Genetic analysis of six communities of Mbyá-Guaraní inhabiting North Eastern Argentina by means of nuclear and mitochondrial polymorphic markers

ANDREA SALA[1], CARINA F. ARGÜELLES[2], MIGUEL E. MARINO[3], CECILIA BOBILLO[1], ALBERTO FENOCCHIO[2] AND DANIEL CORACH[1]

[1] *Servicio de Huellas Digitales Genéticas y Cátedra de Genética y Biología Molecular, Facultad de Farmacia y Bioquímica, Universidad de Buenos Aires, Buenos Aires -Argentina*

[2] *Laboratorio de Citogenética y Genética Humana, Facultad de Ciencias Exactas, Químicas y Naturales, Universidad Nacional de Misiones, Misiones-Argentina*

[3] *Laboratorio de Análisis de ADN, Facultad de Ciencias Médicas, Universidad de Cuyo, Mendoza-Argentina*

KEY WORDS: NATIVE SOUTH AMERINDIAN, MBYÁ-GUARANÍ, Y-STRS, MTDNA, AUTOSOMAL STRS

Corresponding author: Andrea Sala, Junin 956, (1113) Buenos Aires, Argentina- TE/FAX: 0054114964-8281/8282; e-mail: asala@ffyb.uba.ar

Abstract

Autosomal STRs, Y-chromosome markers, and mitochondrial DNA sequences were investigated in six Mbyá-Guaraní villages (North: Fortín M´Bororé, *n*= 52 and Yryapu, *n*=10; mid- West: Tabay, *n*=13 and Kaaguy Poty *n*=24; mid-East: Jejy, *n*=16 and Yaboti, *n*=6) all of them settled within the Province of Misiones, Northeastern Argentina. A total of 121 unrelated individuals were analyzed. The study involved typing of fifteen autosomal STRs, nine Y-STRs and four biallele loci in the non-recombinant region of the Y-chromosome, and mtDNA sequencing of hypervariable regions I and II, as well as detection of the 9 bp ins/del in Region V of mtDNA. All autosomal STRs were in HW equilibrium. The four major Native American mtDNA haplogroups (hg) were represented in the sample investigated. Haplogroups A2 and D1 exhibited the highest frequencies (A2: 40.5% and D1: 36.0%, respectively) while hgB2 and hgC1 appeared to be less frequent (17.5% and 6.0%, respectively). The Native American hg Q1a3a was observed in a relevant proportion (88.8%). In addition, a nine Y-STR haplotype (DYS19/13, DYS389-I/14, DYS389-II/31, DYS390/24, DYS391/11, DYS392/14, DYS393/11, DYS385/14,16) exhibited a frequency of over 36%. Our results indicate that the Argentinean Guaraní individuals analyzed are genetically more closely related to Guaraní from Brazil (genetic distance $\delta\mu^2$=0.48) than with other related tribes, which are geographically close settled. Statistical approaches based on autosomal data do not support the hypothesis of genetic drift previously proposed, however this apparent discrepancy might be due to the lack of sensitivity of the autosomal markers employed herein.

At the time of the first European contact, occurred in the early sixteenth century, the Guaraní aboriginals occupied vast areas of Southern Brazil (Sao Paulo, Rio de Janeiro, Mato Grosso do Sul, Parana, and Rio Grande do Sul States), Eastern-Paraguay and Northeastern and Middle-Eastern Argentina (Misiones, Corrientes and Entre Rios provinces) (Martinez Sarasola 2005). The area occupied ranged from 26° to 33°S and 58° to 52° W.

The Guaraní people arrived in Northeastern Argentina from the Amazon jungle, at a region called Mesopotamia, by following the natural course of the Paraná and Paraguay rivers. Radiocarbon dating studies suggested that the Guaraní had inhabited Misiones province at least 1,000 years before present (Tarragó 2000). The first contact between the Guaraní and the Europeans occurred at the beginning of the sixteenth century. The Spanish Catholic order of the Jesuits founded 30 "reductions" or settlements (later named "Jesuit Missions" in the Paraná and Uruguay river basins of Southeastern Paraguay). The Jesuits ruled the lives of over 150.000 Guaraní aboriginals. The reductions gave rise to a new regional culture characterized by five major aspects: evangelization, integration between the Jesuits and the Guaraní, autonomy regarding the Spanish Crown, no slavery and preservation of aboriginal identity. The missionized Indians became known as Guaraní, while those who avoided conversion to Catholicism became known as Cayua or Cainguá, a name that roughly translates as "men of the forest."

Between 1620 and 1750, the Jesuits were the target of fluctuating economic and political interests between the Portuguese and Spanish Crowns. In 1767, the Jesuitic Order was expelled from America, which caused the final dissolution of the Jesuitic-Guaraní order. The indigenous population largely decreased as a consequence of the volunteer emigration of the Guaraní looking for a new hold. At the beginning of the nineteenth century, the great Guaraní development had completely disappeared (Martinez Sarasola 2005).

The Mbyá-Guaraní language belongs to the Tupí-Guaraní branch of the Tupian linguistic family (Ruhlen and Greenberg 2007). The Guaraní's social habits can be classified into three groups: farmers with highly sedentary behavior, farmers with a certain level of nomadic behavior and nomads. According to Crivos et al. (2007) in a study conducted in two Guaraní villages from Misiones (Kaaguy Poty and Yvy Pyta), the Mbyá, who were characterized by their large- scale migrations, at present exhibit a sedentary tendency with a micro-scale mobility pattern. Their economy is based on traditional crop farming: maize, manioc, sweet potato, calabash, tobacco, etc. Hunting and fishing complement their diet. Basically, their present day social organization is based on relatively large families, with no more than 12 to 15 members. The internal organization in these communities is established by family relationships between members, which give rise to large groups headed by one family leader, usually called Cacique or Paí.

According to Garlet (1997), the Mbyá currently living in the province of Misiones would be the descendants of the Mbyá from the Department of Caaguazú (Eastern Paraguay), who started to migrate towards Argentine territories after the Triple Alliance War, where Argentina, Brazil and Uruguay fought against Paraguay (1864-1870). At present, according to the information provided by the last population census conducted in 2004 by the National Institute of Statistics and Census (INDEC) of Argentina (http/ www.indec.mecon.ar), the estimated population size of the Mbyá-Guaraní inhabiting Misiones province (Northeastern Argentina) was 3,975 habitants. The growing trend towards deforestation and biodiversity reduction as a consequence of soybean cropping, paper, tobacco and wood industries, and new epidemics put the Mbyá-Guaraní survival at risk (Crivos et al. 2007).

