



MtDNA lineages D1g and D1h: new insights into the pioneer peopling of South America's Southern Cone



Martin Bodner^{1*}, Ugo A. Perego^{2,3}, Antonio Salas⁴, Alexander W. Röck¹, Bettina Zimmermann¹, Liane Fendt¹, Gabriela Huber¹, Alberto Gómez-Carballa⁴, Scott R. Woodward³, Hans-Jürgen Bandelt⁵, Alessandro Achilli⁶, Antonio Torroni², Walther Parson¹

¹Institute of Legal Medicine, Innsbruck Medical University, Innsbruck, Austria, ²Department of Genetics and Microbiology, University of Pavia, Pavia, Italy, ³Sorenson Molecular Genealogy Foundation, Salt Lake City, Utah, USA, ⁴Unidade de Xenética, Departamento de Anatomía Patolóxica e Ciencias Forenses, and Instituto de Medicina Legal, Facultad de Medicina, Universidad de Santiago de Compostela, Santiago de Compostela, Galicia, Spain, ⁵Department of Mathematics, University of Hamburg, Hamburg, Germany, ⁶Department of Cellular and Environmental Biology, University of Perugia, Perugia, Italy

Introduction

It is now widely agreed that the Native American founders originated from a Beringian source population ~15-18 thousand years ago (kya) and rapidly populated all of the New World on a Pacific coastal route. The currently identified mtDNA founders comprise the major pan-American lineages A2, B2, C1b, C1c, C1d, and D1, and the minor lineages D2a, D3, X2a, X2g, C4c, and D4h3a [1-4].

Questions concerning the number, route, and timing of human migration events into and across the American double continent are still far from being completely addressed. Several models have been proposed to explain the coexistence of lineages with pan-American and others with much more limited dispersal patterns.

Even though Native American populations were among the first to be examined in detail [5], high quality full mtDNA genome sequences are still scarcely available, thus limiting any genetic research. We provide an additional piece to the puzzle of Native American history by shedding light on the mtDNA composition of the Southern Cone of South America, an area so far only marginally evaluated, and presenting novel data concerning the final phases of the long journey that brought humans from Beringia to the tip of South America.

Material and Methods

To examine the pioneering peopling phase of the South American continent, we screened literature and mtDNA database reports and identified two specific mitochondrial clades of the pan-American founder lineage D1 that are not yet described in the comprehensive mitochondrial phylogeny [6, build 12]. According to highest forensic quality standards [7], we identified, selected and then completely sequenced 43 mtDNA genomes of the two novel lineages that we termed D1g and D1h, constructed phylogenetic trees and performed phylogeographic analyses.

We calculated maximum likelihood (ML) molecular divergences for the novel mtDNA clades using PAML 4.4, as previously reported [3]. The calculations were performed on the entire mtDNA haplotypes; mutational distances were converted into years using the corrected molecular clock [8].

References

- [1] Tamm E et al., PLoS ONE 2007; 2(9):e829.
- [2] Achilli A et al., PLoS ONE 2008; 3(3):e1764.
- [3] Perego UA et al., Curr Biol 2009; 19(1):1-8.
- [4] Perego UA et al., Genome Res 2010; 20(9):1174-1179.
- [5] Schurr T et al., Am J Hum Genet 1990; 46:613-623.
- [6] van Oven M & Kayser M, Hum Mutat 2009; 30(2):E386-E394.
- [7] Fendt L et al., BMC Genomics 2009; 10:139.
- [8] Soares P et al., Am J Hum Genet 2009; 84(6):740-759.
- [9] Erlandson JM et al., J Island Coastal Archaeol 2008; 3(2):277-281
- [10] Surovell TA, Curr Anthropol 2003; 44(4):580-591
- [11] McCulloch RD et al., J Quaternary Sci 2000; 15(4):409-417.

*contact: martin.bodner@i-med.ac.at

Acknowledgements

The authors thank Norman Angerhofer, Daniel Corach, Hovirag Lancioni and Anna Olivieri for their valuable contributions. This project received support from Fondazione Alma Mater Ticinensis, the Italian Ministry of the University: Progetti Ricerca Interesse Nazionale 2009, FIRB-Futuro in Ricerca 2008, Fundación de Investigación Médica Mutua Madrileña (2008/CL444) and Ministerio de Ciencia e Innovación (SAF2008-02971) and the Austrian Science Fund FWF, Translational Research project L397.

Results

The novel haplogroup D1g (represented by 26 entire mtDNA genomes) is defined by the mutational pattern A8116G-C16187T and exhibits at least six major basal branches (D1g1 – D1g6). Haplogroup D1h (17 mtDNA genomes) is characterized by the motif T152C-C16242T-T16311C and currently one prominent subbranch, Dh1a (Fig. 1).

