Genetic Structure of Native Populations from the Gran Chaco Region, South America

Darío A. Demarchi and Angelina García Ministro

Museo de Antropología, Facultad de Filosofía y Humanidades, Universidad Nacional de Córdoba, Córdoba 5000, República Argentina

KEYWORDS Native Americans; South America; population structure; mitochondrial DNA; Y chromosome markers; STRs

ABSTRACT The Gran Chaco region is a vast subtropical plain in South America that covers part of the territories of Argentina, Paraguay, and Bolivia, being the refuge of some of the last hunter-gatherer groups remaining in South America. 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 on a subtropical region inhabited by Tupí-Guaraní tribes. In this paper we present a synthesis of the principal findings of a long term study on the genetic structure of native populations of the Gran Chaco using an assortment of morphological (anthropometrics and dermatoglyphics) and molecular markers (mitochondrial haplogroups and HVR-I sequences, Y-chromosome markers, and nuclear STRs). In general, the biological evidence demonstrates that the Gran Chaco tribes present high genetic diversity and low differentiation among populations, when compared to other regions of South America. The only exception is the Ayoreo, a population with unique genetic and morphological patterns, being an outlier not only in relation to the rest of the Gran Chaco populations but also to any other native group of South America.

INTRODUCTION

The Gran Chaco lies in the center of South America, covering about 650,000 km² and encompassing parts of Argentina, Paraguay, and Bolivia (Fig. 1). The region is a low-lying plain that is drained southward from the Mato Grosso plateau by the Paraguay River and southeastward by the Pilcomayo, Bermejo, and Salado Rivers, which originate near the foothills of the Andes chain. 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, rheas, 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 Gran 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 on 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.

Gran Chaco peoples have been mobile hunter-gatherers (Saeger 2000) who exploit the animals and plants in the watered microenvironments created by the rivers (Mendoza 2003), and who practice some horticulture (Braunstein and Miller 1999). Before the arrival of Europeans, the ancient Gran Chaco consisted of forest, woodland,
savanna, and grassland patches that were occasionally flooded by the rivers and burnt by periodic fires (Alvarsson 1988; Braunstein and Miller 1999).

Because of the geographically intermediate position of the Gran Chaco within South America, the area historically served as a crossroads for people and their trade items. After the arrival of European colonists and their descendants, intense cattle-grazing converted any former natural pastures into woody vegetation (Mendoza 2003). Many peoples in the eastern Chaco adopted use of the horse, thus changing the degree and pattern of subsistence and intergroup relations (Karsten 1967; Metraux 1946). Ways of life began to change more drastically with the arrival of Jesuit missions and a new capitalist economy based on sugar cane harvesting (Cordeau and Siffredi 1971; Saeger 1999). Now, most of the Gran Chaco peoples live in permanent, year-round settlements.

In this paper we present a synthesis of the principal findings of a long term project on the genetic structure of native populations of the Gran Chaco carried out using an assortment of morphological and molecular markers.

**DERMATOGLYPHIC VARIATION**

Our first approach to the population structure of Gran Chaco populations was made through the analysis of digital and palmar dermatoglyphics in tribes of the Argentinean part of the Gran Chaco (Demarchi and Marcellino 1994, 1995;
These studies revealed high genetic differentiation between populations but lack of agreement, at the regional level of analysis, between biological variation and geography or language.

On the other hand, when the analysis was done at continental level a different picture was observed. The biological affinities among South American aboriginal populations, examined by means of a multivariate analysis of finger patterns frequencies (Fig. 2) revealed inter-population relationships in close agreement with Loukotka’s language classification (Demarchi and Marcellino 1998). The UPGMA tree of D² distances revealed two well defined major clusters. The first cluster includes the Gran Chaco populations together with other Paleo-American speaking tribes of Brazil, all of them sharing high incidence of whorls as a distinctive feature. The other cluster includes the Andean and Tropical Forest groups, which present a greater prevalence of arches and ulnar loops and appear closely related to each other suggesting a more recent common origin and/or substantial gene flow. Interpreted in terms of a branching model, those results suggest an earlier separation of the PaleoAmerican speaking tribes from the original colonizer population of South America and a relatively recent separation between the Tropical Forest and Andean populations.

