*Ann. Hum. Genet.* (2002), **66**, 255–259 © *University College London*
DOI: 10.1017/S0003480002001148 *Printed in the United Kingdom*



# Y-chromosome biallelic polymorphisms and Native American population structure

M.-C. BORTOLINI[1,2], F. M. SALZANO[1], C. H. D. BAU[1], Z. LAYRISSE[3], M. L. PETZL-ERLER[4], L. T. TSUNETO[4], K. HILL[5], A. M. HURTADO[5], D. CASTRO-DE-GUERRA[3], G. BEDOYA[6] AND A. RUIZ-LINARES[2,6]

[1] *Departamento de Genética, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil*
[2] *Galton Laboratory, Department of Biology, University College London, UK*
[3] *Instituto Venezolano de Investigaciones Cientificas, Caracas, Venezuela*
[4] *Departamento de Genética, Universidade Federal do Paraná, Curitiba, Brazil*
[5] *Department of Anthropology, University of New Mexico, Albuquerque, USA*
[6] *Laboratorio de Genética Molecular, Universidad de Antioquia, A.A. 1226 Medellín, Colombia*

(Received 31.8.01. Accepted 30.4.02)

SUMMARY

It has been proposed that women had a higher migration rate than men throughout human evolutionary history. However, in a recent study of South American natives using mtDNA restriction fragment polymorphisms and Y-chromosome microsatellites we failed to detect a significant difference in estimates of migration rates between the sexes. As the high mutation rate of microsatellites might affect estimates of population structure, we now examine biallelic polymorphisms in both mtDNA and the Y-chromosome. Analyses of these markers in Amerinds from North, Central and South America agree with our previous findings in not supporting a higher migration rate for women in these populations. Furthermore, they underline the importance of genetic drift in the evolution of Amerinds and suggest the existence of a North to South gradient of increasing drift in the Americas.

## INTRODUCTION

It has recently been proposed that women had a higher migration rate than men throughout human evolution (Seielstad *et al.* 1998). This suggestion was based on the observation, in a world-wide survey and in some specific areas, that molecular diversity on the Y-chromosome was more geographically structured than in mtDNA (Seielstad *et al.* 1998). We recently tested this hypothesis in South America by typing a uniform set of Y-chromosome microsatellites (STRs) and mtDNA RFLPs in five Native Colombian populations, as well as by analysing published data for these markers in other South Amerind groups (Mesa *et al.* 2000). Both sets of uniparental markers revealed similar levels of population structure, as assessed using $G_{ST}$: about 0.17 for the Colombian populations and about 0.30 for South America (Mesa *et al.* 2000).

As STRs have a relatively high mutation rate, and as they are prone to recurrent mutation, it is possible that estimates of population structure based on them might be unreliable (Seielstad, 2000). To assess the dependence of our previous conclusions on the type of Y-chromosome marker examined here, we report analyses of an expanded dataset covering Amerind populations from North to South America and including

Correspondence: Andres Ruiz-Linares, Department of Biology (Wolfson House), University College London, 4 Stephenson Way, London NW1 2HE, UK. Tel/Fax: +44 (0)20 7679 5049/7383 2048.
E-mail: andres.ruiz@ucl.ac.uk

Table 1. *Y-chromosome haplotype frequencies* (%) *in Native Americans and in other human groups*[a]

