This contrasts with the findings of Carvalho-Silva et al. (1999), who reported an extremely low variability at DYS19 in native Americans in comparison to five other Y-chromosome-linked tetranucleotide repeat loci (DYS389I, DYS389II, DYS390, DYS391, and DYS393). These investigators found only one variant, other than allele 13, at DYS19 (a Mayan carrying the *15 allele) in 95 Y chromosomes scored in Amerindians from South, Central, and North America (Surui, Karitiana, Muskoke, Maya, Ticuna) as well as others from the Amazonian region previously reported by Santos et al. (1995). Four different DYS19 alleles (*12, *13, *14, *15) are present in Gran Chaco chromosomes and, although the three smaller alleles have been reported consistently (Bianchi et al. 1998; Ruiz-Linares et al. 1999), the *DYS19*15* allele is rare. Four of the five *DYS19*15* alleles in the present study are found in one tribe, the Wichí of Formosa, but all in unique haplotypes. Bianchi et al. (1998) reported one Toba individual with the *DYS19*15* allele, but in a haplotype (*15,24,10,14,13*; order DYS19, DYS390, DYS391, DYS392, DYS393) different from those in the present study (see Appendix).

Table 1. Y-Chromosome Microsatellite Frequencies in DYS199T Individuals from Five Gran Chaco Populations, Argentina

Allele	Pilagá	Wichí, Formosa	Toba, Formosa	Wichí, Chaco	Toba, Chaco	Total
DYS19						
12			0.154	0.077		0.050
13	0.778	0.714	0.692	0.538	0.818	0.700
14	0.222	0.072	0.077	0.385	0.182	0.183
15		0.214	0.077			0.067
DYS390						
22				0.077		0.017
23	0.222		0.154	0.154	0.182	0.133
24	0.778	0.714	0.461	0.692	0.818	0.683
25		0.214	0.385	0.077		0.150
26		0.071				0.017
DYS391						
9				0.077		0.017
10	0.889	0.929	0.923	0.769	0.909	0.883
11	0.111	0.071	0.077	0.154	0.091	0.100
DYS392						
11			0.154	0.154		0.067
13	0.111	0.214		0.077	0.091	0.100
14	0.778	0.357	0.769	0.461	0.909	0.633
15	0.111	0.429	0.077	0.308		0.200
DYS393						
12			0.077	0.154		0.050
13	1.000	0.857	0.923	0.769	1.000	0.900
14		0.143		0.077		0.050
DYS437						
8	0.444	1.000	0.846	1.000	1.000	0.900
9	0.556		0.154			0.100
DYS438						
10		0.286	0.231			0.117
11	0.778	0.714	0.692	1.000	1.000	0.833
12	0.222		0.077			0.050
DYS439						
11		0.357	0.154	0.077		0.133
12	0.333	0.643	0.538	0.769	1.000	0.667
13	0.667		0.308	0.154		0.200

Estimates of genetic diversity by population for mtDNA and Y-chromosome STRs are given in Table 3. The Pilagá and the Toba of Chaco present the highest and the lowest intragroup variation, respectively, based on Y-chromosome markers, with the other three populations showing intermediate values. The estimates of intragroup variation found in the same population samples for mtDNA haplogroups are, in general, slightly lower, with the lowest variation found among the Wichí of Chaco. In contrast, the Toba of Chaco exhibit a value almost threefold higher than that for the Y-chromosome analysis.

F_{ST} values for both Y-chromosome haplotypes and mtDNA haplogroups in the five tribes plus additional F_{ST} values obtained from the analysis of other

Table 2. Estimates of Gene Diversity and Gene Flow by Locus

Locus	h^a	G_{ST}	N_m
DYS19	0.484	0.079	5.80
DYS390	0.468	0.097	4.65
DYS391	0.255	0.016	31.78
DYS392	0.571	0.137	3.15
DYS393	0.182	0.087	5.23
DYS437	0.177	0.425	0.68
DYS438	0.285	0.143	2.99
DYS439	0.492	0.254	1.47
Mean	0.364	0.151	2.82

a. Nei's (1987) gene diversity.