ANTHROPOMETRICS AND BLOOD GROUPS

In the study of Demarchi et al. (1998) we examined the population structure of the Gran Chaco from unpublished anthropometric data collected by Jehan Vellard, and published data on blood groups and language. A total of 853 males belonging to ten different tribes (Ayoreo, Toba, Pilaga, Wichí, Chorote, Chulupí, Chiriguano, Chane, Tapiete, and Lengua) from Paraguay and Argentina, were studied for ten anthropometric measures. The variables studied included ten dimensions: four from head (head length, head breadth, head height, bizygomatic breadth) and six from body (biacromial breadth, forearm length, arm length, sitting height, thigh length, and leg length). Given that body measurements are supposed to be more environmentally influenced than head and face dimensions, the two sets of variables were analyzed separately. Five blood groups (ABO, MNSs, Rh, P, and Diego) with 16 alleles, extracted from the literature (Matson et al. 1968, 1969; Salzano et al. 1978; Pages-Larraya et al. 1978) were employed for the genetic analysis. The population structure analysis was carried out using John Relethford’s Rmet 3.0 program, which performs R-matrix analysis for quantitative traits (Relethford and Blangero 1990). The same populations were

![Fig. 2. UPGMA tree based on D² distances obtained from the first 5 Principal Components derived from individual fingers pattern type frequencies among the native South American populations (males). Modified after Demarchi and Marcellino (1998).](image-url)
investigated using traditional genetic markers through R-matrix analysis (Harpending and Jenkins 1973).

The results showed high differentiation among the Gran Chaco populations in head dimensions ($F_{ST} = 0.142$). Ayoreo, from Paraguay, the least acculturated and most isolated tribe of the region, contributed most to such a value ($r_i = 0.648$). The differentiation among populations in body measures was much lower ($F_{ST} = 0.042$). The differentiation among populations estimated from blood groups was intermediate between those calculated from body and head measurements but still remarkably high ($F_{ST} = 0.063$). In agreement with the head measurements, the most differentiated population was, by far, the Ayoreo ($r_i = 0.175$).

The relationships among the Gran Chaco populations based on head dimensions are presented in Figure 3. The most obvious finding is the clear separation of the Ayoreo along the first axis, which is reflecting head breadth and length. The second axis separates the Lenguas, who exhibit the lowest head height values, from the other samples. Figure 4 shows the relationships among the Gran Chaco populations derived from blood groups analysis. Ayoreo is again clearly differentiated from the other populations, which are placed close to each other, at low genetic distances. The close relationships observed between Chiriguano and Chané is coincident with their relative position in the anthropometric plot. A conclusion of that study was that the Gran Chaco populations are markedly differentiated in their head shape. The level of differentiation is of the same order as that of craniofacial data among major geographic groups (Relethford 1994). The analysis of the horizontal cephalic index gives a clear picture of the pattern of that differentiation. Chiriguano and Chané, two Tropical Forest groups that migrated to the Gran Chaco around the XV century, are brachycephalic populations ($CI = 82.2$). The ‘typical’ Paleoamerican speakers, Chaco tribes, fall within mesocephalic values ($CI = 77.2-80.9$), whereas the Ayoreo appears in the extreme of the dolicocephalic range ($CI = 71.8$).

The results of the analysis of blood groups confirm the extreme differentiation of the Ayoreo in relation to its neighbors of the Gran Chaco. Body measurements, on the other hand, show low intergroup variation. Given that all the tribes inhabit the same biome or ecological area and that body dimensions (principally linear measurements) present more phenotypic plasticity, this result could be reflecting the homogenizing effect of the environment.

