6-2-2014

Mitochondrial DNA variability among six South-American Amerindian villages from the Pano linguistic group

Celso T. Mendes-Junior
Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Brazil

Aguinaldo L. Simoes
Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Brazil

Mitochondrial DNA variability among six South-American Amerindian villages from the Pano linguistic group

Celso T. Mendes-Junior,[1] Aguinaldo L. Simoes[2]

[1] Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Brazil.

[2] Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Brazil.

Keywords: mtDNA, genetic diversity, Amerindians, South America

Suggested running head: "mtDNA diversity in Pano Amerindians from Brazil"

Abstract. Although scattered throughout a large geographic area, the members of the *Pano* linguistic group present strong ethnic, linguistic and cultural homogeneity, a feature that causes them to be considered as components of a same "*Pano*" tribe. Nevertheless, the genetic homogeneity between *Pano* villages has not been examined before. To study the genetic structure of the *Pano* linguistic group, four major Native American mitochondrial DNA (mtDNA) founder haplogroups were analyzed in 77 Amerindians from six villages of four

Pano tribes (*Katukina*, *Kaxináwa*, *Marúbo*, and *Yaminawa*) located in the Brazilian Amazon. The central position of these tribes in the continent makes them relevant for attempts to reconstruct population movements in South America. Except for a single individual that presented an African haplogroup L, all remaining individuals presented one of the four Native American haplogroups. Significant heterogeneity was observed between the six *Pano* villages. Although Amerindian populations are usually characterized by considerable interpopulational diversity, the high heterogeneity level observed is unexpected if the strong ethnic, linguistic and cultural homogeneity of the *Pano* linguistic group is taken into account. The present findings indicate that the ethnic, linguistic and cultural homogeneity does not imply genetic homogeneity. Even though the genetic heterogeneity uncovered may be a female-specific process, the most probable explanation for that is the joint action of isolation and genetic drift as major factors influencing the genetic structure of the *Pano* linguistic group.

Introduction

The Central West region of the Brazilian Amazon attracts attention due to its central position in South America, which makes it relevant for attempts to reconstruct the population movements and biological relationships of South American indigenous populations. Many tribes inhabit the western part of the northern region of Brazil, including several that are classified as *Pano*. They

descend from populations that used to inhabit the eastern Andean slopes of Ecuador. At the end of the 17th century, they experienced some migrational events that culminated in their distribution along the Juruá and Purus rivers in the Brazilian territory, and in adjacent portions of Peru (Mohrenweiser et al. 1979; Salzano and Jacques 1979).

At that time, they suffered a series of fissions that gave rise to the current tribes (Salzano and Jacques 1979). In spite of their large geographic distribution, the *Pano* group attracts attention for its remarkable ethnic homogeneity, reinforced by outstanding cultural and linguistic cohesion (Erikson 1998). For instance, a linguistic analysis based on the cognate density disclosed a high degree of mutual intelligibility (Mohrenweiser et al. 1979), even greater than that encountered among the various subdivisions of the Yanomama (Spielman et al. 1974). Therefore, the several *Pano* groups are sometimes considered as members of a single *Pano* tribe (Mohrenweiser et al. 1979; Salzano and Jacques 1979).

The genetic structure of the *Pano* group has not been comprehensively studied. So far, the analysis of classical genetic markers, including both blood groups and erythrocyte and serum proteins (*ABO, Kell, MNSs, P, Rh, Duffy, Kidd, Diego, Haptoglobin, GC, Transferrin and Transferrin, Lewis, Ceruloplasmin, Gm, Km, Phosphoglucomutase, Acid Phosphatase, Galactose-1-Phosphate Uridyltransferase* and Esterase D), revealed that the Brazilian and Peruvian *Kaxináwa* (one of the *Pano* tribes) can be grouped in a common gene pool, while

Kaxináwa and *Katukina* (another *Pano* tribe) are somewhat more heterogeneous (Salzano and Jacques 1979). The only DNA markers that were exhaustively studied in *Pano* populations were two *HLA-G* polymorphisms: a cytosine deletion (ΔC) at codon 130 in exon 3 that determines a premature stop codon (Mendes-Junior et al. 2007a) and a 14-bp insertion/deletion polymorphism at exon 8 of the *HLA-G* gene (Mendes-Junior et al. 2007b). While the ΔC was not observed in the four *Pano* tribes, the 14-bp insertion/deletion polymorphism revealed heterogeneous insertion frequencies that ranged from 36.11% (*Marúbo*) to 77.78% (*Kaxináwa*) (Mendes-Junior et al. 2007a; Mendes-Junior et al. 2007b).

Mitochondrial DNA (mtDNA) haplogroups, which present a matrilineal pattern of inheritance, have yet to be extensively analyzed among the *Pano* Indians. Although mtDNA markers have already been analyzed in many South American populations (Bisso-Machado et al. 2012; Salzano 2002), the Central West region of the Brazilian Amazon remains poorly studied, suggesting that further efforts regarding mtDNA typing in Amazonian isolated populations are warranted (Mendes-Junior and Simoes 2009). In addition, there is an almost complete absence of Paleoindian archeological sites in the Amazon basin (Fiedel 2000; Salzano and Callegari-Jacques 1988), rendering the genetic analysis of contemporary isolated populations the only way of retrieving information for making reliable inferences about ancient migration routes in that region.

The aim of this study was to analyze a set of mtDNA markers in 77 individuals spread across six *Pano* villages from four tribes (*Katukina*, *Kaxináwa*, *Marúbo*, and *Yaminawa*) sampled in an expedition that took place in 1976, when admixture with non-indians reached only very low levels. This approach would allow an examination of the female genetic structure of this group of tribes and the quantification of the European and African contributions to the feminine gene pool of isolated indigenous populations of the Central Amazon.

In view of the historic records that indicate a recent origin, associated with linguistic, cultural and ethnic homogeneity (Erikson 1998; Mohrenweiser et al. 1979; Salzano and Jacques 1979), considerable genetic homogeneity would also be expected. Notwithstanding that, highly significant differences between the six *Pano* villages were observed.

