Ann. Hum. Genet. (2000), **64**, 491–506
Printed in Great Britain

Diversity of mtDNA lineages in Portugal: not a genetic edge of European variation

L. PEREIRA[1,2], M. J. PRATA[1,2] AND A. AMORIM[1,2]
[1] *Instituto de Patologia e Imunologia Molecular da Universidade do Porto (IPATIMUP), R. Dr. Roberto Frias s/no, 4200 Porto, PORTUGAL*
[2] *Faculdade de Ciências da Universidade do Porto, Pr. Gomes Teixeira, 4050 Porto, PORTUGAL*

(Received 11.1.00. Accepted 26.7.00)

SUMMARY

The analysis of the hypervariable regions I and II of mitochondrial DNA in Portugal showed that this Iberian population presents a higher level of diversity than some neighbouring populations. The classification of the different sequences into haplogroups revealed the presence of all the most important European haplogroups, including those that expanded through Europe in the Palaeolithic, and those whose expansion has occurred during the Neolithic. Additionally a rather distinct African influence was detected in this Portuguese survey, as signalled by the distributions of haplogroups U6 and L, present at higher frequencies than those usually reported in Iberian populations. The geographical distributions of both haplogroups were quite different, with U6 being restricted to North Portugal whereas L was widespread all over the country. This seems to point to different population movements as the main contributors for the two haplogroup introductions. We hypothesise that the recent Black African slave trade could have been the mediator of most of the L sequence inputs, while the population movement associated with the Muslim rule of Iberia has predominantly introduced U6 lineages.

INTRODUCTION

Since the description of the mitochondrial DNA sequence by Anderson *et al.* (1981), this peculiar genome, which is maternally inherited, non-recombining and fast-evolving, has been intensively investigated and applied to population studies. The initial screening based on restriction fragment length polymorphisms spread all over the molecule, was soon enlarged by direct sequencing of two hypervariable regions located in the control region (D-Loop): HVRI and HVRII.

Correspondence: Luísa Pereira, Instituto de Patologia e Imunologia Molecular da Universidade do Porto (IPATIMUP), R. Dr. Roberto Frias s/n, 4200 Porto PORTUGAL. Tel: +351 22 5570700; Fax: +351 22 5570799.
E-mail: lpereira@ipatimup.pt

Present day sequence variation of mtDNA is a valuable tool for making what Avise *et al.* (1987) referred to as phylogeographic inferences. MtDNA sequences can be used to construct networks or used in other methodological approaches which afford information about prehistoric population size and patterns of gene flow. The evolutionary history of haplogroups, their inferred origin and expansion through the world, provide the basis for reconstructing and dating major prehistoric and historic population movements.

Many published studies based on HVRI sequence diversity focus on the history of European populations (Richards *et al.* 1996; Côrte-Real *et al.* 1996; Richards *et al.* 1998). Richards *et al.* (1998), applying a phylogeographic approach to western Europe mtDNA diversity, concluded that the majority (85%) of

European sequences must have originated during the Upper Palaeolithic and suffered a considerable post-glacial expansion; about 15% of the sequences reflect a restricted Neolithic input, from the Near East toward the West of Europe, and only 1% of the sequences represent more recent influences of Asian and African mtDNA pools.

Focussing on the westernmost edge of Europe, the Iberian Peninsula, some studies (Côrte-Real *et al*. 1996; Salas *et al*. 1998) have pointed to a common origin of all Iberian populations in the Upper Palaeolithic. For most populations, diversity levels were found to be lower than the values reported for central European countries, a feature that was thought to support the expansion model of modern humans from the Middle East in the direction of Western Europe. The lowest Iberian diversity value was observed in Basques, reflecting the uniqueness of this Iberian population (Côrte-Real *et al*. 1996).

Another peculiarity of the Iberian mitochondrial pool is the presence of sequences belonging to the U6 group (Richards *et al*. 1998), signalling a North African influence that has not been detected elsewhere in other European populations.

In this work we have analysed HVRI and HVRII diversity in Portugal, the westernmost country of the Iberian Peninsula, with the aim of obtaining a better characterisation of European mtDNA variability. We have considered three main regions in Portugal: North, Central and South. This was done in parallel with a study of Y chromosome biallelic markers that has revealed statistical differences between the south compared to the north and central regions (Pereira *et al*. 2000).

Our main approach regarding the analysis of mtDNA diversity was the evaluation of patterns of mismatch distribution within the major haplogroups found in Portugal. We intended to assess whether inferences regarding the history of haplogroups were in agreement with those previously published based on network analysis, and, simultaneously, to deepen the evolutionary picture of mtDNA lineages in Europe.

Material and methods

Population samples

Three population samples from Portugal were analysed: 100 unrelated individuals from the North, 82 from the Central region and 59 from the South, according to the country division by the major rivers Douro and Tagus. A total of 15 μl of blood was used to extract DNA by the resin Chelex-100 method (Lareu *et al*. 1994).

