Rapid coastal spread of First Americans: Novel insights from South America's Southern Cone mitochondrial genomes

Martin Bodner, Ugo A. Perego, Gabriela Huber, et al.

Genome Res. 2012 22: 811-820 originally published online February 14, 2012
Access the most recent version at doi:10.1101/gr.131722.111

Rapid coastal spread of First Americans: Novel insights from South America's Southern Cone mitochondrial genomes

Martin Bodner,[1] Ugo A. Perego,[2,3] Gabriela Huber,[1] Liane Fendt,[1] Alexander W. Röck,[1] Bettina Zimmermann,[1] Anna Olivieri,[3] Alberto Gómez-Carballa,[4] Hovirag Lancioni,[5] Norman Angerhofer,[2] Maria Cecilia Bobillo,[6] Daniel Corach,[6] Scott R. Woodward,[2] Antonio Salas,[4] Alessandro Achilli,[5] Antonio Torroni,[3] Hans-Jürgen Bandelt,[7] and Walther Parson[1,8]

[1]*Institute of Legal Medicine, Innsbruck Medical University, 6020 Innsbruck, Austria;* [2]*Sorenson Molecular Genealogy Foundation, Salt Lake City, Utah 84115, USA;* [3]*Department of Genetics and Microbiology, University of Pavia, 27100 Pavia, Italy;* [4]*Unidade de Xenética, Departamento de Anatomía Patolóxica e Ciencias Forenses, and Instituto de Medicina Legal, Facultade de Medicina, Universidad de Santiago de Compostela, Santiago de Compostela, 15782, Galicia, Spain;* [5]*Department of Cellular and Environmental Biology, University of Perugia, 06123 Perugia, Italy;* [6]*Servicio de Huellas Digitales Genéticas, Facultad de Farmacia y Bioquímica, Universidad de Buenos Aires, 1113-Buenos Aires, Argentina;* [7]*Department of Mathematics, University of Hamburg, 20146 Hamburg, Germany*

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, probably mainly following the Pacific coastal route. However, details about the migration into the Americas and the routes pursued on the continent still remain unresolved, despite numerous genetic, archaeological, and linguistic investigations. To examine the pioneering peopling phase of the South American continent, we screened literature and mtDNA databases and identified two novel mitochondrial DNA (mtDNA) clades, here named D1g and D1j, within the pan-American haplogroup D1. They both show overall rare occurrences but local high frequencies, and are essentially restricted to populations from the Southern Cone of South America (Chile and Argentina). We selected and completely sequenced 43 D1g and D1j mtDNA genomes applying highest quality standards. Molecular and phylogeographic analyses revealed extensive variation within each of the two clades and possibly distinct dispersal patterns. Their age estimates agree with the dating of the earliest archaeological sites in South America and indicate that the Paleo-Indian spread along the entire longitude of the American double continent might have taken even <2000 yr. 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 and for clarification of Paleo-Indian migrations, by targeting, if possible, both the general mixed population of national states and autochthonous Native American groups, especially in South America.

[Supplemental material is available for this article.]

After more than two decades of genetic research using mitochondrial DNA (mtDNA), the questions concerning the number, route, and timing of human migration events into America are still far from being completely addressed (Bandelt et al. 2003; Perego et al. 2009, 2010; Hubbe et al. 2010), even though Native American populations were among the first human groups to be examined in detail. Four distinct haplogroup clusters were recognized in early studies (Schurr et al. 1990; Torroni et al. 1993) and initially named A, B, C, and D; haplogroup X was identified a few years later (Forster et al. 1996; Scozzari et al. 1997; Brown et al. 1998). Together, they encompass almost all extant Native American diversity (Schurr 2004; Tamm et al. 2007). Additional mtDNA founders were detected subsequently and currently comprise the so-called major pan-American lineages A2, B2, C1b, C1c, C1d, and D1, and generally less frequent (so-called minor) lineages such as D2a, D3, and X2a, which are restricted to northern North America, C4c and D4h3a with more widespread occurrence, and the very rare X2g (Kemp et al. 2007; Tamm et al. 2007; Achilli et al. 2008; Perego et al. 2009; Malhi et al. 2010; Hooshiar Kashani et al. 2012). Most recently, up to 15 surviving founding mtDNA lineages were recognized (Perego et al. 2010).

Numerous models for migration into and across the American double continent have been proposed to explain the coexistence of lineages with such different dispersal patterns. Because all Native American mtDNA founders originate from stocks in southern Siberia and Mongolia (Goebel et al. 2008; Ray et al. 2010)—with the exception of X2a and X2g (Reidla et al. 2003; Perego et al. 2009), for which no close modern relatives in East Asia were found—but are distinct from their sister clades in Asia, a direct colonization has not been favored in the recent literature. It is generally agreed that population groups that had moved to Beringia before the Last Glacial Maximum (LGM) persisted there, probably forced by an ecological barrier, and migrated into America before the

[8]**Corresponding author.**
E-mail walther.parson@i-med.ac.at.
Article published online before print. Article, supplemental material, and publication date are at http://www.genome.org/cgi/doi/10.1101/gr.131722.111.

end of the last glacial phase (Schurr 2004; Tamm et al. 2007; Achilli et al. 2008; Fagundes et al. 2008a; Perego et al. 2009, 2010). The incubation period in Beringia was long enough to alter the genetic composition and generate American-specific founder mutational motifs. In spite of that, the available data on modern native circumarctic populations of the former Beringia do not show the expected level of mtDNA variation (mainly A2 and D2 lineages). However, even if archaeological investigations in former Beringia are rare due to inundation, frozen ground, and limited settlement activities (Hoffecker et al. 1993; Vasil'ev et al. 2002; Mulligan et al. 2008), recent findings of putative B2, D1, and D3 mtDNA haplotypes in ancient bone samples from Alaska support the source population model (Raff et al. 2010). Several factors may have contributed to this discontinuity: The climatic changes that led to the deglaciation and the inundation of central Beringia between ~13 and 10 kya (Schurr et al. 1999; Zlojutro et al. 2006; Kitchen et al. 2008; Volodko et al. 2008) and continued through the present interstadial (Hu et al. 2001) probably caused a great proportion of the only less than 2000 effective source population individuals (Mulligan et al. 2008) to move from the harsh conditions once a viable alternative in the south was present (Kunz and Reanier 1994; Kitchen et al. 2008; Mulligan et al. 2008), likely giving rise to multiple further replacements, extinctions, and (re-)expansions that reshaped the genetic variation of the population groups in different refugia (Schurr et al. 1999; Schurr 2004; Zlojutro et al. 2006; Volodko et al. 2008). Native tribes have been decimated on both rims of the former Beringia due to oppression, warfare, epidemics, forced relocations, and admixture since the 18th Century (Schurr et al. 1999; Zlojutro et al. 2006, 2009; Rubicz et al. 2010).

In a rapid expansion from Beringia, a small number of founders are thought to have colonized all of America with no nested hierarchy—giving rise to pan-American patterns of mtDNA and other genetic markers (Schurr 2004; Schroeder et al. 2007; Tamm et al. 2007; Fagundes et al. 2008b; Perego et al. 2010). This quasi-concomitant initial migration likely occurred preferentially along the coastal Pacific route (Hurst 1943; Fladmark 1979; Schurr 2004; Wang et al. 2007; Achilli et al. 2008; Perego et al. 2010; Yang et al. 2010; Erlandson and Braje 2011). The migration into America, however, seems to have been a more complex process than a "single arrow on a map" (Bandelt et al. 2003; Perego et al. 2009; Rothhammer and Dillehay 2009; Hubbe et al. 2010; Erlandson and Braje 2011). Additional migration events could explain the distinct dispersal patterns of the minor Native American lineages (Tamm et al. 2007; Achilli et al. 2008; Perego et al. 2009, 2010). Other paths such as the ice-free corridor between the Laurentide and Cordilleran ice sheets in North America have been proposed as early as 1933 (cf. Rothhammer and Dillehay 2009). Later genetic exchange, probably in both directions, is corroborated by nuclear (Rasmussen et al. 2010; Ray et al. 2010), morphometric (González-José et al. 2008), and linguistic (Greenberg et al. 1986; Fortescue 1998) data.

Demographic events after the initial colonization as described for Beringia (see above) may explain the regional genetic patterns also in other areas of America. Tribalization limited genetic exchange (Malhi et al. 2002; Schurr 2004; Tamm et al. 2007; Perego et al. 2010). Population expansions and eastward migrations have been suggested for the spread of people over the interior continental masses of both Americas (Keyeux et al. 2002; Malhi et al. 2002; Fagundes et al. 2008b).