Since the last decades, uniparentally genetic markers such as mitochondrial DNA and the non-recombining region of the Y chromosome have been used to determine population ancestry as well as population evolution (Ginther et al. 1993; Torroni et al. 1994; Bianchi et al. 1997;

Tarazona-Santos et al. 2001) in Native South Americans. The investigations of such markers allowed to find ethnic-specific mitochondrial haplogroups and Y-chromosome polymorphisms. The identification of one of the four mtDNA Native American haplogroups (Torroni et al. 1992, 1993) and the C→T transition at locus DYS199 (Underhill et al. 1996), now named M3, allowed to trace maternal and paternal Native American origin in an individual.

Altuna and coworkers (2006) analyzed 37 samples from the Mbyá-Guaraní by means of six Y-STRs finding a particular high frequent haplotype and propose that it could be a consequence of a recent founder effect.

We investigated six Mbyá-Guaraní communities inhabiting Misiones province, northeastern Argentina. Autosomal STRs (Short Tandem Repeats), Y-STRs, Y-biallele markers, as well as mitochondrial DNA sequences were analyzed in order to: 1-determine the genetic structure of the extant Guaraní inhabiting Argentina; 2- investigate the genetic contributions of maternal and paternal lineages, 3-test the hypothesis of a recent genetic drift process occurred in this Native American group and 4- to compare it with other ethnically and geographically-related communities.

MATERIALS AND METHODS

Populations

Blood or buccal mucosa swab samples were collected from 121 unrelated volunteer donors (98 men and 23 women) belonging to six Mbyá-Guaraní communities from three province Departments: Iguazú in the North (Fortín M´Bororé, *n*= 52 and Yryapu, *n*=10); Libertador General San Martin in mid- West (Tabay, *n*=13 and Kaaguy Poty, *n*=24) and Guaraní Department in mid-East (Jejy, *n*=16 and Yaboti, *n*=6). Figure 1 depicts the geographic location of the communities. Sampling was carried out with permission of the Bureau of Guaraní Affairs, an official organism of the government of Misiones province. All participants, including the community leaders, signed written consent statements (available upon request to the corresponding author of the paper), previously approved by the Bioethics Committee of the School of Pharmacy and Biochemistry of Buenos Aires University.

Autosomal STR analysis

Due to the scarce DNA amounts recovered, only 63 samples yielded results for autosomal STRs typing (49 from Fortín M´Bororé, 5 from Yaboti and 9 from Yryapu). Fifteen autosomal STRs: D18S51, D21S11, TH01, D3S1358, FGA, TPOX, D8S1179, vWA, CSF1PO, D16S539, D7S820, D13S317, D5S818, Penta E and Penta D, were analyzed following the manufacturer protocol (PowerPlex 16, Promega Corp. USA). Electrophoresis analysis was carried out in an automated sequencer ABI3100- *Avant* (Applied Biosystems, Foster City, USA).

MtDNA analysis

D-Loop sequencing: Hypervariable Regions I and II were amplified separately using primers: L15971 and H16410 for HVR I, and L15 and H484 for HVRII (Brandstätter et al. 2004). Amplicons were purified using columns (Qiaquick PCR purification System-Qiagen) or ExoSap (USB Corp. USA). The sequencing reactions were performed with the Big Dye Terminator System v1.1 (Applied Biosystems, Foster City, USA) according to the supplier's protocol. Sequencing reaction products were purified from residual dye terminators by means of ethanol precipitation. Electrophoresis separation was carried out on ABI 310 or ABI3100 *Avant* capillary sequencers (Applied Biosystems, Foster City, USA). All sequences were performed with both forward and reverse primers and electropherograms were visualized and edited with Seq Scape (Applied Biosystems, Foster City, USA) and/or Sequencher v4.8 (Gene Codes Corporation, USA) softwares.

Mt DNA Region V 9 bp ins/del was amplified as previously described (Wrischnik et al. 1987) using forward primer labeled with TET. The detection was carried out with the help of automated sequencers described above, using TAMRA as internal size standard.

Y-chromosome analysis

Amplification of nine Y-STRs (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS385a and DYS385b) were carried out as previously published (Corach et al. 2001; Sala et al. 2004).

Y-SNP: C/T transition that defines the M3-Q1a3a subhaplogroup (YCC, 2002) was evaluated by an allele-specific amplification assay according to Underhill et al. (1996). M242 (Seielstad et al. 2003) and SRY4064 were analyzed by Real Time PCR followed by High Resolution Melting Analysis (HRMA) using Syto 9 (Molecular Probes, Invitrogen, USA) as fluorescent dye and a RotorGen 6000 (Corvette, Australia) RT-PCR equipment.

Y- Ins/del: Y-Chromosome Alu Polymorphism (YAP) was typed according to Hammer (1995) by primer specific amplification followed by agarose gel electrophoresis and Ethidium Bromide staining.

Data analysis

Autosomal STRs analysis: Allele frequencies were obtained with PowerStats V1.2. Hardy-Weinberg equilibrium, gene diversity, heterozygosity and Garza-Williamson index were calculated with Arlequin v 3.1 (Excoffier et al. 2005). Goldstein genetic distance $(\delta\mu)^2$ was estimated from allele frequency (Goldstein et al. 1995). Multi-dimensional scaling plots were obtained with the help of software XLSTAT (Addinsoft Corp.). Mantel test was performed with XLSTAT software.

Drift effect or bottleneck hypothesis was tested using BOTTLENECK 1.2.01 (Cornuet and Luikart 1996). Two models of mutation were used to calculate the expected equilibrium gene diversity (*H*eq): the Stepwise Mutation Model (SMM) (Ohta and Kimura 1973) and Two-phase Model (TPM) (Di Rienzo et al. 1994). The methods employed were based on allele frequency and heterozygosity levels. For TPM computation 90% of SMM proportion was used. Three test were computed: Sign test and Standardized Differences test (Cornuet and Luikart 1996) and Wilcoxon test (Luikart et al. 1998).

Mitochondrial DNA analysis: sequences were aligned with DNA Alignment Software v1.1.3.0 (www.fluxus –engineering.com). Haplotype frequencies and molecular diversity indices were computed with Arlequin v3.1 Median-joining networks were obtained with the Network 4.5.1.0 program (Bandelt et al. 1995; www.fluxus-engineering.com). Star contraction was employed for

haplogroup A in order to reduce the large data set. Transversions were weighted three times over transitions. Since no sample exhibited the modal haplotype that define the hg D1, it was included in the MJ analysis with the name of "D1".

