



# University of HUDDERSFIELD

## University of Huddersfield Repository

Pereira, Luísa, Richards, Martin B., Goios, Ana, Alonso, Antonio, Albarrán, Cristina, Garcia, Oscar, Behar, Doron M., Gölge, Mukaddes, Hatina, Jiří, Al-Gazali, Lihadh, Bradley, Daniel G., Macaulay, Vincent and Amorim, António

High-resolution mtDNA evidence for the late-glacial resettlement of Europe from an Iberian refugium

### Original Citation

Pereira, Luísa, Richards, Martin B., Goios, Ana, Alonso, Antonio, Albarrán, Cristina, Garcia, Oscar, Behar, Doron M., Gölge, Mukaddes, Hatina, Jiří, Al-Gazali, Lihadh, Bradley, Daniel G., Macaulay, Vincent and Amorim, António (2005) High-resolution mtDNA evidence for the late-glacial resettlement of Europe from an Iberian refugium. *Genome Research*, 15 (1). pp. 19-24. ISSN 1088-9051

This version is available at <http://eprints.hud.ac.uk/id/eprint/21358/>

The University Repository is a digital collection of the research output of the University, available on Open Access. Copyright and Moral Rights for the items on this site are retained by the individual author and/or other copyright owners. Users may access full items free of charge; copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational or not-for-profit purposes without prior permission or charge, provided:

- The authors, title and full bibliographic details is credited in any copy;
- A hyperlink and/or URL is included for the original metadata page; and
- The content is not changed in any way.

For more information, including our policy and submission procedure, please contact the Repository Team at: [E.mailbox@hud.ac.uk](mailto:E.mailbox@hud.ac.uk).

<http://eprints.hud.ac.uk/>



## High-resolution mtDNA evidence for the late-glacial resettlement of Europe from an Iberian refugium

Luísa Pereira, Martin Richards, Ana Goios, et al.

*Genome Res.* 2005 15: 19-24

Access the most recent version at doi:[10.1101/gr.3182305](https://doi.org/10.1101/gr.3182305)

---

**References** This article cites 21 articles, 5 of which can be accessed free at:  
<http://genome.cshlp.org/content/15/1/19.full.html#ref-list-1>

**Creative Commons License** This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <http://genome.cshlp.org/site/misc/terms.xhtml>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 3.0 Unported License), as described at <http://creativecommons.org/licenses/by-nc/3.0/>.

**Email Alerting Service** Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

---

The advertisement features the Gene Link logo on the left, which consists of three green diamond shapes. To the right is a photograph of a firefly. The text reads: 'Fluorescent Molecular Probes' in large white font, followed by 'Gene Link synthesizes fluorescent DNA, RNA or chimeric probes with complex modifications for duplex stability & nuclease resistance.' in smaller white font.

---

To subscribe to *Genome Research* go to:  
<http://genome.cshlp.org/subscriptions>

---

# High-resolution mtDNA evidence for the late-glacial resettlement of Europe from an Iberian refugium

Luísa Pereira,<sup>1</sup> Martin Richards,<sup>2</sup> Ana Goios,<sup>1</sup> Antonio Alonso,<sup>3</sup> Cristina Albarrán,<sup>3</sup> Oscar Garcia,<sup>4</sup> Doron M. Behar,<sup>5</sup> Mukaddes Gölge,<sup>6</sup> Jiří Hatina,<sup>7</sup> Lihadh Al-Gazali,<sup>8</sup> Daniel G. Bradley,<sup>9</sup> Vincent Macaulay,<sup>10,12</sup> and António Amorim<sup>1,11</sup>

<sup>1</sup>IPATIMUP (Instituto de Patologia e Imunologia Molecular da Universidade do Porto), 4200-465 Porto, Portugal; <sup>2</sup>Schools of Biology and Computing, University of Leeds, Leeds, LS2 9JT, United Kingdom; <sup>3</sup>Instituto de Toxicología, Sección de Biología, 28002 Madrid, Spain; <sup>4</sup>Area de Laboratorio Ertzaintza, Gobierno Vasco, 48950 Bilbao, Spain; <sup>5</sup>Bruce Rappaport Faculty of Medicine and Research Institute, Technion and Rambam Medical Center, Haifa 31096, Israel; <sup>6</sup>Department of Physiology, University of Kiel, 24118 Kiel, Germany; <sup>7</sup>Charles University, Medical Faculty in Pilsen, Institute of Biology, CZ-301 66 Pilsen, Czech Republic; <sup>8</sup>Department of Paediatrics, Faculty of Medicine and Health Sciences, UAE University, Dubai; <sup>9</sup>Smurfit Institute of Genetics, Trinity College, Dublin 2, Ireland; <sup>10</sup>Department of Statistics, University of Glasgow, Glasgow G12 8QQ, United Kingdom; <sup>11</sup>Faculdade de Ciências da Universidade do Porto, 4099-002 Porto, Portugal

