

# A New Subhaplogroup of Native American Y-Chromosomes From the Andes

Marilza S. Jota,<sup>1</sup> Daniela R. Lacerda,<sup>1</sup> José R. Sandoval,<sup>1</sup> Pedro Paulo R. Vieira,<sup>1,2</sup> Simone S. Santos-Lopes,<sup>1</sup> Rafael Bisso-Machado,<sup>3</sup> Vanessa R. Paixão-Cortes,<sup>3</sup> Susana Revollo,<sup>4</sup> Cesar Paz-Y-Miño,<sup>5</sup> Ricardo Fujita,<sup>6</sup> Francisco M. Salzano,<sup>3</sup> Sandro L. Bonatto,<sup>7</sup> Maria Cátira Bortolini,<sup>3</sup> Fabrício R. Santos,<sup>1\*</sup> and The Genographic Consortium

<sup>1</sup>*Departamento de Biologia Geral, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil*

<sup>2</sup>*Centro de Tecnologia (COPPE), Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil*

<sup>3</sup>*Departamento de Genética, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil*

<sup>4</sup>*Facultad de Bioquímica, Universidad Mayor de San Andrés, La Paz, Bolivia*

<sup>5</sup>*Instituto de Investigaciones Biomédicas, Universidad de las Américas, Quito, Ecuador*

<sup>6</sup>*Instituto de Genética y Biología Molecular, Universidad San Martín de Porres, Lima, Peru*

<sup>7</sup>*Departamento de Biodiversidade e Ecologia, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, Brazil*

**KEY WORDS** Y chromosome; SNP; South America; Andes

**ABSTRACT** The human Y chromosome contains highly informative markers for making historical inferences about the pre-Columbian peopling of Americas. However, the scarcity of these markers has limited its use in the inference of shared ancestry and past migrations relevant to the origin of the culturally and biologically diverse Native Americans. To identify new single nucleotide polymorphisms (SNPs) and increase the phylogenetic resolution of the major haplogroup Q found in the Americas, we have performed a search for new polymorphisms based on sequencing divergent Y chromosomes identified by microsatellite haplotype analysis.

Using this approach, a new Y-SNP (SA01) has been identified in the Andean populations of South America, allowing for the detection of a new sublineage of Q1a3a. This sublineage displays a less complex phylogeographic network of associated microsatellites and more restricted geographic occurrence, and is given the designation Q1a3a4. This result indicates that our approach can be successfully used to identify sublineages of interest in a specific region that allow the investigation of particular histories of human populations. *Am J Phys Anthropol* 146:553–559, 2011. © 2011 Wiley Periodicals, Inc.

Pioneering studies of Native American Y chromosomes (Pena et al., 1995; Santos et al., 1995, 1996) identified a single major lineage in the Americas, which today is recognized as haplogroup Q, and now defined by the SNP M242 (Seielstad et al., 2003). Subsequently, a new C (T transition (M3 or *DYS199*) was found to identify a autochthonous sublineage of haplogroup Q in Native Americans and Siberian Inuits (Underhill et al., 1996; Lell et al., 2002; Bortolini et al., 2003) that is today known as subhaplogroup Q1a3a (Karafet et al., 2008). Most of the remaining individuals in the Americas who belong to haplogroup Q, but not to subhaplogroup Q1a3a, are now recognized as belonging to the paragroup Q1a3\*: they bear a derived allele at M346 (Sengupta et al., 2006; Karafet et al., 2008; Bailliet et al., 2009) and ancestral allele at M3. However, further studies are needed to confirm the distribution and occurrence of SNPs in unresolved Q lineages of Americas, like a recent one identifying a Q1a\* paragroup in a Paleoeskimo (Rasmussen et al., 2010).

Another SNP at P39 identifies a distinct low frequency lineage (C3b) that has only been seen in North America (Zegura et al., 2004), but some of our recent data indicate the existence of C3\* in South America (unpublished information). Therefore indigenous Y chromosomes from the Americas can be unambiguously classified into three different lineages, C3, Q1a\* Q1a3\* and Q1a3a (and their sublineages), while the other unrelated lineages detected so far are most easily explained by recent

admixture (Schurr and Sherry, 2004; González-José et al., 2008).

Haplogroup Q lineages in the Americas, Q1a3a and Q1a3\* (and likely Q1a\*), make up more than 80% of indigenous Y chromosomes, and are present in all clearly non-admixed South American Indians (Underhill et al., 1996; Lell et al., 1997; Bortolini et al., 2003; Schurr and Sherry, 2004; Santos et al., 2007). Thus, a detailed analysis of this lineage will reveal most of the paternal history of the peopling of Americas. In this regard, Q1a3a today represents the great majority of all native Y chro-

Additional Supporting Information may be found in the online version of this article.