Materials and Methods

Populations. The *Pano* represents a large group of Amerindians distributed along the Brazilian (southwestern portion of the Amazonas State and western half of Acre State), Peruvian, and Bolivian territories (Povos Indígenas no Brasil, http://pib.socioambiental.org/en/c/quadro-geral). At the time of sample collection (1976), there were about 18 thousand *Pano* members (Gershowitz and Neel 1978;

Mohrenweiser et al. 1979). They speak Macro-Panoan languages of the Ge-Pano-Carib phylum (Greenberg 1987).

The samples analyzed in this study were from six villages located in the Central West region of the Brazilian Amazon (Fig. 1), visited during the Alpha Helix expedition that took place in the summer of 1976. The 77 chosen Amerindians, without direct kinship (siblings or parents/sibs), belong to four tribes [*Katukina* (Morada Nova and Sete Estrelas villages), *Kaxináwa* (Cana Brava and Paredão villages), *Marúbo* (Vida Nova village), and *Yaminawa* (nameless village)] distributed along the States of Acre and Amazonas, Brazil. At the time of sample collection these four tribes were composed of about 700, 2000, 400 and 410 indigenous members, respectively. Even though these Amerindians live in somewhat isolated communities, they have had a long history of contact with the outside world. Despite the many years of contact with non-Indians, the number of persons of mixed or non-Indian ancestry living in these communities is generally small (Salzano and Jacques 1979). Moreover, low levels (below 1%) of interethnic admixture assessed by genetic markers have been reported (Gershowitz and Neel 1978; Mohrenweiser et al. 1979; Salzano and Callegari-Jacques 1988). Regarding social organization, while matrilinearity holds in *Katukina*, *Marúbo*, and *Yaminawa*, the *Kaxináwa* tribe is organized into patrilinear groups (Mohrenweiser et al. 1979) (Povos Indígenas no Brasil, http://pib.socioambiental.org/en). Detailed descriptions of these populations such

as geographic distribution and demographic information can be found elsewhere (Gershowitz and Neel 1978; Luizon et al. 2008; Mestriner et al. 1980; Mohrenweiser et al. 1979; Salzano and Jacques 1979).

This study was approved by the Comitê de Ética em Pesquisa of this Institution (Hospital das Clínicas - FMRP, USP), and by the Comissão Nacional de Ética em Pesquisa, according to process HCRP nº 7869/2004 and CONEP nº 25000.120707/2004–02, respectively.

Laboratory analysis. Blood samples were collected in 1976 and processed as described elsewhere (Luizon et al. 2008; Mendes-Junior et al. 2007a). Four polymorphisms that define the Native American A (+663 *Hae*III), B (9bp deletion), C (-13259 *Hinc*II), and D (-1571 *Alu*I) haplogroups were analyzed according to previously reported conditions (Keyeux et al. 2002; Stone and Stoneking 1993). Moreover polymorphisms that define the X (+1715 *Dde*I) and L (+3592 *Hpa*I) haplogroups were also analyzed (Keyeux et al. 2002; Stone and Stoneking 1993) in a single sample that was not assigned to any of the traditional Amerindian mtDNA haplogroups. PCR or PCR-RFLP products were analyzed by non-denaturing PAGE (8%) followed by silver staining (Sanguinetti et al. 1994).

Statistical analysis. The frequencies of each mtDNA haplogroup were computed via the direct counting method. The intrapopulational genetic diversities (h_{Sk}) and

their standard deviations (Nei 1987) were estimated in each village by the ARLEQUIN 3.5.1.2 software (Excoffier and Lischer 2010). This software was also used to obtain F_{ST} estimates between two or more populations (Weir and Cockerham 1984) and to perform the Analysis of Molecular Variance (AMOVA) (Excoffier et al. 1992) and the exact test of population differentiation based on haplogroup frequencies (Raymond and Rousset 1995). Since the pairwise F_{ST} and the exact test of population differentiation between pairs of villages represent 15 statistical comparisons, the Bonferroni correction was used to adjust the significance level for multiple testing, resulting in $\alpha = 0.0033$ (*i.e.*, .05/15).

Given that F_{ST} estimates rely on two hidden assumptions (that expected gene identities and hence, effective population sizes are equal for all subpopulations, and divergence between all pairs of populations is equal and independent), their violation may strongly bias the outcome of the analyses and lead to a failure to identify important differentiation among groups (Long and Kittles 2003). To counteract such limitations, given that the violations of both assumptions usually occur in nature, population-specific F_{ST} [$F_{ST(k)}$] values were also estimated according to (Long and Kittles 2003). A $F_{ST(k)}$ value measures the gain in gene identity in the k^{th} village in relation to the gene identity in the whole tribe. It represents the degree of evolution of the k^{th} village from a common ancestral population that would have split into all the villages considered.

The geographic distances between the 15 pairs of *Pano* villages were calculated from the geographic coordinates (latitude and longitude) by the *Calculator for Distances between Geographical Locations* applet (Bogan 1998). Spearman rank (r_S) correlations between genetic and geographic distances were performed with the GraphPad InStat 3.01 software (GraphPad Software Inc, San Diego, USA).

Results

Intrapopulational diversity. The four Native American mitochondrial DNA (mtDNA) founder haplogroups typical of South-American autochthonous populations were found in the present sample (Table 1). Except for a single individual from Morada Nova (*Katukina*) that presented an African haplogroup L, all remaining individuals presented one of the four Native American haplogroups. Such individual was excluded from all analysis performed in this study. No European haplogroup was found. Each village presented at least two haplogroups, with haplogroup B being the least frequent and restricted to only three villages (Table 1). The intrapopulational diversity ranged from 0.3626 in Sete Estrelas (*Katukina*) to 0.6818 in Cana Brava (*Kaxináwa*).

Interpopulational diversity. Two independent tools were employed to evaluate the extent of differentiation between the pairs or groups of villages: F_{ST} and the exact test of population differentiation based on haplogroup frequencies. Although these analyses have the same purpose and may result in a similar outcome, both were performed to provide more reliable and robust conclusions, in case one of the methods is more or less sensible (or conservative) than the other.