The advent of complete mitochondrial DNA (mtDNA) sequence data has ushered in a new phase of human evolutionary studies. Even quite limited volumes of complete mtDNA sequence data can now be used to identify the critical polymorphisms that define sub-clades within an mtDNA haplogroup, providing a springboard for large-scale high-resolution screening of human mtDNAs. This strategy has in the past been applied to mtDNA haplogroup V, which represents <5% of European mtDNAs. Here we adopted a similar approach to haplogroup H, by far the most common European haplogroup, which at lower resolution displayed a rather uninformative frequency distribution within Europe. Using polymorphism information derived from the growing complete mtDNA sequence database, we sequenced 1580 base pairs of targeted coding-region segments of the mtDNA genome in 649 individuals harboring mtDNA haplogroup H from populations throughout Europe, the Caucasus, and the Near East. The enhanced genealogical resolution clearly shows that sub-clades of haplogroup H have highly distinctive geographical distributions. The patterns of frequency and diversity suggest that haplogroup H entered Europe from the Near East ~20,000–25,000 years ago, around the time of the Last Glacial Maximum (LGM), and some sub-clades re-expanded from an Iberian refugium when the glaciers retreated ~15,000 years ago. This shows that a large fraction of the maternal ancestry of modern Europeans traces back to the expansion of hunter-gatherer populations at the end of the last Ice Age.

[The sequence data from this study have been submitted to GenBank under accession nos. AY776364–AY778959.]

Haplogroup H accounts for 40%–50% of the mtDNA pool in most of Europe, and ~20%–30% in the Near East and the Caucasus region (Richards et al. 2000; Achilli et al. 2004; Loogväli et al. 2004). It is thought to have evolved in the vicinity of the Near East ~23,000–28,000 years ago, and to have spread into Europe ~20,000 years ago. Founder analysis, based on control-region HVS-I sequences and limited restriction typing of coding-region markers in about 3000 samples, suggested that, in a manner similar to haplogroup V, some or all of European haplogroup H may then have re-expanded from a European glacial refuge ~15,000 years ago (Torrioni et al. 1998, 2001; Richards et al. 2000). However, the phylogenetic resolution of haplogroup H mitochondrial DNAs (mtDNAs) identified by control-region sequencing is poor, and the haplogroup as a whole does not display the marked frequency gradient present in haplogroup V. This meant that the use of haplogroup H as a whole to locate the likely refuge(s) was not possible.

We therefore decided upon a screening strategy based on

complete mtDNA sequence information. Recent complete sequence data have indicated a number of highly informative coding-region polymorphisms that resolve distinct sub-clades of haplogroup H. Finnilä et al. (2001) defined two sub-haplogroups, H1 and H2, based, respectively, on the polymorphisms G3010A and A4769G (along with A1438G) in a sample of Finns. Herrnstadt et al. (2002), studying a large sample of haplogroup H individuals from the UK and US, described two further sub-groups, H3 and H4. H3 was the next most common sub-haplogroup after H1, characterized by the polymorphism T6776C. H4 was rarer, with eight polymorphisms (including C3992T and A4024G) separating it from the root of the haplogroup. Quintáns et al. (2004) further identified H5, defined by T4336C, H6, defined by G3915A (and a further sub-branch by A4727G), and H7, defined by A4793G, and additional sub-clades were recently proposed by Loogväli et al. (2004) and Achilli et al. (2004).

## Results

### Distribution of the major sub-clades within haplogroup H

We sequenced 1580 base pairs of coding region in 649 samples belonging to haplogroup H from 20 populations from Europe,

<sup>12</sup>Corresponding author.

E-mail [vincent@stats.gla.ac.uk](mailto:vincent@stats.gla.ac.uk); fax 44 141 330 4814.

Article and publication are at <http://www.genome.org/cgi/doi/10.1101/gr.3182305>.

the Caucasus, and the Near East (Table 1) and combined them with published data (Finnilä et al. 2001; Herrnstadt et al. 2002). A phylogenetic network for the variation scored in the 894 haplogroup H mtDNAs is shown in Figure 1. In addition to the seven major clades defined previously, we identified a further minor sub-haplogroup defined by A4745G (recently labeled as H13 by Achilli et al. 2004). The two most frequent sub-haplogroups, H1 and H3, each show a rather star-like phylogeny. We refer to the paraphyletic collection of H mtDNAs outside these eight main sub-clades as H\*.

The frequencies of haplogroup H as a whole, and its sub-haplogroups, are reported in Table 1 for the 22 populations analyzed here, with age estimates in Table 2. The majority of the European populations have an overall haplogroup H frequency of 40%–50%. Frequencies decrease in the southeast of the continent, reaching ~20% in the Near East and Caucasus, and <10% in the Gulf (Fig. 2A). Thus, haplogroup H as a whole displays a broadly southeast-northwest frequency pattern, reminiscent of the first principal component of classical marker frequencies (Cavalli-Sforza et al. 1994). However, genealogical dissection into sub-clades reveals a quite different sub-structure, showing that this overall pattern is something of a chimera.

The distribution of H1, the largest sub-clade, displays two peaks, one in Iberia and another in Scandinavia (Fig. 2B). However, the Norwegian sample size is low ( $n = 18$ ) and haplogroup H is overrepresented (~70%, while larger data sets for Norway point to a frequency of ~50%: Richards et al. 2000). When we removed the Norwegian sample, the Scandinavian peak disappeared, and the picture showed only the decreasing frequency of sub-haplogroup H1 from the southwest to the north and east. H1 is almost exclusively European, with its only incursion into the Near East being a few Palestinian individuals bearing the most common haplotype. This absence of derived lineages in the Near

East sample suggests that the H1 sub-clade had its origin in Europe. H1 has an age of ~14,000 years (SE 4000) using coding-region data and ~16,000 years (SE 3500) using HVS-I. No significant difference between its diversity in western and eastern Europe was manifest.