| Population | Haplotype[1] (%) 1 | 2 | 3 | 4 | 5 | 6 | 7 | No. of chromosomes | Linguistic subfamily | Country |
|---|---|---|---|---|---|---|---|---|---|---|
| Native American | | | | | | | | | | |
| South Amerind | | | | | | | | | | |
| Ache | | 100 | | | | | | 53 | Ge-Pano-Carib | Paraguay |
| Bari | | 100 | | | | | | 12 | Equatorial-Tucano | Venezuela |
| Cinta-Larga | | 100 | | | | | | 15 | Equatorial-Tucano | Brazil |
| Guarani | | 76 | 18 | | 2 | | 4 | 61 | Equatorial-Tucano | Brazil |
| Ingano | | 10 | 80 | | | 10 | | 10 | Andean | Colombia |
| Krahó | | 100 | | | | | | 9 | Ge-Pano-Carib | Brazil |
| Mekranoti | | 72 | 14 | | | | 14 | 7 | Ge-Pano-Carib | Brazil |
| Paccás Novos | | 100 | | | | | | 15 | Equatorial-Tucano | Brazil |
| Parakanã | | 100 | | | | | | 20 | Equatorial-Tucano | Brazil |
| Ticuna | 67 | 22 | 11 | | | | | 35 | Equatorial-Tucano | Colombia |
| Urubu-Kaapor | | 100 | | | | | | 16 | Equatorial-Tucano | Brazil |
| Warao | | 100 | | | | | | 12 | Chibchan-Paezan | Venezuela |
| Wayuu | 10 | 47 | 28 | | 10 | | 5 | 18 | Equatorial-Tucano | Colombia |
| Xikrin | | 100 | | | | | | 8 | Ge-Pano-Carib | Brazil |
| Yagua | | 86 | 14 | | | | | 7 | Ge-Pano-Carib | Colombia |
| Yupca | | 100 | | | | | | 12 | Ge-Pano-Carib | Venezuela |
| Zenu | | 46 | 41 | | 11 | | 2 | 46 | —[b] | Colombia |
| Total | 7 | 78 | 11 | | 2 | 1 | 1 | 356 | | |
| Central Amerind (6) | | 80 | 15 | | 3 | 2 | | 110 | | |
| North Amerind (4) | | 22 | 58 | 7 | 1 | 1 | 1 | 122 | | |
| Eskimo/Aleut (2) | | 45 | 46 | | 9 | | | 66 | | |
| Na-Dene (2) | | 37 | 51 | 8 | 4 | | | 68 | | |
| Total | | | | | | | | 256 | | |
| Asian | | | | | | | | | | |
| Central (4) | | | 45 | 40 | 6 | 9 | | 229 | | |
| East (5) | | | 61 | 13 | 4 | 22 | | 269 | | |
| North (15) | | 1 | 63 | 34 | 2 | | | 572 | | |
| Total | | 1 | 59 | 30 | 3 | 7 | | 1070 | | |
| European (4) | | | 48 | | 37 | 15 | | 145 | | |
| Sub-sahara African (6) | | | 2 | | 32 | 13 | 53 | 312 | | |

[a] Data for populations other than South Amerinds were obtained from Karafet *et al.* (1999). The number of populations considered in each region is indicated in parentheses. The M19 polymorphism has not been detected in a survey of North American Natives (Underhill *et al.* 1997 and pers. comm.). Since this polymorphism seems to have a restricted distribution within South America, we have assumed it to be the ancestral allele for the populations examined by Karafet *et al.* (1999).

[b] The Zenu currently speak only Spanish. It is unclear what might have been the linguistic affiliation of their language (Mesa *et al.* 2000).

data for mtDNA and Y-chromosome biallelic polymorphisms.

## Subjects and methods

Y-chromosome typings were performed in 356 unrelated male samples from 17 Native South American populations: 8 from Brazil, 5 from Colombia, 3 from Venezuela and one from Paraguay (Table 1). This population sample includes representatives of the four major linguistic subfamilies present in South America (Andean, Chibchan-Paezan, Equatorial-Tucano and Ge-Pano-Carib). Genomic DNA was extracted from whole blood, plasma or RBC samples using the Qiagen Kit, following the manufacturer's instructions.

Seven biallelic markers (DYS199, M19, 92R7, M9, DYS271 or YAP, SY81, and RPS4Y711) were examined using the experimental conditions reported by Karafet *et al.* (1999) and Ruiz-Linares *et al.* (1999). These novel Y-chromosome data were integrated with available data from other populations for which information on the same Y-markers, as well as on mtDNA polymorphisms, was available. Genetic diversity was

assessed using Nei's estimators (Nei, 1986), with $G_{ST}$ corrected for the number of populations sampled (thus corresponding to Nei's $G_{ST}'$). Assuming migration-drift equilibrium and ignoring mutation, effective migration rate estimates were obtained as $N_m = (1/G_{ST}) - 1$.

## RESULTS AND DISCUSSION

The Y-chromosome markers examined allow the identification of seven haplotypes. The evolutionary relatedness of these haplotypes and the allelic states defining them are shown in Figure 1. Table 1 shows the frequencies of these seven haplotypes in Native Americans and other representative human populations. The presence of haplotypes 5, 6 and 7 in some Amerindian populations is most likely the result of admixture with recent immigrants, as these haplotypes are frequent amongst Europeans and/or Africans, but are generally rare in Native Americans (Underhill *et al.* 1997; Karafet *et al.* 1999). Haplotype 2 corresponds to a lineage originating at about the time of the initial colonization of the New World (Lell *et al.* 1997; Karafet *et al.* 1997), while its derived haplotype 1 has a restricted distribution in South America (we only detected it in the Ticuna and the Wayuu). Haplotypes 3 and 4 correspond to founder lineages introduced in America from Asia, although a fraction of haplotype 3 chromosomes also result from post-Columbian admixture (Karafet *et al.* 1999; Ruiz-Linares *et al.* 1999).