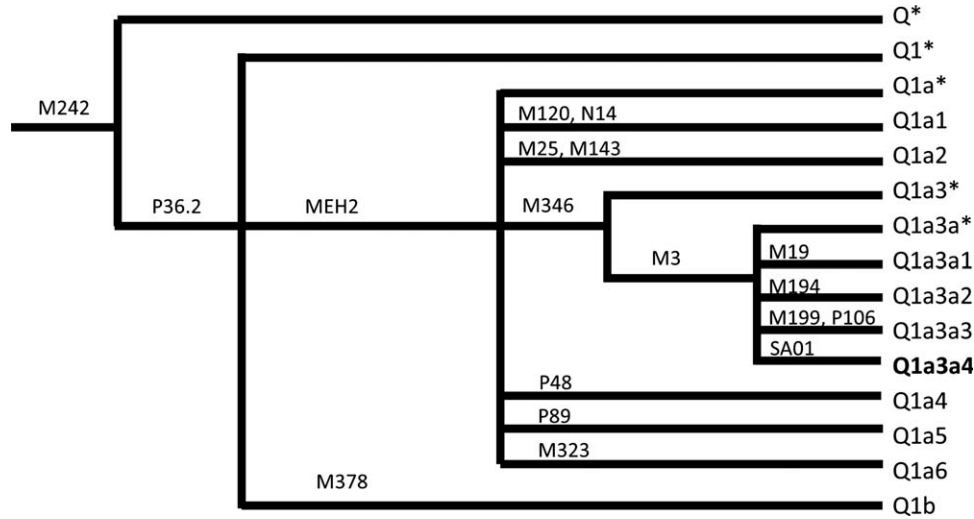
Grant sponsors: FAPEMIG (MG, Brazil), FAPERGS (MG, Brazil), CNPq (Brazil), National Geographic Society (USA).

\*Correspondence to: Fabrício R. Santos, Associate Professor, Departamento de Biologia Geral, ICB, Universidade Federal de Minas Gerais, Av. Antônio Carlos, 6627, C.P. 486, 31270-010, Belo Horizonte, MG, Brazil. E-mail: fsantos@icb.ufmg.br

Received 21 December 2010; accepted 31 January 2011

DOI 10.1002/ajpa.21519

Published online 13 September 2011 in Wiley Online Library (wileyonlinelibrary.com).



**Fig. 1.** Haplogroup Q and its derivatives. The new SNP (SA01) identifies a new sublineage, here named Q1a3a4 and highlighted in bold (after Karafet et al., 2008). Other SNP markers are depicted above the lineage branch identified by them.

mosomes in America, including those from tribes of all major linguistic (and cultural) groups, and also indicates the existence of a founder effect during the initial peopling and a common origin for all native peoples of America (Santos et al., 2007).

This finding supports a model that proposes a founder population occupying Beringia during the last glaciation, where autochthonous Native American lineages arose (Fagundes et al., 2008; González-José et al., 2008). After a population expansion, which could have occurred concomitantly with their entry southwards into America, more recent circum-arctic gene flow would have enabled the dispersion of northeast Asian-derived characters and some genetic lineages/alleles from East Asia to America and vice-versa, including the minor Y lineage C3 (Zegura et al., 2004; González-José et al., 2008). However, a higher resolution of Y chromosome haplogroup Q lineage diversity is needed to investigate paternal genetic aspects of more recent histories, such as the peopling of South America and the origin of peoples speaking different languages within the same linguistic family.

Recently, the Y Chromosome Consortium published a revised Y haplogroup tree depicting 311 distinct paragroups/haplogroups defined by approximately 600 binary markers (SNPs) (Karafet et al., 2008). Within clade Q, 13 paragroups/haplogroups characterized by 17 SNPs were recognized (Karafet et al., 2008). However, excluding sublineage Q1a3a1 detected by the SNP M19 in the Ticuna and Wayuu Indians (Bortolini et al., 2003), no other sublineage has been found in significant numbers among Native Americans, particularly in South America (unpublished data).