The distribution of the second most frequent sub-clade, H3 (Fig. 2C), shows a very similar pattern, again suggesting a European origin. The frequency difference between west and east is highly significant ( $\chi^2 = 28.2$ ;  $P < 0.000001$ ), as it is also for H1 ( $\chi^2 = 137.1$ ;  $P < 0.000001$ ). H3 is exclusively European, with no Near Eastern representatives, and is ~9000 years old (SE 3000) based on the coding-region data and ~11,000 years old (SE 3000) using HVS-I.

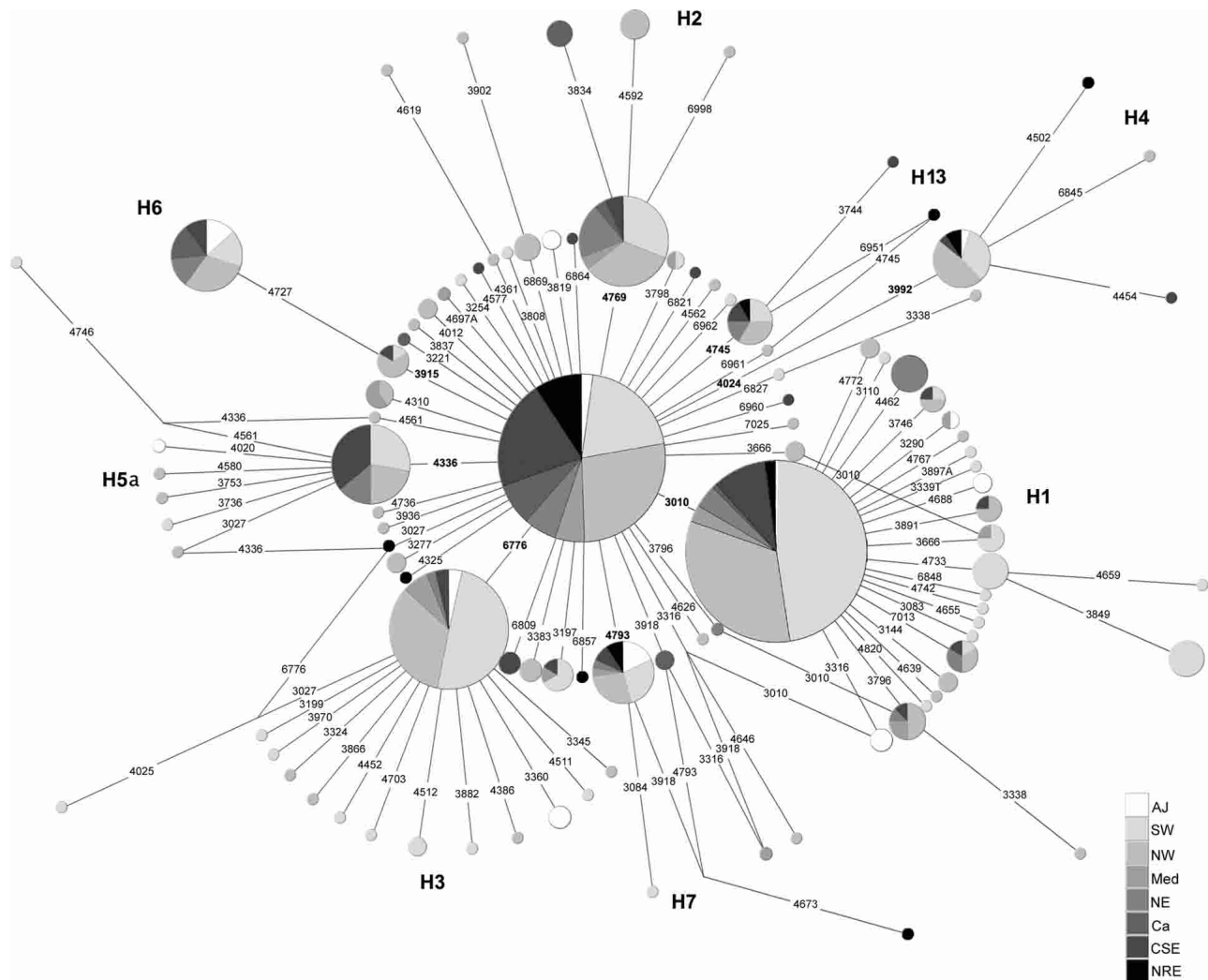
### Minor sub-clades within haplogroup H