MtDNA amplification and sequencing

MtDNA was amplified using the primers L15997 (5′-CACCATTAGCACCC AAAGCT-3′) and H16401 (5′-TGATTTCACGGAGGATGGTG-3′) for HVRI, and L48 (5′-CTCACGGGAGCTCTCCATGC-3′) and H408 (5′-CTGTTAAAAGTGCATACCG CCA-3′) for HVRII. The temperature profile was 95 °C for 10 sec., 60 °C for 30 sec. and 72 °C for 30 sec., for 35 cycles of amplification.

The amplified samples were purified with Microspin™ S-300 HR columns (Pharmacia Biotech), according to the manufacturer's specifications. The sequencing reactions were carried out using the Kit Big-Dye™ Terminator Cycle Sequencing Ready Reaction (Perkin-Elmer), with one of the above described primers, in both forward and reverse directions.

A protocol based on $MgCl_2$/ethanol precipitation was used for post-sequence reaction purification of samples, which were then applied to a 6% PAGE gel and run in an automatic sequencer ABI 377.

Genetic analysis

The nucleotide positions considered for analysis were between bp 16024 and 16383 for HVRI and between 73 and 340 for HVRII (in the numbering system of Anderson *et al*. 1981).

Sequence classification into haplogroups was based on HVRI and position 00073 of HVRII, and the nomenclatures of Richards *et al*. (1998), Macaulay *et al*. (1999) and Rando *et al*. (1999) were followed for European, Sub-Saharan and North African clusters, respectively. Sequences are available in GenBank (accession

nos. AF277997–AF278237 and AF278238–AF278478).

Molecular diversity indexes and mismatch distributions were executed using the software ARLEQUIN 1.1 (Schneider *et al.* 1997).

RESULTS AND DISCUSSION

HVRI and HVRII diversity in Portugal

Some diversity parameters obtained in the three Portuguese regions studied for HVRI and/or II are presented in Table 1. HVRI presented a higher mean number of nucleotide differences than region II. However, when corrected for fragment sizes both regions showed a similar mean number of nucleotide pairwise differences: HVRII/HVRI mean ratios were 0.87, 0.90 and 1.05 in North, Central and South Portugal, respectively.

The proportion of polymorphic sites (Table 2) was higher in HVRI than in HVRII, averaging 17.68% and 11.20%, respectively. However, region II presented a slightly faster mutation rate than the former, a conclusion that was based on the comparison of estimates of the τ parameter of Rogers & Harpending (1992) in both regions. This parameter consists in $\tau = \mu lt$, where μ is mutation rate per nucleotide, l is sequence length and t is time in generations after a population expansion. Since t is equal in HVRI and HVRII in a certain population, and knowing the length of both regions, we can obtain μ_{II}/μ_I from the ratio τ_{II}/τ_I. The μ_{II}/μ_I ratios for Portugal were 1.044, 0.984 and 1.273, in North, Central and South respectively, with a mean value of 1.10, a value close to that reported by Salas *et al.* (2000) for other European populations (0.998 for British, 1.216 for Austrian and 0.943 for Tuscan), excepting Galicia, where the very high value found (1.845) is related to the low mean of nucleotide differences reported for HVRI in that population.

Combining both sets of observations it is clear that in HVRI mutations tend to occur more homogeneously along different nucleotide positions than in HVRII. Therefore, our data are in agreement with previous studies (Meyer *et al.* 1999; Torroni *et al.* 1996) that have described fast-mutating positions in the HVRII region: positions 146, 150, 152 and 195 are very prone to substitution events whereas at position 309 a length polymorphism is frequently found.

The differential mutational behaviour of HVRI and HVRII is also reflected in the pattern of mismatch distributions. The observed and expected numbers of pairwise differences were analysed in the three geographic regions, but as the distribution patterns were basically identical in the three regions, only those corresponding to the overall Portuguese sample will be presented (Figure 1). For HVRII, the observed distribution closely matches the expected distribution. However, the observed distributions for HVRI and HVRI + HVRII are characterised by the presence of slight shoulders and higher number of nucleotide differences compared to the expected values.

It is widely accepted that the mismatch distribution retains valuable information about demographic episodes undergone during the history of a population. Unimodal curves with modes at a small number of differences have been observed in western European populations (Côrte-Real *et al.* 1996; Salas *et al.* 1998) and were interpreted as signatures of relatively recent population expansions. By contrast, the tendency of African populations (Mateu *et al.* 1997) to display ragged and multimodal distributions has supported the idea of their being more ancient and stationary, or more diversified. It is worth mentioning that making population demographic inferences from the analysis of mismatch distributions must be tentative since several factors have to be considered: time of expansion, size before the expansion, gene flow between populations and sub-structuring of populations. Some simulation studies have shown that the statistical effects of some factors can be quite convergent (Marjoram & Donnelly, 1994).