Thus, although a number of new markers and lineages have been described (see Fig. 1), there is still insufficient resolution of the haplogroup Q such that a detailed analysis of Native American patrilineages is not possible. To overcome this paucity of variation, Y microsatellite markers have been used to investigate the recent history and demography of Native Americans. For example, Tarazona-Santos et al. (2001) used six microsatellite markers to analyze the recent paternal history within Q1a3a Y chromosomes in South

American Indians. On the basis of microsatellite haplotype distributions, they observed high intrapopulation diversity and low interpopulation divergence in Andean populations, contrasted with their observations in the Amazon Basin. They suggested this pattern was a consequence of a higher effective population size and gene flow in the Andes, in comparison to populations in the lowlands of South America.

In another study, Bortolini et al. (2003) used Y-SNPs and microsatellites to investigate questions related to Amerindian dispersions in South America revealed by lineage Q1a3a (detected by M3) and its sublineage Q1a3a1 (detected by M19). They estimated a minimum age for the Q1a3a lineage at around 7,000–8,000 years, using microsatellite haplotype diversity and published mutation rates. These results suggested that population isolation and possibly the process of tribalization of Native Americans could be very ancient in this region (Bortolini et al., 2003; Santos et al., 2007).

Although Y microsatellites can easily be combined with Y-SNPs to increase the resolution of lineages, there are some limitations on their use. For example, in phylogeographic studies designed to reveal past migrations, most haplotypes appear connected in very complicated networks (Tarazona-Santos et al., 2001; Bortolini et al., 2003). This observation is likely related to the high levels of microsatellite homoplasy (Santos and Tyler-Smith, 1996). Thus, new Y markers that are able to subdivide Q1a3a or to detect new monophyletic sublineages among Q1a3\* or Q1a\* chromosomes can be very informative for inferences about the history of pre-Columbian populations.

In this study we have searched for new Y-SNP markers in order to characterize new informative Q sublineages in Native South Americans from the Andes. Our approach was based on sequencing 15 Y-chromosomal segments from individuals selected on the basis of microsatellite haplotype analysis to represent divergent Y chromosomes. We detected a new SNP marker that identifies an autochthonous sublineage from the Andes, here called Q1a3a4, and thereby increased the phylogeographic resolution within clade Q1a3a (see Fig. 1).

TABLE 1. Geographic localization and linguistic affiliation of the Native South American populations used to search for new Y SNPs

Population	<i>n</i>	Country	Geographic coordinates	Language (Language family)
South America				
Arequipa	19	Peru	71°32'W; 16°23'S	Quechua (Quechua)
Coyaima	13	Colombia	74°05'W; 03°50'N	Coyaima (Carib)
San Martin de Pangoa	33	Peru	74°29'W; 11°25'S	Quechua (Quechua) and Nomatsiguenga (Arawak)
Tayacaja	33	Peru	74°34'W; 12°24'S	Quechua (Quechua)

## MATERIALS AND METHODS

### Samples

Our main sample set used to survey for new SNPs consisted of 98 Native American individuals from three Andean populations from Peru and one from Colombia, all belonging to Y haplogroup Q (Table 1). The Andean samples have been described in other studies (Tarazona-Santos et al., 2001; Fuselli et al., 2007) with their linguistic affiliations and geographical coordinates being shown in Table 1. We subsequently used a second sample of 1,754 Q1a3a native individuals (carrying the derived allele at SNP M3) from Peru, Bolivia, Ecuador and Brazil to genotype the new Y-SNP found in this study. These DNA samples were either obtained from previous studies (Córdova et al., 2008) or extracted using standard procedures from mouth swabs collected during The Geographic Project. Ethical approval for the present study was provided by the Brazilian National Ethics Commission (CONEP Resolution number 763/2009), as well as by local ethical committees in the countries where the non-Brazilian samples were collected.

### Y-chromosome genotyping

All samples had either been typed previously with the SNP markers M242 and M3 using a TaqMan RT-PCR assay (see below) to identify Q\* and Q1a3a chromosomes, or were typed in this study. We also typed M346 by sequencing all Q\* chromosomes (see details below) to confirm their inclusion into Q1a3\* as previously suggested (Bailliet et al., 2009). The main Native American sample (Table 1) belonging to Q1a3\* ( $n = 8$ ) and Q1a3a ( $n = 90$ ) chromosomes from different locations in the Peruvian and Colombian Andes were genotyped with 17 Y-microsatellites (*DYS389a*, *DYS389b*, *DYS390*, *DYS456*, *DYS19*, *DYS385*, *DYS458*, *DYS437*, *DYS438*, *DYS448*, *GATA\_H4*, *DYS391*, *DYS392*, *DYS393*, *DYS439*, *DYS635*) using the Y-filer™ PCR Amplification kit on an ABI3130xl with the GeneScan™ -500 LIZ® Size Standard (Applied Biosystems), or through a non-multiplex analysis on a MegaBACE and scored with Fragment Profiler (GE Healthcare). We have also used published microsatellite data for 19 Greenland Inuits (de Maat et al., 1999) to generate a network together with our data from South Amerindians.