The remaining sub-clades occur at low frequency, and it is difficult to detect any geographical patterns (Fig. 2D). Within haplogroup H HVS-I lineages, the most frequent sub-clade (~4% of Europeans) is defined by T16304C (Richards et al. 2000). This sub-clade encompasses all of H5 and a fraction of H\* lineages, indicating that the T16304C transition may have happened only once within haplogroup H (although see Loogväli et al. 2004) and occurred before the H5-defining transition at np 4336. Thus sub-haplogroup H5 can be broadened to include the 16304 transition, as suggested by Loogväli et al. (2004), within which T4336C defines a further sub-clade, H5a. The frequency of H5a appears to be highest on the central European plain (Table 1), and dates to ~7000–8000 years (Table 2). It is fairly evenly distributed at low levels across Europe but is absent from the Caucasus and the Near East, again suggesting a European origin. In contrast, the H5 clade is present at low levels (1%–3%) throughout the Near East and may have evolved there, spreading later

**Table 1.** Distribution of H sub-haplogroups

Population	<i>n</i>	Number and proportion of H mtDNAs (%)	No. of H mtDNAs sub-typed	Reference	Frequency of H sub-haplogroups									
					H1	H2	H3	H4	H5a	H6	H7	H13	H*	
Portugal	549	225 (41.0)	194	Pereira et al. 2004	0.188	0.016	0.078	0.009	0.013	0.013	0.004	0.011	0.078	
Spain (Madrid)	119	58 (48.7)	36	Present study	0.269	0.025	0.017	0.017	0.042	0.000	0.000	0.025	0.092	
Basque country	201	101 (50.2)	75	Present study	0.254	0.060	0.080	0.015	0.025	0.000	0.005	0.000	0.063	
France	72	26 (36.1)	23	Richards et al. 2000	0.139	0.000	0.028	0.000	0.000	0.014	0.000	0.000	0.180	
Ireland	102	43 (42.2)	32	Richards et al. 2000	0.147	0.000	0.069	0.029	0.010	0.039	0.010	0.010	0.108	
UK/US	561	214 (38.1)	214	Herrnstadt et al. 2002	0.128	0.029	0.043	0.016	0.018	0.016	0.009	0.004	0.119	
Italy	50	11 (22.0)	9	Richards et al. 2000	0.080	0.000	0.080	0.000	0.000	0.000	0.000	0.000	0.060	
Sardinia	48	22 (45.8)	12	Richards et al. 2000	0.188	0.042	0.083	0.000	0.000	0.000	0.000	0.042	0.103	
Crete	41	17 (41.5)	17	Richards et al. 2000	0.049	0.024	0.024	0.000	0.000	0.000	0.000	0.000	0.318	
Romania	94	32 (34.0)	30	Richards et al. 2000	0.032	0.000	0.000	0.000	0.074	0.000	0.011	0.000	0.223	
Czech Republic	89	35 (39.3)	31	Richards et al. 2000	0.157	0.000	0.022	0.000	0.034	0.034	0.000	0.000	0.146	
Bulgaria	111	44 (39.6)	44	Richards et al. 2000	0.099	0.027	0.009	0.009	0.009	0.009	0.018	0.018	0.198	
Poland	38	19 (50.0)	8	Richards et al. 2000	0.132	0.000	0.000	0.053	0.132	0.000	0.000	0.000	0.183	
Russia	57	18 (31.6)	17	Richards et al. 2000	0.105	0.018	0.000	0.000	0.018	0.000	0.018	0.000	0.157	
Norway	18	13 (72.2)	11	Richards et al. 2000	0.444	0.000	0.056	0.000	0.000	0.000	0.056	0.000	0.166	
Finland	480	188 (39.2)	31	Finnilä et al. 2001	0.177	0.050	0.025	0.000	0.038	0.000	0.013	0.000	0.089	
Palestine	110	29 (26.4)	19	Richards et al. 2000	0.069	0.000	0.000	0.014	0.000	0.000	0.014	0.028	0.139	
Kurd	79	11 (13.9)	9	Richards et al. 2000	0.000	0.000	0.000	0.013	0.000	0.000	0.000	0.000	0.126	
Caucasus	213	47 (22.1)	36	Richards et al. 2000	0.014	0.038	0.000	0.000	0.000	0.033	0.000	0.014	0.122	
Chuvash, Russia	36	11 (30.6)	11	Richards et al. 2000	0.028	0.111	0.000	0.000	0.028	0.111	0.000	0.028	0.000	
Gulf states	75	6 (8.0)	6	Richards et al. 2000	0.000	0.000	0.000	0.013	0.000	0.000	0.000	0.000	0.067	
Ashkenazi Jews	565	119 (21.0)	29	Behar et al. 2004	0.051	0.000	0.044	0.007	0.000	0.028	0.000	0.028	0.052	
Total	3708	1289	894											

Three sequences from Ingman et al. (2000), included in some analyses, are excluded here.



**Figure 1.** Reduced-median network (Bandelt et al. 1995) for coding-region polymorphisms found in 894 samples belonging to haplogroup H, in the four segments described in the text. Mutations that define the H sub-haplogroups are shown in bold, and sub-clades are labeled. The area of each circle is proportional to the number of mtDNAs in the total sample harboring the corresponding haplotype. Geographic regions are as follows: AJ, Ashkenazi Jews; SW, Iberia; NW, France, Ireland, UK/US sample, Norway; Med, mainland Italy, Sardinia, Crete; NE, Finland, Russia, Chuvash; Ca, North Caucasus; CSE, Poland, Czech Republic, Romania, Bulgaria; NRE, Near East. Further details of samples are given in Table 1.

into Europe. Its age based on HVS-I variation is 11,500 (SE 2700) years, and its ancestor was identified as a putative late-glacial founder type by Richards et al. (2000). However, the HVS-I database indicates that it is common (>4%) not only in Iberia but also in central, eastern, and southeast Europe, and rather less frequent in northwest Europe.