Table 3. Estimates of Genetic Diversity by Population

Population	Y-Chromosome STRs	mtDNA Haplogroups
Pilagá	0.36	0.35
Wichí, Formosa	0.37	0.32
Toba, Formosa	0.43	0.36
Wichí, Chaco	0.41	0.30
Toba, Chaco	0.13	0.32
Total	0.38	0.34

biological variables for native Chaco populations are presented in Table 4. The intergroup variation based on Y-chromosome STRs is only moderately high (10.7%) but is greater than that for mtDNA (4%). For comparative purposes, F_{ST} values were computed for the same ethnic groups using the blood group data of Matson et al. (1969) and Pages-Larraya et al. (1978). The F_{ST} value, 0.056, is similar to the minimum estimates obtained from dermatoglyphic and anthropometric traits for the region, although not for the same population samples (Demarchi 2000; Demarchi et al. 1998).
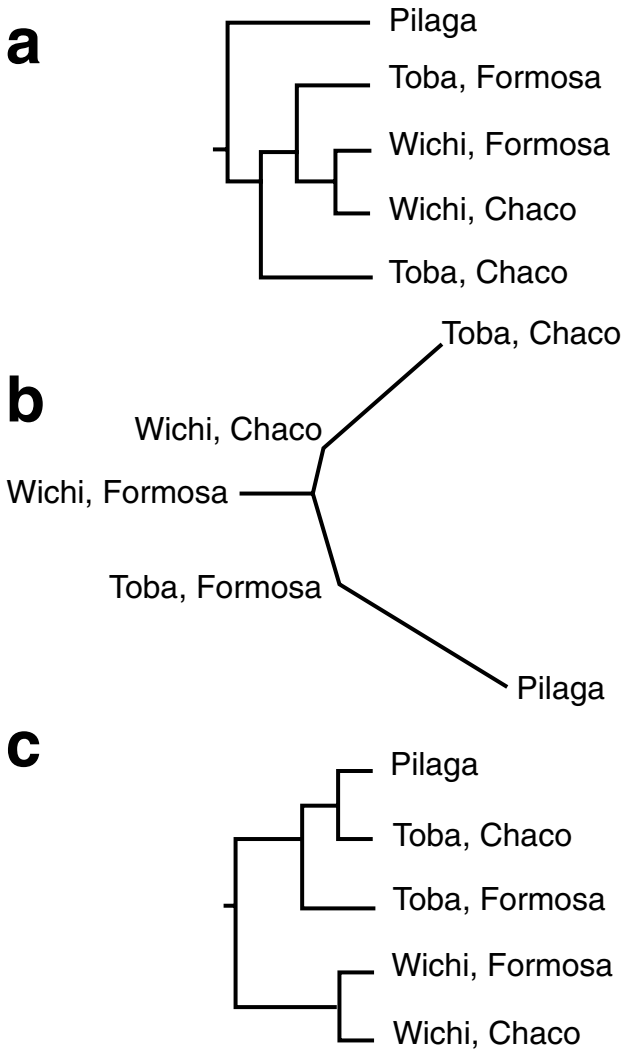
An UPGMA and N-J tree derived from Reynold's F_{ST} distances representing relationships among the five population samples are presented in Figure 2. The UPGMA tree (Figure 2a) shows no clear or consistent relationship of the five tribes to geography (see Figure 1) or to language. The Pilagá and the Toba

Table 4. Genetic Variation in the Argentinian Chaco: Coefficient of Differentiation (F_{ST}) Among Populations Based on Different Biological Markers

Variable	F_{ST}
Eight Y-chromosome STRs (T individuals)	0.107
mtDNA haplogroups	0.040
Blood groups	0.056
Dermatoglyphics (palmar traits)	0.058[a]
Anthropometrics (head measurements)	0.053[a]

a. Minimal estimates (Relethford and Blangero 1990).

a

Pilaga

Toba, Formosa

Wichi, Formosa

Wichi, Chaco

Toba, Chaco

b

Toba, Chaco

Wichi, Chaco

Wichi, Formosa

Toba, Formosa

Pilaga

c

Pilaga

Toba, Chaco

Toba, Formosa

Wichi, Formosa

Wichi, Chaco

Figure 2. (a) UPGMA and (b) N-J tree of Reynold's F_{ST} distances, depicting relationships among five native American populations from the Argentinean part of the Gran Chaco, based on eight Y-chromosome STR haplotypes; (c) UPGMA tree based on five Y-chromosome STR haplotypes (DYS19, DYS390, DYS391, DYS392, and DYS393).

of Chaco (both Guaycuru speaking) are the most distant, and the Wichí of Formosa and Chaco (Mataco speakers) show the smallest distance. That the two Wichí groups inhabit the banks of the Bermejo River reinforces their close relationship.