Both the exact test of population differentiation based on haplogroup frequencies ($p < 0.0001$) and the F_{ST} estimate ($F_{ST} = 0.1985$; $p < 0.0001$) revealed the existence of a highly significant difference between the six *Pano* villages. The analysis of the pairwise F_{ST} matrix reveals a large range of variation in F_{ST} values: from -0.1014 to 0.4690 (Table 2). Seven (46.67%) of the fifteen pairs of *Pano* villages differ significantly at the 5% significance level (Table 3). However, when the conservative Bonferroni correction is used to adjust for multiple testing, the significance threshold becomes $\alpha=0.0033$, and thus only three (20.00%) of the fifteen pairs of *Pano* villages remain statistically different at the 5% significance level. A similar outcome was observed by the matrix composed of non-differentiation probability values obtained through the exact test of population differentiation (Table 3): eight (53.33%) pairs of *Pano* villages differing significantly at the 5% significance level (Table 3), five of them remaining significantly different even after the application of the conservative Bonferroni correction. There was no correlation between F_{ST} values and geographic distances

(shown in Table 2) between the 15 pairs of *Pano* villages (r_S = 0.1828; p = 0.5143).

In view of the significant heterogeneity between the *Pano* villages, it is necessary to find out if the *Pano* villages are equally different from one other, or if there is a set of homogeneous villages together with one or two heterogeneous outliers that contribute differently to the conventional F_{ST}. In order to identify these potential outlying villages, population-specific F_{ST} values were estimated (Table 4).

The $F_{ST(k)}$ values ranged from 0.0879 (Cana Brava) to 0.5086 (Sete Estrelas). Besides Sete Estrelas, Paredão [$F_{ST(Paredão)}$ = 0.2852] and Yaminawa [$F_{ST(Yaminawa)}$ = 0.2347] were also more divergent than the others (Table 4). The conventional F_{ST} estimate between all villages excluding Sete Estrelas (*Katukina*) resulted in a lower but still significant value (F_{ST} = 0.1026; p = 0.0137). If both Sete Estrelas (*Katukina*), Paredão (*Kaxináwa*), and *Yaminawa* were excluded from the conventional F_{ST} estimate, a much lower and non-significant F_{ST} value (F_{ST} = 0.0195; p = 0.2743) would result. A simple observation of the haplogroup frequencies in these three highly divergent villages (Table 1) reveals that Sete Estrelas presents the lowest level of genetic diversity and is the only one in which haplogroup C was not sampled; Paredão, as well as Sete Estrelas, does not exhibit haplogroup D, which was sampled in the other villages. Moreover, these three villages show the lower levels of genetic diversity when compared to the three

other villages, which may be due to isolation coupled with genetic drift. Finally, both the pairwise F_{ST} and the exact test of population differentiation between pairs of villages (Table 3) evidence the strong heterogeneity between Sete Estrelas and the remaining villages.

Analysis of molecular variance. Where no obvious criterion exists for the definition of groups of populations, the investigation of the genetic structure in a set of populations may be difficult. The strongest structure of populations would be represented by groups of populations that are maximally differentiated from each other (i.e. those for which the proportion of total genetic variance due to differences between groups is maximum) (Dupanloup et al. 2002). We have tested six hierarchical structures in an attempt of defining the strongest structure. Tribal affiliation (for structure 1) and population-specific differentiation ($F_{ST(k)}$) (for structures 2-6) were used as criteria to define these six groups of populations for AMOVA (Table 5).

In the absence of a meaningful structure, it would be expected a positive correlation of F_{CT} (proportion of total genetic variance due to differences between groups) with the number of groups that compose a given structure. This would be due to the reduction of F_{SC} (proportion of variance due to differences between populations within each group) in the structures with larger number of groups. Although we have tested hierarchical structures that comprises up to four groups,

the structure that maximizes the variance due to differences between groups and minimizes the variance due to differences between populations within each group was the only one composed of three groups (structure 5). In such structure, the first group encompasses the three villages that presented the lowest $F_{ST(k)}$ values, the second one encompasses the village (Sete Estrelas) that presented the largest $F_{ST(k)}$ value, and the third one encompasses two highly homogeneous villages (Paredão and *Yaminawa*) with intermediate $F_{ST(k)}$ values (Tables 2-4). Differences between the three groups account for 25.17% of the variance, whereas there is no variance (-0.03%) as consequence of differences between the villages that belong to a same group (Table 5).

It should be emphasized that tribal affiliation (structure 1) does provide the weakest structure among the six verified (Table 5), with the lowest proportion of variance among groups (-12.60%) and highest variance among populations within each group (31.49%).

Discussion

Inter-ethnic admixture. Studies of classical markers in this same set of villages have reported admixture levels that ranged from 0.3% (immunoglobulin allotypes) to 0.4% (blood groups and proteins). We have observed that only one of the 77 individuals (1.3%) studied carried a non-indigenous haplogroup (Table 1). Taken

together, these low inter-ethnic admixture levels indicate that the *Pano* is one of the few Amerindian groups that were able to preserve their genetic identity by taking advantage of the difficult access to colonizers provided by the Amazon rainforest.

It should be emphasized that the non-Amerindian mtDNA haplogroup found is the African haplogroup L, and all previously observed non-Amerindian immunoglobulin allotypes in *Pano* Indians were of African origin (Gershowitz and Neel 1978). Moreover, such haplogroup was found in Morada Nova (*Katukina*), which is the village in which the greatest number of persons of mixed (5%) or non-Indian (12%) ancestry was reported (Salzano and Jacques 1979).

Given that three of the four *Pano* tribes studied (*Katukina*, *Marúbo*, and *Yaminawa*) are characterized by matrilineal and matrilocal social organizations, it is possible that a slightly higher male-mediated European interethnic admixture would have occurred, particularly in Morada Nova (*Katukina*). This would be in agreement with other historic and genetic data, which suggest that recent admixture in Amerindians is mainly characterized by an asymmetric pattern of gene flow involving immigrant men (mostly European) and native women, but also native men and African women (Mesa et al. 2000; Wang et al. 2008).