As for South America, there is no general consensus on a peopling model neither from Y-chromosomal, nor from mtDNA data (cf. Lewis et al. 2007; Lewis 2010). Envisioning the topography of South America with the Andes, a mountain range spanning from ~10°N to ~55°S latitude attaining an altitude of >6900 m, as a dominant feature (Masello et al. 2011), two main plausible scenarios can be outlined: The population groups expanding from the north could either have solely progressed on the coastal (or continental) side of the Andes and later crossed the cordillera at different latitudes, or alternatively, a random split of the source population could have occurred in a northern area of South America, resulting in separated coastal and continental population groups. Several genetic studies (using autosomal, X-, Y-chromosomal, and mtDNA markers), as well as craniometric and linguistic investigations support the latter hypothesis of an early split into an Andean (western) population and a smaller, less diverse continental (eastern) population, with limited subsequent exchange (Cavalli-Sforza et al. 1994; Luiselli et al. 2000; Rothhammer et al. 2001; Tarazona-Santos et al. 2001; Keyeux et al. 2002; Pucciarelli et al. 2006; Wang et al. 2007; Rothhammer and Dillehay 2009; Yang et al. 2010). A bifurcate coastal migration to the south and east after passing through the Isthmus of Panama has also been postulated (Gruhn 1994). Other models proposed multiple separate migrations from the north into South America with different patterns of dispersal (cf. Lewis et al. 2007). Extant Andean populations exhibited higher diversity and lower differentiation, showing signs of expansion and a larger long-term effective population size. Eastern South American populations, in contrast, revealed lower diversity, greater differentiation, and signs of either a recent bottleneck or a smaller effective size during expansion (Wang et al. 2007; Lewis 2010; Yang et al. 2010). Differential genetic drift and gene flow in western and eastern populations have been suggested (Fuselli et al. 2003; Lewis et al. 2005). However, separate source populations for the eastern and western regions have been questioned due to a previously underestimated variation in eastern South America (Lewis and Long 2008). Much later, the genetic composition of both Americas was heavily influenced by contact with Europeans and the *trans*-Atlantic slave trade (Salas et al. 2004; Mendizabal et al. 2008; Catelli et al. 2011; Gómez-Carballa et al. 2012).

In addition, the arrival time of the ancestors of Native Americans remains under debate. It has been broadly estimated to be ~10–40 kya (Schurr 2004). Recent publications agreed on an average coalescence time of 15–18 kya for the pan-American haplogroups (Tamm et al. 2007; Achilli et al. 2008; Perego et al. 2009, 2010). For the minor founder haplogroups, similar ages have been calculated (Fagundes et al. 2008b; Perego et al. 2010; Hooshiar Kashani et al. 2012). Mainly, differences are due to the various calculation methods and mutation rates applied and the amount of sequence information available. The estimates have recently been narrowed by archaeological findings: The traditional, but controversial Clovis-first (or Single Origin) model of 1937 (cf. Rothhammer and Dillehay 2009) was not supported by recent genetic research (e.g., Achilli et al. 2008; Perego et al. 2009) and became obsolete by excavations in Texas that provide evidence of human presence by 15.5 kya (Waters et al. 2011). The colonization time for South America is archaeologically estimated to 14.1–14.6 kya (Fig. 2, Monte Verde site, Chile; Dillehay et al. 2008; Erlandson et al. 2008).

Only when all founder lineages are analyzed at a high phylogenetic resolution in their entire distribution range, will a more comprehensive conclusion on migration and timing be feasible. In this study, 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 evaluated marginally (e.g., Horai et al. 1993; Moraga et al. 2000; Bobillo et al. 2010), and providing novel data concerning the final phases of the

long journey that brought humans from Beringia to the tip of South America.

Results

Refining the phylogeny of D1: subhaplogroups D1g and D1j

Although D1 is a major pan-American founder lineage, its phylogeny is still poorly resolved, mainly due to the paucity of high-quality full mtDNA genome sequence data. By screening the Sorenson Molecular Genealogy Foundation (SMGF, http://www.smgf.org) mtDNA control region (CR) database, the European DNA Profiling Group Mitochondrial Population Database (EMPOP, http://www.empop.org), an in-house database of more than 7000 Native American mtDNA CR sequences (A. Salas), and literature for sequences (partially) matching the D1 motif, we identified two major subsets of D1 mtDNAs that could not be assigned to any known D1 subhaplogroup of the current complete mtDNA tree (van Oven and Kayser 2009) (Build 13). The members of one subset harbored a transition at nucleotide position (np) C16187T, while the other group was defined by the distinguishing mutational motif T152C–C16242T–T16311C. To clarify whether these control region motifs defined novel subhaplogroups, we completely sequenced 45 selected D1 mtDNAs, almost all displaying either one or the other of the mutational motifs. We applied highest quality standards during the sequencing analysis to ensure precise base-calling, which is especially important when investigating lineages with limited prior coverage, because novel polymorphisms can be expected. The complete mtDNA sequences formed two clades, both with considerable internal variation and geographically restricted to southern South America. They were termed "D1g" and "D1j" (Fig. 1), thereby expanding the established mtDNA tree (van Oven and Kayser 2009) (Build 13) and modifying an earlier proposal (Bobillo et al. 2010). Two of the fully sequenced D1 mtDNAs (#44 and #45) did not belong to either D1g or D1j, including one harboring a transversion (C to A) instead of a transition at np 16187. This novel lineage, so far only represented by one fully sequenced mtDNA genome (Fig. 1), is further supported by CR data (see below).

Haplogroup D1g, represented by 26 full mtDNA genomes, is defined by the mutational pattern A8116G–C16187T. This haplogroup exhibits ample diversity with at least six major basal branches (D1g1–D1g6) (Fig. 1). Haplogroup D1j encompasses 17 entire mtDNA genomes, and it is characterized by the CR mutational pattern T152C–C16242T–T16311C. Most of our completely sequenced D1j mtDNAs (15 out of 17) harbored an additional transition at np C15868T, thus forming a subclade that we termed "D1j1." Again, all but two of the D1j1 samples cluster into a prominent branch (D1j1a), characterized by the mutational motif A4212G–T5004C–A15644G (Fig. 1).

The origins of the sample donors for the 45 completely sequenced mtDNAs are listed along with references in Supplemental Table S1. The complete sequences are illustrated in Supplemental Table S2 and have been deposited in GenBank under the accession numbers JN253391–JN253435.

Phylogeographic patterns of D1g and D1j

Plotting the sample donors' origins on a map (of the entire genomes considered in the present study; see Supplemental Table S1), haplogroup D1g appeared widespread over Chile and Argentina (Fig. 2). All six basal branches were present on both sides of the Andean Cordillera, and, when two or more samples were found, they were not restricted to specific regions of the countries. The greatest proportion (12/26) of donors derived from the Argentinean province of Rio Negro, in agreement with previously published CR data (Bobillo et al. 2010).

Haplogroup D1j mtDNAs showed a less broad dispersal in Chile (Fig. 2), where only three samples were found; and it is virtually absent from the southern part of the Southern Cone. The majority of samples and the more derived clusters were found in northern and eastern Argentina. Most donors originated from the Buenos Aires province of Argentina, which again corresponds to the high local frequency previously reported for the corresponding CR motif (Bobillo et al. 2010). Two additional samples were from southern Brazil.

To better evaluate the geographical distribution of the novel haplogroups, we surveyed published and unpublished CR data from a wide range of populations (Ginther et al. 1993; Horai et al. 1993; Alves-Silva et al. 2000; Moraga et al. 2000; Lalueza-Fox et al. 2001; García-Bour et al. 2004; Tajima et al. 2004; Cabana et al. 2006; Álvarez-Iglesias et al. 2007; Tamm et al. 2007; Carvalho et al. 2008; Salas et al. 2008; Bobillo et al. 2010; Catelli et al. 2011; Gayà-Vidal et al. 2011; Prieto et al. 2011; Sans et al. 2011; MC Bobillo, unpubl.). This allowed the identification of 103 putative D1g and D1j mtDNAs (Supplemental Table S3). The 54 potential D1g samples in these earlier reports confirm the spread of this subhaplogroup throughout Argentina and Chile; the 49 potential D1j samples broaden the geographic range of this lineage in Argentina and Brazil. Further D1j matches were found in Bolivia, Uruguay, and, intriguingly, the Dominican Republic (see below). Seven partial sequences matching the mutational motif of sample #44 (transversion at np 16187) were also identified; notably six of them in Brazil and one in an Argentinean sample (Supplemental Table S3). Therefore, these samples most likely represent another rare South American subclade of D1.