In contrast, H2 and H6 are both common in eastern Europe and the Caucasus, although there are hints that they may have dispersed from western Europe. In particular, the basal type of H6 is exclusively European, and there is a single derived type that is common in eastern Europe and the Caucasus. Neither H2 nor H6 are found in our Near Eastern sample. The infrequent sub-clades H4, H7, and H13 occur in both Europe and the Near East, and the latter is also present in the Caucasus.

### Origins of haplogroup H

The paraphyletic ancestral cluster, H\*, is the main Near Eastern representative of haplogroup H, in agreement with the suggestion that the haplogroup evolved in the Near East and spread

subsequently into Europe. Its distribution is to some extent the inversion of the distributions for H1 and H3: It is most frequent in east-central Europe and the Balkans, but is also well represented on the western fringes of Europe, including Iberia and Ireland. The age of H\* is best estimated as the age of the haplogroup as a whole, which comes to 29,900 (SE 7700) years using the present coding-region data set (excluding the 3010 variant which renders the tree very non-star-like). Using the complete coding-region sequence data of Finnilä et al. (2001) and Herrstadt et al. (2002), the age estimate of H (including 3010) is rather less, at 17,600 (SE 2200) years. This may be because no Near Eastern lineages are included, or it may simply reflect the high uncertainty of the estimate from our coding-region segments.

### Discussion

It seems likely, on the basis of this evidence, that haplogroup H entered Europe not much more than ~20,000–25,000 years ago,

**Table 2.** Ages for subclades of haplogroup H based on coding region (1580bp) and HV5-I (nps 16090–16365) variation (excluding UK/US data)

Haplogroup (sample size)	Age in years (SE)		Haplogroup (sample size)	Age in years (SE)	
	Coding region	HVS-I		Coding region	HVS-I
H1 (n = 247)	14,000 (4,000)	16,000 (3,500)	H4 (n = 15)	7,000 (5,000)	16,000 (6,000)
H1 Western Europe (n = 195)	14,000 (5,000)	16,000 (3,500)	H5a (n = 32)	8,000 (4,000)	7,000 (2,000)
H1 Iberia (n = 148)	13,000 (6,000)	14,000 (3,000)	H6 (n = 27)	43,000 (43,000)	16,000 (5,000)
H1 Eastern Europe (n = 52)	15,000 (6,000)	18,000 (5,000)	H7 (n = 19)	8,000 (4,000)	16,000 (5,000)
H2 (n = 39)	11,600 (8,000)	21,000 (9,000)	H13 (n = 12)	8,000 (6,000)	18,000 (6,000)
H3 (n = 75)	9,000 (3,000)	11,000 (3,000)	All H less H1 and H3 (n = 360)	33,000 (10,000)	23,000 (4,000)

and dispersed rapidly to the southwest of the continent. Although this was at the peak of the last Ice Age, a passage into Europe at this time is not implausible from an archaeological perspective, since there is evidence for extensive contacts between people of the Badegoulian culture of east-central Europe and those of southwest Europe. Indeed, it now seems likely that the west European Magdalenian culture had its roots in the Badegoulian, and not in the local Solutrean of the western glacial refugium. It is the Magdalenian culture that is seen to expand dramatically from the Iberian refugium from ~15,000 years ago in the radiocarbon record for western Europe, although Europe was probably never completely depopulated during the LGM (Housley et al. 1997; Terberger and Street 2002; Gamble et al. 2004).

Haplogroup V was identified, on the basis of control-region sequences, as a likely marker of a human dispersal in Late Pleistocene Europe (Torrioni et al. 1998). Higher phylogenetic resolution of the lineages concerned clarified the geographic pattern by distinguishing the more derived haplogroup V from its ancestor, pre-V, which could now be seen to display a quite distinct phylogeographic pattern (Torrioni et al. 2001). Haplogroup pre-V appeared to have entered Europe from the east sometime around 20,000–25,000 years ago, at the time of the LGM. However, the diversity and frequency of the derived haplogroup V suggested that it had evolved from pre-V in western Europe, with its age suggesting an expansion from a glacial refuge in Iberia ~15,000 years ago, accompanying the Magdalenian expansion.

It is clear that the phylogeographic patterns displayed by sub-haplogroups H1 and H3 both closely resemble that of haplogroup V. The star-like phylogenies, geographic distribution, and estimated ages of all three clades suggest that they all took part in a major expansion from southwest to northeast Europe ~12,000–14,000 years ago. Between them H1 and H3 amount to around half of the haplogroup H samples in our coding-region database. They comprise ~65% of haplogroup H lineages in Iberia, ~46% in the northwest, ~27% in central and eastern Europeans, and ~5%–15% in the Near East/Caucasus, falling to zero in the Gulf. It is notable that the diversity does not fall within H1 moving from west to east, unlike the situation with haplogroup V (Torrioni et al. 2001), but a rapid expansion within the time-frame of the Magdalenian would in fact not be expected to result in a west-east diversity gradient. The cline seen in haplogroup V

diversities most likely has its explanation in more recent founder events in the east.