Y-chromosome polymorphism analysis: haplotype data was analyzed with Arlequin v3.1. Haplotype diversity was calculated according to Nei (1987). Median-joining networks were obtained with Network 4.5.1.0. Since it was not possible to discriminate between the variants DYS385a and b, without sequencing, it was excluded from Network analysis. Concerning locus DYS389, the length of DYS389II, expressed in number of repeat units, was subtracted to that of DYS389I in order to avoid entering redundant information in the network.

The Y Haplotype Reference Database (www.yhrd.org) was used to search worldwide haplotype frequency distribution. Multi-dimensional scaling plots from Slatkin linearized F*st* distances were obtained with the software XLSTAT (Addinsoft Corp.).

AMOVA analysis was conducted using Arlequin v3.1 software in order to evaluate genetic substructure in the Guarani population, taking into account the three geographical regions included in this study (North, middle-East and middle-West). Tests were performed on Y-chromosome and mitochondrial haplotype data; autosomal genotypes were insufficient to be used for this test.

Results obtained from Guaraní data including autosomal, mitochondrial and Y-chromosome polymorphisms were compared with others Native American groups (Kohlraush et al. 2005; Marino et al. 2007; Crossetti et al. 2008).

RESULTS and DISCUSSION

Six Guaraní populations inhabiting Misiones province, North Eastern Argentina, were analyzed by means of DNA polymorphisms located in nuclear (autosomal and Y-STRs markers) and mitochondrial genomes.

Fifteen STRs were analyzed in 63 samples. Since 58 of these individuals belonged to communities inhabiting the northern region, the results are presented as a single dataset, without referring to geographic origin. Frequency distribution and observed and expected heterozygosity are depicted in Table 1. All the STRs except TPOX and D18S51 met Hardy-Weinberg expectations, however after Bonferroni correction (1936) none of the analyzed markers departed from equilibrium. The observed heterozygosity ranged from 0.525 (TPOX) to 0.859 (PentaE). The average ± SD gene diversity over loci was 0.696 ± 0.364. The mean ± SE number of alleles per locus was 7.2 ± 2.9, similar to that observed in other Native American groups (Hutz et al. 2002; Kohlraush et al. 2005; Crossetti et al. 2008).

Thirteen systems (all systems analyzed except PentaE and PentaD) were used to establish comparison with other Native American groups previously reported (Kohlrausch et al. 2005; Crossetti et al. 2008). Genetic distances ($\delta\mu^2$) were calculated from allele frequencies and represented by means of a multidimensional scaling plots (Figure 2). The genetic distance between Guaraní from Misiones (Gu-M) and Guaraní from Brazil (Gu-Br) was the lowest (0.48) whereas the groups who exhibited the highest distance from Gu-M were the Ayoreo (1.43) and Caingang (1.03). Gu-Br is close to Gran Chaco Argentinean groups whereas the Aché are more distant from this cluster and also from Gu-M (0.73). Mantel´s correlation between genetic and linear geographical distances was not statistically significant ($p > 0.80$).

In order to test the hypothesis proposed by Altuna et al. (2006) based on male lineages that suggests a recent genetic drift effect that might have occurred in the Misiones Guaraní population, a possible excess of heterozygosity was evaluated through autosomal STRs data. Results of BOTTLENECK software are shown in Table 2. Since population bottlenecks induce a transient excess of heterozygosity, the fact of finding an observed heterozygosity that is higher than the expected (equilibrium) heterozygosity, for a large majority of loci, suggest that this population may have recently experienced a genetic bottleneck (Cornuet and Luikart 1996). By considering SMM we detected a high heterozygosity deficiency in both methods employed (see Material and Methods). This fact could be explained by, on the one hand, over sensitivity of this model for microsatellite mutation and, on the other hand, the fact that heterozygosity deficiency can occur with loci evolving under the strict one step mutation model (Cornuet and Luikart 1996). In that sense, an intermediate model such as TPM shows a significant heterozygosity deficiency only with the Heterozygosity method. The null hypothesis tested for heterozygosity excess using Wilcoxon test was accepted under both SMM and TPM models, implying that the population has not experienced any recent genetic bottleneck. Additionally, the mode-shift indicator test was employed to detect genetic bottleneck. The assumption behind this test is that a population under mutation-drift equilibrium is expected to have a larger proportion of alleles with low frequencies. The allele frequency distribution as revealed by the test was L-shaped, indicating a larger proportion of low frequency allele classes (Figure 3). On the other hand, Garza-Williamson index obtained, close to one (0.969 ± 0.155), reinforces the previously described results.

MtDNA Hypervariable Regions I and II sequences of 121 samples from six Guaraní communities are summarized in Table 3. Twenty three different lineages were observed, with a mean number of pairwise differences of 8.174 ± 3.817. All the samples analyzed belonged to one

of the four Native American hgs (Torroni et al. 1992, 1993; Forster et al. 1996; Bandelt et al. 2003). The haplogroup distribution showed that hgs A2 and D1 are the most frequent (40.5% and 36.0%, respectively) whereas hgs B2 and C1 are the less frequent (17.5% and 6.0%, respectively). No sample belonged to hg X was found (Brown et al. 1998; Dornelles et al. 2005). The sequence analysis was further extended to HVRII (Table 3), which allowed detecting specific point mutations in association with the founder haplogroups. Transitions at positions 114C→T and 143G→A were observed in 95% of the samples belonging to hg D1 (these transitions are present in subhaplogroup D1a –Achilli et al., 2008); the transitions: 146T→C, 153A→G, 207G→A and 235A→G were observed in the most frequent hg A2 and substitution 103G→A was highly frequent in hg B2 samples.

The average substitution rate was 4.8% for HVRI and 6.4% for HVRII. The most frequent substitutions observed were transitions, in a total number of 32, and only two transversions (16182 A→C and 16183 A→C, in hg B2) were detected. Haplotype diversity was 0.822 ± 0.016 and nucleotide diversity (average over loci) was 0.009± 0.004. The analysis of the HVR I and II allowed us to detect a common sequence in individuals belonging to hgs D1 and A2. Eighty nine percent of the "D1" samples share a common haplotype (MT19) while 63% of the "A2" samples share the haplotype MT1 (Table 3). All samples assigned to hg B2 (21 out of 121) by sequencing were confirmed by detecting the 9 bp COII/$tRNA^{Lys}$ deletion. Previously published reports demonstrated that this hg was present in 90% in the Aché group (Schmidt et al. 2004); 5% in the Caingang and was absent in the Guaraní from Brazil (Marrero et al. 2007a). Thirty eight percent of the hg B samples (present study) denoted a substitution in position 16241A→G, also observed in association with hg B samples in the Gaucho from Brazil, and has been proposed as a derivation of an Amazonian native sequence (Marrero et al. 2007b).