Frequency patterns of D1g and D1j

A survey of the 9567 CR sequences contained in the EMPOP database (v.2.1, release 5) yielded only the 17 D1g (0.18%) and 10 D1j (0.10%) samples previously reported from Argentina (Bobillo et al. 2010; MC Bobillo, unpubl.), most of which were completely sequenced here (Supplemental Tables S1, S3). A CR-based search of the 38,460 American samples contained in the SMGF mtDNA database (as of July 2011) revealed an overall frequency of 0.33% and 0.03% for D1g and D1j, respectively. All samples derived from populations of South America (Argentina, Bolivia, Brazil, Chile, Peru, and Uruguay). Only a singleton from North America (Michigan, USA) matched the D1g CR motif but could be traced back to a recent immigration from central-western Argentina based on the available genealogical data. The countrywide frequencies of D1g in the South American countries ranged between 0.22% (Brazil) and 2.33% (Argentina) in the SMGF database, with the exception of Chile, where it appears to be very common (16.60%). D1j samples exhibited lower frequencies spanning from 0.05% (Peru) to 9.30% (Argentina), but the latter high frequency was observed in a very small sample. In another in-house database of 5732 South American mitochondrial genome hypervariable segment I sequences (A. Salas, as of May 2011), haplogroup D1g samples appeared at a frequency of 1.19% and D1j samples at a frequency of 0.87%, and none of them was found outside South America in the more than 130,000 worldwide samples in this database. These results for D1g and D1j are concordant with a recent study (Bobillo

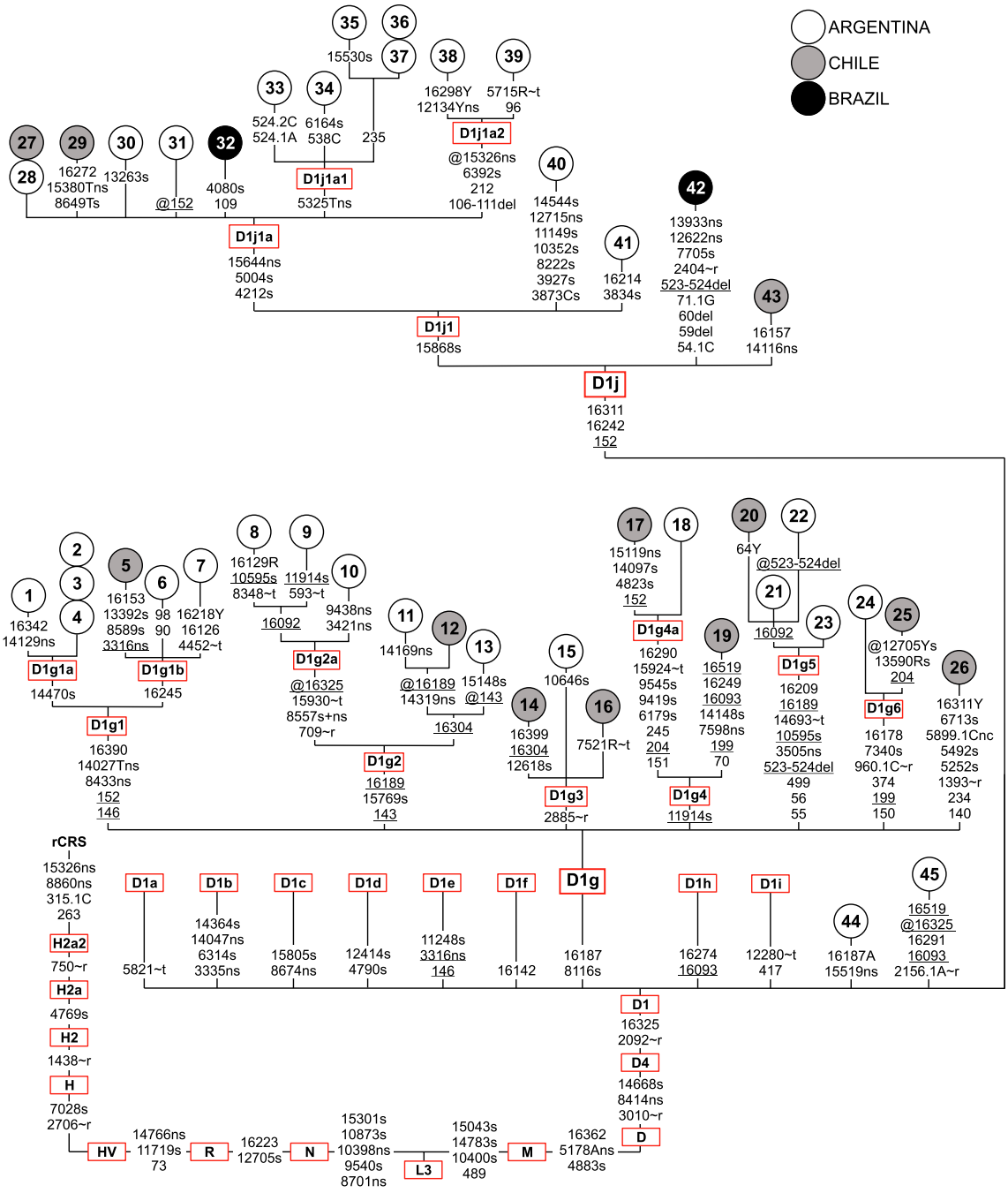

Figure 1. Phylogenetic tree of southern South American haplogroups D1g and D1j. This tree includes 43 new complete mtDNA sequences belonging to D1g and D1j and illustrates subhaplogroup affiliations. The position of the revised Cambridge Reference Sequence (rCRS) (Andrews et al. 1999) is indicated for reading off sequence motifs. Samples #1–#26 belong to D1g, while samples #27–#43 belong to D1j. The country of origin of the sample donors is indicated by circle color (see legend). All SNPs and indels are shown on the branches except for cytosine insertions at np 309. In the case of transversions, insertions, or heteroplasmic mutations, the base is indicated according to the IUPAC nucleotide code. The prefix @ indicates the reversion of a mutation occurring earlier in the phylogeny. The suffixes "s" and "ns" indicate synonymous and nonsynonymous substitutions, respectively, while "~t" and "~r" indicate affected positions in tRNA and rRNA loci, respectively. Recurrent mutations within the phylogeny are underlined. The mutational motifs of the previously described D1 subhaplogroups (D1a–D1f, D1h, D1i) are also shown (van Oven and Kayser 2009) (Build 13). D1 sequences #44 and #45 are also new. Additional information regarding each of the 45 novel mitochondrial genomes is available in Supplemental Tables S1 and S2.

et al. 2010), in which the lineage here termed "D1g" was found at an overall frequency of 4.73% in Argentina (almost exclusively in the south, at a local high of 27.7%) and the lineage now termed "D1j" at 2.66%. These figures differ only slightly from the ones in another recent report (Catelli et al. 2011), in which the frequency of D1g in Argentina was 1.3% and that of D1j 3.9%. For details on database frequencies, see Supplemental Table S4.

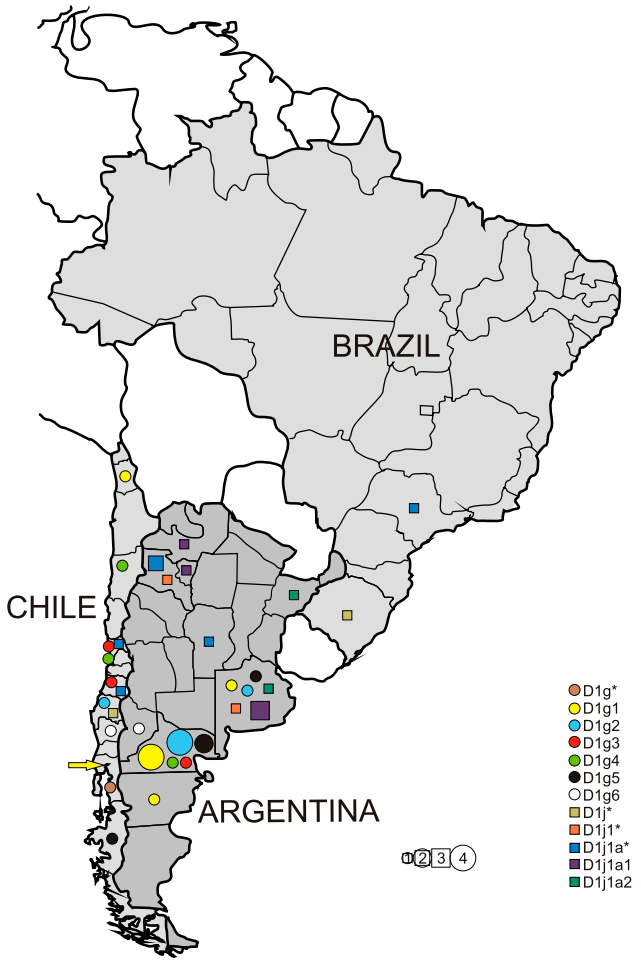

Figure 2. Origin of D1g and D1j sample donors. The provinces/states of origin of the sample donors of the fully sequenced mtDNA haplogroup D1g (circles) and D1j (squares) samples included in this study (see Fig. 1) are shown. Details are listed in Supplemental Table S1. The sizes of the circles and squares correspond to the number of samples assigned to a certain province/state, while their colors indicate subhaplogroups (see legend). The yellow arrow points at the location of the Monte Verde site (Chile).

In brief, the literature and available database reports indicate that D1g and D1j are relatively rare lineages restricted to the Southern Cone of South America. Almost all of the samples derive from Argentina, Chile, and Brazil (and neighboring countries). The few exceptions are two samples from the Dominican Republic and a singleton from the United States of Argentinean origin. To the best of our knowledge, these lineages were not found outside America in any database or publication.