The N-J tree (Figure 2b) shows only a slightly better fit. The lack of congruence of the tribal relationships with geography and/or language most probably

reflects that random drift has been the most powerful mechanism acting on these tribes as a result of their ecology. The seasonal fluctuations of resources in the Gran Chaco compelled these tribal groups to be nomadic for much of the year and especially in winter months, when there was a high dependence on hunting and fishing. At this time the bands were small, and those groups without access to riverine resources had to split up into very small units to survive the harsh conditions. Accordingly, these nomadic groups were isolated from each other for much of the year but were in contact during the summer months of richer resources. Some of this contact, however, involved war and the kidnapping of women from other tribes.

If the UPGMA tree includes only the five classic microsatellites (Figure 2c), it shows a much closer agreement with geographic and linguistic distances. The Wichí from Chaco and Formosa cluster and are clearly separated from the three Guaycurú-speaking tribes (Pilagá and Toba), which form the remaining cluster. One explanation of the difference between the UPGMA trees (Figures 2a and 2c) is that the complex repeat structure of the DYS437 and DYS439 loci and the high variability of DYS437, DYS438, and DYS439 compared with the other five loci may be unduly influencing the pattern of relationships and adding background noise.

Patterns of Paternal Lineages Variation in the Argentinean Chaco.   The results of two-dimensional scaling of Reynold's F_{ST} distances based on five Y-chromosome STR haplotypes (DYS19, DYS390, DYS391, DYS392, and DYS393) of DYS199T individuals from eight populations of the Argentinean Chaco, including data on the Chorote (another Mataco-speaking tribe), the Wichí, and the Toba, all from Salta province, reported by Bianchi et al. (1998), are shown in Figure 3. The relative positions of the populations correlate well with their language affiliation and geographic location. The first axis separates the two linguistic families; the Guaycurú groups fall on the left-hand side of the plot and the Mataco-speaking populations lie on the right-hand side. The smallest distance is that between the Pilagá and the Toba of Salta, which reflects the ease of movement between them by means of the Pilcomayo River. On the other hand, the Toba of Formosa are the tribe most distant from all samples, reflecting their residence in Formosa City, which is situated far from the other aboriginal groups (Figure 1). Although the relationships revealed by this analysis fit geography and language, the degree of differentiation, as measured by F_{ST}, decreases from 0.107 (when eight microsatellites were used for the five tribes) to 0.042, a value similar to that obtained from mtDNA haplogroups.

We also examined the degree of association between the distance matrices derived from analysis of both mtDNA haplogroups and Y-chromosome haplotypes as well as each of these with geography and language. No association was found between maternal and paternal lineage distributions among the eight samples ($r = 0.040$). Maternal lineages also show no correlation with geography
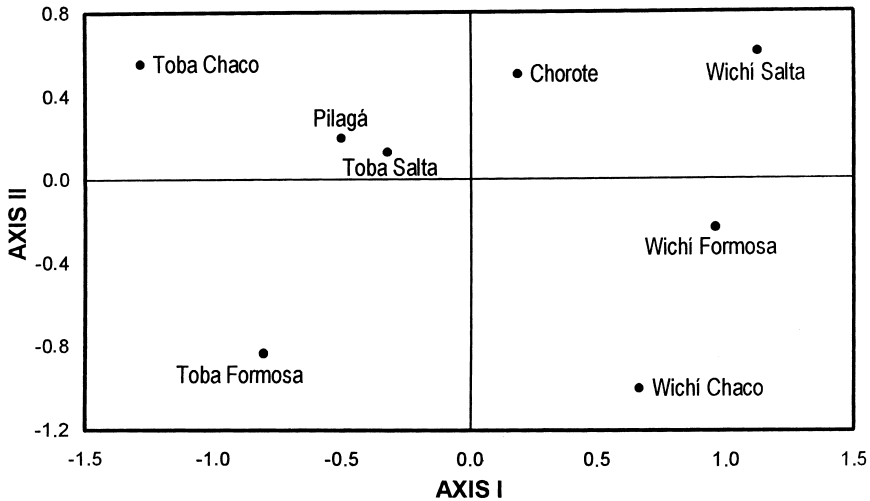
Figure 3. Two-dimensional scaling of Reynold's F_{ST} distances based on five Y-chromosome STR haplotypes (DYS19, DYS390, DYS391, DYS392, and DYS393) of DYS199T individuals from eight populations of the Argentinean Chaco.