In the presence of mismatch distributions showing some deviations from regular characteristic unimodal distributions, as we have found for HVRI and HVRI + HVRII in Portugal, we may be facing a population in which one or more

Table 1. *MtDNA diversity parameters in North (NP), Central (CP) and South (SP) Portugal, considering HVRI and/or HVRII*

	% of different haplotypes	Mean no. of nucleotide pairwise differences	Nucleotide diversity
HVRI			
NP	67.0	4.78	0.013
CP	75.6	4.87	0.014
SP	69.5	4.54	0.013
HVRII			
NP	47.0	3.09	0.012
CP	50.0	3.26	0.012
SP	61.0	3.55	0.013
HVRI+HVRII			
NP	84.0	7.87	0.012
CP	92.7	8.13	0.013
SP	91.5	8.09	0.013

Table 2. *HVRI and HVRII variability in North (NP), Central (CP) and South (SP) Portugal*

	HVRI (360 bp)	HVRII (268 bp)
Polymorphic sites (%)		
NP	71 (19.7)	28 (10.5)
CP	66 (18.3)	31 (11.6)
SP	54 (15.0)	31 (11.6)
No. of substitutions (%)		
NP	73 (100.0)	25 (89.3)
CP	69 (100.0)	30 (93.8)
SP	54 (100.0)	29 (90.6)
No. of transitions (%)		
NP	66 (90.4)	24 (96.0)
CP	62 (89.9)	28 (90.3)
SP	50 (92.6)	27 (93.1)
No. of transversions (%)		
NP	7 (9.6)	1 (4.0)
CP	7 (10.1)	2 (6.7)
SP	4 (7.4)	2 (6.9)
No. of indels (%)		
NP	0 (0.0)	3 (10.7)
CP	0 (0.0)	2 (6.3)
SP	0 (0.0)	3 (9.4)

of those factors has played a significant role. Even without discriminating which factors these might be, the mismatch distribution observed for HVRII suggests that this sequence stretch is more resistant to their effects. We cannot exclude the possibility that the regular unimodal distribution for HVRII could be the result of an (unknown) selective effect. However, if this was the case, this selection would also affect HVRI (since they are linked on the same molecule). Another explanation is related to the fact that, as already mentioned, mutational events in HVRII are more heterogeneously dispersed. Since some HVRII sites present very high mutation rates the nucleotide variation does not tend to be so strongly associated in blocks as happens in the HVRI region. Within HVRI, substitutions are more often found to be specific to certain haplogroups and, by turn, it is easier to classify the observed variation into haplogroups, which manifestly tend to show geographic clustering. Attending to all these features, we suggest that HVRI and HVRI +HVRII distributions are more sensitive to demographic factors and consequently their study seems to be more informative for making population demographic inferences.

Diversity comparison with other populations

For population comparison purposes, we have not considered the diversity values registered for HVRII, since population data for this region are more scarce.

Table 3 presents some diversity estimates for several populations. The Portuguese samples studied here display a mean number of nucleotide pairwise differences typical of European populations, which are, in general, lower than the values characteristic of Asian and African populations (Comas *et al.* 1998; Mateu *et al.* 1997). However, compared to a neighbouring Iberian population, namely Galicia (Salas *et al.* 1998), the populations analysed here present a higher level of diversity, even when assessed excluding the African L sequences (data not shown), which, as will be further discussed, are found at very low frequency in other Iberian populations.

It is difficult to explain the higher mtDNA diversity in Portugal compared to other Iberian neighbouring populations. The data might suggest an ancient settlement of this region of the Iberian Peninsula, a hypothesis that is not supported by historical or palaeonthological records. Alternatively, the use of North Portugal as a refuge zone during the last glacial maximum, as has been suggested for Andalusia (Côrte-Real