**MITOCHONDRIAL HAPLOGROUPS**

It has been shown that the mitochondrial DNA (mtDNA) of most Native Americans belongs to one of four haplogroups, designated as A, B, C, and D, representing four separate maternal lineages. The presence of a *HaeIII* site at nucleotide pair (np) 663 defines haplogroup A, the absence of a *HincII* site at np 13,259 defines haplogroup C, the absence of an *A1uI* site at np 5,176 characterizes haplogroup D, and the presence of a 9-bp deletion in the mitochondrial region V identifies the haplogroup B (Schurr et al. 1990; Torroni et al. 1993). Another haplogroup (named X) has been identified in several Native American populations (Easton et al. 1996), but not yet in South America (Dornelles et al. 2005). The large amount of data collected from a relatively large number of widely dispersed populations makes these genetic markers as useful tools for investigating the patterns of maternal lineage distribution among Native American populations. Demarchi et al. (2001) typed mitochondrial DNA from 141 individuals to examine the distribution of the founding mtDNA lineage haplogroups in four population samples representing three tribes (Wichí, Pilagá, and Toba) that currently inhabit the Argentinean part of the Gran Chaco. This study also included published data from another three populations of the region. All the four lineages were identified in the four population samples typed in this study as well as in those described by Bianchi et al. (1995), whereas among the Wichí of the province of Chaco scored by Torroni et al. (1993), lineage C was absent. Disregarding the ethnic or geographic origin, haplogroups B and D exhibit the highest incidence. Haplogroup C is also well-represented among the Pilagá and Wichí of Formosa. The Toba of Formosa show a relatively high percentage of lineage A, which is frequent in natives from North and Central America, but not in South Amerindians, except among the Piaroa, Kraho, and Quechua. The position of the populations in a neighbor-joining tree of $F_{ST}$ distances (not shown) corresponds approximately to its current geographic location, which also corresponds to its historic distribution.
Fig. 3. Centroids of the Gran Chaco populations projected onto the first two axes (Relethford-Blangero analysis) for head measurements.

Fig. 4. Principal coordinates plot of Nei’s standard distances showing the relationships among Gran Chaco populations based on traditional genetic markers.
However, only 6 of the 21 pairwise genetic distances were statistically significant at the 5% level, which indicates a very low level of differentiation between populations.

The calculation of Nei’s diversity estimates of mtDNA haplogroups distribution for the different geographic-ecological regions of South America (Andes, Tropical Forests, Gran Chaco, and Tierra del Fuego/Patagonia) reveals that total diversity (Ht) ranges from 0.515 (Tierra del Fuego/Patagonia) to 0.75 (Tropical Forest). The low Ht observed in Tierra del Fuego/Patagonia aboriginals is due to the absence of lineages A and B. On the other hand, the high value found in the Tropical Forest group reflects the heterogeneity of this group, which includes populations which are widely dispersed and with different linguistic backgrounds (i.e., Tropical Forests and Paleoamericans speakers). The highest Hs value is observed among the Gran Chaco tribes (0.64), whereas the lowest value corresponds to Tierra del Fuego/Patagonia (0.36). The largest amount of genetic variation attributable to interpopulational diversity (G’ST) was that observed among the recently extinct Tierra del Fuego/Patagonia tribes (39%). Intermediate values are observed for the Tropical Forest (29%) and the Andean (15%) groups, whereas the Chaco tribes exhibit, by far, the lowest (6%) intergroup diversity. That is, the Gran Chaco tribes present the highest intragroup and the lowest intergroup diversity when compared to other regions of South America. The observed pattern suggested considerable amount of gene flow not only among the Chaco tribes but also with the neighboring populations from outside the region.

**Y CHROMOSOME MARKERS**

To further investigate the population structure and variation in this region’s indigenous population, males from the same population samples investigated for mtDNA variation were also scored for 8 Y-chromosome microsatellites (DYS19, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, and DYS439) and the unique Native American single nucleotide polymorphism, DYS199 (Demarchi and Mitchell 2004).