($r = -0.054$) and present a negative (but insignificant) association with language ($r = -0.245$). On the other hand, the paternal lineages present high positive correlation with language ($r = 0.465$, $p < 0.05$) and also with geography ($r = 0.352$). In this latter case the value does not reach significance, most likely because of small sample size. AMOVA reveals that at this level of analysis only 3.7% of the total variation is due to intergroup differences.

Diversity of Paternal Lineages in South America. The AMOVA was extended to other South Amerindian populations for which Y-chromosome haplotype data, based on either four or five microsatellites (in DYS199T individuals), are available, to investigate variation among other regions of South America (Table 5). Tropical forest groups show the highest F_{ST} values and the lowest intragroup variation. Andean populations show the second highest F_{ST} value but much higher intragroup diversity. The Gran Chaco populations, on the other hand, exhibit intragroup variation as low as the tropical forest groups, but their intergroup diversity is almost seven times lower than that observed for the tropical forest region. Patagonia, here represented only by the Mapuche and the Tehuelche groups, presents the highest intragroup variation and the lowest F_{ST} value. However, the calculations were made on the basis of only 10 individuals, which makes this estimate of little value. The AMOVA for South America, at three levels, shows that the among-regions component of the total variation (5%) is statistically insignificant and that most of the intergroup variation is concentrated in the within-region diversity (14%) ($p < 0.01$).

**Table 5.** Paternal Lineage Diversity in South American Amerindians (DYS199T Males Only)

Population	Number of Individuals	Molecular Diversity	Variation Among Populations (%)
Andes	85	0.426	8.99[b]
Arequipa	15	0.507	
Humahuaca	10	0.400	
Susque	16	0.277	
Tayacaja	44	0.438	
Tropical forests	110	0.366	29.19[b]
Cayapa	26	0.442	
Gavião-Zoro-Suruí	34	0.229	
Karitiana	8	0.063	
Ticuna	32	0.276	
Wai Wai	5	0.200	
Xavante	5	0.150	
Gran Chaco	91	0.368	4.46[a]
Ayoreo	5	0.240	
Chorote	5	0.240	
Lengua	5	0.200	
Wichí, Chaco	12	0.561	
Wichí, Formosa	14	0.409	
Wichí, Salta	12	0.342	
Pilaga	9	0.283	
Toba, Chaco	11	0.204	
Toba, Formosa	13	0.382	
Toba, Salta	5	0.240	
Patagonia	10	0.507	2.78[NS]
Tehuelche	7	0.533	
Mapuche	3	0.400	
South America	296	0.391	
Among regions			5.05[NS]
Within regions			13.89[b]
Within populations			81.06[b]

NS = non significant
a. $P < 0.05$.
b. $P < 0.001$.