The remaining haplogroup H lineages present a more complex pattern. The explanation must include the evolution of haplogroup H from its ancestor haplogroup HV, probably in the vicinity of the Near East (Richards et al. 2000; Loogväli et al. 2004), and subsequent founder events in Europe, seen in H\*. Minor sub-clades found in both Europe and the Near East (H4, H7, and H13) may also have entered Europe around the LGM, and/or during later dispersals from the Near East, such as the Neolithic. H must have given rise to H1 and H3 in the western refuge (analogous to ancestral lineages within haplogroup pre-V giving rise to haplogroup V; Torrioni et al. 2001), and itself appears very likely to have been partly redistributed alongside them by the late-glacial re-expansion, since an Atlantic European cluster clearly forms part of the H\* phylogeny. Several

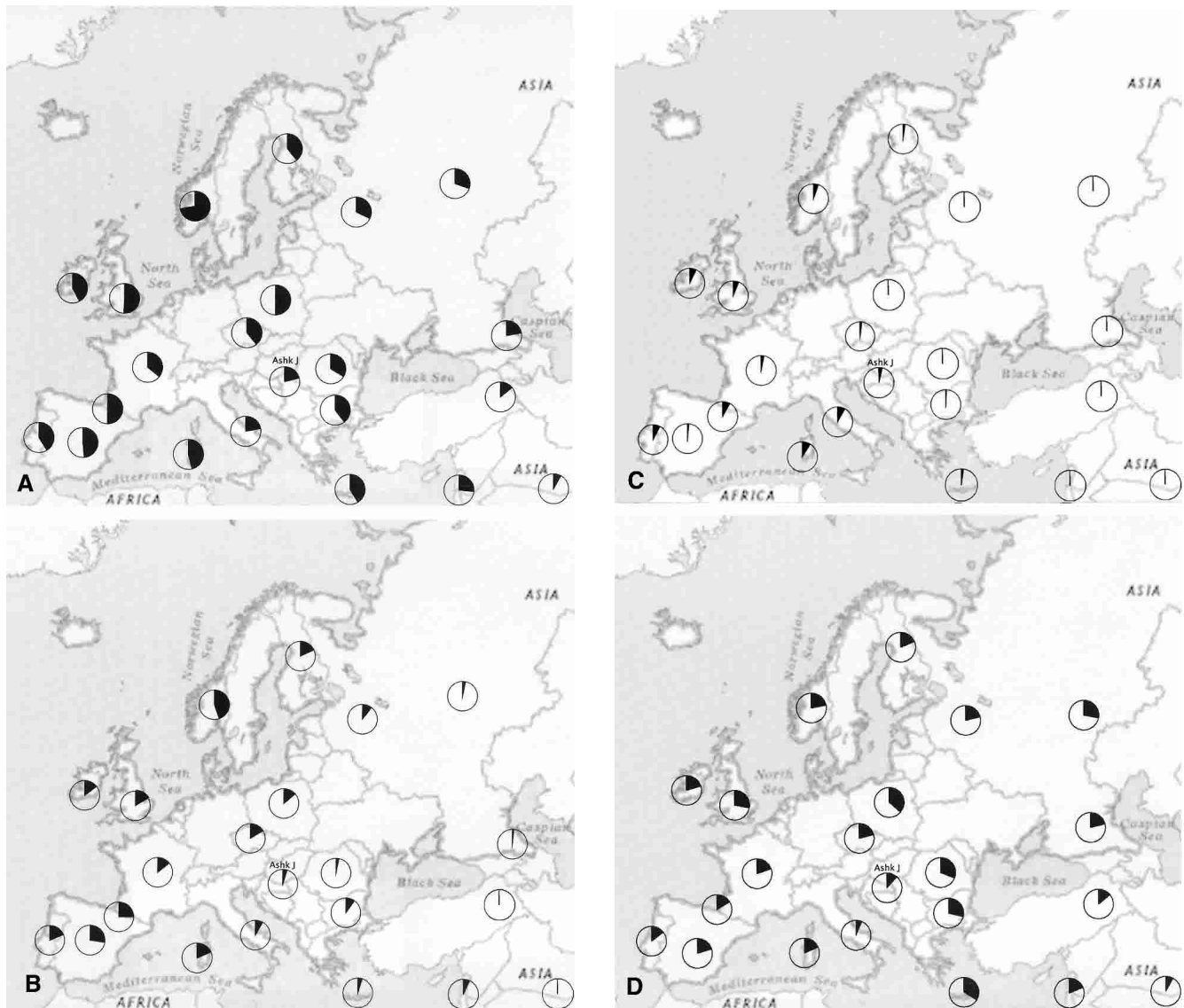
other minor sub-clades (H2, H5a, H6) also seem likely to have taken part in this process, and may also have evolved in western Europe: More data will be needed to trace their phylogeographic patterns more closely. Interestingly, however, the frequency profile of H5a suggests that, if indeed it has largely been distributed by late-glacial dispersals, this sub-haplogroup may trace a distinct dispersal route into central and eastern Europe. In contrast, H1 and H3 appear at least in part to have spread northwards fairly close to the Atlantic coastline, into the British Isles.