Table 2 shows estimates of within ($H_S$) and between ($G_{ST}$) population diversity of Amerinds from North, Central and South America based on Y-chromosome biallelic markers and mtDNA haplogroups. We focus these analyses on the Amerind linguistic family, as this group is thought to represent the initial paleoindian migration into the New World (Cavalli-Sforza, Menozzi & Piazza, 1994).

For both mtDNA and Y chromosome markers we observe increasing levels of population structure from North to South America, with Central America showing intermediate values. In the case of Y-chromosome markers the increase in $G_{ST}$ values is associated with a marked reduction of intrapopulation diversity, with South American populations showing on average less than half the diversity of North Amerinds (0.015 vs. 0.042). At the mitochondrial level South Amerind populations are also less diverse than North Amerind populations, although the difference is less pronounced than for the Y chromosome (0.523 vs. 0.609). Central Amerinds show the lowest mean mtDNA diversity of the three regions (0.4), possibly because the data available includes several Chibchan groups that seem to have undergone recent population bottlenecks (Kolman *et al.* 1995, 1997). The observation of a trend of greater population structure from North to South America with both mtDNA and Y-chromosome markers, accompanied by a decrease of intra-population diversity, suggests a Southward gradient of increasing drift in the Americas. A higher level of genetic drift for South Amerinds relative to North Amerinds has been previously suggested based on autosomal polymorphisms (Cavalli-Sforza, Menozzi & Piazza, 1994). Possible explanations for such an increase in drift include a greater population fragmentation imposed by geography in South America, and a higher level of drift at the front of the wave of advance in the population that initially expanded to the Americas (Cavalli-Sforza, Menozzi & Piazza, 1994).

Interestingly, the trend towards a southward increase in population structure is more pronounced for the Y-chromosome than for mtDNA. Estimates of $N_m$ obtained from Y-chromosome data decrease eightfold between North and South America, while the corresponding values for mtDNA are only halved (Table 2). This leads to an inversion of the relative male/female migration rate between North and South America from about 3/1 to 1/2. It is not obvious why the southward increase in population structure is more pronounced for Y-chromosome markers than for mtDNA. This trend could relate to a southward increase in patrilocality, but ethnographic surveys indicate that patrilocality seems to be more common in North America than in South America (Burton *et al.* 1996).

| HAPLOTYPE | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| M19 | A | T | T | T | T | T | T |
| DYS199 | T | T | C | C | C | C | C |
| RPS4Y711 | C | C | C | T | C | C | C |
| YAP | - | - | - | - | - | + | + |
| SY81 | A | A | A | A | A | A | G |
| M9 | G | G | G | C | C | C | C |
| 92R7 | T | T | T | C | C | C | C |

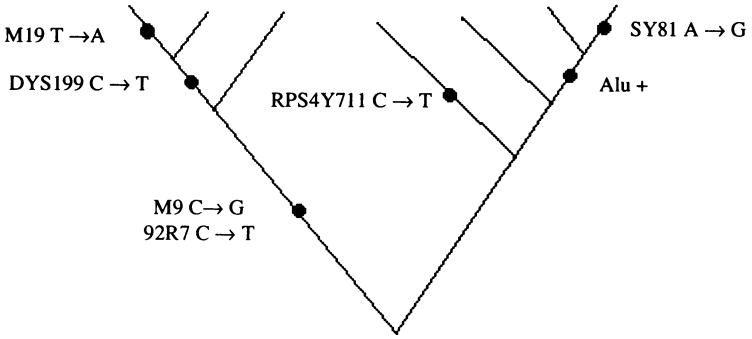


Fig. 1. Polymorphisms defining the seven Y-chromosome haplotypes examined here and their phylogenetic relationships.