Network Analysis and Time Estimates of DYS199T Chromosomes. The network of eight-locus haplotypes in the five population samples (not shown, because there are more than 40 haplotypes) indicates that there are many male lineages represented in the Gran Chaco. Although there is some clustering, there are numerous other lineages quite distinct from any cluster. The network based on the six microsatellites with a tetranucleotide motif is shown in Figure 4, and more structure is evident. The diversity reflects the high mobility of males, principally from the Andes, in this transitional ecological area of central South America. At the individual tribe level the four haplotypes in the Chaco Toba likely represent descendants of a single male founder who carried haplotype 19
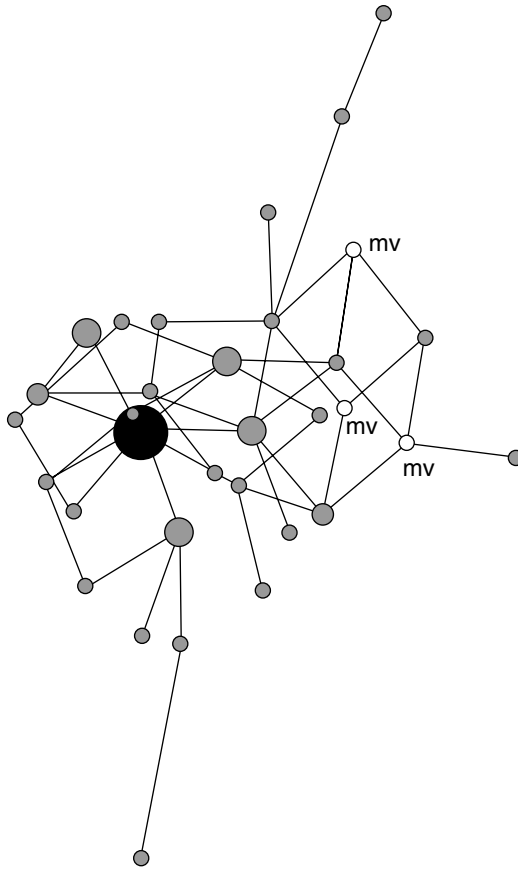
Figure 4. Median-joining network of Y-chromosome haplotypes based on six tetranucleotide microsatellites.

(see Appendix) but no other tribe shows any equivalent clustering of all haplotypes. When all the Toba are networked, there is a cluster comprising 9 of the 14 haplotypes, but the remainder are distant. The networks of the two linguistic groups, the Mataco (comprising the Wichí of both Formosa and Chaco) and Guaycurú (comprising the Toba and Pilagá), show only weak clustering.

The ancestral haplotype among a specific Y-chromosome haplogroup or lineage is determined from its frequency and location in the center of the network (Hurles et al. 1999). In the Gran Chaco populations this is haplotype *13,24,10,14,13,8,11,12* (loci order DYS19, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, and DYS439) (haplotype 19 in the Appendix). This haplotype matches (when loci can be compared between studies) the ancestral haplotype of another study of South Americans (Bianchi et al. 1998). Using a generation interval of 30 years, the TMRCA of the DYS199T haplotypes is about

21,000 years (with a standard deviation of about 4,000 years) and if 35 years is used, the TMRCA is about 26,000 years (with a standard deviation of about 5,000 years). If only the six microsatellites with a tetranucleotide motif are used in the analysis, the time estimate is about 13,000 years (SD about 3,000 years) if a generation interval of 30 years is used and 16,000 years (SD about 3,600 years) if 35 years is used. Although these estimates have a considerable range, the dates are reasonably close to those that Forster et al. (2000) estimated for the colonization of South America after reanalyzing the data of Bianchi et al. (1998), which used seven Y-chromosome microsatellites.

Although the sample sizes representing the Gran Chaco tribes are relatively small, the findings make it clear that there is no suggestion of a restricted male gene pool colonizing the Gran Chaco region of Argentina. This genetic diversity of the colonizers appears to have occurred despite the hostility of the environment and the recent date of human settlement. Further, the contemporary Y-chromosome lineages reflect a history as long as the settlement of the South American continent itself.

Literature Cited

Ayub, Q., A. Mohyuddin, R. Qamar et al. 2000. Identification and characterization of novel human Y-chromosomal microsatellites from sequence database information. *Nucleic Acids Res.* 28(2):e8.

Bandelt, H.-J., P. Forster, B.C. Sykes et al. 1995. Mitochondrial portraits of human populations using median networks. *Genetics* 141:743–753.

Bianchi, N.O., C.I. Catanesi, G. Bailliet et al. 1998. Characterization of ancestral and derived Y-chromosome haplotypes of New World native populations. *Am. J. Hum. Genet.* 63:1,862–1,871.