Genetic diversity among Pano tribes. Due to the ethnic, linguistic and cultural homogeneity that characterized the populations of the Central *Pano* linguistic

group at least in previous decades, the tribes that speak such language were usually referred as members of the *Pano* "tribe". Even though, the genetic homogeneity between *Pano* villages has not been examined before. Although sample size may seem small at a first glance, it in fact reflects the actual population diversity of the *Pano* villages considered in this study. The present sample of 77 individuals without direct kinship corresponds to approximately 10% to 15% of the inhabitants of the considered villages at the time of sample collection (1976) and does consist in a representative sample. Moreover, assuming that the carefully obtained relatedness information during interviews is reliable (i.e. reported paternity and maternity relationships are accurate) it is possible to infer the mtDNA lineages of more than 300 sampled individuals; this inference exercise does not result in significant differences between observed and inferred haplogroup frequencies.

At first sight, an atypical pattern of populational differentiation is disclosed: while the two *Kaxináwa* villages (Cana Brava and Paredão) have shown to be extremely homogeneous (F_{ST} = -0.0573; p = 0.6721), the two *Katukina* villages (Morada Nova and Sete Estrelas) presented the highest F_{ST} value observed in the 15 pairwise comparisons performed (F_{ST} = 0.4690; p = 0.0002) (Tables 2–3). The homogeneity of the *Kaxináwa* villages was expected not only for they being members of the same tribe, but also for they being closely geographically interconnected (inhabitants of one used to have relatives in the

other) (Salzano and Jacques 1979). Genetic markers previously analyzed support this observation (Luizon et al. 2008; Mendes-Junior et al. 2007b; Mohrenweiser et al. 1979). On the other hand, the heterogeneity of the *Katukina* villages may be explained by the fact that they are separated by about 135 kilometers (Table 2), which represents a considerable distance in the Amazon rainforest when the difficulties in locomotion are taken into account. In addition, Morada Nova is actually a village established by the recent admixture of Amerindian people from three different *Pano* tribes (*Katukina*, *Kaxináwa*, and *Yaminawa*) (Salzano and Jacques 1979).

Significant heterogeneity was observed between the six *Pano* villages (F_{ST} = 0.1985; $p < 0.0001$). Amerindian populations are usually characterized by high interpopulational diversity (Tarazona-Santos et al. 2001), but this high heterogeneity level is unexpected if the strong ethnic, linguistic and cultural homogeneity of the members of the *Pano* linguistic group (Erikson 1998; Mohrenweiser et al. 1979; Salzano and Jacques 1979) is taken into account. As previously stated (Torroni et al. 2006), a prerequisite for future progress in mtDNA evolutionary studies is the use of the information from the entire DNA molecule, which would improve the phylogeographic resolution and the interpretation of intra- and inter-population genetic diversity patterns. However, relevant information can be achieved by means of comparisons concerning low-resolution haplotype diversity. For instance, we have shown that eight villages

from the *Tikúna* tribe, whose geographic limits of distribution can be set from 3º to 5º South and from 68º to 70º West and which presents remarkable degrees of genetic and linguistic isolation, presented a much lower degree of heterogeneity (F_{ST} = 0.0481; p = 0.0058) (Mendes-Junior and Simoes 2009). Moreover, F_{ST} estimates based on mtDNA haplogroup frequencies from three other sets of at least six Amerindian villages pertaining to a same tribe (Table 6) ranged from 0.0548 (p = 0.0004) among six *Quechua* villages to 0.0643 (p = 0.0001) among ten *Aymara* villages.

In order to further explore the intratribal patterns of population differentiation in a broader context, intrapopulational and interpopulational diversities concerning these different sets of populations were evaluated by means of h_{sk} and F_{ST}. If standard deviations are taken into account, *Pano´*s intrapopulational diversity (h_{sk} = 0.6944 ± 0.0257) does not differ from that of *Mapuche* (h_{sk} = 0.6893 ± 0.0093), *Quechua* (h_{sk} = 0.6570 ± 0.0239) and *Tikúna* (h_{sk} = 0.6981 ± 0.0116) tribes (Table 6). However, *Aymara´*s intrapopulational diversity (h_{sk} = 0.5110 ± 0.0289) is definitely lower than the diversities of those tribes. Therefore, one may conclude that the high levels of heterogeneity observed among the *Pano* villages is not related to their overall genetic diversity. In order to assess whether intratribal variability exceed between-tribe variability, average F_{ST} values from all pairwise comparisons of villages within a same tribe ($F_{ST\text{-}W}$), as well as average F_{ST} values from all pairwise comparisons of villages between

different tribes (F_{ST-B}) were estimated for the five considered tribes (Table 7). Very high differentiation can be seen between *Pano* villages and villages from the other four Amerindian tribes from South America (average $F_{ST-B} = 0.2596$) (Table 7), similar to the value observed for the *Aymara* villages (average $F_{ST-B} = 0.2668$), and higher than those observed for *Mapuche*, *Quechua* and *Tikúna*. Moderate to high differentiation can be seen between *Pano* villages (average $F_{ST-W} = 0.1495$), while moderate differentiation was observed between *Aymara* villages (average $F_{ST-W} = 0.1133$). On the other hand, little differentiation was observed between *Mapuche*, *Quechua* or *Tikúna* villages (average $F_{ST-W} < 0.05$). Although in no instance intratribal variability exceed between-tribe variability, the highest F_{ST-W} / F_{ST-B} ratio was observed for *Pano* villages (57.57%). Taken together, these observations corroborate the conclusion of the unexpectedly high heterogeneity level between populations of the Pano linguistic group. Although at first glance this heterogeneity finding among the *Pano* may seem quite surprising, it corroborates other results concerning *HLA-G* diversity (Mendes-Junior et al. 2007b; Mendes-Junior et al. 2013), which suggest that the strong ethnic, linguistic and cultural homogeneity that characterizes the *Pano* linguistic group does not imply genetic homogeneity.