Coalescence age calculation

Clade ages were estimated using the ρ statistics and maximum likelihood (ML) molecular divergence, considering the entire mtDNA genome with both approaches. The ML point estimate calculated for D1g was 18.3 ± 2.4 thousand years (ky). Its subclades grouped into older ones—D1g2 and D1g4, whose ages were 14.8 ± 3.1 ky and 16.9 ± 2.5 ky, respectively—and much younger subclades with point estimates between 1.2 ± 1.4 ky and 7.4 ± 2.1 ky. As in other studies based on entire mtDNA sequences (Perego et al. 2009, 2010; Ottoni et al. 2010), the age estimates calculated from the ρ statistics were comparable with the ML results, but generally with larger standard error (SE) (Table 1). Taking into account that the ML ages were slightly smaller than the ρ estimates, there is no indication for an underestimation by the ρ statistics here, notwithstanding a potential effect that would become clearly noticeable with low-resolution data (Cox 2008).

Haplogroup D1j seems slightly younger than D1g with an ML point estimate of 13.9 ± 2.9 ky. The age of its major subclade D1j1 is dated at the beginning of the Holocene (11.5 ± 2.6 ky). The pronounced subclade D1j1a of D1j1 reflects a much more recent development at an age of 4.0 ± 1.0 ky; its subclades yielded point estimates of 2.5 ± 1.0 (D1j1a1) and 0.8 ± 0.6 ky (D1j1a2) (see Table 1).

Finally, a potential founder age was also calculated treating the 43 samples of both haplogroups D1g and D1j as if they derived from the same root (which would reflect a common founder event). This reduced the SE and yielded remarkably similar overall ages for both methods applied, overlapping in the range of 16.9 ± 1.6 ky (ML: 16.5 ± 2.0, ρ: 17.8 ± 2.5) (Table 1).

Discussion

D1g and D1j: Southern Cone lineages present in the rapid pioneer settlement

The two novel mitochondrial lineages D1g and D1j appear almost exclusively distributed in the South American Southern Cone. Together with their ample diversity, this indicates that these haplogroups probably arose in South America, while their age in the range of 16.9 ± 1.6 ky implies that they originated at the very early stages of human colonization of this area. Thus, they were most probably carried by the pioneer settlers of the South American continent and present in the mtDNA pool of the human groups that founded the most ancient known settlement, the Monte Verde site in Chile, dated to 14.1–14.6 kya (Dillehay et al. 2008; Erlandson et al. 2008).

Table 1. Molecular divergences and age estimates obtained by maximum likelihood and ρ statistics for D1g, D1j, and their major subclades

		All nucleotide substitutions							
Haplogroup	*N*[a]	ML[b]	SE[c]	*T* (ky)[d]	Δ*T* (ky)[d]	ρ	σ	*T* (ky)[d]	Δ*T* (ky)[d]
D1g	26	6.7	0.8	18.3	2.4	7.2	1.0	19.7	3.0
>D1g1	7	2.8	0.8	7.4	2.1	2.4	0.9	6.4	2.3
>D1g2	6	5.5	1.1	14.8	3.1	5.0	1.4	13.4	3.7
>D1g3	3	1.8	0.9	4.7	2.4	1.3	0.7	3.4	1.7
>D1g4	4	6.2	0.9	16.9	2.5	8.7	2.2	23.8	6.3
>D1g5	4	0.5	0.5	1.2	1.4	0.8	0.8	1.9	1.9
>D1g6	2	0.6	0.5	1.4	1.4	0.5	0.5	1.3	1.3
D1j	17	5.2	1.0	13.9	2.9	5.5	1.7	14.9	4.7
>D1j1	15	4.3	0.9	11.5	2.6	4.9	1.6	13.0	4.4
>D1j1a	13	1.5	0.4	4.0	1.0	1.9	0.6	5.0	1.5
>D1j1a1	5	1.0	0.4	2.5	1.0	1.2	0.7	3.1	1.8
>D1j1a2	2	0.3	0.3	0.8	0.6	0.5	0.5	1.3	1.3
D1g and D1j[e]	43	6.1	0.7	16.5	2.0	6.6	0.9	17.8	2.5

[a]The number of complete mtDNA sequences.
[b]The maximum likelihood molecular divergence.
[c]Standard error.
[d]Age estimates using the corrected molecular clock proposed by Soares et al. (2009).
[e]D1g and D1j sequences treated as if they were deriving from the same root.

The divergence ages calculated for D1g and D1j further confirm that the human spread along the entire longitude of the American double continent (15,400 km in a conservative estimate) (Surovell 2003) was extremely fast and might have taken even less than 2000 yr, because the coalescence time for the entire haplogroup D1, ranging between 13.9 and 18.3 kya in recent reports (Achilli et al. 2008; Perego et al. 2009; Soares et al. 2009), sets a close upper limit. Such a rapid movement is consistent with the results of three simulation studies: a best-case model based on demography and ecology resulting in a time of 1.5 ky from the estimated starting point in North America to South America and ~2.6 ky to Monte Verde (Surovell 2003); a simulation study based on genetic data resulting in ~3 ky in a conservative approach for the colonization along the North and South American coasts (Fix 2005); and the best-fitting model of an anisotropic diffusion simulation, assuming human dispersal in South America to have taken 1.5–2 ky from a starting point in the north (Martino et al. 2007). This appears indeed feasible under similar ecological resources, where little adaptation is required, such as the "Pacific Rim Highway" along the coast providing diverse plant and animal foods (Erlandson and Braje 2011). The human capacity to survive in such littoral conditions with a minimal material culture has been ethnographically demonstrated by the example of the Yaghan people of Tierra del Fuego (Gruhn 1994). Because the oldest archaeological findings in South America were made in coastal areas and indicate a coastal subsistence pattern (Dillehay et al. 2008), a coastal colonization route is further supported.

Scenarios for the peopling of the South American continental interior

Besides shedding light on the initial colonization process of South America, the mitochondrial haplogroups D1g and D1j might further contribute to the clarification of the peopling of the continental interior. While there is still no common consent about the number and nature of the migrations (see above) (cf. Lewis et al. 2007; Lewis 2010), their footprints could be visible from mtDNA patterns in extant South American populations: (1) A haplogroup composition exhibiting overlapping and ample basal diversity on both sides of the Andes (with likely some additional variation developed in situ) would support a common source population with late separation (i.e., after all major subclades had developed). (2) Diverse and diverged haplogroup spectra would indicate an early separation or even separate migrations into South America. (3) A basal haplogroup diversity restricted to one side of the cordillera and the presence of only derived lineages on the other would indicate a colonization from the former (as shown for other species, e.g., Masello et al. 2011). However, extensive genetic exchange in later migrations could have blurred all initial patterns.

The results from analyzing the available D1g and D1j sequences are neither supporting the separation into a western and an eastern population group with limited gene flow that was favored in many earlier studies, nor a founding population restricted to one side of the Andes with restricted exchange after colonization of the other. Most likely, extensive *trans*-Andean migrations explain the present-day widespread dispersal of all basal subclades of D1g on both sides of the Andes (Fig. 2). The extant known dispersal of D1j does not clearly pinpoint any particular migration route. The most derived clades, D1j1a1 and D1j1a2, appear restricted to the eastern regions of the Southern Cone, where D1j is more frequent and widespread, which would possibly indicate its eastern origin, probably after a split; but both D1j* and derived D1j1a* samples are also found on the coastal side of the Andes (Fig. 2; Supplemental Table S3). Because D1j is even rarer than D1g (Supplemental Table S4), we suspect that our current view may be somewhat afflicted by sampling bias.

Appreciating the similar age estimates for D1g and D1j and their apparent dispersal pattern, we hypothesize a Pacific coastal route into South America followed by extensive gene flow across the cordillera as the most likely migration model. Only a coastal route, facilitated by little need for adaptation, can explain the speed of the migration from Beringia to Monte Verde at 13,400 km estimated distance (Surovell 2003). Strikingly, the peopling of the continental interior across the mountains that likely followed after the initial colonization became possible at nearly the same time that Monte Verde (41°30′S latitude) was founded: The first deglaciation and warming period (corresponding to the Bølling in Europe) after the LGM in southern South America, which opened several "low" (<2500 m) and "intermediate" (2500–4000 m) Andean passes south of 30°S latitude (Masello et al. 2011) as potential entrance routes, is dated to 14.6–14.3 kya (McCulloch et al. 2000). By 14 kya, a moderate increase in effective moisture and temperature could be inferred from vegetation and fire history in southern Patagonia (Chile, 46°S latitude) (Markgraf et al. 2007).

Another potential interpretation of our results would appear much less plausible: the origin of all D1g and D1j lineages in a common source population that separated in the north with little or no later migration over the mountain barrier. This would involve a split of these founder groups after all subclades present on both Andean sides had developed (≤5 kya), and thus a very recent start of the southward movement. Furthermore, this model would not conform to the presence of humans at the Monte Verde site at ~14 kya. Hence, a common source population that split into an eastern and a western group would be likely only with extended migrations, as described above, starting or continuing after the youngest lineages had differentiated. In addition, this demographic model would find more difficulties to explain the total absence of D1g and D1j in northern South America.

Figure 3 illustrates the three alternative migration models into and on the South American continent described above along with the timing of events resulting from our calculations. More mitochondrial sequences for D1g and D1j are needed to confirm our hypotheses and to fully support one or the other migration scenario.