Figure 4 a, b, c and d show the median-joining network connecting the different haplotypes. A color code was employed to differentiate the three geographical regions of Misiones where the samples were obtained (North, middle-East and mid-West). Arrows indicate the modal haplotypes. Figure 4a shows a MJ network obtained with the most frequent A2 haplotypes (only samples whose haplotype frequencies >1 were considered). The highly frequent haplotype (MT1), detected in the three regions, is connected to the modal sequence (MT5) by means of two mutation steps (16291T and 207A), whereas MT6 and MT7 only differ from MT5 by mutations in HVRII. In hg B2 (Fig. 4b), the mutation points in HVRII connect the different haplotypes, while in hg C1 and hg D1 (Fig. 4c and 4d, respectively), the main differences are observed in HVRI. MT19 (Fig. 4d), the most frequent haplotype in hg D1 samples, is represented in the three geographical regions.

As it can be observed, the communities in the Middle-Eastern region exhibited the lowest variability across the different hgs. This finding could be related to the fact that these communities are the most isolated in terms of their geographical location. However, no statistically significant differences were obtained when AMOVA analysis was conducted comparing the three geographical areas ($p = 0.062 \pm 0.007$)

The investigations of nine Y-STRs in the Guaraní sample (N=98) allowed to identify 35 different haplotypes, 23 of which occurred only once. The haplotype diversity was 0.848 and the average gene diversity was 0.514 ± 0.281. Y-STR haplotype and typing results for M3, SRY4063 and YAP are shown in Table 4. Due to poor DNA recovery, incomplete haplotypes were obtained for three samples. The Y-STR haplotype defined as: DYS19/13, DYS389 I/14, DYS389 II/31, DYS390/24, DYS391/11, DYS392/14, DYS393/11, DYS385a/14, DYS385b/16 (Y4, Table 4) denoted high frequency in the Guaraní (36.73%). This particular haplotype is combined with derived state of M242 (data not shown) and M3 (Native American- specific type) and is absent in

other Argentinean aboriginal tribes such as Mapuche, Tehuelche, Pilaga, Toba and Wichi as well as in the general Misiones population (Marino et al. 2007). On the other hand, only two samples out of 82,251 haplotypes included in the worldwide database (YHRD) denoted this combination (http//www.yhrd.org, release 32). At the time this paper was written, this haplotype had only been reported in the Brazilian urban populations of Rio de Janeiro (Dominguez et al. 2007) and Manaus (http//www.yhrd.org, accession number YP000577). Comparing our results for 5 Y-STRs haplotypes (DYS19, DYS390, DYS391, DYS392 and DYS393), summarized in Table 4, with previously published data, we noted that haplotypes Y4, Y5, Y6, Y8, Y9, Y10, Y12, Y15 and Y30, were reported in some Brazilian communities, such as: Guaraní (Bortolini et al. 2003; Leite et al. 2008), Caingang (Leite et al. 2008) and Gaucho (Marrero et al. 2007b).

Figure 5 depicts the median-joining network connecting the haplotypes belonging to Q1a3a individuals. Since locus DYS385 was not included for this calculation and a special consideration was employed for the locus DYS389 (see Material and Method), some nodes included more than one haplotype of those described in Table 4 (for example, Y4 includes Y4+Y10). The M-J shows the relationship between the different haplotypes with the modal haplotype, represented by Y0, proposed by Bianchi et al. (1998). The most frequent Guaraní haplotype, Y4, appears to be connected to the modal haplotype by means of three mutational events. Color codes were used to differentiate the three geographical areas to which the communities belonged. The AMOVA analysis showed no significant differences between them ($p = 0.855 \pm 0.010$).

Eighty four percent of the sample denoted the DYS19/13 variant which agrees with other Native American groups studied (Bianchi et al. 1997; Ruiz-Linares et al. 1999; Forster et al. 2000). On the other hand variant DYS393/11, is highly frequent in the Guaraní samples (41.8%),

but rare in other populations. In other Latin American populations, this variant is present with a frequency of less than 1% in admixed population from Colombia (Yunis et al. 2005); Ecuador (Baeza et al. 2007); Perú (Iannacone et al. 2005) and in Brazilian populations (de Souza Góes et al. 2005; Grattapaglia et al. 2005; Domingues et al. 2007).

Eighty eight percent of the Guaraní samples (87/98) displayed the derivate state of M3 indicating that these samples belong to Q1a3a-M3 Native American specific subhaplogroup (Karafet et al. 2008). Eleven samples exhibited the ancestral state at M3 (Table 4). Three out of eleven C samples showed a derivate state of YAP and SRY 4064 and could belong to Clade E, highly frequent in Africa, and moderately frequent in Middle East and southern Europe (Karafet et al. 2008).

Figure 6 shows a multidimensional scaling plot of Slatkin linearized F*st* distance values of six populations: Gu-M (Guaraní from Misiones, present study); Gu-Br (Guaraní from Brazil, Bortolini et al. 2003); Pilaga, Wichi and Toba (Northern Argentina, Marino et al. 2007; http://www.yhrd.org, Pop. ID YP000016, YP000018 and YP000019), and Aché (Eastern Paraguay, Bortolini et al. 2003). In this comparison, only Q1a3a-M3 samples were included, and only five Y-STRs systems were taken into account (DYS19, DYS390, DYS391, DYS392 and DYS393). Although all distances were significant ($p < 0.05$), except between Pilaga and Toba, both belonging to the Guaicurú linguistic sub branch (Ruhlen and Greenberg 2007), the figure depicts a clear genetic proximity between the Gu-M respect to the Gu-Br. Aché group appears to be isolated, partly due to a unique haplotype found in this population.

In our analysis, it was observed that samples within the Argentinean Guaraní populations exclusively exhibited Native American mitochondrial motives and non-Native American Y-

chromosome haplogroups was restricted to as low as 11%. In contrast, no sample exhibited Native Amerincan Y-chromosome in combination with non- Native American mtDNA, showing that mating between aboriginal female with non-aboriginal male is more frequent than the alternative. The percentage of indigenous women contribution was estimated as 100% on the basis of the proportion of typically Native American haplogroups, whereas the male autoctonous contribution to the gene pool was 89%.

Statistically significant genetic distances were observed between Guaraní from Misiones and all of the Native American groups compared. Nevertheless, Guaraní from Brazil, which belongs to the same linguistic group, exhibited the lowest distance. Previously published genetic studies suggested that the Aché represents a group of admixed biological origin, ancestrally related to the Tupí-Guaraní and also to the Jé linguistic family (Gaspar et al. 2002; Callegari-Jacques et al. 2008). However, beside the significant genetic distances between Aché and Guaraní from Misiones, the Y-STR haplotypes and mitochondrial haplogroups distributions previously reported for this group (Bortolini et al. 2003; Schmitt et al. 2004) exhibited a different figure than the observed in Guaraní from North-Eastern Argentina.