Carvalho-Silva, D.R., F.R. Santos, M.H. Hutz et al. 1999. Divergent human Y-chromosome microsatellite evolution rates. *J. Mol. Evol.* 49(2):204–214.

Demarchi, D.A. 2000. Variation in palmar quantitative dermatoglyphics among Argentinian aboriginal populations. *Homo* 51:19–29.

Demarchi, D.A., S.E. Colantonio, and A.J. Marcellino. 1998. Population structure in the Gran Chaco and the uniqueness of the Ayoreo [abstract]. *Am. J. Hum. Biol.* 10:120–121.

Demarchi, D.A., G.M. Panzetta-Dutari, M. López de Basualdo et al. 2001. Mitochondrial DNA haplogroups in Amerindian populations from the Gran Chaco. *Am. J. Phys. Anthropol.* 115:199–203.

Excoffier, L., P. Smouse, and J. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* 131:479–491.

Forster, P., R. Harding, A. Torroni et al. 1996. Origin and evolution of Native American mtDNA variation: A reappraisal. *Am. J. Hum. Genet.* 59:935–945.

Forster, P., A. Rohl, P. Lunnemann et al. 2000. A short tandem repeat based phylogeny for the human Y chromosome. *Am. J. Hum. Genet.* 67:182–196.

Grignani, P., G. Peloso, P. Fattorini et al. 2000. Highly informative Y-chromosomal haplotypes by the addition of three new STRs, DYS437, DYS438, and DYS439. *Int. J. Legal Med.* 114:126–129.

Gusmao, L., C. Alves, and A. Amorim. 2001. Molecular characteristics of four human Y-specific microsatellites (DYS434, DYS437, DYS438, DYS439) for population and forensic studies. *Ann. Hum. Genet.* 65:285–291.

Hernández, I. 1992. *Los indios de Argentina.* Madrid: MAPFRE.

Hurles, M., R. Veitia, E. Arroyo et al. 1999. Recent male-mediated gene flow over a linguistic barrier in Iberia, suggested by an analysis of a Y-chromosomal DNA polymorphism. *Am. J. Hum. Genet.* 65:1,437–1,448.

Kayser, M., A. Casglia, D. Corach et al. 1997. Evaluation of Y chromosome STRs: A multicenter study. *Int. J. Legal Med.* 110:125–133.

Kayser, M., L. Roewer, M. Hedman et al. 2000. Characteristics and frequencies of germline mutations at microsatellite loci from the Y chromosome as revealed by direct observations in fathers/son pairs. *Am. J. Hum. Genet.* 66:1,580–1,589.

Kruskal, J.B. 1964. Nonmetric multidimensional scaling: A numerical method. *Psychometrika* 29:115–129.

Loukotka, C. 1968. *Classification of South American Indian Languages.* Los Angeles: Latin American Center, University of California.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27:209–220.

Matson, A., H. Sutton, J. Swanson et al. 1969. Distribution of hereditary blood groups among Indians in South America. VII. In Argentina. *Am. J. Phys. Anthropol.* 30:61–84.

Morell, V. 1990. Confusion in earliest America. *Science* 241:439–441.

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

Pages-Larraya, F., N. Wyller de Contardi, and E. Servi. 1978. Marcadores geneticos de la población aborigen del Chaco argentino. *Rev. Inst. Antropol.* (Cordoba, Argentina), 6:217–241.

Pena, S.D.J., F.R. Santos, N.O. Bianchi et al. 1995. A major founder Y-chromosome haplotype in Amerindians. *Natur. Genet.* 11:15–16.

Relethford, J.H., and J. Blangero. 1990. Detection of differential gene flow from patterns of quantitative variation. *Hum. Biol.* 62:5–25.

Reynolds, J., B.S. Weir, and C.C. Cockerham. 1983. Estimation on the coancestry coefficient: Basis for a short-term genetic distance. *Genetics* 105:767–779.

Ruiz-Linares, A., D. Ortiz-Barrientos, M. Figueroa et al. 1999. Microsatellites provide evidence for Y chromosome diversity among the founders of the New World. *Proc. Natl. Acad. Sci. USA* 96:6,312–6,317.