The mtDNA legacy of the Mapuche

Information on the Native groups in which an mtDNA haplogroup is present can help to clarify migrational routes and explain extant dispersal patterns. The ethnic affiliation of sample donors of the fully sequenced mitochondrial genomes presented in this study was only available for five Argentinean samples. One D1g sample was obtained from the Mapuche, three D1j samples from the Diaguita, and one from the Kolla/Coya (Supplemental Table S1). The partial sequences compiled from the literature (Supplemental Table S3) revealed more information on the ethnic origin of the donors: Potential D1g samples have been reported in Argentinean Mapuche (Ginther et al. 1993); in three native Chilean populations (Moraga et al. 2000), namely, Mapuche, Pehuenche Mapuche, and Yaghan/Yamana; and in (19th Century) skeletal remains of the Kaweskar and Yaghan/Yamana from Tierra del Fuego (García-Bour et al. 2004). Haplogroup D1j was matched by sequences obtained from Argentinean Mapuche (Ginther et al. 1993), Kolla/Coya (Álvarez-Iglesias et al. 2007), Pilagá and Wichi/Mataco (Cabana et al. 2006), Bolivian Quechua (Gayà-Vidal et al. 2011), and extant

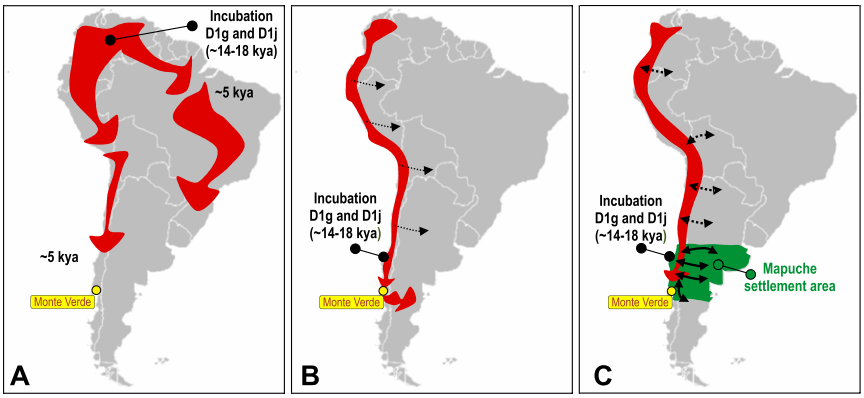

Figure 3. Migration models into the South American continent. The three migration models discussed are illustrated along with a timing of events resulting from our calculations: the incubation of population groups arriving from the north in a northern area of South America and a late split into coastal and continental population groups after the full development of all major D1g and D1j subclades (*A*); a coastal southward migration followed by the colonization of the continental interior by *trans*-Andean migrations, with limited later exchange along the cordillera (*B*); and a coastal southward migration and *trans*-Andean colonization of the continental interior, followed by extensive *trans*-Andean migrations, especially in the south, with the Mapuches favoring bidirectional gene flow between west and east in the Southern Cone (*C*). The latter is the most likely model from our results.

Dominican and Dominican Taino (Lalueza-Fox et al. 2001; Tajima et al. 2004), as well as African Brazilian (Carvalho et al. 2008) and self-identified Basque descendants from Uruguay (Sans et al. 2011), revealing maternal admixture in the latter two.

The novel mtDNA lineages D1g and D1j could constitute ancient Mapuche lineages, because their area of dispersal largely coincides with the territory occupied by the ancestors of the Mapuche people: the Southern Cone of South America, including the south of present-day Argentina and Chile. Unlike other indigenous peoples in South America, the Mapuche ("people of the earth") have maintained their independence from the European Conquistadores for more than three centuries, in the land they had occupied for thousands of years (Fig. 3). The pact of Quilín in 1641 guaranteed their independence south of the Bío-Bío river until the second half of the 19th Century, when the Mapuche began to suffer massive loss of land, followed by slaughter, and were finally defeated, becoming progressively poorer and disadvantaged (Carter 2010). Many were brought to the urban areas as servants and slaves. This deportation has also spread their genetic heritage. Today, the Mapuche constitute an admixed ethnic group and are the largest of all Argentinean and Chilean indigenous ethnicities (cf. Argentinean National Institute of Statistics and Census, http://www.indec.gov.ar; and Chilean National Statistical Institute, http://www.ine.cl), making up the majority of the population in some counties (Carter 2010).

The Mapuche settlement area could have enabled the two scenarios that are visible from our results (albeit to be confirmed with more data): The ancestor population of the Mapuche, possibly living in an area north of Chile, could represent the common source population compatible with the split scenario, where the incubation time before was long enough to develop all of the variation that is observed on both sides. The continuous extensive bidirectional gene flow across the mountain barrier after the initial coastal migration and differentiation postulated in the other scenario could have been mediated by the Mapuche ancestors that inhabited the areas on both sides and thereby served as a long-term genetic *trans*-Andean link (Fig. 3).

The identification of distinct haplogroups that compose an ethnic group's mtDNA pool (although they may have a somewhat broader distribution), such as obviously D1g and D1j for the Mapuche, can serve as a proxy for investigating the proportion of matrilineal ancestry from these groups in the extant mixed population. Extended studies on ethno-specific haplogroups will be meaningful toward a clarification of the population history of Native Americans and also shed light on historical (such as the differentiation or separation of ethnic groups) and more recent demographic events.

D1j in the Caribbean Islands?

Two independent studies reported partial CR sequences possibly belonging to haplogroup D1j in the Dominican Republic. While in one publication extant donors were sampled (Tajima et al. 2004), for which recent genetic inputs could be argued, the other (Lalueza-Fox et al. 2001) analyzed mtDNA sequences from skeletal remains of pre-Columbian Tainos, a Native Caribbean population group that was drastically and quickly reduced and became extinct soon after the European contact (Lalueza-Fox et al. 2001; Mendizabal et al. 2008). If the D1j status of these samples could be confirmed by either complete mtDNA sequencing or a survey of diagnostic coding region markers, excluding contaminations and other errors, a more complex migration scenario into, in, and from South America should be envisioned. A South American origin, probably a migration from the Orinoco Valley to the Caribbean Islands ~3 kya, has been historically favored for the ancestral Tainos, but also North and Central American contributions have been postulated (Lalueza-Fox et al. 2001; Mendizabal et al. 2008). Therefore, if confirmed, the presence of D1j mtDNAs in the Dominican Republic could represent the genetic echo of a truly South American source population's input into the Caribbean, supporting the hypothesis of a peopling of the Caribbean Islands from the southeast to the northwest.

Conclusion

In this study, we have revealed two new subclades of the pan-American founder haplogroup D1 that are exclusively limited to the Southern Cone of South America, and have added 45 new full mitochondrial genome sequences of southern South American origin to the clarification of the most basal variation in Native American mtDNA. The limited geographic dispersal but ample diversity of the novel haplogroups D1g and D1j and the coalescence ages calculated at the highest resolution together indicate a coastal and rapid initial colonization of South America on the way from Beringia to Tierra del Fuego that was succeeded by extensive *trans*-Andean migrations. The two lineages could represent the genetic heritage of the pioneer settlers of South America that is probably conserved in today's Mapuche people.

The phylogeny of haplogroup D1 considerably expanded by the novel sequences presented in this study and the expected diversity yet to uncover clearly show that more sampling of Native and mixed South American populations is highly desirable, which will also deliver a more detailed picture of the migration history

into and within South America than is currently obtained with the still only scarcely available complete mtDNA sequences.

The full mitochondrial genome constitutes the maximum of information available from the maternal side, but insights gained into early South American population history by studying haplogroups D1g and D1j have to be assessed against other local haplogroups with similar or contrasting dispersal patterns. Last but not least, because "the best chance for obtaining a higher resolution of population history will be to examine many independently inherited loci" (Lewis and Long 2008), further Y-chromosomal, autosomal, and archaeological investigations (cf. O'Rourke and Raff 2010) can confirm the hypotheses from the maternal side and complete the complex picture of South American colonization.

Methods

Sample collection and analysis of mtDNA sequence variation

A set of 45 mtDNAs from Argentina, Brazil, and Chile was selected on the basis of their CR variation following a detailed survey of literature and databases. DNA samples were provided by multiple collaborators, thus no common methods for DNA extraction were used. All donors gave their informed consent and were fully anonymized. Whole mitochondrial genome sequences were generated ensuring highest forensic standards by (1) applying a strategy that aims to minimize sequencing errors (by the minimum double-strand coverage of each nucleotide and the use of sequencing primers with optimal signal-to-noise ratio) in a sequencing workflow avoiding artificial recombination (Fendt et al. 2009); and (2) a posteriori data inspection, including (filtered) phylogenetic analyses, performed by two independent analysts and reviewed by a third. The sequences were aligned to the revised Cambridge Reference Sequence (rCRS) (Andrews et al. 1999) using Sequencher v.4.8 (GeneCodes) and following updated nomenclature guidelines for mtDNA (Bandelt and Parson 2008). Point heteroplasmic positions were repeatedly sequenced until clarity was obtained. Phylogeny construction was performed as previously described (Achilli et al. 2008).