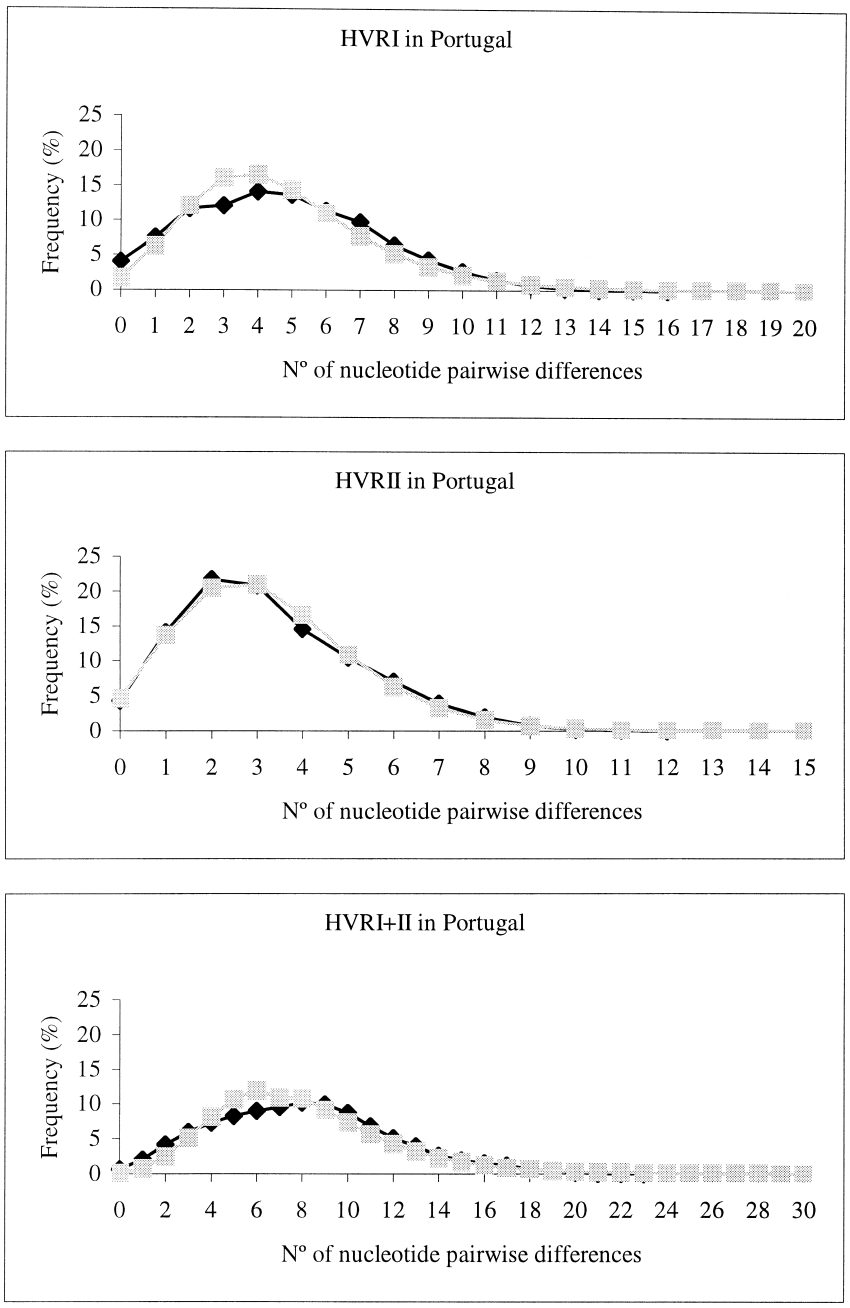

Fig. 1. Observed (black) and expected (grey) mismatch distributions for HVRI, HVRII and HVRI+HVRII in Portugal.

et al. 1996), maintaining the level of diversity while population size was reduced in other Iberian regions, is a possible explanation, but in contradiction with the known palaeoclimatic data (Mellars, 1998). A more likely explanation is the occurrence at different times of several significant influxes of different mtDNA lineages, a hypothesis that is in accordance with the multiple contacts with different people that have characterised the recent history (i.e. the last thousand years) of the region.

Haplogroup diversity analysis

The haplogroup frequencies observed in North, Central and South Portugal are summarised in Table 4. In HVRII, the CRS sequence was never found because two positions, not polymorphic in Portugal, were always different from the ref-

Table 3. *Sequence diversity observed in HVRI in several populations. N, sample size; K, number of different sequences found; A, number of variable nucleotides positions; B, mean nucleotide pairwise differences; C, percentage average pairwise difference per nucleotide; π, nucleotide diversity.*

	N	K	A	B	C	π	References
Basque	106	52	52	2.95	0.82	0.008	1, 2
Galician	92	53	56	3.13	0.87	0.009	3
Portuguese	54	38	46	3.60	1.00	0.010	2
Catalonian	15	11	16	3.73	1.04	0.010	2
British	100	71	67	4.45	1.24	0.012	4
S. Portuguese	59	41	54	4.51	1.25	0.013	This study
N. Portuguese	100	67	71	4.78	1.33	0.013	This study
C. Portuguese	82	62	66	4.87	1.35	0.014	This study
Spanish	89	70	69	5.02	1.39	0.014	2, 5
Tuscan	49	40	55	5.03	1.40	0.014	6
Turkish	96	79	82	5.45	1.51	0.015	7, 8
Middle-Eastern	42	38	59	7.08	1.97	0.020	9
S. Tomean	50	32	53	7.56	2.10	0.021	10

[1] Bertranpetit *et al.* (1995); [2] Côrte-Real *et al.* (1996); [3] Salas *et al.* (1998); [4] Piercy *et al.* (1996); [5] Pinto *et al.* (1996); [6] Francalacci *et al.* (1996); [7] Calafell *et al.* (1996); [8] Comas *et al.* (1996); [9] Di Rienzo *et al.* (1991); [10] Mateu *et al.* (1997).