Table 2. *Genetic diversity and effective number of migrants* ($N_m$) *in Amerind populations*[a]

| | #Populations | $H_T$ | $H_S$ | $G_{ST}$ | $N_m$ |
|---|---|---|---|---|---|
| South Amerind | | | | | |
| Y-chromosome | 17 | 0.027 | 0.015 | 0.47 | 1.1 |
| MtDNA | 35 | 0.749 | 0.526 | 0.31 | 2.2 |
| Central Amerind | | | | | |
| Y-chromosome | 7 | 0.025 | 0.022 | 0.15 | 5.7 |
| MtDNA | 10 | 0.512 | 0.400 | 0.24 | 3.2 |
| North Amerind | | | | | |
| Y-chromosome | 4 | 0.046 | 0.042 | 0.11 | 8.1 |
| MtDNA | 9 | 0.745 | 0.609 | 0.20 | 4.0 |

Values calculated for the Y-chromosome exclude lineages which are most likely of non-Amerind origin (haplotypes 5 to 7). The mtDNA data set was compiled from the literature (Easton *et al.* 1996; Forster *et al.* 1996; Bonato & Salzano, 1997; Mesa *et al.* 2000 and references therein).

[a] See Table 1.

Higher levels of population structure for the Y-chromosome could relate to it having a smaller effective population size relative to mtDNA, due to a higher variance in male reproductive success (Cavalli-Sforza & Bodmer, 1971). This phenomenon is well documented in some Amerind populations and it might have been common amongst the initial colonizers of the continent (Salzano *et al.* 1967). Finally, differences in rates of non-Amerind admixture in the populations examined could also be contributing to the greater southward increase in population structure for the Y-chromosome relative to mtDNA. This is particularly so since recent admixture in Amerinds usually involved Native women and immigrant men, and (as mentioned above) a fraction of haplotype 3 Y-chromosomes could be of non-Amerind origin (Mesa *et al.* 2000; Carvajal-Carmona *et al.* 2000).

Our previous observations in South America (Mesa *et al.* 2000) are consistent with the enlarged dataset examined here using Y-chromosome biallelic markers. Although the novel $G_{ST}$ estimate in South America is about 50% higher than that obtained previously using Y-STR loci (0.30), the difference is well below the 3–5 fold increase in $G_{ST}$ values between STR vs. biallelic markers reported for world-wide population comparisons (Jorde *et al.* 2000; Quintana-Murci *et al.* 1999). The observation of higher levels of population structure with mtDNA than Y-chromosome markers for North and Central America further questions the generality of the proposed higher migration rate of females during human evolution (Seielstad *et al.* 1998). The parallel analysis of mtDNA and Y-chromosome markers further highlights the importance of drift in the evolution of Native American populations, particularly in South America.

Thanks are due to the Brazilian Fundação Nacional do Índio for authorising this study and for logistic assistance. The subjects of this investigation were informed about the aims of the study and their consent to participate is gratefully acknowledged. Some of the samples examined here were collected many years ago as part of a long-term association that F. M. Salzano had with the late Prof. James V. Neel and with Prof. Francis L. Black. The Brazilian National Ethics Commission approved this investigation (CONEP Resolution no. 123/98). Financial help in Brazil was provided by the 'Programa de Apoio a Núcleos de Excelência, Financiadora de Estudos e Projetos, Conselho Nacional de Desenvolvimento Científico e Tecnológico', and the 'Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul'. Laboratory work in the UK was funded by a Wellcome Trust Travelling Research Fellowship (Grant 059763 to M.-C Bortolini and A. Ruiz-Linares). M.-C Bortolini is grateful to Profs. T. A. Weimer and M. H. Hutz for encouragement and support. We also thank L. Chikhi for helpful comments on the manuscript.

## REFERENCES

Bonatto, S. L. & Salzano, F. M. (1997). Diversity and age of the four major mtDNA haplogroups, and their implications for the peopling of the New World. *Am. J. Hum. Genet.* **61**, 1413–1423.

Burton, M. L., Moore, C. C., Whiting, I. W. M. & Romney, A. K. (1996). Regions based on social structure. *Curr. Anthropol.* **37**, 87–123.

Carvajal-Carmona, L. G., Soto, I. D., Pineda, N., Ortiz-Barrientos, D., Duque, C., Ospina-Duque, J., McCarthy, M., Montoya, P., Alvarez, V. M., Bedoya, G. & Ruiz-Linares, A. (2000). Strong Amerind/White Sex Bias and a Possible Sephardic Contribution among the Founders of a Population in Northwest Colombia. *Am. J. Hum. Genet.* **67**, 1287–1295.

Cavalli-Sforza, L. L. & Bodmer, W. F. (1971). The genetics of human populations. W. H. Freeman, San Francisco.

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

Easton, R. D., Merriwether, D. A., Crews, D. E. & Ferrel, R. E. (1996). mtDNA variation in the Yanomami: evidence for additional New World founding lineages. *Am. J. Hum. Genet.* **59**, 213–225.