Genetic evidence including a reduced mitochondrial nucleotide and Y-chromosome gene diversities, and a highly frequent Y-STR haplotype could be in agreement with the hypothesis of a founder effect (Altuna et al. 2006) in Argentinean Guarani people. However, the analysis based on autosomal STRs does not support the hypothesis or was insufficient to prove it. Further investigation using ancestry sensitive markers could provide stronger evidences for testing the proposed hypothesis.

The present results allowed detecting some genetic peculiarities in Guaraní of Argentina. Its characteristic allowed us to distinguish them from other Native American groups

geographically, ethnically or linguistically related. Further studies would be necessary to explain and to date the peopling process, occurred along their migration from Amazonia or more recently, their re-entry into Northeastern Argentina.

Acknowledgments This work was supported, in part, by grants PIP6114-CONICET and UBACyT.B-047, to DC, and grant Res.1106-CONICET to AS. DC and AS are members of Consejo Nacional de Investigaciones Científicas y Tecnológicas. We thank Dr. MC Bortolini for kindly providing Aché sample data and Dr. Michael Forster for critical revision of the median-joining analysis.

LITERATURE CITED

Achilli, A., U.A. Perego, C.M. Bravi et al. 2008. The Phylogeny of the Four Pan-American MtDNA Haplogroups: Implications for Evolutionary and Disease Studies. PLoS ONE 3(3): e1764. doi:10.1371/journal.pone.0001764.

Altuna, M.E., N.M. Modesti, D.A. Demarchi. 2006.Y-Chromosomal Evidence for a Founder Effect in Mbyá-Guaraní Amerindians from Northeast Argentina. *Hum Biol* 78:635–639.

Baeza, C., R. Guzman, M. Tirado et al. 2007. Population data for 15 Y-chromosome STRs in a population sample from Quito (Ecuador). *Forensic Sci Int* 173:214-219.

Bandelt, H.J., and P. Forster. 1995. *Network: Phylogenetic Network Software.* Available at http://www.fluxus-engineering.com/netwinfo.htm.

Bandelt, H.J., C. Herrnstadt, Y.G. Yao et al. 2003. Identification of Native American founder mtDNAs through the analysis of complete mtDNA sequences: some caveats. *Ann Hum Genet* 67:512-524.

Bianchi, N.O., G. Baillet, C.M. Bravi et al. 1997. Origin of Amerindian Y-Chromosomes as Inferred by the Analysis of Six Polymorphic Markers. *Am J Phys Anthropol* 102:79–89.

Bianchi, N.O., C. Catanesi, G. Bailliet et al. 1998. Characterization of ancestral and derived Y-chromosome haplotypes of New World native populations. *Am. J. Hum. Genet*. 63:1862-1871.

Bonferroni, C. E. 1936. Teoria statistica delle classi e calcolo delle probabilit `a. Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze 8, 3-62.

Bortolini, M.C., F.M. Salzano, M.G. Thomas et al. 2003. Y-Chromosome Evidence for Differing Ancient Demographic Histories in the Americas. *Am J Hum Genet* 73:524- 539.

Brandstätter, A., H. Niederstätter, W. Parson W. 2004. Monitoring the inheritance of heteroplasmy by computer-assisted detection of mixed basecalls in the entire human mitochondrial DNA control region. *Int J Legal Med*. 118: 47-54.

Brown, M.D., S.H. Hosseini, A. Torroni et al. 1998. mtDNA haplogroup X: an ancient link between Europe/Western Asia and North America?. *Am J Hum Genet* 63:1852-1861.

Callegari-Jacques, S.M., K. Hill, A.M. Hurtado et al. 2008. Genetic clues about the origin of Aché hunter-gatherers of Paraguay. *Am J Hum Biol* 20:735-737.

Corach, D., L. Filgueira Risso, M. Marino et al. 2001. Routine Y-STR typing in forensic casework. *Forensic Sci Int* 118:131-135.

Cornuet, J.M. and G.Luikart . 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 144:2001-2014.

Crossetti, S.G., D.A. Demarchi, P.E. Raimann et al. 2008. Autosomal STR genetic variability in the Gran Chaco Native Population: Homogeneity or Heterogeneity?. *Am J Hum Biol* 20: 704-711.

Crivos, M., M.R. Martínez, M.L. Pochettino et al. 2007. Pathways as "signatures in landscape": towards an ethnography of mobility among the Mbya-Guaraní (Northeastern Argentina). *J Ethnobiol Ethnomed* 5;3:2.

de Souza Góes, A.C., E.F. de Carvalho, I. Gomes et al. 2005. Population and mutation analysis of 17 Y-STR loci from Rio de Janeiro (Brazil).*Int J Legal Med* 119:70-76.

Di Rienzo A., A. C. Peterson, J. C. Garza et al. 1994. Mutational processes of simple-sequence repeat loci in human populations. *Proc. Natl. Acad. Sci*. USA 91: 3166-3170.

Domingues, P.M., L. Gusmao, D.A. Silva et al. 2007. Sub-Saharan Africa descendents in Rio de Janeiro (Brazil): population and mutational data for 12 Y-STR loci. *Int J Legal Med* 121:238-241.

Dornelles, C.L., S.L. Bonatto, L.B. De Freitas et al. 2005. Is haplogroup X present in extant South American Indians?. *Am J Phys Anthropol* 127: 439-448.

Excoffier, L., G. Laval, S. Schneider. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evol Bioinform Online* 1:47-50.

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

Forster, P., A. Röhl, P. Lünnemann et al. 2000. A short Tandem Repeat-Based Phylogenny for the Human Y Chromosome. *Am J Hum Genet* 67: 182-196.

Garlet, I. 1997. Mobilidade Mbya: Historia e Significacao.Masters thesis. Catholic Pontifice University of Rio Grande do Sul.

Gaspar, P.A., M.H. Hutz F.M. Salzano et al. 2002. Polymorphisms of CYP1a1, CYP2e1, GSTM1, GSTT1, and TP53 genes in Amerindians. *Am J Phys Anthropol*.119:249-256.

Goldstein, D.B., A. Ruiz-Linares, L.L. Cavalli-Sforza et al. 1995. Genetic absolute dating based on microsatellites and the origin of modern humans. *Proc Natl Acad Sci* USA 92:6723-6727.

Grattapaglia, D., S. Kalupniek, C. Guimaraes et al. 2005. Y-chromosome STR haplotype diversity in Brazilian populations. *Forensic Sci Int* 149:99-107.