Table 4. *MtDNA haplogroup distributions (no. of individuals and % values in parenthesis) in North (NP), Central (CP) and South (SP) Portugal*

	NP	CP	SP
H	40+1? (41.00)	31 (37.81)	25+1? (50.0)
I	1 (1.00)	—	1 (1.70)
J*	2 (2.00)	4 (4.88)	3 (5.09)
J1	2 (2.00)	—	—
J1b	—	—	2 (3.39)
J2	2 (2.00)	1 (1.22)	—
K	3 (3.00)	6 (7.32)	4 (6.78)
L1b	—	1 (1.22)	1 (1.70)
L2	3 (3.00)	2 (2.44)	1 (1.70)
L3*	2 (2.00)	5 (6.10)	2 (3.39)
M1	—	1 (1.22)	—
T*	3+1? (4.00)	8 (9.76)	6 (10.17)
T1	6+1? (7.00)	1 (1.22)	—
U*	1+3? (4.00)	2+1? (3.66)	1? (1.70)
U2	—	2 (2.44)	—
U3	2 (2.00)	—	1 (1.70)
U4	2 (2.00)	2 (2.44)	—
U5	1 (1.00)	—	—
U5a	2 (2.00)	1 (1.22)	1 (1.70)
U5a1	2 (2.00)	1 (1.22)	1 (1.70)
U5a1a	1 (1.00)	2 (2.44)	1 (1.70)
U5a/b	1 (1.00)	—	—
U5b	1 (1.00)	1 (1.22)	1 (1.70)
U6	1 (1.00)	—	—
U6a	4 (4.00)	—	—
U6b	2 (2.00)	—	—
U7	—	—	1 (1.70)
V	8 (8.00)	3+3? (7.32)	4 (6.78)
W	2 (2.00)	1 (1.22)	–
X	–	3 (3.66)	1+1? (3.39)

Note: Haplogroups where the classification presented ambiguity are assigned with a question mark (?).

erence sequence: at position 00263 we found a G, and at position 00311 a length polymorphism with one more C. A recent revision of the CRS sequence (Andrews *et al.* 1999) confirmed the presence of an A in 00263 and 5C in 00311, both allelic states occurring at very low frequency.

Comparisons between North, Central and South Portugal did not reveal statistical differences between the three regions with respect to mtDNA variability (p values of F_{ST} pairwise genetic distances between populations were 0.327, 0.673 and 0.921 for North-Central, North-South and Central-South, respectively). Accordingly, the sequences of the three regions could be merged in a global Portuguese sample, for the analysis that will be presented next.

In order to evaluate if the overall diversity in Portugal was due to a high diversity within particular haplogroups or, alternatively, to a high diversity of haplogroups, after classifying the different sequences into haplogroups we have performed mismatch distribution analysis within each of the major haplogroups. We will only present the results based on HVRI diversity (Figure 2), as when the less informative HVRII region was used (data not shown) all clusters exhibited clear unimodal curves with low modes. In this section we will try to correlate the inferences taken from the analysis of haplogroups mismatch distributions with those made by Richards *et al.* (1998) based on network analysis.