Sixty (77%) of the total sample carried the DYS199*T chromosome, and those samples were the focus of further analysis. Unlike most other Native Americans, Gran Chaco males show a moderate level of diversity at the DYS19 locus but still less than what is seen in non-Americans. The FST value for Y markers in Gran Chaco (0.107) was more than double that found for mtDNA haplogroups in the same tribes, but not particularly high compared with other Y chromosome studies. Phylogenetic trees based on all 8 microsatellites showed relatively poor correlation with either geography or language. On the contrary, the UPGMA tree based on 5 microsatellites (DYS19, 390, 391, 392, and 393), showed a more positive relationship, suggesting that DYS437, 438 and 439 may behave differently from the other microsatellites. No association was found between maternal and paternal lineage distributions. The analysis of molecular variance was extended to other South Amerindian populations for which Y-haplotype data, based on either 4 or 5 microsatellites (in DYS199*T individuals), are available, to investigate variation among other regions of South America. Tropical forests populations show the highest FST values (0.291) as well as the lowest intragroup diversity (0.366). Andean populations show the second highest FST value (0.089) but much higher intergroup diversity (0.426). The Gran Chaco populations, on the other hand, exhibit intergroup variation as low as Tropical Forest groups (0.368) but inter-group differentiation almost seven times lower (0.046) than the latter. Patagonia, represented only by the Mapuche and the Tehuelche, presents the highest intragroup variation (0.507) and the lowest FST value (0.028). However, since the calculations were based on only 10 individuals, this estimate is of little value.

The analysis of molecular variance for South America, shows that among regions component of the total variation (5%) is statistically insignificant, and most of the intergroup variation is concentrated within the regions (14%, p < 0.01).

The network of 8 loci haplotypes in the five population samples indicates that there are many male lineages (>40 haplotypes) represented in the Gran Chaco (not shown). Though there is some clustering, there are numerous unique lineages, which are quite distinct from any cluster. The network based on the 6 microsatellites with a tetrancleotide motif is shown in Figure 5. The diversity reflects high mobility of males, principally from the Andes, within this transitional ecological area of central South America. Although the sample sizes representing the Gran Chaco tribes are relatively small the results do not reflect
GENETIC STRUCTURE IN THE GRAN CHACO

restricted male gene pool colonizing the Gran Chaco region of Argentina. We conclude that the male lineages of the contemporary Gran Chaco population reflect the level of diversity found in South America as a whole and that the male founders of this region did not carry a restricted gene pool.

MITOCHONDRIAL HVS-I SEQUENCES

Cabana et al. (2006) reevaluated the hypothesis of the unique pattern of genetic diversity of the Gran Chaco populations with higher within- and lower between-population mitochondrial DNA (mtDNA) haplogroup frequency variation compared to populations in other South American regions as proposed by Demarchi et al. (2001). To test this hypothesis of regional uniqueness, we applied analytical and simulation methods to mtDNA hypervariable segment I (HVS-I) sequence data from a broad set of comparative South and Central American population samples. When the scale of analysis was limited to the Central Chaco, that is including the same populations studied in Demarchi et al. (2001), the results conform to their earlier findings. On the other hand, when the Ayoreo population was included in the analysis, the estimate of genetic differentiation among populations increased dramatically, from \( \text{fst} = 0.03 \) to \( \text{fst} > 0.30 \).

Fig. 5. Median-joining network of Y chromosome haplotypes of Gran Chaco populations based on 6 tetranucleotide microsatellites (from Demarchi and Mitchell 2004).

Fig. 6. Two-dimensional scaling plot of Kimura 2P distances based on HVS-I sequences among South American native populations.
In Figure 6 we present a bi-dimensional plot based on multidimensional scaling of the genetic distances depicting relationships among a number of South American native populations based on HVS-I sequences. The Central Chaco populations (Wichí, Toba, and Pilagá), cluster together close to the centroid. The Ayoreo, on the other hand, appears as outlier not only in relation to the Chaco populations but also to the other native groups of South America. This is because of the extremely reduced mtDNA variability, being classifiable into just two haplogroups (C and D). Seventy-six out of 91 mtDNA HVS-I sequences analyzed for this populations (83.5%) belong to haplogroup C, which is present at low frequencies in the other Gran Chaco populations (Dornelles et al. 2004). All these C sequences possess the unique transition C to T at the nucleotide position 16357, and sixty-five of them share the same single haplotype (Fig. 7).