A detailed analysis of the *Pano* village by means of the population-specific F_{ST} estimates [$F_{ST(k)}$], which could be indicative of special evolutionary constraints in some of these populations, confirm that the *Pano* villages are

somewhat equally heterogeneous (Table 4). For instance, when this same analysis was performed in the study of mtDNA in eight *Tikúna* villages, it revealed an intratribal genetic heterogeneity pattern characterized by two highly homogeneous *Tikúna* groups that differed considerably from each other (Mendes-Junior and Simoes 2009). In order to clarify some aspects of this complex pattern of genetic heterogeneity of the *Pano* linguistic group, AMOVA was employed to evaluate some hierarchical structures (Table 5) defined on basis of tribal affiliation and population-specific differentiation [$F_{ST(k)}$]. The strongest structure (structure 5), based on $F_{ST(k)}$ values, reinforces the heterogeneity of *Pano* villages, particularly Sete Estrelas (*Katukina*), while tribal affiliation (structure 1) provides the weakest structure among the six verified (Table 5).

In spite of the recent origin of these villages (Salzano and Jacques 1979), a plausible explanation for this finding is that the gene flow between villages was not being strong enough to maintain the original homogeneity and to counteract the differentiation effects of genetic drift. Alternatively, based on the evidences that most of the studied *Pano* tribes are organized into matrilineal societies (Mohrenweiser et al. 1979) (Povos Indígenas no Brasil, http://pib.socioambiental.org/en), which usually result in matrilocality, a sex-specific scenario in which gene flow may be mediated almost exclusively by men may be proposed. This would lead to a lower female migration rate and differentiation of the mtDNA gene pool. The analysis of autosomal and Y-linked

microsatellites would be required in order to verify whether one or both hypothesis are correct.

In conclusion, the present findings indicate that the strong ethnic, linguistic and cultural homogeneity that characterizes the *Pano* linguistic group does not imply genetic homogeneity. Moreover, the lack of correlation between genetic (F_{ST}) and geographic distances suggests that geography is not a major factor influencing the genetic structure of the *Pano* tribes. Even though this may be a female-specific process, the most probable explanation for the genetic heterogeneity uncovered is the joint action of isolation and genetic drift as major factors influencing the genetic structure of the *Pano* linguistic group.

Acknowledgments. We are grateful to Mrs. Maria do Carmo Tomitão Canas and Mrs. Ana Lucia Pimentel for technical assistance in the laboratory analyses and to Dora Yovana Barrios Leal for assistance in the map design. This research was supported by grants from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Fundação de Apoio ao Ensino, Pesquisa e Assistência do HCFMRP-USP (FAEPA), and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). C.T.M.J. (305493/2011-6) and A.L.S. (312547/2009-9) are currently supported by Research fellowships from CNPq/Brazil. The authors declare no conflict of interest.

Literature Cited

Bailliet, G., F. Rothhammer, F. R. Carnese et al. 1994. Founder mitochondrial haplotypes in Amerindian populations. *Am. J. Hum. Genet.* 55:27–33.

Bert, F., A. Corella, M. Gene et al. 2001. Major mitochondrial DNA haplotype heterogeneity in highland and lowland Amerindian populations from Bolivia. *Hum. Biol.* 73:1–16.

Bianchi, N. O., G. Bailliet, and C. M. Bravi. 1995. Peopling of the Americas as inferred through the analysis of mitochondrial DNA. *Braz. J. Genet.* 18:661–668.

Bisso-Machado, R., M. C. Bortolini, and F. M. Salzano. 2012. Uniparental genetic markers in South Amerindians. *Genet. Mol. Biol.* 35:365–387.

Bogan, L. 1998. Calculator for distances between geographical locations. Available at http://www.nature1st.net/bogan/soaring/jslatlng.html.

Dupanloup, I., S. Schneider, and L. Excoffier. 2002. A simulated annealing approach to define the genetic structure of populations. *Mol. Ecol.* 11:2571–2581.

Erikson, P. 1998. Uma singular pluralidade: a etno-história pano. In *História dos Índios no Brasil*, M. C. Cunha, ed. São Paulo, Brazil: Companhia das Letras, 239–252.

Excoffier, L., and H. E. Lischer. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Res.* 10:564–567.

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

Fiedel, S. J. 2000. The peopling of the New World: present evidence, new theories, and future directions. *J. Archaeol. Res.* 8:39–103.

Fuselli, S., E. Tarazona-Santos, I. Dupanloup et al. 2003. Mitochondrial DNA diversity in South America and the genetic history of Andean highlanders. *Mol. Biol. Evol.* 20:1682–1691.

Gershowitz, H., and J. V. Neel. 1978. The immunoglobulin allotypes (Gm and Km) of twelve Indian tribes of Central and South America. *Am. J. Phys. Anthropol.* 49:289–301.

Ginther C., D. Corach, A. Penacino et al. 1993. Genetic variation among the Mapuche Indians from the Patagonian region of Argentina: mitochondrial DNA sequence variation and allele frequencies of several nuclear genes. In *DNA Fingerprinting: State of the Science*, S. D. J. Pena, R. Chakraborty R, J. T. Epplen, and A. J. Jeffreys, eds. Basel, Switzerland: Birkhauser Verlag, 211–219.

Goicoechea, A.S., F. R. Carnese, A.L. Caratini et al. 2000. Demography, genetic diversity, and population relationships among Argentinean Mapuche Indians. *Gen. Mol. Biol.* 23:513–518.

Greenberg, J. H. 1987. *Language in the Americas*. Stanford, CA: Stanford University Press.

Horai, S., R. Kondo, Y. Nakagawa-Hattori et al. 1993. Peopling of the Americas, founded by four major lineages of mitochondrial DNA. *Mol. Biol. Evol.* 10:23–47.

Keyeux, G., C. Rodas, N. Gelvez et al. 2002. Possible migration routes into South America deduced from mitochondrial DNA studies in Colombian Amerindian populations. *Hum. Biol.* 74:211–233.

Lewis, C. M., R. Y. Tito, B. Lizarraga et al. 2004. Land, language, and loci: mtDNA in Native Americans and the genetic history of Peru. *Am. J. Phys. Anthropol.* 127:351–360.

Long, J. C., and R. A. Kittles. 2003. Human genetic diversity and the nonexistence of biological races. *Hum Biol* 75:449–471.