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

Jorde, L. B., Watkins, W. S., Bamshad, M. J., Dixon, M. E., Ricker, C. E., Seiesltad, M. T. & Batzer, M. A. (2000). The distribution of human genetic diversity: a comparison of mitochondrial, autosomal, and Y-chromosome data. *Am. J. Hum. Genet.* **66**, 979–989.

Karafet, T., Zegura, S. L., Vuturo-Brady, J., Posukh, O., Osipova, L., Wiebe, V., Romero, F., Long, J. C., Harihara, S., Jin, F., Dashnyam, B., Gerelsaikhan, T., Omoto, K. & Hammer, M. F. (1997). Y chromosome markers and Trans-Bering Strait dispersals. *Am. J. Phys. Anthropol.* **102**, 301–14.

Karafet, T. M., Zegura, S. L., Posukh, O., Osipova, L., Bergen, A., Long, J., Goldman, D., Klitz, W., Harihara, S., de Knijff, P., Wiebe, V., Griffiths, R. C., Templeton, A. R. & Hammer, M. F. (1999). Ancestral Asian source(s) of New World Y-chromosome founder haplotypes. *Am. J. Hum. Genet.* **64**, 817–831.

Kolman, C. J., Bermingham, E., Cooke, R., Ward, R. H., Arias, T. D. & Guionneau-Sinclair, F. (1995). Reduced mtDNA diversity in the Ngone Amerinds of Panamá. *Genetics* **140**, 275–283.

Kolman, C. J. & Bermingham, E. (1997). Mitochondrial and nuclear DNA diversity in the Choco and Chibcha Amerinds of Panama. *Genetics* **147**, 1289–1302.

Lell, J. T., Brown, M. D., Schurr, T. G., Sukernik, R. I., Starikovskaya, Y. B., Torroni, A, Moore, L. G., Troup, G. M. & Wallace, D. C. (1997). Y chromosome polymorphisms in native American and Siberian populations: identification of native American Y chromosome haplotypes. *Hum. Genet.* **100**, 536–43.

Mesa, N. R., Mondragón, M. C., Soto, I. D., Parra, M. V., Duque, C., Ortíz-Barrientos, D., García, L. F., Velez, I. D., Bravo, M. L., Múnera, J. G., Bedoya, G., Bortolini, M. C. & Ruiz-Linares, A. (2000). Autosome, mtDNA, and Y-chromosome diversity in Amerinds: Pre- and Pos-Columbian patterns of gene flow in South America. *Am. J. Hum. Genet.* **67**, 1277–1286.

Nei, M. (1986). Definition and estimation of fixation indices. *Evolution* **40**, 643–645.

Quintana-Murci, L., Semino, O., Poloni, E. S., Liu, A., Van Gijn, M., Passarino, G., Brega, A., Nasidze, I. S., Maccioni, L., Cossu, G., Al-Zahery, N., Kidd, J. R., Kidd, K. K. & Santachiara-Benerecetti, A. S. (1999). Y-chromosome specific YCAII, DYS19 and YAP polymorphisms in human populations: a comparative study. *Ann. Hum. Genet.* **63**, 153–66.

Ruiz-Linares, A., Ortiz-Barrientos, D., Figueroa, M., Mesa, N., Munera, J. G., Bedoya, G., Velez, I. D., Garcia, L. F., Perez-Lezaun, A., Bertranpetit, J., Feldman, M. W. & Goldstein, D. B. (1999). Microsatellites provide evidence for Y chromosome diversity among the founders of the New World. *Proc. Natl. Acad. Sci. USA.* **96**, 6312–6317.

Salzano, F. M., Neel, J. V. & D. Maybury-Lewis, D. (1967). Further studies on the Xavante Indians: I. Demographic data on two additional villages: genetic structure of the tribe. *Am. J. Hum. Genet.* **19**, 463–489.

Seielstad, M. T., Minch, E. & Cavalli-Sforza, L. L. (1998). Genetic evidence for a higher female migration rate in humans. *Nature Genetics* **20**, 278–280.

Seiesltad, M. T. (2000). Asymmetries in the maternal and paternal genetic histories of Colombian populations. *Am. J. Hum. Genet.* **67**, 1062–1066.

Underhill, P. A., Jin, L., Lin, A. A., Mehdi, S. Q., Jenkins, T., Vollrath, D., Davis, R. W., Cavalli-Sforza, L. L. & Oefner, P. J. (1997). Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. *Genome Res.* **10**, 996–1005.