A novel SNP found in this study was genotyped using a TaqMan Genotyping Assay in a 7900HT Fast Real Time PCR System (Applied Biosystems), with PCR primers SA01F GCAACCACTGTGTATTGTTCAGT and SA01R CTCTGGCCCCTAACAAACCTAATTA. The two alleles were detected with probes VIC-ATGTAAATGAC-CaCCACTGC and 6FAM-AAATGACCgCCACTGC. To

score these alleles we have used the AutoCaller Software V1.1 (Applied Biosystems).

### Y-chromosome SNP search using DNA sequencing

We surveyed sequence variation in 15 Y-chromosomal loci (M19, M25, M120, M143, M194, M199, M323, M346, M378, MEH2, N14, P106, P292, P48, and P89) after amplification using the primers and conditions in Karafet et al. (2008). Although these loci carry known SNPs in Native Americans (see Fig. 1), they have not been thoroughly screened by sequencing in South American individuals (Karafet et al., 2008). Except the SNP M19 that is present in Ticuna and Wayuu Indians (Bortolini et al., 2003), and SNPs at MEH2 and M346 that seem to be present in most of Native American haplogroup Q chromosomes (see Fig. 1), all other known SNPs were absent or very rare in South America (this study and unpublished data). All fragments were sequenced in 12 selected individuals (8 Q1a3a and 4 Q1a3\*; selection criteria below) to identify new SNPs as well as to screen for variation at known SNPs. The PCR products were either purified following Exonuclease I and Alkaline Phosphatase (Amersham Biosciences) treatment or PEG/NaCl precipitation. Both DNA strands were sequenced using a MegaBACE 1000 sequencer (GE Healthcare) or ABI3130xl Genetic Analyzer (Applied Biosystems). High quality sequences were aligned and checked for quality and accuracy using the Phred, Phrap, and Consed programs (<http://www.genome.washington.edu>) to visualize and check manually all chromatograms for the presence of new polymorphisms. Whenever a putative polymorphism was identified, the sample was re-amplified and resequenced for confirmation.

### Analytical methods

Using the Median Joining method in the Network 4.5.1.0 software (Bandelt et al., 1999), a haplotype network was constructed with data from six microsatellites (*DYS19*, *DYS390*, *DYS391*, *DYS393*, *DYS389a*, and *DYS389b*) from 98 American Indians with chromosomes Q1a3\* ( $n = 8$ ) or Q1a3a ( $n = 90$ ) originating from different locations in the Peruvian and Colombian Andes (Table 1). Twelve Andean individuals were selected based on microsatellite haplotype analysis because of their peripheral phylogeographic positions in the network (data not shown).

Estimates of the time to the most recent common ancestor (TMRCA) of chromosomes bearing a derived allele at SNP SA01 (Q1a3a4) were determined using the mean effective mutation rate ( $6.9 \times 10^{-4}$ /locus/25 years) for Y microsatellite loci (Zhivotovsky et al., 2004) and

TABLE 2. Microsatellite haplotypes for 16 individuals showing the derived allele for the SA01 marker