The mtDNA evidence therefore correlates well with Y-chromosome evidence for late-glacial expansions from a southwest European refugium (Semino et al. 2000; Rootsi et al. 2004). It indicates that the major demographic signal in the modern European mtDNA pool is the result of the expansion of hunter-gatherer populations at the end of the Palaeolithic, although this has not entirely erased the traces of earlier processes.

## Methods

### Samples and sequencing

We dissected haplogroup H variation in 649 samples from 20 populations from Europe, the Caucasus, and the Near East (see Table 1) previously analyzed only for HV5-I sequence variation and some haplogroup-diagnostic RFLPs. We sequenced four mtDNA coding-region segments encompassing the principal diagnostic positions in haplogroup H samples: 3001–3360, 3661–4050, 4281–4820, and 6761–7050 (a total of 1580 base pairs) (Andrews et al. 1999). Primers used were, respectively: L2978, 5'-GTCCATATCAACAATAGGGT-3' and H3361 5'-CGTTCGGTA AGCATTAGGAA-3'; L3640, 5'-TCTAGCCACCTCTAGCCTAG-3' and H4051 5'-TAGAGTTCAGGGGAGAGTGC-3'; L4264, 5'-CATTCCCCTCAAACCTAAG-3' and H4821 5'-AGAGGGGTGC CTTGGGTAAC-3'; L6740, 5'-TGGTCTGAGCTATGATATCA-3' and H7051 5'-GATGGCAAATACAGCTCCTA-3'. The temperature profiles for the PCR were: 95°C for 10 sec, 64°C for 30 sec, and 72°C for 30 sec, for 35 cycles, for the third pair of primers, and the same except 58°C as annealing temperature for the others. We carried out automated sequencing in an ABI 3100, using the Kit Big-Dye Terminator Cycle Sequencing Ready Reaction (AB Applied Biosystems).



**Figure 2.** Frequency distributions of haplogroups H (A), H1 (B), H3 (C), and H less H1 and H3 (D) in Europe, the Caucasus and the Near East.

### Genetic analysis

Including 31 complete sequences from Finland (Finnilä et al. 2001), as well as (for the phylogenetic analyses) the US/UK coding-region data of Herrnstadt et al. (2002), we analyzed a total of 894 haplogroup H mtDNAs for these coding-region segments. The three haplogroup H sequences from Ingman et al. (2000) were also included in some analyses. The inclusion of both the control-region and 1580-base pair coding-region segments in the majority of individuals in our database allowed us to estimate clade ages using the  $\rho$  statistic (Saillard et al. 2000) in two ways, using a calibration of 1 transition per 20,180 years for HVS-I (Forster et al. 1996) and 1 substitution per 50,200 years ( $\mu = 1.26 \times 10^{-8}$  substitutions per year per base; Mishmar et al. 2003) for the coding region. We constructed reduced-median networks (Bandelt et al. 1995) separately for the coding-region segments and HVS-I (between positions 16090–16365) and estimated ages of clades using the program “Network” (Shareware Phylogenetic Network Software, version 4.0).

### Accession numbers

The new sequences generated for this work have been deposited in GenBank, accession nos. AY776364–AY778959. The published complete mtDNA sequences used in this analysis from Finnilä et al. (2001) are available in GenBank, accession nos. AY339402–AY339432. Coding-region sequences from Herrnstadt et al. (2002) are available at <http://www.mitokor.com/science/560mtdnas.php>

### Acknowledgments

We thank Sergei Rychkov, Oksana Rychkov, Pierre-Marie Danze, Dimitar Dimitrov, Peter and Anna Katharina Forster, Ariella Oppenheim, and Gheorghe Stefanescu for collecting and extracting DNA samples. This work was partially supported by a research grant to L.P. (SFRH/BPD/7121/2001) from Fundação para a Ciência e a Tecnologia and IPATIMUP by Programa Operacional Ciên-

cia, Tecnologia e Inovação (POCTI), Quadro Comunitário de Apoio III. This study is included in the Project POCTI/ANT/45139/2002 financed by Fundação para a Ciência e a Tecnologia (Eixo 2, Medida 2.3 do POCTI, QCA III).