Ginther, C., D. Corach, G.A. Penacino et al. 1993. Genetic variation among the Mapuche Indians from the Patagonian region of Argentina: mitochondrial sequence variation and allele

frequencies of several nuclear genes. In: Pena SDJ, Chakraborty R, Epplen JT, Jeffreys AJ, editors. DNA Fingerprinting: State of the Science. Berlin: Birkhäuser Verlag. p 211-219.

Hammer, M.F. 1995. A recent common ancestry for human Y chromosomes. *Nature* 378: 376-378.

Hutz, M.H., S.M. Callegari-Jacques, S.E.M. Almeida SEM et al. 2002. Low levels of STRP variability are not universal in Amerindians. *Hum Biol* 74:791-806.

Iannacone, G.C., R.Y. Tito, P.W. Lopez et al. 2005. Y-chromosomal haplotypes for the PowerPlex Y for twelve STRs in a Peruvian population sample. *J Forensic Sci* 50:239-242.

INDEC: Instituto Nacional de Estadísticas y Censos. Encuesta Complementaria de Pueblos Indígenas (ECPI) 2004-2005. (http/ www. Indec.mecon.ar).

Karafet, T.M., F.L. Mendez, M.B. Meilerman et al. 2008. New binary polymorphisms reshape and increase resolution of the human Y Chromosomal haplogroup tree. *Genome Res*.18: 830-838

Kohlrausch, F.B., S.M. Callegari-Jacques, L.T. Tsuneto et al. 2005. Geography Influences Microsatellite Polymorphism Diversity in Amerindians. *Am J Phys Anthropol* 126:463–470.

Luikart, G., W.F. Sherwin, B.M. Steele et al. 1998. Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. *Molecular Ecology* 7: 963-974.

Leite, F.P., S.M. Callegari-Jacques, B.A. Carvalho et al. 2008. Y-STR Analysis in Brazilian and South Amerindian Populations. *Am J Hum Biol* 20:359–363.

Marino, M., A. Sala, D. Corach. 2007. Genetic attributes of the YHRD minimal haplotype in 10 provinces of Argentina. *Forensic Sci Int Genet* 1:129-133.

Marrero, A.R., W.A. Silva-Junior, C.M.Bravi et al. 2007a. Demographic and evolutionary trajectories of the Guaraní and Kaingang natives of Brazil. *Am J Phys Anthropol* 132:301-310.

Marrero, A.R., C.M. Bravi, S. Stuart et al. 2007b. Pre- and post-Columbian gene and cultural continuity: the case of the Gaucho from southern Brazil. *Hum Hered* 64:160-171.

Martinez Sarasola, C. 2005. Nuestros paisanos los indios. Emecé Edts. Buenos Aires. Argentina.

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

Ohta, T., and M. Kimura. 1973. A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. *Genet. Res. Camb*. 22:201-204.

Ruhlen, M, J. Greenberg. 2007. An Amerindian Etymological Dictionary. Stanford University press.

Ruiz-Linares, A., D. Ortíz-Barrientos, M. Figueroa et al. 1999. Microsatellites provide evidence for Y chromosome diversity among founders of the New World. *Proc Natl Acad Sci* 96:6312-6317.

Sala, A., M.Marino, C.F. Arguelles et al. 2004. Uniparentally inherited genetic markers as tools for ethnic and geographical origin detection of forensic samples. *Progress in Forensic Genetics 10*, International Congress Series 1261:625-627.

Schmitt, R., S.L. Bonatto, L.B. Freitas et al. 2004. Extremely limited mitochondrial DNA variability among the Aché Natives of Paraguay. *Ann Hum Biol* 31: 87-94.

Seielstad, M., N. Yuldasheva, N. Singh et al. 2003. A novel Y-chromosome variant puts an upper limit on the timing of first entry into the Americas. *Am J Hum Genet* 73:700-705.

Tarragó, M.N. 2000. Nueva Historia Argentina. Los pueblos originarios y la conquista. Edts Sudamericana. Buenos Aires.

Tarazona-Santos, E., D.R. 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* 68:1485-1496.

The Y Chromosome Consortium. 2002. A nomenclature system for the tree of human Y-chromosome binary haplogroups. *Genome Res*. 12:339-348.

Torroni, A., T.G. Schurr, C.C. Yang et al. 1992. Native American mitochondrial DNA analysis indicates that the Amerind and the Nadene populations were founded by two independent migrations. *Genetics* 130:153-162.

Torroni, A., T.G. Schurr, M.F. Cabell et al. 1993. Asian affinities and continental radiation of the four founding Native American mtDNAs. *Am J Hum Genet* 53: 563-590.

Torroni, A. J.V. Neel, R. Barrantes et al. 1994. Mitochondrial DNA "clock" for the Amerinds and its implications for timing their entry into North America. *Proc Natl Acad Sci* 91:1158-1162.

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* 93:196-200.

Wrischnik, L.A., R.G. Higuchi, M. Stoneking et al. 1987. Length mutations in human mitochondrial DNA: direct sequencing of enzymatically amplified DNA. *Nucleic Acids Res* 15:529-542.

Yunis, J.J., L.E. Acevedo, D.S. Campo et al. 2005. Population data of Y-STR minimal haplotypes in a sample of Caucasian-Mestizo and African descent individuals of Colombia. *Forensic Sci Int* 151:307-313.

Table 1: Allele frequency distribution in Guaraní population from Misiones, Argentina

Allele	D3S1358	THO-1	D21S11	D18S51	PentaE	D5S818	D3S1358	D7S820	D16S539	CSF1PO	PentaD	vWA	D8S1179	TPOX	FGA
4					0,009										
5					0,009										
6		0,443												0,009	
7		0,295			0,018	0,421	0,016								
8		0,008			0,009		0,008	0,008			0,161			0,302	
9		0,057				0,032	0,238	0,087	0,270		0,214			0,009	
9,3		0,197													
10					0,018	0,024	0,008	0,127	0,198	0,170	0,161		0,063	0,009	
11					0,027	0,413	0,357	0,460	0,167	0,304	0,018		0,063	0,422	
12				0,183	0,255	0,032	0,143	0,286	0,198	0,473	0,188		0,063	0,250	
13				0,225	0,045	0,079	0,190	0,032	0,167	0,036	0,188		0,310		
14	0,008			0,500	0,027		0,032			0,018	0,063	0,065	0,444		
15	0,365			0.033	0,118		0,008					0,065	0,040		
16	0,421			0,042	0,082							0,202	0,016		
17	0,103				0,082						0,009	0,540			
18	0,071			0,008								0,129			
19	0,032			0,008	0,055										0,060
20					0,064										0,009
21					0,145										0,069
22					0,036										0,026
23															0,336
24															0,164
25															0,207
26															0,112
28			0,040												
29			0,278												0,017
30			0,365												
31			0,063												
31,2			0,079												

32,2			0,151												
33,2			0,024												
N	63	61	63	60	55	63	63	63	63	56	56	62	63	58	58
Ho	0.619	0.730	0.809	0.550	0.859	0.698	0.777	0.666	0.714	0.66	0.807	0.714	0.666	0.525	0.745
He	0.672	0.688	0.762	0.663	0.883	0.649	0.766	0.687	0.799	0.659	0.834	0.633	0.698	0.680	0.799
p	0.596	0.076	0.615	0.013	0.809	0.081	0.470	0.09	0.111	0.522	0.899	0.939	0.746	0.009	0.075

N: Number of genotypes; Ho: Observed heterozygosity¸He: Expected heterozygosity, *p*: probability.