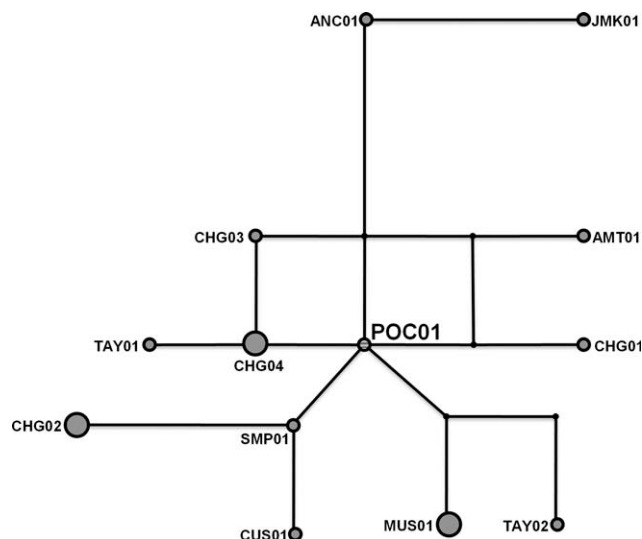
Codes <sup>a</sup>	Y microsatellite haplotypes															# Individuals
	19	389a <sup>b</sup>	389b <sup>b</sup>	390	391	392	393	437	438	439	448	456	458	635	GATA H4	
JMK01	13	13	17	21	10	14	13	14	11	12	22	14	16	22	11	1
AMT01	13	13	18	20	10	14	13	14	11	12	21	14	17	22	11	1
CHG01	13	13	18	21	10	15	13	14	11	13	21	14	17	22	11	1
CHG02	13	14	17	21	10	15	13	14	11	13	22	14	17	22	11	2
CHG03	13	14	18	21	10	14	13	14	11	12	21	15	17	22	11	1
CHG04	13	14	18	21	10	14	13	14	11	13	21	15	17	22	11	2
SMP01	13	14	17	21	10	14	13	14	11	13	21	14	17	22	11	1
CUS01	13	14	17	22	10	14	13	14	11	13	21	14	17	22	11	1
ANC01	13	14	18	21	10	14	13	14	11	12	22	14	16	22	11	1
MUS01	13	14	19	21	10	14	13	14	11	13	21	14	17	22	12	2
TAY01	13	14	18	21	10	14	13	14	11	13	21	15	18	22	11	1
TAY02	14	14	19	21	10	14	13	14	11	13	21	14	18	22	11	1
POC01	13	14	18	21	10	14	13	14	11	13	21	14	17	22	11	1

<sup>a</sup> Codes for haplotypes found in South American populations: JMK, Janko Marka, Bolivia ( $n = 1$ ); AMT, Amantani, Peru ( $n = 1$ ); CHG, Chogo, Peru ( $n = 6$ ); SMP, San Martín de Pangoa, Peru ( $n = 1$ ); CUS, Cusco, Peru ( $n = 1$ ); ANC, Ancahuasi, Peru ( $n = 1$ ); TAY, Tayacaja, Peru ( $n = 2$ ); MUS, MusucanCHA, Peru ( $n = 2$ ); POC, Puerto Ocopa, Peru ( $n = 1$ ). See Figure 2 for sample locations.

<sup>b</sup> DYS389a-b: Allele designations according to Tarazona-Santos et al. (2001).



Fig. 2. Sample locations for individuals bearing the derived allele at marker SA01.



**Fig. 3.** Median joining network of 15 Y microsatellites for 16 Q1a3a4 chromosomes (derived allele at SA01). Circles areas are proportional to the number of individuals bearing each haplotype (see Table 2).

the *rho* statistic implemented in the Network program (Bandelt et al., 1999). The ancestral haplotype was inferred using the modal allele at each microsatellite locus as no clear inference of a founder could be made using outgroup haplotypes (Q1a3a\*).

## RESULTS

Using primer pairs for 15 Y-chromosomal segments (Materials and Methods), we were theoretically able to sequence 9,176 bp of the Y chromosome from each individual. Because of the failure of the PCR primers for M120 and M143 and the poor DNA quality of a few samples, not all regions were successfully sequenced in all individuals. In the end, we generated up to 5.7 Kb of sequence data per individual, around 46,000 nucleotides in all, for SNP screening.

This approach allowed us to find one new Y-SNP, which was called SA01 (from South America). It consists of a C>T transition located 166 bp downstream from SNP M194 (Y chromosome Build 36.1 position 13,524,110), and identifies a subset of Q1a3a chromosomes (bearing also SNP M3), which we designate Q1a3a4 (see Fig. 1). In the first round of sequencing, the SA01 SNP was confirmed in three individuals from Peru. After a TaqMan assay was custom designed (Janet Ziegler, Applied Biosystems) and detected by RealTime PCR, we screened 1,754 haplogroup Q1a3a individuals from Peru, Bolivia, Ecuador and Brazil (see Supporting Information for all ethnic groups surveyed).

We found 16 individuals carrying the derived state at SNP SA01 (samples AMT 1/42; SMP 1/33; MUS 2/5; ANC 1/20; POC 1/20; TAY 2/33; CUS 1/6; JMK 1/3; CHG 6/9), and 15 of these originated from eight different Peruvian indigenous communities, while one (JMK01) was from Bolivia (Table 2). All locations showed one or two individuals exhibiting the derived state for SA01, except the Chogo which presented six individuals with this new SNP (Table 2, Fig. 3). The frequency of individuals with SA01 derived allele varied from 2.4 to 66.7% for populations Amantani and Chogo, respectively. Furthermore,

disregarding a single individual found in Puerto Ocopa at the border of the Andes, who belonged to the Ashaninka (Arawak language) indigenous community, all individuals spoke Andean languages (14 Peruvian Quechuas and a single Aymara from northern Bolivia). The fact that its geographic range was restricted mainly to the Andes (see Fig. 2) indicates a relatively recent origin for the SA01 SNP.

