CENTER FOR THE STUDY OF THE FIRST AMERICANS
Department of Anthropology
Texas A&M University

November 21, 2011

Virginia Ramallo
Instituto Multidisciplinario de Biología Celular (MBICE)
P.O. Box 403
Calle 526 e/10 y 11
1900 La Plata
Argentina

Dear Virginia,

This letter certifies that the paper "Native Male Founder Lineages of America", by the authors Virginia Ramallo, Marina Muzzio, María R. Santos, Josefina M. B. Motti, Laura S. Jurado Medina, Claudio M. Bravi, and Graciela Bailliet, has been accepted for publication in a special volume of *Current Research in the Pleistocene*, the scientific journal of the Center for the Study of First Americans, Texas A&M University, USA.

The volume is co-edited by Laura Miotti, Nora Flegenheimer and Mónica Salemme. Ted Goebel serves as the editor for *Current Research in the Pleistocene* and will be editing the English version of your paper and managing the production of the volume.

Publication of the volume is estimated for mid-2012. By January 2012 you should receive a corrected version of your manuscript, in English, for your review. Later, in March or April, you will receive page proofs for review as well.

Congratulations! If you have any questions about the publishing process, please do not hesitate to contact me at goebel@tamu.edu.

Sincerely,

Ted Goebel
Professor of Anthropology &
Associate Director, Center for the Study of the First Americans

NATIVE MALE FOUNDER LINEAGES OF AMERICA

Virginia Ramallo[1,2], Marina Muzzio[1,3,4], María R. Santos[1,3,5], Josefina M.B. Motti[1,3,6], Laura S. Jurado Medina[1,7], Claudio M. Bravi[1,3,8] , Graciela Bailliet[1,9]

[1] IMBICE, P.O. Box 403, Calle 526 e/10 y 11, 1900 La Plata, Argentina.

Tel/Fax +54 221 421 0112.

[3] FCNyM, UNLP. Calle 60 y 122, 1900 La Plata, Argentina.

[2]vramallo@yahoo.com, [4]marinamuzzio@yahoo.com.ar, [5]mritasantos@yahoo.com.ar, [6] josemotti@yahoo.com.ar, [7] laurajurado87@hotmail.com, [8]cmbravi@yahoo.com.ar, [9]gbailliet@imbice.org.ar

Key words: Y-chromosome lineages – South America – Regional differentiation

Abstract

Native-American males carry a Y-chromosome lineage characterized by one base-pair polymorphism (M3). All populations from Alaska to the Magellan's Strait present this lineage in frequencies higher than 60%. This lineage was considered the "founder", but further information shows that M3 occurred in America, or Siberia shortly before migration to America and so it is actually considered autochthonous. It belongs to the Q haplogroup, shows derivate state for M242, M346, and M3 polymorphisms, and is named subhaplogroup Q1a3a. Another lineage, paragroup Q1a3*, entered America showing derived states for M242 and M346 polymorphisms, and ancestral for M3. It is present in Eurasia and is more frequent in North than in South America. Though found

at low frequency, it has not been possible to rule out a genetic bottleneck occurrence during the migration to South America.

Introduction

Advances in the human genome knowledge impacted on the comprehension of the Y-Chromosome male-specific region. During the last years, almost 600 bialelic polymorphisms have been described for it, allowing the construction of a standard nomenclature system and a solid phylogeny (YCC 2002, Karafet et al., 2008). Interestingly, the distribution of the major clades from the phylogeny is consistent with the continental distribution of lineages.

Most South American Native males carry a Y-chromosome lineage characterized by a one base pair polymorphism (M3). All populations from Alaska to the Magellan's Strait present this lineage, in average frequencies of around 60% or higher (Bianchi et al., 1998, Bortolini et al., 2003). This lineage was initially considered to be the "founder" (Underhill et al, 1996), and further information showed that M3 occurred either in America or in Siberia shortly before migrating (Lell et al, 2002), thus it is considered "autochthonous". It belongs to the Q haplogroup, bears the derivate state for the M242, M346, and M3 polymorphisms, and was named Q1a3a subhaplogroup (Karafet et al., 2008). Another lineage, the Q1a3* para-haplogroup, entered America revealing a derivate state for M242 and M346, and the ancestral state for M3. Since it is present in Eurasia, it has been considered as a "founder lineage" for America, being more frequent in North than in South America (Bortolini et al., 2003; Bolnick et al., 2006, Bailliet et al., 2009).

The aim of this paper was to analyze the Q1a3a and Q1a3* haplotypes in order to identify the relatedness of American populations, and their regional and chronological differentiation.

Materials and methods

We analyzed 137 individual samples corresponding to the Q1a3a sub-haplogroup, and 13 individuals belonging to the Q1a3* sub-haplogroup, selected from 759 samples of voluntary male donors from 16 populations in Argentina: Susques (18), Rinconada (4), Cochinoca (4), Humahuaca (10), San Salvador de Jujuy (6), Salta (14), Catamarca (6), Tucumán (1), Córdoba (1), Wichis from Salta (15), Wichis from Formosa (26), Toba (4), Chorote (5), Mocovi (3), Mapuche (8), Tehuelche (9), and Paraguay: Lengua (12), Ayoreo (7). All Q1a3a and Q1a3* samples were analyzed for seven Y-chromosome short tandem repeat (Y-STR) polymorphic loci (DYS389I/II, DYS390, DYS19, DYS393, DYS391, DYS392) and four biallelic markers (SNPs). The SNPs analyzed by PCR-RFLP were M3 (Bianchi et al., 1998), M242 (Seielstad et al., 2003), M346 (Sengrupta et al., 2006), and P27 (Karafet et al. 1999).