Santos, F.R., M.H. Hutz, C.E.A. Coimbra et al. 1995. Further evidence for the existence of a major founder Y chromosome haplotype in Amerindians. *Braz. J. Genet.* 18:669–672.

Santos, F.R., A. Pandya, C. Tyler-Smith et al. 1999. The central Siberian origin for Native American Y chromosomes. *Am. J. Hum. Genet.* 64:619–628.

Schneider, S., J.-M. Kueffer, D. Roesslie et al. 1997. *Arlequin, Version 1.1: A Software for Population Genetic Data Analysis.* Geneva, Switzerland: Genetics and Biometry Laboratory, University of Geneva.

Slatkin, M., and N.H. Barton. 1989. A comparison of three indirect methods for estimating average levels of gene flow. *Evolution* 43:1,349–1,368.

Steward, J., and L. Faron. 1959. *Native Peoples of South America.* New York: McGraw-Hill.

Tarrazona-Santos, E., D. Carvalho-Silva, D. Pettener et al. 2001. Genetic differentiation in South Amerindians is related to environmental and cultural diversity: Evidence from the Y chromosome. *Am. J. Hum. Genet.* 68:1,485–1,496.

Tremblay, M., and H. Vezina. 2000. New estimates of intergenerational time intervals for the calculation of age and origins of mutations. *Am. J. Hum Genet.* 66:651–658.

Underhill, P., L. Jin, R. Zemans et al. 1996. A pre-Columbian Y-chromosome-specific transition and its implications for human evolutionary history. *Proc. Natl. Acad. Sci. USA* 93:196–200.

Underhill, P., G. Passarino, A.A. Lin et al. 2001. The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. *Am. J. Hum. Genet.* 65:43–62.

Appendix. Y-Chromosome Eight-Locus Haplotype Distribution in Five Gran Chaco Populations

Haplotype[a]	Pilagá	Wichí, Formosa	Toba, Formosa	Wichí, Chaco	Toba, Chaco	Total
12,24,10,15,13,8,11,13		1				1
12,24,9,14,13,8,11,12				1		1
12,25,10,14,13,8,11,13			1			1
13,22,10,15,13,8,11,13				1		1
13,23,10,14,13,8,11,13	1					1
13,23,10,14,12,9,10,11			1			1
13,23,10,15,14,8,11,12				1		1
13,23,10,13,12,8,11,11				1		1
13,23,10,14,13,8,11,12					2	2
13,24,10,14,13,9,12,12	1					1
13,24,10,14,13,8,12,12			1			1
13,24,10,14,14,8,11,12		1				1
13,24,10,15,13,8,11,11		1				1
13,24,10,15,13,8,10,12		1				1
13,24,10,14,13,8,11,11		1				1
13,24,10,14,13,8,10,12		1				1
13,24,10,14,13,8,11,13	2		2			4
13,24,10,14,13,9,11,13	3					3
13,24,10,14,13,8,11,12		1	1	2	6	10
13,24,10,15,13,8,11,12		2				2
13,24,11,14,13,8,11,12				1	1	2
13,25,10,15,14,8,11,11		1				1
13,25,10,14,13,8,11,12			3			3
13,25,10,14,13,8,10,12		1	1			2
14,23,11,15,13,9,11,12	1					1
14,23,11,11,13,8,10,12			1			1
14,24,10,13,13,8,12,12	1					1
14,24,10,13,13,8,10,11		1				1
14,24,11,15,13,8,11,13			1			1
14,24,10,15,13,8,11,12			1			1
14,24,10,14,12,8,11,12			1			1
14,24,10,11,13,8,11,12			1			1
14,24,10,14,13,8,11,12					1	1
14,24,10,13,13,8,11,12					1	1
14,25,10,11,13,8,11,12			1			1
15,24,10,11,13,8,10,11		1				1
15,24,10,13,13,8,11,12		1				1
15,24,10,15,13,8,11,12		1				1
15,25,11,13,13,8,11,11		1				1
15,26,10,13,13,8,11,12		1				1

a. The order of loci in each haplotype is DYS19, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439.