Luizon, M. R., C. T. Mendes-Junior, S. F. de Oliveira et al. 2008. Ancestry informative markers in Amerindians from Brazilian Amazon. *Am. J. Hum. Biol.* 20:86–90.

Mendes-Junior, C. T., E. C. Castelli, A. L. Simoes et al. 2007a. Absence of the HLA-G*0105N allele in Amerindian populations from the Brazilian Amazon Region: a possible role of natural selection. *Tissue Antigens* 70:330–334.

Mendes-Junior, C. T., E. C. Castelli, R. T. Simoes et al. 2007b. HLA-G 14-bp polymorphism at exon 8 in Amerindian populations from the Brazilian Amazon. *Tissue Antigens* 69:255–260.

Mendes-Junior, C. T., and A. L. Simoes. 2009. Mitochondrial DNA variability among eight *Tikúna* villages: evidence for an intratribal genetic heterogeneity pattern. *Am. J. Phys. Anthropol.* 140:526–531.

Mendes-Junior, C. T., E. C. Castelli, D. Meyer et al. 2013. Genetic diversity of the HLA-G coding region in Amerindian populations from the Brazilian Amazon: a possible role of natural selection. *Genes Immun.* doi: 10.1038/gene.2013.47.

Merriwether, D. A., F. Rothhammer, and R. E. Ferrell. 1995. Distribution of the four founding lineage haplotypes in Native Americans suggests a single wave of migration for the New World. *Am. J. Phys. Anthropol.* 98:411–430.

Mesa, N. R., M. C. Mondragon, I. D. Soto, et al. 2000. Autosomal, mtDNA, and Y-chromosome diversity in Amerinds: pre- and post-Columbian patterns of gene flow in South America. *Am. J. Hum. Genet.* 67:1277–1286.

Mestriner, M. A., A. L. Simoes, and F. M. Salzano. 1980. New studies on the esterase D polymorphism in South American Indians. *Am. J. Phys. Anthropol.* 52:95–101.

Mohrenweiser, H., J. V. Neel, M. A. Mestriner et al. 1979. Electrophoretic variants in three Amerindian tribes: the Baniwa, Kanamari, and Central Pano of western Brazil. *Am. J. Phys. Anthropol.* 50:237–246.

Moraga, M., E. Aspillaga, C. Santoro et al. 2001. mtDNA analysis of mummies from northern Chile endorse the hypothesis of an Amazonian origin of Andean populations. *Rev. Chil. Hist. Nat.* 74:719–726.

Moraga, M. L., P. Rocco, J. F. Miquel et al. 2000. Mitochondrial DNA polymorphisms in Chilean aboriginal populations: implications for the peopling of the southern cone of the continent. *Am. J. Phys. Anthropol.* 113:19–29.

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

Raymond, M., and F. Rousset. 1995. An exact test for population differentiation. *Evolution* 49:1280–1283.

Rocco, P. P., C. G. Morales, M. V. Moraga et al. 2002. Genetic composition of the Chilean population. Analysis of mitochondrial DNA polymorphisms. *Rev. Med. Chile* 130:125–131.

Rodriguez-Delfin, L. A., V. E. Rubin-de-Celis, and M. A. Zago. 2001. Genetic diversity in an Andean population from Peru and regional migration patterns of Amerindians in South America: data from Y chromosome and mitochondrial DNA. *Hum. Hered.* 51:97–106.

Salzano, F. M. 2002. Molecular variability in Amerindians: widespread but uneven information. *An. Acad. Bras. Cienc.* 74:223–263.

Salzano, F. M., and S. M. Callegari-Jacques. 1988. South American Indians: a Case Study in Evolution. New York, NY: Oxford University Press.

Salzano, F. M., and S. M. Jacques. 1979. Genetic demography of the Central Pano and Kanamari Indians of Brazil. *Hum. Biol.* 51:551–564.

Sanguinetti, C. J., E. Dias Neto, and A. J. Simpson. 1994. Rapid silver staining and recovery of PCR products separated on polyacrylamide gels. *Biotechniques* 17:914–921.

Spielman, R. S., E. C. Migliazza, and J. V. Neel. 1974. Regional linguistic and genetic differences among Yanomama indians. *Science* 184:637–644.

Stone A. C., and M. Stoneking. 1993. Ancient DNA from a pre-Columbian Amerindian population. *Am. J. Phys. Anthropol.* 92:463–471.

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.

Torroni, A., A. Achilli, V. Macaulay, et al. 2006. Harvesting the fruit of the human mtDNA tree. *Trends Genet*. 22:339–345.

Wang, S., N. Ray, W. Rojas et al. 2008. Geographic patterns of genome admixture in Latin American Mestizos. *PLoS Genet.* 4:e1000037.

Weir, B., and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38:1358–1370.

TABLE 1. Mitochondrial DNA haplogroup distributions and intrapopulational gene diversity (h_{Sk}) in four Pano tribes (six villages) studied.

Pano tribes and villages	*n* [1]	*Haplogroups*					h_{Sk} [2]
		A	*B*	*C*	*D*	L	
Katukina	25	12	3	5	4	1	0.6920 ± 0.0717
Morada Nova	11	1		5	4	1	0.6444 ± 0.1012
Sete Estrelas	14	11	3				0.3626 ± 0.1302
Kaxináwa	19	8	1	9	1		0.6257 ± 0.0669
Cana Brava	12	4	1	6	1		0.6818 ± 0.1019
Paredão	7	4		3			0.5714 ± 0.1195
Marúbo (Vida Nova)	18		3	9	6		0.6471 ± 0.0691
Yaminawa	15	9		5	1		0.5619 ± 0.0954
Pano (Total)	77	29	7	28	12	1	0.6944 ± 0.0257

[1] *n* = number of sampled individuals.

[2] h_{Sk} was estimated taking into account only frequencies from haplogroups A, B, C, and D.

TABLE 2. Matrix of pairwise F_{ST} values based on mtDNA haplogroup frequencies (under the diagonal) and geographic distances between localities expressed in kilometers (above the diagonal) for the six Pano villages (four tribes) analyzed in the present study. Statistically significant F_{ST} values are in boldface ($p < 0.05$) or italicized boldface ($p < 0.01$). Statistically significant values at a 5% significance level after Bonferroni correction are marked with an asterisk ($p < 0.0033$).