To survey the internal divergence time of SA01-derived chromosomes, we typed 15 Y microsatellites (the 17 from the Y-filer<sup>TM</sup> PCR Amplification kit, excluding the two *DYS385* markers) in these samples and detected 13 different haplotypes that were closely connected in a phylogeographic network (see Fig. 3). Ten out of the 13 microsatellite haplotypes were present in single individuals, while three were found in two persons, always belonging to the same population. This network was then used to estimate a time for the origin of the SA01 mutation event.

By inferring the ancestral haplotype using the modal allele (Zerjal et al., 1997) at each locus (Table 2), we identified the likely root as haplotype POC01, which was also the most central node in the network (see Fig. 3). From this, we estimated a TMRCA for the chromosomes carrying the derived allele of the SA01 SNP of around 5,280 years ago, with a standard deviation of 1,330 years.

## DISCUSSION AND CONCLUSION

We have employed an optimized approach to search for new SNP markers on the human Y chromosome and discovered one new polymorphism useful for phylogeographic studies due to its presence among multiple populations from a restricted area in the Andes. Because our focus was on finding additional recent sublineages within indigenous haplogroup Q chromosomes in the Andes from South America, we based our search on chromosomes chosen after microsatellite haplotype divergence analysis. This approach revealed a recently evolving sublineage (Q1a3a4), which displayed a quite straightforward geographic signature around the Andes. Its discovery further highlights the importance of identifying new markers for phylogeographic studies.

The finding of a single SNP in about 46,000 sequenced nucleotides from selected individuals with divergent Y microsatellite haplotypes indicates the relative paucity of markers among closely related Y chromosomes, as might be expected. The restricted geographical distribution of the new SNP and the closely linked microsatellite haplotypes are consistent with its relatively recent origin, long after the peopling of South America. The estimated date of around 5,300 years ago for the origin of the SA01 SNP suggests that it occurred in the preceramic age of the Andes (Duncan et al., 2009), just before the spread of maize agriculture in the region (Perry et al., 2006). Thus, this strategy to find new SNPs can potentially be useful for generating new markers that will allow a more extensive investigation of recent genealogical and historical events.

## ACKNOWLEDGMENTS

The authors of the Genographic Consortium are: Syama Adhikarla<sup>1</sup>, Christina J. Adler<sup>2</sup>, Danielle A. Badro<sup>3</sup>, Elena Balanovska<sup>4</sup>, Oleg Balanovsky<sup>4</sup>, Jaume Bertranpetit<sup>5</sup>, Andrew C. Clarke<sup>6</sup>, David Comas<sup>5</sup>, Alan

Cooper<sup>2</sup>, Clio S. I. Der Sarkissian<sup>2</sup>, Matthew C. Dulik<sup>7</sup>, Jill B. Gaieski<sup>7</sup>, ArunKumar GaneshPrasad<sup>1</sup>, Wolfgang Haak<sup>2</sup>, Marc Haber<sup>3</sup>, Angela Hobbs<sup>8</sup>, Asif Javed<sup>9</sup>, Li Jin<sup>10</sup>, Matthew E. Kaplan<sup>11</sup>, Shilin Li<sup>10</sup>, Begoña Martínez-Cruz<sup>5</sup>, Elizabeth A. Matisoo-Smith<sup>6</sup>, Marta Melé<sup>5</sup>, Nirav C. Merchant<sup>11</sup>, R. John Mitchell<sup>12</sup>, Amanda C. Owings<sup>7</sup>, Laxmi Parida<sup>9</sup>, Ramasamy Pitchappan<sup>1</sup>, Daniel E. Platt<sup>9</sup>, Lluís Quintana-Murci<sup>13</sup>, Colin Renfrew<sup>14</sup>, Ajay K. Royyuru<sup>9</sup>, Theodore G. Schurr<sup>7</sup>, Himla Soodiyall<sup>8</sup>, David F. Soria Hernanz<sup>15</sup>, Pandikumar Swamikrishnan<sup>16</sup>, Chris Tyler-Smith<sup>17</sup>, Arun Varatharajan Santhakumari<sup>1</sup>, Miguel G. Vilar<sup>7</sup>, R. Spencer Wells<sup>15</sup>, Pierre A. Zalloua<sup>3</sup>, and Janet S. Ziegler<sup>18</sup>.