Table 2: Results of the Bottleneck analysis

Methods	Models	Sign test	Standardized Diferences Test	Wilcoxon test
(A)Frequency method	SMM	Hee= 8.91	T_2= -3.016	*P* (one tail for H deficiency): 0.007
		Hd= 12	*P*= 0.001	*P* (one tail for H excess): 0.993
		He= 3		*P* (two tails for H excess and deficiency): 0.015
		P= 0.002		
	TPM	Hee=8.96	T_2= -0.897	*P* (one tail for H deficiency): 0.126
		Hd=10	*P* = 0.184	*P* (one tail for H excess): 0.885
		He=5		*P* (two tails for H excess and deficiency): 0.252
		P= 0.035		
(B) Heterozygosity method	SMM	Hee= 8.88	T_2= -5.517	*P* (one tail for H deficiency): 0.004
		Hd= 12	*P* = 0.000	*P* (one tail for H excess): 0.998
		He= 3		*P* (two tails for H excess and deficiency): 0.008
		P= 0.002		
	TPM	Hee=8.93	T_2= -2.741	*P* (one tail for H deficiency): 0.047
		Hd=9	*P* = 0.003	*P* (one tail for H excess): 0.958
		He=6		*P* (two tails for H excess and deficiency): 0.094
		P= 0.101		

Hee: heterozygosity excess expected; Hd: heterozygosity deficiency; He: heterozygosity excess; *P*: probability; SMM: stepwise mutation model; TPM: two-phase model. Parameters for TPM: variance= 30.00; proportion of SMM in TPM= 90%; estimation based on 1000 replications.

Table 3: Distribution of mtDNA haplotypes in a Guaraní sample (N=121)

				1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1																		
				6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6																		
				0	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	3	3	3	3			1	1	1	1	1	1	2	2	2	2	2	2	2	3	3	3
				5	0	1	4	7	7	8	8	8	8	1	2	4	6	9	9	9	9	1	2	2	6	6	7	0	1	4	4	5	5	0	0	3	4	6	9	9	0	0	1
				1	8	1	2	2	9	2	3	7	9	7	3	1	6	0	1	5	8	9	5	7	2	4	3	3	4	3	6	2	3	4	7	5	9	3	0	1	9	9	5
																																									.	.	.
																																									1	2	1
Mt	Hg	N	r	A	C	C	C	T	C	A	A	C	T	T	C	A	C	C	C	C	T	G	T	C	T	C	A	A	C	G	T	T	A	T	G	A	A	A	A	A	-	-	-
1	A	31		.	.	T	.	.	.	.	.	.	.	.	T	.	.	T	T	.	.	A	.	.	C	T	G	.	.	.	C	.	G	.	A	G	.	G	.	.	C	.	C
2		1		.	.	T	.	.	.	.	.	.	.	.	T	.	T	T	.	.	.	A	.	.	C	T	G	.	.	.	C	.	.	.	.	G	.	G	.	.	C	C	C
3		1		.	.	T	.	.	.	.	.	.	.	.	T	.	.	T	T	.	.	A	.	.	C	T	G	.	.	.	C	.	G	.	A	G	.	G	.	.	.	.	C
4		1		.	.	T	.	.	.	.	.	.	.	.	T	.	.	T	T	.	.	A	.	.	C	T	G	.	.	.	C	.	G	.	A	G	.	G	.	.	C	C	C
5		6		.	.	T	.	.	.	.	.	.	.	.	T	.	.	T	.	.	.	A	.	.	C	T	G	.	.	.	C	.	G	.	.	G	.	G	.	.	C	.	C
6		2		.	.	T	.	.	.	.	.	.	.	.	T	.	.	T	.	.	.	A	.	.	C	T	G	.	.	.	C	.	G	.	.	G	.	G	.	.	C	C	C
7		4		.	.	T	.	.	.	.	.	.	.	.	T	.	.	T	.	.	.	A	.	.	C	T	G	.	.	.	C	.	.	C	.	G	.	G	.	.	.	.	C
8		1		.	T	T	.	.	.	.	.	.	.	.	T	.	.	T	.	.	.	A	.	.	C	T	G	.	.	.	C	.	.	C	.	G	.	G	.	.	.	.	C
9		1		.	T	T	.	.	.	.	.	.	.	.	T	.	.	T	.	.	.	A	.	.	C	T	G	.	.	.	C	.	G	C	.	.	.	G	.	.	.	.	C
10		1		.	.	T	.	.	.	.	.	.	.	.	T	.	.	T	T	.	.	A	.	.	C	T	G	.	.	.	C	.	G	.	.	G	.	G	.	.	C	.	C
		49																																									

11	B	8	. . . . C . C C . G . . . . G G . . C C . . . . . G . . C . C
12		6	. . . . C . C C . . . . G . . . . G . . C . C
13		1	. . . . . . C C . C C . . . . G . . . . G . . C C C
14		6	. . . . C . C C . . . . G . . . . G . . C C C
		21	
15	C	4	. . . . T . . T . . . . . C . C T . . G . . . . d G d d C . C
16		1	. . . . C . T . . . . . C . C T . . G . . . . d G d d C . C
17		1	G . . . C . . . . . . T . . . . . C . C T . . G . . . . d G d d C . C
18		1	. . . . T . . . . C T . . G . . . . d G d d C . C
		7	
19	D	39	. . . . T . . . . C . C . G . T A . . . . G . . C . C
20		2	. . . . C . T . . . . C . C . G . T A . . . . G . . C . C
21		1	. . . T . T . . . . . T . . . . T . . C . C . G . . . . G . . C . C
22		1	. . . . T . . . . C . C . G . . A . . . . G . . C . C
23		1	. . . . T . . . . C . C . G . T A . . . . G . . C C C
		44	