	Morada Nova	*Sete Estrelas*	*Cana Brava*	*Paredão*	*Vida Nova*	*Yaminawa*
Morada Nova (*Katukina*)	-	134.87	5.20	5.20	248.17	134.87
Sete Estrelas (*Katukina*)	***0.4690****	-	138.96	138.96	177.28	0.00
Cana Brava (*Kaxináwa*)	0.0290	***0.2821***	-	0.00	248.91	138.96
Paredão (*Kaxináwa*)	0.1620	0.1765	-0.0573	-	248.91	138.96
Vida Nova (*Marúbo*)	-0.0472	***0.4660****	0.0630	**0.2162**	-	177.28
Yaminawa	**0.1961**	**0.1238**	0.0112	-0.1014	***0.2530****	-

TABLE 3. Matrix of probabilities associated with pairwise F_{ST} values presented in Table 2 (under the diagonal) and non-differentiation probabilities obtained by means of exact tests of population differentiation based on haplogroup frequencies (above the diagonal) for the six Pano villages (four tribes) analyzed in the present study. Statistically significant values are in boldface ($p < 0.05$) or italicized boldface ($p < 0.01$). Statistically significant values at a 5% significance level after Bonferroni correction are marked with an asterisk ($p < 0.0033$).

	Morada Nova	*Sete Estrelas*	*Cana Brava*	*Paredão*	*Vida Nova*	*Yaminawa*
Morada Nova (*Katukina*)	-	***0.0000****	0.2258	0.0807	0.4544	**0.0202**
Sete Estrelas (*Katukina*)	***0.0002****	-	***0.0025****	0.0504	***0.0000****	***0.0094***
Cana Brava (*Kaxináwa*)	0.2556	***0.0053***	-	0.8553	**0.0458**	0.5158
Paredão (*Kaxináwa*)	0.0769	0.1445	0.6721	-	***0.0025****	1.0000
Vida Nova (*Marúbo*)	0.6609	***0.0000****	0.1020	**0.0130**	-	***0.0000****
Yaminawa	**0.0132**	**0.0399**	0.2907	0.7642	***0.0023****	-

TABLE 4. Population-specific F_{ST} estimates [$F_{ST(k)}$] in the six Pano villages (four tribes). For the $F_{ST(k)}$ estimate in each village, the corresponding gene identity (J_k) and the total gene identity in the Pano linguistic group (J_T) were used.

Pano village	J_k	J_T	$F_{ST(k)}$
Morada Nova (*Katukina*)	0.4200	0.3148	0.1536
Sete Estrelas (*Katukina*)	0.6633	0.3148	0.5086
Cana Brava (*Kaxináwa*)	0.3750	0.3148	0.0879
Paredão (*Kaxináwa*)	0.5102	0.3148	0.2852
Vida Nova (*Marúbo*)	0.3889	0.3148	0.1082
Yaminawa	0.4756	0.3148	0.2347

TABLE 5. Analysis of Molecular Variance (AMOVA) for Amerindian mtDNA haplogroup frequencies, according to six ways of grouping the six Pano villages considered in the present study.

Hierarchical structure	*Groups composing the hierarchical structure* [a]	*Variance*		
		among groups (F_{CT})	*among populations within groups (F_{SC})*	*within populations (F_{ST})*
1	G1: MN (*Kat*), SE (*Kat*) G2: CB (*Kax*), P (*Kax*) G3: VN (*Mar*) G4: *Yam*	-12.60% (p = 0.8221 ± 0.0041)	31.49% (p = 0.0011 ± 0.0003)	81.11% (p = 0.0000 ± 0.0000)
2	G1: MN (*Kat*), CB (*Kax*), P (*Kax*), VN (*Mar*), *Yam* G2: SE (*Kat*)	21.98% (p = 0.1650 ± 0.0037)	9.03% (p = 0.0146 ± 0.0012)	68.99% (p = 0.0000 ± 0.0000)
3	G1: MN (*Kat*), CB (*Kax*), VN (*Mar*), *Yam* G2: SE (*Kat*), P (*Kax*)	16.44% (p = 0.2016 ± 0.0038)	10.16% (p = 0.0053 ± 0.0007)	73.40% (p = 0.0000 ± 0.0000)

4	G1: MN (*Kat*), CB (*Kax*), VN (*Mar*) G2: SE (*Kat*), P (*Kax*) , *Yam*	24.11% ($p = 0.0992 \pm 0.0029$)	3.31% ($p = 0.1299 \pm 0.0035$)	72.57% ($p = 0.0001 \pm 0.0001$)
5	G1: MN (*Kat*), CB (*Kax*), VN (*Mar*) G2: SE (*Kat*) G3: P (*Kax*) , *Yam*	25.17% ($p = 0.0328 \pm 0.0016$)	-0.03% ($p = 0.4837 \pm 0.0048$)	74.86% ($p = 0.0328 \pm 0.0016$)
6	G1: MN (*Kat*), CB (*Kax*), VN (*Mar*) G2: SE (*Kat*) G3: P (*Kax*) G4: *Yam*	20.94% ($p = 0.0979 \pm 0.0026$)	2.57% ($p = 0.2728 \pm 0.0048$)	76.50% ($p = 0.0000 \pm 0.0000$)

[a] MN (*Kat*): Morada Nova (*Katukina*); SE (*Kat*): Sete Estrelas (*Katukina*); CB (*Kax*): Cana Brava (*Kaxináwa*); P (*Kax*): Paredão (*Kaxináwa*); VN (*Mar*): Vida Nova (*Marúbo*); *Yam*: *Yaminawa*

TABLE 6. Mitochondrial DNA haplogroup distributions and intrapopulational gene diversity (h_{Sk}) in four Amerindian tribes. F_{ST} values (with their respective p-values) and non-differentiation probabilities obtained by means of exact tests of population differentiation based on mtDNA haplogroup frequencies. Statistically significant values are in italicized boldface ($p < 0.01$).