Analysis of close maternal relatedness

To avoid a biased representation of lineages, subjects with identical mtDNA sequences were inspected for maternal relatedness. By typing autosomal STR loci and the amelogenin sex-related length polymorphism, pedigree construction, and calculation of likelihood ratios (LR) using the STR allele frequencies found in the population, closely maternally related individuals (i.e., mother–child and sibling constellations) can be identified (Egeland et al. 2000; Bodner et al. 2011). Based on 10 STR loci and allele frequencies previously reported (Sala et al. 1999; Gangitano et al. 2002; Berardi et al. 2003; Marino et al. 2006; Borosky et al. 2009), none of the donor pairs with identical mtDNA sequences revealed close maternal relatedness when applying a cut-off LR of 1000.

Age estimates

To obtain maximum likelihood (ML) molecular divergences, we used PAML 4.4 (Yang 2007), assuming the HKY85 mutation model with gamma-distributed rates, as previously reported (Perego et al. 2009). The ML estimates were then compared with those directly obtained from the averaged distance (ρ) of the haplotypes of a clade to the respective root haplotype, accompanied by a heuristic estimate of the standard error (σ) calculated from an estimate of the genealogy. These calculations were performed on the entire mtDNA haplotypes, while mutational distances were converted into years using the corrected molecular clock proposed by Soares et al. (2009). Concerns about the accuracy of the estimation via ρ statistics have been raised (Cox 2008), however, only considering CR, a small and hypervariable fraction of the mitochondrial genome. This may easily explain why studies based on entire mtDNA sequences (Perego et al. 2009, 2010; Ottoni et al. 2010) found good concordance between ML and ρ results, especially for multifurcating clades.

Data access

The GenBank (http://www.ncbi.nlm.nih.gov/genbank) accession numbers for the 45 novel complete mtDNA sequences reported in this paper are JN253391–JN253435.

Acknowledgments

We are grateful to the individuals who donated their DNA for research and thank Daniela Niederwieser (Institute of Legal Medicine, Innsbruck Medical University) for excellent technical assistance. This project received support from Fondazione Alma Mater Ticinensis (to A.T.); the Italian Ministry of the University: Progetti Ricerca Interesse Nazionale 2009 (to A.A. and A.T.), FIRB-Futuro in Ricerca 2008 (to A.A.); Sorenson Molecular Genealogy Foundation (to U.A.P. and S.R.W.); Fundación de Investigación Médica Mutua Madrileña (2008/CL444) and Ministerio de Ciencia e Innovación (SAF2008-02971) (to A.S.); and the Austrian Science Fund FWF, Translational Research project L397 (to M.B., A.W.R., and W.P.).

References

Achilli A, Perego UA, Bravi CM, Coble MD, Kong QP, Woodward SR, Salas A, Torroni A, Bandelt HJ. 2008. The phylogeny of the four pan-American mtDNA haplogroups: Implications for evolutionary and disease studies. *PLoS ONE* **3:** e1764. doi: 10.1371/journal.pone.0001764.

Álvarez-Iglesias V, Jaime JC, Carracedo Á, Salas A. 2007. Coding region mitochondrial DNA SNPs: Targeting East Asian and Native American haplogroups. *Forensic Sci Int Genet* **1:** 44–55.

Alves-Silva J, da Silva Santos M, Guimares PEM, Ferreira ACS, Bandelt HJ, Pena SDJ, Prado VF. 2000. The ancestry of Brazilian mtDNA lineages. *Am J Hum Genet* **67:** 444–461.

Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N. 1999. Reanalysis and revision of the Cambridge Reference Sequence for human mitochondrial DNA. *Nat Genet* **23:** 147. doi: 10.1038/13779.

Bandelt HJ, Parson W. 2008. Consistent treatment of length variants in the human mtDNA control region: A reappraisal. *Int J Legal Med* **122:** 11–21.

Bandelt HJ, Herrnstadt C, Yao YG, Kong QP, Kivisild T, Rengo C, Scozzari R, Richards M, Villems R, Macaulay V, et al. 2003. Identification of Native American founder mtDNAs through the analysis of complete mtDNA sequences: Some caveats. *Ann Hum Genet* **67:** 512–524.

Berardi G, Toscanini U, Raimondi E. 2003. STR data for PowerPlex 16 System from Buenos Aires population, Argentina. *Forensic Sci Int* **134:** 222–224.

Bobillo MC, Zimmermann B, Sala A, Huber G, Röck A, Bandelt HJ, Corach D, Parson W. 2010. Amerindian mitochondrial DNA haplogroups predominate in the population of Argentina: Towards a first nationwide forensic mitochondrial DNA sequence database. *Int J Legal Med* **124:** 263–268.

Bodner M, Irwin JA, Coble MD, Parson W. 2011. Inspecting close maternal relatedness: Towards better mtDNA population samples in forensic databases. *Forensic Sci Int Genet* **5:** 138–141.

Borosky A, Catelli L, Vullo C. 2009. Analysis of 17 STR loci in different provinces of Argentina. *Forensic Sci Int Genet* **3:** e93–e95.

Brown MD, Hosseini SH, Torroni A, Bandelt HJ, Allen JC, Schurr TG, Scozzari R, Cruciani F, Wallace DC. 1998. MtDNA haplogroup X: An ancient link between Europe/Western Asia and North America? *Am J Hum Genet* **63:** 1852–1863.

Cabana GS, Merriwether DA, Hunley K, Demarchi DA. 2006. Is the genetic structure of Gran Chaco populations unique? Interregional perspectives on native South American mitochondrial DNA variation. *Am J Phys Anthropol* **131:** 108–119.