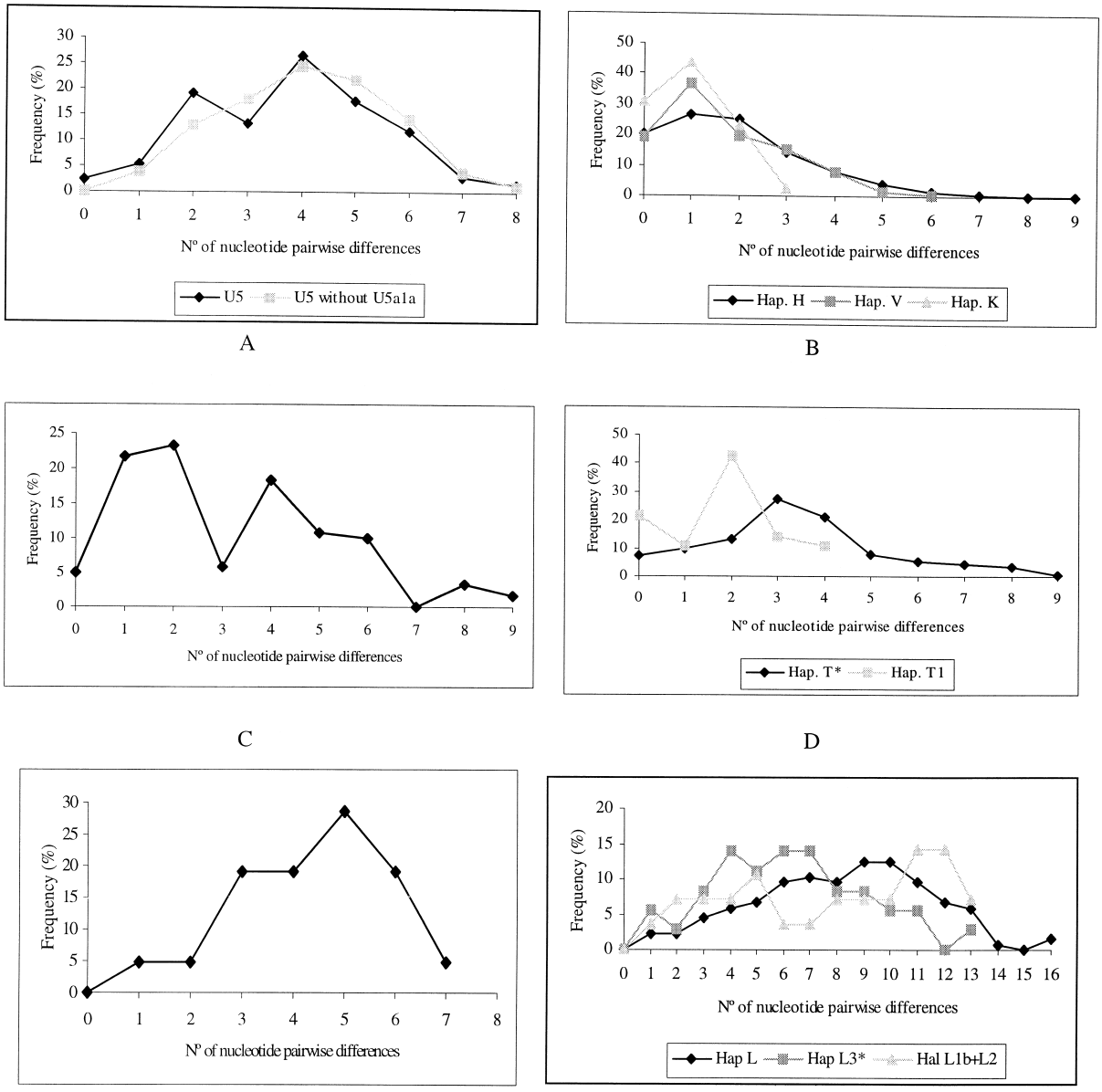

Fig. 2. Mismatch distributions for the haplogroups observed in Portugal (North, Central and South) considering only HVRI diversity. A- haplogroup U5 with and without U5a1a sequences. B- haplogroups H, V and K. C- haplogroup J. D- haplogroups T* and T1. E- haplogroup U6. F- haplogroups L1b + L2, L3* and all simultaneously.

Haplogroup U5 had the highest mode for the number of pairwise differences distribution, and showed a regular unimodal pattern (Figure 2A). Both features are in accordance with its being the oldest haplogroup in Europe which has registered a regional development. As expected, and depicted in Figure 2A, a slight bimodality appeared when the more recent sub-haplogroup U5a1a was considered in the analysis.

The three haplogroups H, V and K, all considered to be post-glacially expanded European haplogroups (Figure 2B), showed clear unimodal distributions but with low means, reflecting the shorter period since the accumulation of variation began.

Haplogroups J and T, said to have had a common origin in the Near East, presented very distinct mismatch distribution patterns. Haplogroup J, which seems to have been recently introduced into Europe during the Neolithic,