The ML point estimate of 16.9 ± 1.6 ky for both sublineages taken together (18.3 ± 2.4 ky for D1g, 13.9 ± 2.9 kya for D1h) agrees with the age of the most ancient known settlement in South America at the Monte Verde site, dated to 14.1-14.6 kya [9]. Phylogeographic analyses revealed a widespread distribution of younger and older (≤ 5-17 ky) subclades of D1g and D1h on both the Pacific coastal and the continental side of the Andes, not restricted to specific regions (Fig. 2).

In the available database and literature reports, both D1g and D1h show overall relatively rare occurrences (≤ 5% countrywide), but local high frequencies, and are essentially restricted to populations from the Southern Cone of South America (Chile, Argentina and Brazil).

Discussion

The ample diversity of the novel mtDNA haplogroups D1g and D1h, limited to the Southern Cone of South America, suggests that these two lineages arose in this area, while their age implies that they originated at the very early stages of human colonization of the continent and were most probably part of the mtDNA pool of the pioneer settlers. Our findings indicate a very early arrival of Paleo-Indians in the Southern Cone after a rapid movement along the entire American double continent (~15,400 km [10]), which appears feasible along the postulated Pacific coastal route under similar ecological resources, requiring little adaptation [9].

Integrating the age estimates for D1g and D1h and their apparent dispersal, we hypothesize an extremely fast Pacific coastal migration in probably less than 2,000 years, followed by continuous and extensive gene flow over the Andean cordillera, as the most likely scenario for the peopling of the South American continent. Migrations over the mountain barrier could have been facilitated by several deglaciation and warming periods after 14.6 kya [11], i.e. about the time when Monte Verde was founded [9]. The Mapuche people, in which the mtDNA haplogroups D1g and D1h are present, have been settling on both sides of the cordillera in the Southern Cone for thousands of years, and could have served as a genetic trans-Andean link.

This study confirms that major sampling and sequencing efforts are mandatory for uncovering all of the most basal variation in the Native American mtDNA haplogroups by targeting both the mixed population of national states and autochthonous Native American groups, especially in South America.

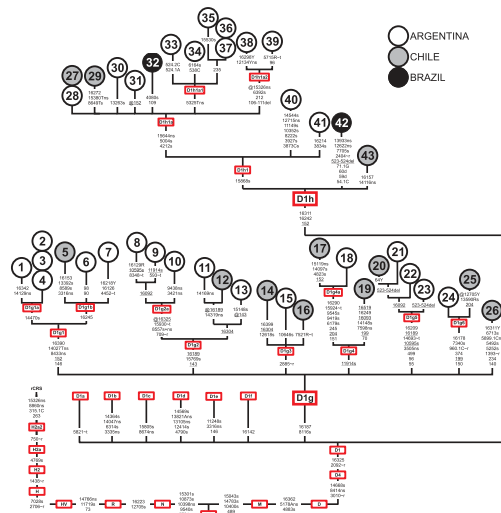


Figure 1. Phylogenetic tree of the two novel Southern South American mtDNA haplogroups D1g and D1h. This tree includes 43 new complete mtDNA sequences. The country of origin of the sample donors is indicated by the circle colour. The prefix @ indicates the reversion of a mutation occurring earlier in the phylogeny. The suffixes s and ns indicate synonymous and non-synonymous substitutions, respectively, while -t and -r indicate affected positions in tRNA and rRNA loci, respectively. Recurrent mutations within the phylogeny are underlined. All mutations are indicated except cytosine insertions at np 309. Haplogroups except D1g and D1h as in [6, build 12].

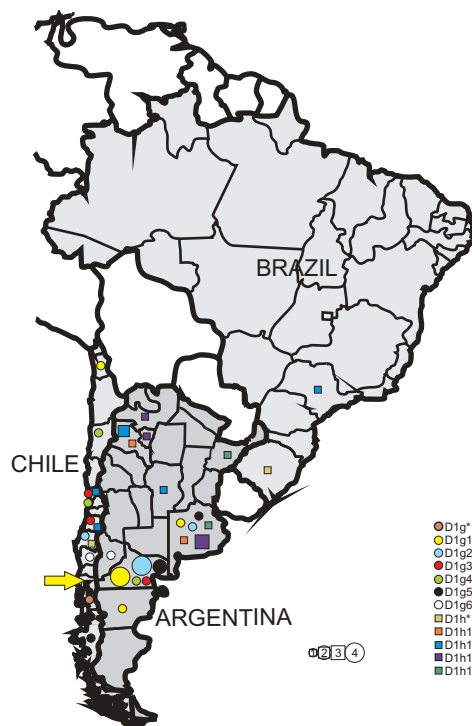


Figure 2. Origin of D1g and D1h sample donors. The provinces/states of origin of the sample donors of the fully sequenced mtDNA haplogroup D1g (represented by circles) and D1h (represented by squares) samples are shown. Sizes of circles and squares correspond to the number of samples, their colours indicate subhaplogroups. The arrow points at the location of the Monte Verde site [9].