Carter D. 2010. Chile's other history: Allende, Pinochet, and redemocratisation in Mapuche perspective. *Stud Ethn & Nationalism* **10:** 59–75.
Carvalho BM, Bortolini MC, dos Santos SEB, Ribeiro-dos-Santos ÂKC. 2008. Mitochondrial DNA mapping of social-biological interactions in Brazilian Amazonian African-descendant populations. *Genet Mol Biol* **31:** 12–22.
Catelli ML, Álvarez-Iglesias V, Gómez-Carballa A, Mosquera-Miguel A, Romanini C, Borosky A, Amigo J, Carracedo Á, Vullo C, Salas A. 2011. The impact of modern migrations on present-day multi-ethnic Argentina as recorded on the mitochondrial DNA genome. *BMC Genet* **12:** 77. doi: 10.1186/1471-2156-12-77.
Cavalli-Sforza LL, Menozzi P, Piazza A. 1994. *The history and geography of human genes*. Princeton University Press, Princeton, NJ.
Cox MP. 2008. Accuracy of molecular dating with the ρ statistic: Deviations from coalescent expectations under a range of demographic models. *Hum Biol* **80:** 335–357.
Dillehay TD, Ramírez C, Pino M, Collins MB, Rossen J, Pino-Navarro JD. 2008. Monte Verde: Seaweed, food, medicine, and the peopling of South America. *Science* **320:** 784–786.
Egeland T, Mostad PF, Mevåg B, Stenersen M. 2000. Beyond traditional paternity and identification cases. Selecting the most probable pedigree. *Forensic Sci Int* **110:** 47–59.
Erlandson JM, Braje TJ. 2011. From Asia to the Americas by boat? Paleogeography, paleoecology, and stemmed points of the Northwest Pacific. *Quat Int* **239:** 28–37.
Erlandson JM, Braje TJ, Graham MH. 2008. How old is MVII? Seaweeds, shorelines, and the pre-Clovis chronology at Monte Verde, Chile. *J Isl & Coast Archaeol* **3:** 277–281.
Fagundes NJ, Kanitz R, Bonatto SL. 2008a. A reevaluation of the Native American mtDNA genome diversity and its bearing on the models of early colonization of Beringia. *PLoS ONE* **3:** e3157. doi: 10.1371/journal.pone.0003157.
Fagundes NJ, Kanitz R, Eckert R, Valls ACS, Bogo MR, Salzano FM, Smith DG, Silva WA Jr, Zago MA, Ribeiro-dos-Santos AK, et al. 2008b. 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.
Fendt L, Zimmermann B, Daniaux M, Parson W. 2009. Sequencing strategy for the whole mitochondrial genome resulting in high quality sequences. *BMC Genomics* **10:** 139. doi: 10.1186/1471-2164-10-139.
Fix AG. 2005. Rapid deployment of the five founding Amerind mtDNA haplogroups via coastal and riverine colonization. *Am J Phys Anthropol* **128:** 430–436.
Fladmark KR. 1979. Routes: Alternate migration corridors for early man in North America. *Am Antiq* **44:** 55–69.
Forster P, Harding R, Torroni A, Bandelt HJ. 1996. Origin and evolution of Native American mtDNA variation: A reappraisal. *Am J Hum Genet* **59:** 935–945.
Fortescue M.D. 1998. *Language relations across Bering Strait: Reappraising the archaeological and linguistic evidence*. Cassell, London.
Fuselli S, Tarazona-Santos E, Dupanloup I, Soto A, Luiselli D, Pettener D. 2003. Mitochondrial DNA diversity in South America and the genetic history of Andean highlanders. *Mol Biol Evol* **20:** 1682–1691.
Gangitano DA, Garófalo MG, Juvenal GJ, Budowle B, Lorente JA, Padula RA. 2002. STR data for the PowerPlex 16 loci in Buenos Aires population (Argentina). *J Forensic Sci* **47:** 418–420.
García-Bour J, Pérez-Pérez A, Álvarez S, Fernández E, López-Parra AM, Arroyo-Pardo E, Turbón D. 2004. Early population differentiation in extinct aborigines from Tierra del Fuego-Patagonia: Ancient mtDNA sequences and Y-chromosome STR characterization. *Am J Phys Anthropol* **123:** 361–370.
Gayà-Vidal M, Moral P, Saenz-Ruales N, Gerbault P, Tonasso L, Villena M, Vasquez R, Bravi CM, Dugoujon JM. 2011. MtDNA and Y-chromosome diversity in Aymaras and Quechuas from Bolivia: Different stories and special genetic traits of the Andean altiplano populations. *Am J Phys Anthropol* **145:** 215–230.
Ginther C, Corach D, Penacino GA, Rey JA, Carnese FR, Hutz MH, Anderson A, Just J, Salzano FM, King MC. 1993. Genetic variation among the Mapuche Indians from the Patagonian region of Argentina: Mitochondrial DNA sequence variation and allele frequencies of several nuclear genes. *EXS* **67:** 211–219.
Goebel T, Waters MR, O'Rourke DH. 2008. The Late Pleistocene dispersal of modern humans in the Americas. *Science* **319:** 1497–1502.
Gómez-Carballa A, Ignacio-Veiga A, Álvarez-Iglesias V, Pastoriza-Mourelle A, Ruíz Y, Pineda L, Carracedo Á, Salas A. 2012. A melting pot of multicontinental mtDNA lineages in admixed Venezuelans. *Am J Phys Anthropol* **147:** 78–87.
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.
Greenberg JH, Turner CG II, Zegura SL. 1986. The settlement of the Americas: A comparison of the linguistic, dental and genetic evidence. *Curr Anthropol* **27:** 477–497.
Gruhn R. 1994. The Pacific coast route of initial entry: An overview. In *Method and theory for investigating the peopling of the Americas* (ed. R Bonnichsen, DG Steele), pp. 249–256. Center for the Study of the First Americans, Corvallis, OR.
Hoffecker JF, Powers WR, Goebel T. 1993. The colonization of Beringia and the peopling of the New World. *Science* **259:** 46–53.
Hooshiar Kashani B, Perego UA, Olivieri A, Angerhofer N, Gandini F, Carossa V, Lancioni H, Semino O, Woodward SR, Achilli A, et al. 2012. Mitochondrial haplogroup C4c: A rare lineage entering America through the ice-free corridor? *Am J Phys Anthropol* **147:** 35–39.
Horai S, Kondo R, Nakagawa-Hattori Y, Hayashi S, Sonoda S, Tajima K. 1993. Peopling of the Americas, founded by four major lineages of mitochondrial DNA. *Mol Biol Evol* **10:** 23–47.
Hu FS, Ito E, Brown TA, Curry BB, Engstrom DR. 2001. Pronounced climatic variations in Alaska during the last two millennia. *Proc Natl Acad Sci* **98:** 10552–10556.
Hubbe M, Neves WA, Harvati K. 2010. Testing evolutionary and dispersion scenarios for the settlement of the New World. *PLoS ONE* **5:** e11105. doi: 10.1371/journal.pone.0011105.
Hurst CT. 1943. A Folsom site in a mountain valley of Colorado. *Am Antiq* **8:** 250–253.
Kemp BM, Malhi RS, McDonough J, Bolnick DA, Eshleman JA, Rickards O, Martinez-Labarga C, Johnson JR, Lorenz JG, Dixon EJ, et al. 2007. Genetic analysis of early holocene skeletal remains from Alaska and its implications for the settlement of the Americas. *Am J Phys Anthropol* **132:** 605–621.
Keyeux G, Rodas C, Gelvez N, Carter D. 2002. Possible migration routes into South America deduced from mitochondrial DNA studies in Colombian Amerindian populations. *Hum Biol* **74:** 211–233.
Kitchen A, Miyamoto MM, Mulligan CJ. 2008. A three-stage colonization model for the peopling of the Americas. *PLoS One* **3:** e1596. doi: 10.1371/journal.pone.0001596.
Kunz ML, Reanier RE. 1994. Paleoindians in Beringia: Evidence from arctic Alaska. *Science* **263:** 660–662.
Lalueza-Fox C, Calderón FL, Calafell F, Morera B, Bertranpetit J. 2001. MtDNA from extinct Tainos and the peopling of the Caribbean. *Ann Hum Genet* **65:** 137–151.
Lewis CM Jr. 2010. Hierarchical modeling of genome-wide Short Tandem Repeat (STR) markers infers Native American prehistory. *Am J Phys Anthropol* **141:** 281–289.
Lewis CM Jr, Long JC. 2008. Native South American genetic structure and prehistory inferred from hierarchical modeling of mtDNA. *Mol Biol Evol* **25:** 478–486.
Lewis CM Jr, Tito RY, Lizárraga B, Stone AC. 2005. Land, language, and loci: mtDNA in Native Americans and the genetic history of Peru. *Am J Phys Anthropol* **127:** 351–360.
Lewis CM Jr, Lizárraga B, Tito RY, López PW, Iannacone GC, Medina A, Martínez R, Polo SI, De La Cruz AF, Cáceres AM, et al. 2007. Mitochondrial DNA and the peopling of South America. *Hum Biol* **79:** 159–178.
Luiselli D, Simoni L, Tarazona-Santos E, Pastor S, Pettener D. 2000. Genetic structure of Quechua-speakers of the Central Andes and geographic patterns of gene frequencies in South Amerindian populations. *Am J Phys Anthropol* **113:** 5–17.
Malhi RS, Eshleman JA, Greenberg JA, Weiss DA, Schultz Shook BA, Kaestle FA, Lorenz JG, Kemp BM, Johnson JR, Smith DG. 2002. The structure of diversity within New World mitochondrial DNA haplogroups: Implications for the prehistory of North America. *Am J Hum Genet* **70:** 905–919.
Malhi RS, Cybulski JS, Tito RY, Johnson J, Harry H, Dan C. 2010. Mitochondrial haplotype C4c confirmed as a founding genome in the Americas. *Am J Phys Anthropol* **141:** 494–497.
Marino M, Sala A, Corach D. 2006. Genetic attributes of 15 autosomal STRs in the population of two Patagonian provinces of Argentina. *Forensic Sci Int* **160:** 84–88.
Markgraf V, Whitlock C, Haberle S. 2007. Vegetation and fire history during the last 18,000 cal yr B.P. in southern Patagonia: Mallín Pollux, Coyhaique, Province Aisén (45°41′30″S, 71°50′30″W, 640 m elevation). *Palaeogeogr Palaeocl* **254:** 492–507.
Martino LA, Osella A, Dorso C, Lanata JL. 2007. Fisher equation for anisotropic diffusion: Simulating South American human dispersals. *Phys Rev E* **76:** 031923. doi: 10.1103/PhysRevE.76.031923.
Masello JF, Quillfeldt P, Munimanda GK, Klauke N, Segelbacher G, Schaefer HM, Failla M, Cortes M, Moodley Y. 2011. The high Andes, gene flow and a stable hybrid zone shape the genetic structure of a wide-ranging South American parrot. *Front Zool* **8:** 16. doi: 10.1186/1742-9994-8-16.
McCulloch RD, Bentley MJ, Purves RS, Hulton NRJ, Sugden DE, Clapperton CM. 2000. Climatic inferences from glacial and paleoecological