Affiliations for the Geographic Consortium members: <sup>1</sup>Madurai Kamaraj University, Madurai, Tamil Nadu, India; <sup>2</sup>University of Adelaide, South Australia, Australia; <sup>3</sup>Lebanese American University, Chouran, Beirut, Lebanon; <sup>4</sup>Research Centre for Medical Genetics, Russian Academy of Medical Sciences, Moscow, Russia; <sup>5</sup>Universitat Pompeu Fabra, Barcelona, Spain; <sup>6</sup>University of Otago, Dunedin, New Zealand; <sup>7</sup>University of Pennsylvania, Philadelphia, PA, USA; <sup>8</sup>National Health Laboratory Service, Johannesburg, South Africa; <sup>9</sup>IBM, Yorktown Heights, NY, USA; <sup>10</sup>Fudan University, Shanghai, China; <sup>11</sup>University of Arizona, Tucson, AZ, USA; <sup>12</sup>La Trobe University, Melbourne, Victoria, Australia; <sup>13</sup>Institut Pasteur, Paris, France; <sup>14</sup>University of Cambridge, Cambridge, UK; <sup>15</sup>National Geographic Society, Washington, DC, USA; <sup>16</sup>IBM, Somers, NY, USA; <sup>17</sup>The Wellcome Trust Sanger Institute, Hinxton, UK; <sup>18</sup>Applied Biosystems, Foster City, CA, USA.