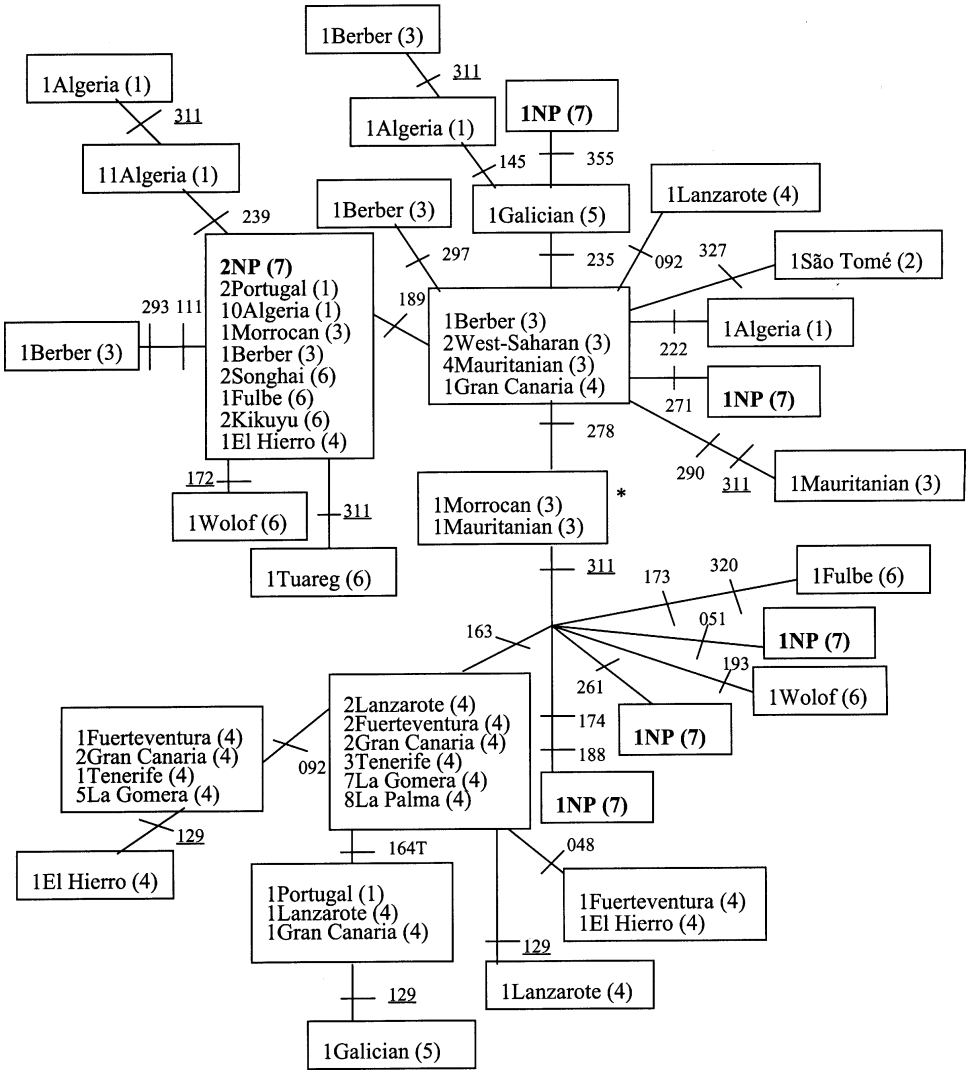

Fig. 3. A most parsimonious tree of sequences belonging to cluster U6. Root motif 172-219 indicated with an asterisk. Branches are labelled by the nucleotide positions in HRVI (minus 16000) to designate transitions; transversions are further specified and positions underlined represent parallel mutations. Numbers in brackets represent bibliographic reference (1) Côrte-Real *et al.* (1996), (2) Mateu *et al.* (1996), (3) Rando *et al.* (1998), (4) Rando *et al.* (1999), (5) Salas *et al.* (1998), Watson *et al.* (1996), (7) this work.

showed a very irregular curve with a ragged shape (Figure 2C). This pattern additionally suggests a high level of diversification of the founder sequences and also expresses the heterogeneity within the haplogroup: some well-defined J sub-haplogroups are distinctively spread through Europe, paralleling the clines observed for Y chromosomal, and some autosomal, markers.

In contrast, haplogroup T presented a unimodal and bell-shaped curve (Figure 2D), which can be explained by a more ancient introduction into Europe with subsequent accumulation, *in loco*, of homogenising mutations, starting from a diversity level that must also have been lower relatively to the initial package of J sequences. When the sub-haplogroup T1 was analysed separately, a slight bimodality and a low mean of pairwise differences were registered. Those features are quite compatible with its more recent introduction, although we cannot exclude some bias due to the low number of sequences considered.

In the Portuguese sample analysed two haplo-

groups, U6 and L, that have been reported as occurring sporadically in other European populations, were detected with comparatively high frequency. Both haplogroups were characterised by high levels of diversity and displayed very irregular mismatch distributions (Figure 2 E and F). Moreover, haplogroup U6 was found to be restricted to the North region of the country, whereas the L sequences were spread all over the country.

These haplogroups have been reported to be characteristic of African populations, where their frequency is inversely correlated with the North-South axis: the frequency of U6 is high in North Africa and decreases in a southerly direction, being almost absent south of the equator; the L cluster has an opposite distribution (Rando *et al.* 1998, 1999; Watson *et al.* 1996; Mateu *et al.* 1996).

In Portugal, as well as generally in Iberia, many migration waves from both North and sub-Saharan African populations are well documented. The geographical proximity of North Africa and the Iberian Peninsula certainly afforded many opportunities for mutual population contacts. Among them, we stress the movement of Berbers and Arabs that took place during the very recent Muslim rule of Iberia (from the 8th century to the end of the 15th, in some regions). In addition, many sub-Saharan individuals entered the region during the slave trade period, from its very beginning (middle 15th century) until its total ban in the late 19th century.

As it would be interesting to find out the origin of the L and U6 sequences detected in Portugal, we have tried to compare the motifs of the sequences observed in Portugal with those described in the literature for several populations (Figures 3 and 4). However most of the matches found for the Portuguese sequences were with sequences widely distributed in Africa, and no clear pattern of geographic clustering was detected.

A striking aspect observed for the U6 haplogroup was that 5 out of 7 of the Portuguese sequences were unique to Portugal, not allowing, therefore, any accurate assignment of their geographical origin. The Canarian characteristic sub-haplogroup U6b1 (Rando *et al.* 1999), observed in other Iberian samples, was not detected in the present study.

Admitting that U6 sequences could have been at least partially introduced by Berber people during the Muslim rule of Iberia, it is strange to find them restricted to North Portugal. As a matter of fact, most historical sources document a deeper influence of Berber (as well as Arab) people in Central and particularly South Iberia (as judged from toponyms and general cultural affinities), compared to North Iberia where the Muslim presence is recorded to have been more ephemeral and consequently to have made less cultural and demographic impact. The data does not exclude the possibility that U6 introductions could have been additionally reinforced by later sub-Saharan inputs mediated by the African slave trade. Even if this mixed scenario is plausible, the presence of U6 sequences exclusively in North Portugal is a question that deserves further analysis. The hypothesis of an earlier introduction in the region does not seem to be favoured, neither by its presence in a restricted geographical area, nor by the high level of heterogeneity that characterises the set of sequences that were found among this haplogroup.