Mt: refers to the different haplotypes observed; Hg: refers to the different haplogroups observed, r: Cambridge Reference Sequence. Vertical numbers refers to mtDNA position

Table 4: Y-STR Haplotype distribution

M3	YAP	406 4	Y-STR Haplotype[1]	Kaaguy Poty (15)[2]	Tabay (8)	Jejy (7)	Yaboti (6)	Yriapu (10)	Mborere (52)	Total (98)	Code[3]
*	-	-	13 13 29 24 10 14 13 12 20	1	0	0	0	0	0	1	Y1
*	-	-	13 13 30 25 10 14 13 14 16	1	0	0	0	0	0	1	Y2
*	-	-	13 13 31 22 10 14 14 16 18	4	1	0	1	0	3	9	Y3
*	-	-	13 14 31 24 11 14 11 14 16	3	7	2	3	3	18	36	Y4
*	-	-	13 13 29 24 10 14 13 13 20	2	0	0	0	0	0	2	Y5
*	-	-	13 13 31 22 10 14 13 15 18	1	0	1	1	0	2	5	Y6
*	-	-	13 13 31 25 10 14 13 14 14	0	0	1	0	1	0	2	Y7
*	-	-	13 13 30 25 10 14 13 14 14	0	0	0	0	0	5	5	Y8
*	-	-	14 12 29 23 10 14 13 13 16	0	0	0	0	1	1	2	Y9
*	-	-	13 14 31 24 11 14 11 15 16	0	0	0	0	2	0	2	Y10
*	-	-	14 12 29 24 10 14 13 13 18	0	0	0	0	1	0	1	Y11
*	-	-	13 14 31 24 11 14 13 14 16	0	0	0	0	1	3	4	Y12
*	-	-	14 12 29 23 10 14 13 13 18	0	0	0	0	1	0	1	Y13
*	-	-	13 13 31 22 10 13 13 15 18	0	0	0	1	0	0	1	Y14
*	-	-	13 12 30 24 10 14 13 14 17	0	0	0	0	0	2	2	Y15
*	-	-	13 12 29 25 10 15 12 14 14	0	0	0	0	0	4	4	Y16
*	-	-	13 13 31 22 10 14 14 15 18	0	0	0	0	0	1	1	Y17
*	-	-	13 12 27 24 10 11 13 14 17	0	0	0	0	0	1	1	Y18
*	-	-	13 14 31 25 12 14 11 14 16	0	0	0	0	0	1	1	Y19
*	-	-	13 14 31 23 11 14 11 14 16	0	0	0	0	0	1	1	Y20

* - -	13 14 31 24 10 14 11 14 16	0	0	0	0	0	1	1	Y21
* - -	14 13 30 25 12 13 13 11 11	1	0	0	0	0	0	1	Y22
* - -	13 15 ? 24 11 14 11 20 22	0	0	1	0	0	0	1	Y23
* - -	14 14 ? 24 11 14 14 14 17	0	0	1	0	0	0	1	Y24
* - -	13 ? ? 22 10 14 14 16 18	0	0	1	0	0	0	1	Y25
- * *	13 12 29 25 10 11 13 17 17	0	0	0	0	0	1	1	Y26
- * *	13 13 31 24 9 11 13 17 19	0	0	0	0	0	1	1	Y27
- * *	15 13 30 24 10 13 13 18 19	0	0	0	0	0	1	1	Y28
- - -	14 13 29 24 10 13 13 11 14	0	0	0	0	0	1	1	Y29
- - -	14 14 30 24 11 13 13 11 14	2	0	0	0	0	0	2	Y30
- - -	14 13 29 23 11 13 13 11 14	0	0	0	0	0	1	1	Y31
- - -	14 14 30 24 11 14 13 11 15	0	0	0	0	0	1	1	Y32
- - -	14 13 31 24 10 13 14 11 14	0	0	0	0	0	1	1	Y33
- - -	14 13 29 25 10 13 13 11 14	0	0	0	0	0	1	1	Y34
- - -	14 14 30 24 11 13 13 11 15	0	0	0	0	0	1	1	Y35

[1]Y-STR haplotype: DYS19, DYS389I/II; DYS390, DYS391, DYS392, DYS393, DYS385a/b;
[2]Sample size are presented between brackets
[3]Code: haplotype name from Y1 to Y35.
Codes for M3, YAP and SRY4064: (*) refers to derivate state, (-) refers to ancestral type

Figure Captions

Figure 1: Map of Misiones province, Northeastern Argentina, showing geographic location of the communities sampled. MB: Fortín M´Borere (54ºW, 25ºS) and YU: Yryapu (54ºW, 25ºS) (Iguazú Department); KP: Kaaguy Poty (55ºW, 27ºS) and TA: Tabay (55ºW, 26ºS) (San Martín Department); JE: Jejy (54ºW, 26ºS) and YA: Yabotí (54ºW, 26ºS) (Guarani Department).

Figure 2: MDS plot of linearized Slatkin distance showing the relationship between Guaraní from Misiones (present study), Caingang and Guaraní from Brasil, Aché and Ayoreo from Paraguay (Kohlrausch et al. 2005); Pilaga, Toba and Wichi from North Argentina (Crossetti et al. 2008), based on 15 autosomal STR. References: Gu-M and Gu-Br: Guaraní from Misiones and Brasil, respectively; To-F and Wi-F: Toba and Wichi from Formosa province; To-Ch and To-F: Toba and Wichi from Chaco province. Kruskal stress 0.141.

Figure 3: L-shaped mode shift graph showing lack of bottleneck in Guaraní population.

Figure 4: Median-joining network for mtDNA haplogroups. A, B, C and D refers to m-j for each Native American hg. Arrows indicate a founder haplotype. Areas of the circles are proportional to the number of individuals. Color codes: black for Northern (MB+YU); white for middle-East (JE+YA) and light gray for middle-West (TA+KA).

Figure 5: Median-joining network for Y-STR haplotypes (Loci DYS19, DYS390, DYS391, DYS392 and DYS393). Areas of the circles are proportional to the number of individuals. Color

codes: black for Northern (MB+YU); white for middle-East (JE+YA) and light gray for middle-West (TA+KA). Y0 indicates the founder haplotype proposed by Bianchi et al.1998.

Figure 6: MDS based on five Y-STRs. Gu-M: Guaraní from Misiones (present study): GuB: Guaraní (from Brazil) and Aché (from Paraguay) (Bortolini et al. 2003); Pilaga, Wichi and Toba from North Argentina (Marino et al. 2007; http://www.yhrd.org, Pop. ID YP000016, YP000018 and YP000019). Kruskal stress 0.00016