## LITERATURE CITED

- Bailliet G, Ramallo V, Muzzio M, García A, Santos MR, Alfaro EL, Dipierri JE, Salceda S, Carnese FR, Bravi CM, Bianchi NO, Bianchi DA. 2009. Brief communication: restricted geographic distribution for Y-Q\* paragroup in South America. *Am J Phys Anthropol* 140:578–582.
- Bandelt H-J, Forster P, Rohl A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16:37–48.
- Bortolini MC, Salzano FM, Thomas MG, Stuart S, Nasanen SP, Bau CH, Hutz MH, Layrisse Z, Petzl-Erler ML, Tsuneto LT, Hill K, Hurtado AM, Castro-de-Guerra D, Torres MM, Groot H, Michalski R, Nymadawa P, Bedoya G, Bradman N, Labuda D, Ruiz-Linares A. 2003. Y chromosome evidence for differing ancient demographic histories in the Americas. *Am J Hum Genet* 73:524–539.
- Córdova JH, Sandoval J, Velásquez M, Távora C, Cotos D, Vásquez J, Barletta C, Fujita R, Descailleaux J. 2008. Poblamiento del Continente Americano y del Perú sugerido de un análisis filogeográfico de haplogrupos del mtDNA en etnias nativas. *Inferencias primarias*. *Archivos de Biología Andina* 14:23–39.
- de Maat MP, Bladbjerg EM, Johansen LG, de Knijff P, Gram J, Klufft C, Jespersen J. 1999. DNA-polymorphisms and plasma levels of vascular disease risk factors in Greenland Inuit—is there a relation with the low risk of cardiovascular disease in the Inuit? *Thromb Haemostasis* 81:547–552.
- Duncan NA, Pearsall DM, Benfer RA Jr. 2009. Gourd and squash artifacts yield starch grains of feasting foods from pre-ceramic Peru. *Proc Natl Acad Sci USA* 106:13202–13206.
- Fagundes NJR, Kanitz R, Eckert R, Valls ACS, Bogo MR, Salzano FM, Smith DG, Silva WA Jr, Zago MA, Ribeiro-dos-Santos AK, Santos SEB, Petzl-Erler ML, Bonatto SL. 2008. Mitochondrial population genomics supports a single pre-Clovis origin with a coastal route for the peopling of the Americas. *Am J Hum Genet* 82:583–592.
- Fuselli S, Gilman RH, Chanock SJ, Bonatto SL, De Stefano G, Evans CA, Labuda D, Luiselli D, Salzano FM, Soto G, Vallejo G, Sajantila A, Pettener D, Tarazona-Santos E. 2007. Analysis of nucleotide diversity of NAT2 coding region reveals homogeneity across Native American populations and high intra-population diversity. *Pharmacogenomics J* 7:144–152.
- González-José R, Bortolini MC, Santos FR, Bonatto SL. 2008. The peopling of America: craniofacial shape variation on a continental scale and its interpretation from an interdisciplinary view. *Am J Phys Anthropol* 137:175–187.
- Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, Hammer MF. 2008. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Res* 18:830–838.
- Lell JT, Brown MD, Schurr TG, Sukernik RI, Starikovskaya YB, Torroni A, Moore LG, Troup GM, Wallace DC. 1997. Y chromosome polymorphisms in Native American and Siberian populations: identification of founding Native American Y chromosome haplotypes. *Hum Genet* 100:536–543.
- Lell JT, Sukernik RI, Starikovskaya YB, Su B, Jin L, Schurr TG, Underhill PA, Wallace DC. 2002. The dual origin and Siberian affinities of Native American Y chromosomes. *Am J Hum Genet* 70:192–206.
- Pena SD, Santos FR, Bianchi NO, Bravi CM, Carnese FR, Rothhammer F, Gerelsaikhan T, Munkhtuja B, Oyunsuren T. 1995. A major founder Y chromosome haplotype in Amerindians. *Nat Genet* 11:15–16.
- Perry L, Sandweiss DH, Piperno DR, Rademaker K, Malpass MA, Umire A, de la Vera P. 2006. Early maize agriculture and interzonal interaction in southern Peru. *Nature* 440:76–79.
- Rasmussen M, Li Y, Lindgreen S, Pedersen JS, Albrechtsen A, Moltke I, Metspalu M, Metspalu E, et al. 2010. Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature* 463:757–762.
- Santos FR, Bonatto SL, Bortolini MC. 2007. Molecular evidence from contemporary indigenous populations to the peopling of America. In: Santos C, Lima M, editors. *Recent advances in molecular biology and evolution: applications to biological anthropology*. Kerala: Research Signpost.
- Santos FR, Hurtz MH, Coimbra CEA., Santos RV, Salzano FM, Pena SD. 1995. Further evidence of existence of major founder Y chromosome haplotype in Amerindians. *Braz J Genet* 18:669–672.
- Santos FR, Rodriguez-Delfin L, Pena SD, Moore J, Weiss KM. 1996. North and South Amerindians may have the same major founder Y chromosome haplotype. *Am J Hum Genet* 58:1369–1370.
- Santos FR, Tyler-Smith C. 1996. Reading the human Y chromosome: the emerging DNA markers and human genetic history. *Braz J Genet* 19:665–670.
- Schurr TG, Sherry ST. 2004. Mitochondrial DNA and y chromosome diversity and the peopling of the Americas: evolutionary and demographic evidence. *Am J Hum Biol* 16:420–439.
- Seielstad M, Yuldasheva N, Singh N, Underhill P, Oefner P, Shen P, Wells RS. 2003. A novel Y chromosome variant puts an upper limit on the timing of the first entry into the Americas. *Am J Hum Genet* 73:700–705.
- Sengupta S, Zhivotovsky LA, King R, Mehdi SQ, Edmonds CA, Chow CE, Lin AA, Mitra M, Sil SK, Ramesh A, Ramesh A, Usha Rani MV, Thakur CM, Cavalli-Sforza LL, Majumder PP, Underhill PA. 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.
- Tarazona-Santos E, Carvalho-Silva DR, Pettener D, Luiselli D, Stefano GF, Labarga CM, Rickards O, Tyler-Smith C, Pena SDJ, Santos FR. 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.
- Underhill PA, Jin L, Zemans R, Oefner PJ, Cavalli-Sforza LL. 1996. A pre-Columbian Y chromosome-specific transition and its implications for human evolutionary history. *Proc Natl Acad Sci USA* 93:196–200.

- Zegura SL, Karafet TM, Zhivotovsky LA, Hammer MF. 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.
- Zerjal T, Dashnyam B, Pandya A, Kayser M, Roewer L, Santos FR, Schiefenhövel W, Fretwell N, Jobling MA, Harihara S, Shimizu K, Semjiddmaa D, Sajantila A, Salo P, Crawford MH, Ginter EK, Evgrafov OV, Tyler-Smith C. 1997. Genetic relationships of Asians and northern Europeans, revealed by Y-chromosomal DNA analysis. *Am J Hum Genet* 60:1174–1183.
- Zhivotovsky LA, Underhill PA, Cinniöglu C, Kayser M, Morar B, Kivisild T, Scozzari R, Cruciani F, Destro-Bisol G, S Pedini G, Chambers GK, Herrera RJ, Yong KK, Gresham D, Tournev I, Feldman MW, Kalaydjieva L. 2004. The effective mutation rate at Y chromosome short tandem repeats, with application to human population divergence time. *Am J Hum Genet* 74:50–61.