evidence at the last glacial termination, southern South America. *J Quaternary Sci* **15:** 409–417.
Mendizabal I, Sandoval K, Berniell-Lee G, Calafell F, Salas A, Martínez-Fuentes A, Comas D. 2008. Genetic origin, admixture, and asymmetry in maternal and paternal human lineages in Cuba. *BMC Evol Biol* **8:** 213. doi: 10.1186/1471-2148-8-213.
Moraga ML, Rocco P, Miquel JF, Nervi F, Llop E, Chakraborty R, Rothhammer F, Carvallo P. 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.
Mulligan CJ, Kitchen A, Miyamoto MM. 2008. Updated three-stage model for the peopling of the Americas. *PLoS ONE* **3:** e3199. doi: 10.1371/journal.pone.0003199.
O'Rourke DH, Raff JA. 2010. The human genetic history of the Americas: The final frontier. *Curr Biol* **20:** R202–R207.
Ottoni C, Primativo G, Hooshiar Kashani B, Achilli A, Martínez-Labarga C, Biondi G, Torroni A, Rickards O. 2010. Mitochondrial haplogroup H1 in North Africa: An early holocene arrival from Iberia. *PLoS ONE* **5:** e13378. doi: 10.1371/journal.pone.0013378.
Perego UA, Achilli A, Angerhofer N, Accetturo M, Pala M, Olivieri A, Kashani BH, Ritchie KH, Scozzari R, Kong QP, et al. 2009. Distinctive Paleo-Indian migration routes from Beringia marked by two rare mtDNA haplogroups. *Curr Biol* **19:** 1–8.
Perego UA, Angerhofer N, Pala M, Olivieri A, Lancioni H, Hooshiar Kashani B, Carossa V, Ekins JE, Gómez-Carballa A, Huber G, et al. 2010. The initial peopling of the Americas: A growing number of founding mitochondrial genomes from Beringia. *Genome Res* **20:** 1174–1179.
Prieto L, Zimmermann B, Goios A, Rodriguez-Monge A, Paneto GG, Alves C, Alonso A, Fridman C, Cardoso S, Lima G, et al. 2011. The GHEP-EMPOP collaboration on mtDNA population data—a new resource for forensic casework. *Forensic Sci Int Genet* **5:** 146–151.
Pucciarelli HM, Neves WA, González-José R, Sardi ML, Rozzi FR, Struck A, Bonilla MY. 2006. East-West cranial differentiation in pre-Columbian human populations of South America. *Homo* **57:** 133–150.
Raff J, Tackney J, O'Rourke DH. 2010. South from Alaska: A pilot aDNA study of genetic history on the Alaska Peninsula and the Eastern Aleutians. *Hum Biol* **82:** 677–693.
Rasmussen M, Li Y, Lindgreen S, Pedersen JS, Albrechtsen A, Moltke I, Metspalu M, Metspalu E, Kivisild T, Gupta R, et al. 2010. Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature* **463:** 757–762.
Ray N, Wegmann D, Fagundes NJR, Wang S, Ruiz-Linares A, Excoffier L. 2010. A statistical evaluation of models for the initial settlement of the American continent emphasizes the importance of gene flow with Asia. *Mol Biol Evol* **27:** 337–345.
Reidla M, Kivisild T, Metspalu E, Kaldma K, Tambets K, Tolk HV, Parik J, Loogväli EL, Derenko M, Malyarchuck B, et al. 2003. Origin and diffusion of mtDNA haplogroup X. *Am J Hum Genet* **73:** 1178–1190.
Rothhammer F, Dillehay TD. 2009. The late Pleistocene colonization of South America: An interdisciplinary perspective. *Ann Hum Genet* **73:** 540–549.
Rothhammer F, Llop E, Carvallo P, Moraga M. 2001. Origin and evolutionary relationships of native Andean populations. *High Alt Med Biol* **2:** 227–233.
Rubicz R, Melton PE, Spitsyn V, Sun G, Deka R, Crawford MH. 2010. Genetic structure of native circumpolar populations based on autosomal, mitochondrial, and Y chromosome DNA markers. *Am J Phys Anthropol* **143:** 62–74.
Sala A, Penacino G, Carnese R, Corach D. 1999. Reference database of hypervariable genetic markers of Argentina: application for molecular anthropology and forensic casework. *Electrophoresis* **20:** 1733–1739.
Salas A, Richards M, Lareu MV, Scozzari R, Coppa A, Torroni A, Macaulay V, Carracedo Á. 2004. The African diaspora: Mitochondrial DNA and the Atlantic slave trade. *Am J Hum Genet* **74:** 454–465.
Salas A, Jaime JC, Álvarez-Iglesias V, Carracedo Á. 2008. Gender bias in the multiethnic genetic composition of central Argentina. *J Hum Genet* **53:** 662–674.
Sans M, Figueiro G, Ackermann E, Barreto I, Egaña A, Bertoni B, Poittevin-Gilmet E, Maytia D, Hidalgo PC. 2011. Mitochondrial DNA in Basque descendants from the city of Trinidad, Uruguay: Uruguayan- or Basque-like population? *Hum Biol* **83:** 55–70.
Schroeder KB, Schurr TG, Long JC, Rosenberg NA, Crawford MH, Tarskaia LA, Osipova LP, Zhadanov SI, Smith DG. 2007. A private allele ubiquitous in the Americas. *Biol Lett* **3:** 218–223.
Schurr TG. 2004. The peopling of the New World: Perspectives from molecular anthropology. *Annu Rev Anthropol* **33:** 551–583.
Schurr TG, Ballinger SW, Gan YY, Hodge JA, Merriwether DA, Lawrence DN, Knowler WC, Weiss KM, Wallace DC. 1990. Amerindian mitochondrial DNAs have rare Asian mutations at high frequencies, suggesting they derived from four primary maternal lineages. *Am J Hum Genet* **46:** 613–623.
Schurr TG, Sukernik RI, Starikovskaya YB, Wallace DC. 1999. Mitochondrial DNA variation in Koryaks and Itel'men: Population replacement in the Okhotsk Sea–Bering Sea region during the Neolithic. *Am J Phys Anthropol* **108:** 1–39.
Scozzari R, Cruciani F, Santolamazza P, Sellitto D, Cole DE, Rubin LA, Labuda D, Marini E, Succa V, Vona G, et al. 1997. MtDNA and Y chromosome-specific polymorphisms in modern Ojibwa: Implications about the origin of their gene pool. *Am J Hum Genet* **60:** 241–244.
Soares P, Ermini L, Thomson N, Mormina M, Rito T, Röhl A, Salas A, Oppenheimer S, Macaulay V, Richards MB. 2009. Correcting for purifying selection: An improved human mitochondrial molecular clock. *Am J Hum Genet* **84:** 740–759.
Surovell TA. 2003. Simulating coastal migration in New World colonization. *Curr Anthropol* **44:** 580–591.
Tajima A, Hamaguchi K, Terao H, Oribe A, Perrotta VM, Baez CA, Arias JR, Yoshimatsu H, Sakata T, Horai S. 2004. Genetic background of people in the Dominican Republic with or without obese type 2 diabetes revealed by mitochondrial DNA polymorphism. *J Hum Genet* **49:** 495–499.
Tamm E, Kivisild T, Reidla M, Metspalu M, Smith DG, Mulligan CJ, Bravi CM, Rickards O, Martinez-Labarga C, Khusnutdinova EK, et al. 2007. Beringian standstill and spread of Native American founders. *PLoS ONE* **2:** e829. doi: 10.1371/journal.pone.0000829.
Tarazona-Santos E, Carvalho-Silva DR, Pettener D, Luiselli D, De 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.
Torroni A, Schurr TG, Cabell MF, Brown MD, Neel JV, Larsen M, Smith DG, Vullo CM, Wallace DC. 1993. Asian affinities and continental radiation of the four founding Native American mtDNAs. *Am J Hum Genet* **53:** 563–590.
van Oven M, Kayser M. 2009. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat* **30:** E386–E394.
Vasil'ev SA, Kuzmin YV, Orlova LA, Dementiev VN. 2002. Radiocarbon-based chronology of the Paleolithic in Siberia and its relevance to the peopling of the New World. *Radiocarbon* **44:** 503–530.
Volodko NV, Starikovskaya EB, Mazunin IO, Eltsov NP, Naidenko PV, Wallace DC, Sukernik RI. 2008. Mitochondrial genome diversity in arctic Siberians, with particular reference to the evolutionary history of Beringia and pleistocenic peopling of the Americas. *Am J Hum Genet* **82:** 1084–1100.
Wang S, Lewis CM Jr, Jakobsson M, Ramachandran S, Ray N, Bedoya G, Rojas W, Parra MV, Molina JA, Gallo C, et al. 2007. Genetic variation and population structure in Native Americans. *PLoS Genet* **3:** e185. doi: 10.1371/journal.pgen.0030185.
Waters MR, Forman SL, Jennings TA, Nordt LC, Driese SG, Feinberg JM, Keene JL, Halligan J, Lindquist A, Pierson J, et al. 2011. The Buttermilk Creek complex and the origins of Clovis at the Debra L. Friedkin site, Texas. *Science* **331:** 1599–1603.
Yang Z. 2007. PAML 4: A program package for phylogenetic analysis by maximum likelihood. *Mol Biol Evol* **24:** 1586–1591.
Yang NN, Mazieres S, Bravi C, Ray N, Wang S, Burley MW, Bedoya G, Rojas W, Parra MV, Molina JA, et al. 2010. Contrasting patterns of nuclear and mtDNA diversity in Native American populations. *Ann Hum Genet* **74:** 525–538.
Zlojutro M, Rubicz R, Devor EJ, Spitsyn VA, Makarov SV, Wilson K, Crawford MH. 2006. Genetic structure of the Aleuts and circumpolar populations based on mitochondrial DNA sequences: A synthesis. *Am J Phys Anthropol* **129:** 446–464.
Zlojutro M, Rubicz R, Crawford MH. 2009. Mitochondrial DNA and Y-chromosome variation in five eastern Aleut communities: Evidence for genetic substructure in the Aleut population. *Ann Hum Biol* **36:** 511–526.

Received September 8, 2011; accepted in revised form January 20, 2012.