## References

- Achilli, A., Rengo, C., Magri, C., Battaglia, V., Olivieri, A., Scozzari, R., Cruciani, F., Zeviani, M., Briem, E., Carelli, V., et al. 2004. The molecular dissection of mtDNA haplogroup H confirms that the Franco-Cantabrian glacial refuge was a major source for the European gene pool. *Am. J. Hum. Genet.* **75**: 910–918.
- Andrews, R.M., Kubacka, I., Chinnery, P.F., Lightowlers, R.N., Turnbull, D.M., and Howell, N. 1999. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet.* **23**: 147.
- Bandelt, H.-J., Forster, P., Sykes, B., and Richards, M. 1995. Mitochondrial portraits of human populations using median networks. *Genetics* **141**: 743–753.
- Behar, D.M., Hammer, M.F., Garrigan, D., Villemes, R., Bonn -Tamir, B., Richards, M., Gurwitz, D., Rosengarten, D., Kaplan, M., Della Pergola, S., et al. 2004. mtDNA evidence for a genetic bottleneck in the early history of the Ashkenazi Jewish population. *Eur. J. Hum. Genet.* **12**: 355–364.
- Cavalli-Sforza, L.L., Menozzi, P., and Piazza, A. 1994. *The history and geography of human genes*. Princeton University Press, Princeton, NJ.
- Finnil , S., Lehtonen, M.S., and Majamaa, K. 2001. Phylogenetic network for European mtDNA. *Am. J. Hum. Genet.* **68**: 1475–1484.
- Forster, P., Harding, R., Torroni, A., and Bandelt, H.-J. 1996. Origin and evolution of Native American mtDNA variation: A reappraisal. *Am. J. Hum. Genet.* **59**: 935–945.
- Gamble, C., Davies, W., Pettitt, P., and Richards, M. 2004. Climate change and evolving human diversity in Europe during the last glacial. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **359**: 243–254.
- Hernstadt, C., Elson, J.L., Fahy, E., Preston, G., Turnbull, D.M., Anderson, C., Ghosh, S.S., Olefsky, J.M., Beal, M.F., Davis, R.E., et al. 2002. Reduced-median-network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian, and European haplogroups. *Am. J. Hum. Genet.* **70**: 1152–1171.
- Housley, R.A., Gamble, C.S., Street, M., and Pettitt, P. 1997. Radiocarbon evidence for the late glacial human recolonisation of northern Europe. *Proc. Prehist. Soc.* **63**: 25–54.
- Ingman, M., Kaessmann, H., P  bo, S., and Gyllensten, U. 2000. Mitochondrial genome variation and the origin of modern humans. *Nature* **408**: 708–713.
- Loogv li, E.-L., Roostalu, U., Malyarchuk, B.A., Derenko, M.V., Kivisild, T., Metspalu, E., Tambets, K., Reidla, M., Tolk, H.-V., Parik, J., et al. 2004. Disuniting uniformity: A pied cladistic canvas of mtDNA haplogroup H in Eurasia. *Mol. Biol. Evol.* **21**: 2012–2021.
- Mishmar, D., Ruiz-Pesini, E., Golik, P., Macaulay, V., Clark, A.G., Hosseini, S., Brandon, M., Easley, K., Chen, E., Brown, M.D., et al. 2003. Natural selection shaped regional mtDNA variation in humans. *Proc. Natl. Acad. Sci.* **100**: 171–176.
- Pereira, L., Cunha, C., and Amorim, A. 2004. Predicting sampling saturation of mtDNA haplotypes: An application to an enlarged Portuguese database. *Int. J. Legal Med.* **118**: 132–136.
- Quint ns, B., Alvarez-Iglesias, V., Salas, A., Phillips, C., Lareu, M.V., and Carracedo, A. 2004. Typing of mitochondrial DNA coding region SNPs of forensic and anthropological interest using SNaPshot minisequencing. *Forensic Sci. Int.* **140**: 251–257.
- Richards, M., Macaulay, V., Hickey, E., Vega, E., Sykes, B., Guida, V., Rengo, C., Sellitto, D., Cruciani, F., Kivisild, T., et al. 2000. Tracing European founder lineages in the Near Eastern mtDNA pool. *Am. J. Hum. Genet.* **67**: 1251–1276.
- Rootsi, S., Magri, C., Kivisild, T., Benuzzi, G., Help, H., Bermisheva, M., Kutuev, I., Barac, L., Pericic, M., Balanovsky, O., et al. 2004. Phylogeography of Y-chromosome haplogroup I reveals distinct domains of prehistoric gene flow in Europe. *Am. J. Hum. Genet.* **75**: 128–137.
- Saillard, J., Forster, P., Lynnerup, N., Bandelt, H.-J., and N rby, S. 2000. mtDNA variation among Greenland Eskimos: The edge of the Beringian expansion. *Am. J. Hum. Genet.* **67**: 718–726.
- Semino, O., Passarino, G., Oefner, P.F., Lin, A.A., Arbuzova, S., Beckman, L.E., De Benedictis, G., Francalacci, P., Kouvatusu, A., Limborska, S., et al. 2000. The genetic legacy of Paleolithic *Homo sapiens sapiens* in extant Europeans: A Y chromosome perspective. *Science* **290**: 1155–1159.
- Terberger, T. and Street, M. 2002. Hiatus or continuity? New results for the question of pleniglacial settlement in Central Europe. *Antiquity* **76**: 691–698.
- Torroni, A., Bandelt, H.-J., D'Urbano, L., Lahermo, P., Moral, P., Sellitto, D., Rengo, C., Forster, P., Savantaus, M.-L., Bonn -Tamir, B., et al. 1998. mtDNA analysis reveals a major late Paleolithic population expansion from southwestern to northeastern Europe. *Am. J. Hum. Genet.* **62**: 1137–1152.
- Torroni, A., Bandelt H.-J., Macaulay, V., Richards, M., Cruciani, F., Rengo, C., Martinez-Cabrera, V., Villemes, R., Kivisild, T., Metspalu, E., et al. 2001. A signal, from human mtDNA, of post-glacial recolonization in Europe. *Am. J. Hum. Genet.* **69**: 844–852.

Received August 23, 2004; accepted in revised form November 2, 2004.