Amerindian tribes and villages [a]	*n* [b]	*Haplogroups*					h_{Sk} [c]	F_{ST} (*p-value*) [c]	*Exact test of population differentiation*[c]
		A	*B*	*C*	*D*	*Others*			
Aymara	**325**	**20**	**217**	**43**	**43**	**2**	**0.5110 ± 0.0289**		
Bolívia [1]	33	-	31	1	1	-	0.1193 ± 0.0756	***0.0643***	***0.0014 ± 0.0010***
Arica - Chile [2,3]	120	9	68	22	19	2	0.6065 ± 0.0387	***(0.0001 ± 0.0001)***	
Caquena - Chile [4]	23	4	14	3	2	-	0.6008 ± 0.0986		
Codpa - Chile [4]	9	1	6	1	1	-	0.5833 ± 0.1833		

Esquina - Chile [4]	14	-	9	-	5	-	0.4945 ± 0.0876		
Guallatiri - Chile [4]	9	-	5	3	1	-	0.6389 ± 0.1258		
Guanacagua - Chile [4]	17	-	17	-	-	-	0.0000 ± 0.0000		
Illapata - Chile [4]	12	-	12	-	-	-	0.0000 ± 0.0000		
Parinacota - Chile [4]	12	1	6	5	-	-	0.6212 ± 0.0867		
Visviri - Chile [4]	76	5	49	8	14	-	0.5421 ± 0.0563		
Mapuche	**316**	**15**	**74**	**100**	**115**	**12**	**0.6893 ± 0.0093**		
Argentina [5]	63	4	23	12	18	6	0.7005 ± 0.0283		
Aguada Guzmán - Argentina [6,7]	32	2	9	6	14	1	0.6925 ± 0.0488	***0.0633***	
Anecón Grande - Argentina [7,8]	39	6	15	8	10	-	0.7395 ± 0.0326	***(0.0000 ± 0.0000)***	***0.0000 ± 0.0000***
Cerro Polícia - Argentina [6,7]	26	1	9	7	4	5	0.7000 ± 0.0563		
Ilha Huapi - Chile [9]	111	-	8	49	54	-	0.5684 ± 0.0206		
Laitec, Yaldad, Quellon, La	45	2	10	18	15	-	0.6929 ± 0.0304		

Mision, Los Galpones, Guinimo, Paraquina, Cocauque and Icalma - Chile [4,10]									
Quechua	**219**	**30**	**111**	**30**	**45**	**3**	**0.6570 ± 0.0239**		
Bolívia [1]	32	5	24	3	-	-	0.4173 ± 0.0957		
Arequipa - Peru [11]	22	2	15	3	2	-	0.5238 ± 0.1157		
Cabana, Chacas and Independencia - Peru [12]	33	3	17	6	7	-	0.6686 ± 0.0623	***0.0548*** ***(0.0004 ± 0.0001)***	***0.0003 ± 0.0003***
Pasco - Peru [13]	52	2	28	9	10	3	0.6088 ± 0.0573		
Peru [4]	19	5	7	1	6	-	0.7310 ± 0.0480		
Tayacaja - Peru [11]	61	13	20	8	20	-	0.7344 ± 0.0213		
Tikúna	**187**	**46**	**11**	**64**	**66**	**-**	**0.6981 ± 0.0116**	***0.0481*** ***(0.0058 ± 0.0004)***	***0.0059 ± 0.0029***
Belém [14]	26	5	-	5	16	-	0.5692 ± 0.0838		

Bom Jardim [14]	18	3	2	10	3	-	0.6601 ± 0.0998
Campo Alegre [14]	23	5	1	9	8	-	0.7075 ± 0.0466
Feijoal [14]	25	12	2	9	2	-	0.6533 ± 0.0611
Marajá [14]	22	7	3	6	6	-	0.7662 ± 0.0374
Nova Itália [14]	26	5	3	9	9	-	0.7385 ± 0.0418
Umariaçu [14]	24	4	-	11	9	-	0.6486 ± 0.0501
Vendaval [14]	23	5	-	5	13	-	0.6126 ± 0.0768

[a] The numbers after population names indicate the following bibliographic references: 1- (Bert et al. 2001); 2- (Moraga et al. 2001); 3- (Rocco et al. 2002); 4- (Merriwether et al. 1995); 5- (Bianchi et al. 1995); 6- (Bailliet et al. 1994); 7- (Goicoechea et al. 2000); 8- (Ginther et al. 1993); 9- (Moraga et al. 2000); 10- (Horai et al. 1993); 11- (Fuselli et al. 2003); 12- (Lewis et al. 2004); 13- (Rodriguez-Delfin et al. 2001); 14- (Mendes-Junior and Simoes 2009).

[b] n = number of sampled individuals.

[c] h_{Sk}, F_{ST} and the exact test of population differentiation were estimated taking into account only frequencies from haplogroups A, B, C, and D.

TABLE 7. Average pairwise F_{ST} values based on comparisons of mtDNA haplogroup frequencies of villages within a same tribe (F_{ST-W}), and of villages between different tribes (F_{ST-B}).

Tribes (number of villages)	F_{ST-W} *(number of pairwise comparisons)*	F_{ST-B} *(number of pairwise comparisons)*	F_{ST-W} / F_{ST-B} *ratio*
Pano (6)	0.1495 (15)	0.2596 (180)	0.5757
Aymara (10)	0.1133 (45)	0.2668 (260)	0.4246
Mapuche (6)	0.0406 (15)	0.1421 (180)	0.2859
Quechua (6)	0.0488 (15)	0.1595 (180)	0.3059
Tikúna (8)	0.0448 (28)	0.2130 (224)	0.2102

F_{ST-W}: Average F_{ST} from all pairwise comparisons of villages within a same tribe.

F_{ST-B}: Average F_{ST} from all pairwise comparisons of villages between different tribes.

Fig. 1. a: Map of South America, indicating the approximate geographic distribution of the Pano villages in the Brazilian territory in 1976. b: Enlargement of part of the border between Amazonas and Acre states, with locations of the villages in which blood samples were obtained: (1) Morada Nova; (2) Sete Estrelas; (3) Cana Brava; (4) Paredão; (5)Vida Nova; (6) *Yaminawa*.

a)

b)