For each data set, we calculated the diversity parameters, an AMOVA (Excoffier et al., 1992) based on the sum of square differences (Rst), by Arlequin 3.1 (Excoffier et. al., 2005). We constructed maps from the genetic distance matrix (Nei 1972) using multiple dimensional scaling (MDS, Kruskal, 1964) by NTSYSpc 2.11 (Applied Biostatistics 2000-2003), and built a Median Joining Network of Q1a3* haplotypes (after MP procedure) employing Network (v 4.2.1.0, www.fluxusengineering.com) (Bandelt et al., 1999).

Results and discussion

Ninety-seven haplotypes were found in 137 Q1a3a individuals. Fixation Index (Fst) for these lineages was 0.112, and the mean gene diversity was 0.501. We observed a great allele frequency differentiation of Q1a3a haplotypes. In the MDS plot (Fig 1), the first axis separates populations by their geographic position, the Northwestern Andean populations as Rinconada, Cochinoca and Humahuaca on the upper left side, and the Northeastern Gran Chaco populations, as Wichi, Toba, Chorote, Ayoreo and Lengua, on the upper right side. The finding of Susques also in this position is probably due to shared haplotypes with the Wichi and suggests some recent interpopulational contact.

Fourteen individuals belonged to the Q1a3* para-haplogroup. This clade was only present in one Mapuche, two Mocoví, one Wichi, one Salta, one Córdoba, seven Lengua, and two Ayoreo. Frequencies were below 6% except for the Lengua (29%) and Ayoreo (22%), although it showed a considerable high allele frequency differentiation (mean gene diversity=0.478). Probably, genetic drift was the phenomenon that caused their numerical restriction.

Network analysis of Q1a3* haplotypes showed three Lengua at the central position, the only haplotypes that differed in one or two allelic changes from these were from Lengua or Ayoreo populations. The rest of haplotypes diverged in more allele changes, while three median vectors (indicating absent haplotypes in the sample) were interposed between the central and derived haplotypes. This is concordant with the hypothesis of severe drift acting over these less frequent haplotypes. Lengua 2 and Lengua 3 carried each other 2 identical haplotypes (Fig. 2).

Evidence supports that the allelic diversity among populations depends on the geographic region analyzed. We expect that this tendency to regional differentiation would be better visualized with the inclusion of more microsatellites for each lineage.

References

Bailliet, G., V. Ramallo, M. Muzzio, et al. 2009 Restricted geographic distribution for Y-Q* paragroup in South America. *Am J Phys Anthrop* 140:578–582.

Bandelt, H-J., P. Forster, A. Röhl 1999. Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol 16:37–48*.

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

Bolnick, D.A., D.I, Bolnick, and D.G. Smith, 2006 Asymmetric male and female genetic histories among native Americans from Eastern North America. *Mol Biol Evol* 23: 2161–2174.

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.

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

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.

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.

Karafet, T.M., S.L., Zegura, O. Posukh et al. 1999 Ancestral Asian source(s) of New World Y-chromosome founder haplotypes. *Am J Hum Genet* 64:817–831.

Kruskal, J. B. 1964. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Pyschometrika 29:1–27*.

Lell, J.T., R.I. Sukernik, Y.B. Starikovskaya et al. 2002 The dual origin and Siberian affinities of Native American Y chromosomes. *Am J Hum Genet* 70:192–206.

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.

Nei, M. 1972. Genetic distance between populations. *Amer. Naturalist,* 106:283-292.

Sengupta, S., L.A. Zhivotovsky, R. King et al. 2006 Polarity and temporality of high-resolution Y- chromosome distributions in India identify both indigenous and exogenous expansions and reveal minor genetic influence of Central Asian pastoralists. *Am J Hum Genet* 78:202–221.

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

YCC (The Y Chromosome Consortium). 2002 A nomenclature system for the tree of Y chromosomal binary haplogroups. *Genome Res* 12:339–348.

Zegura, S.L., T.M. Karafet, L.A. Zhivotovsky et al. 2004 High-resolution SNPs and microsatellite haplotypes point to a single, recent entry of Native American Y chromosomes into the Americas. *Mol Biol Evol* 21:164–175.

Figure legends.

Figure 1. Two-dimensional scaling of genetic distances (Nei, 1972) based on seven Y-chromosome STRs haplotypes belonging to Q1a3a haplogroup from 17 populations from Argentina and Paraguay. SSJ: San Salvador de Jujuy. STRESS1 = 0.11876.

Figure 2. Microsatellite network for South American Q1a3* haplotypes. Circle size is proportional to frequency. L: Lengua, A: Ayoreo, Mo: Mocoví, Ma: Mapuche, Cor: Córdoba. Mv 1, 2 and 3 are median vectors not found in the sample.