**ANALYSIS OF NUCLEAR SHORT TANDEM REPEATS**

The last approach to the study of population structure and variation in the Gran Chaco’s indigenous population was the analysis of 15 short tandem repeats (STRs) or microsatellites (D2S1338, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D19S433, D21S11, CSF1PO, TH01, TPOX, VWA and FGA) in 128 individuals from the same population samples analyzed for mtDNA and Y chromosome markers (Crossetti et al. unpublished results). A non-metric multi-dimensional scaling representation of D_{A} genetic distances between the six Gran Chaco populations (including the Ayoreo) is shown in Figure 8. The first axis clearly separates the Ayoreo from the other populations. The second axis separates the Guaycurú speakers (positive values) from the Mataco speakers (with negative values).

Considering 13 STRs for which data from the literature were available, Nei’s gene diversity coefficients $H_s$ and $G_{st}$ were calculated for the Gran Chaco populations and compared with those of the Amazonian Forest (represented by the Gavião, Suruí, Waiwai, and Zoró tribes) and Native Americans that live in the Savannah and the Sub-tropical forest of South America (Aché, Guarani, Kaingang, and Xavante). The Argentinean Chaco populations exhibit the highest average gene diversity within population ($H_s = 0.70$) and the lowest among populations gene differentiation coefficient ($G_{st} = 0.03$). This value was lower than those obtained using four
Genetic Structure in the Gran Chaco

Fig. 8. Two-dimensional scaling plot of $D_s$ distances, based on 15 nuclear STRs, among Gran Chaco native populations.

HLA STR polymorphisms ($G_{st}' = 0.06$; Dejean et al. 2004) or mtDNA haplogroups, as shown above. This variation could be attributed to the differences in the type and/or number of markers. Nevertheless, again and as noted by Demarchi et al. (2001), the $G_{st}$ of the Argentinean Chaco was lower than that observed in any other region of South America. When the Ayoreo tribe was included in the Gran Chaco region, $H_s$ remained high ($0.68$), but the differentiation among populations showed a considerable increase ($G_{st}' = 0.06$), being higher than the Amazonian value ($G_{st}' = 0.05$) but lower than that for the Savannah and Sub-tropical forest ($G_{st}' = 0.08$) populations. Among the three South American regions, the Gran Chaco region is genetically most homogeneous if the Ayoreo are not considered.

Conclusions

The Gran Chaco population investigated here can be viewed as genetically homogeneous. In terms of genetic variability, the Gran Chaco groups are not much different from those inhabiting the other geographic-ecological regions of South America, although the level of genetic diversity observed is moderate to high. The results suggest the genetic proximity between the Mataco and Guaykurú linguistic families, implying intense gene flow across large geographic distances. Language did not seem to be a barrier for gene flow in this region. The old ways of life of these peoples, related with the ecological conditions of the region, gave rise to seasonal migrations. In a region where food is scarce, the probability of contact between small nomadic groups of different origins would be high especially at the places where water and other supplies were more easily found. The Ayoreo constitutes an exception to this model. Both morphological and molecular data show a reduced genetic variation in this group and extremely large differentiation from the other populations of the region. The same was observed when blood group and protein markers were evaluated (Salzano et al. 1978), although similar differences were not apparent in the Alu insertions investigated by Dornelles et al. (2004), which has been interpreted as probably due to founder effect or random loss. However, the results obtained from anthropometric traits, particularly the head measurements, which are polygenic, hence expected to be evolutionarily most stable, suggest that this tribe could have a different origin.
other than that of the “typical” Gran Chaco groups. Overall, it may be concluded that the genetic homogeneity in the Gran Chaco’s populations is a rule with a solitary exception of the Ayoreo. Nevertheless, it should be noted that this region is geographically extensive and culturally diverse, and the number of populations evaluated so far is small.

ACKNOWLEDGEMENTS

Dario A Demarchi is Career Investigator Member of the Consejo Nacional de Investigaciones Científicas y Técnicas de la República Argentina (CONICET). Angelina García-Ministro is a fellow of the FONCyT (SECRET, República Argentina) Grant sponsor: Third World Academy of Sciences - TWAS; Grant Sponsor: Consejo Nacional de Investigaciones Científicas y Técnicas – CONICET (PEI 6147).

REFERENCES