With respect to the L sequences, it is widely accepted that they have a sub-Saharan origin, excepting some L3* lineages that, as analysis of Figure 4 suggests, might indeed have a non-African origin. The presence of L sequences in North African regions does not allow us to exclude the possibility that population influxes from this region, namely the above referred Berber/Arab movement, have introduced a significant fraction of L sequences into Iberia. However, it seems more likely that most of the L lineages found nowadays in Portugal have been carried by African slaves, since the country was actively involved in the Transatlantic slave trade. Nine out of 17 L sequences found in this study showed matches with widespread African sequences, and with regard to the 8 remaining

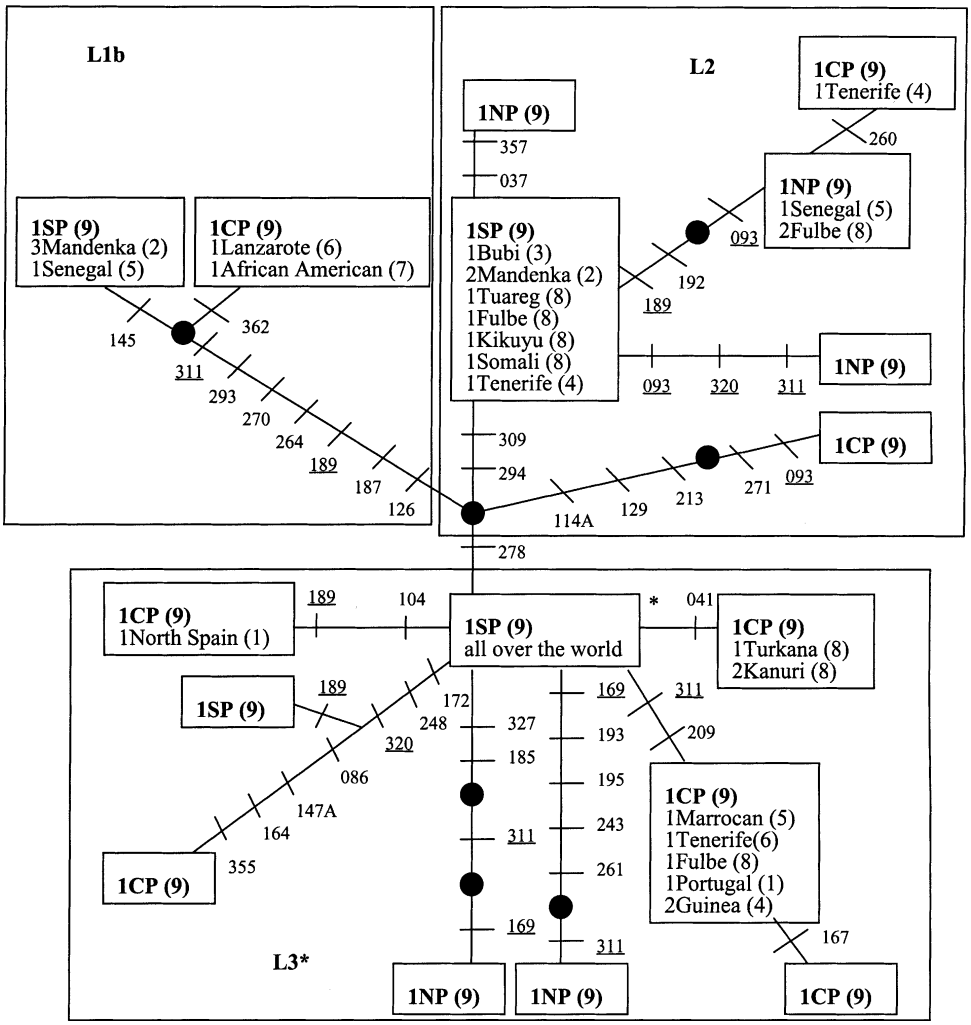

Fig. 4. A phylogeny of Portugese sequences belonging to African clades L1b and L2 and to the default cluster L3* (some members of which may have a non-African origin). The sequence with a transition from the CRS at np 16223 is indicated with an asterisk. Sequence matches in other populations are shown. Numbers in brackets represent bibliographic reference (1) Côrte-Real *et al.* (1996), (2) Graven *et al.* (1995), (3) Mateu *et al.* (1996), (4) Pinto *et al.* (1996), (5) Rando *et al.* (1998), (6) Rando *et al.* (1999), (7) Vigilant *et al.* (1991), (8) Watson *et al.* (1996), (9) this work. Solid circles represent sequences observed in the mtDNA database or branching nodes in the mtDNA phylogeny which aid in the identification of parallel mutations, which are shown with the position underlined.

sequences the absence of matches can be due to the present bias in the description of sub-Saharan mtDNA variability. Broad areas corresponding to Ivory Coast, Angola and Mozambique, which represented very important sources of African slaves, remain uncharacterised.

There were more African slaves in Portugal than in any other European country: in 1550, Lisbon boasted 10000 resident slaves in a population of 100000, and Portugal as a whole probably had over 40000 (Thomas, 1998). In the mid-sixteenth century the birth of slaves' children was stimulated in Portugal for internal traffic purposes. Inter-breeding between autochthonous individuals and African slaves certainly occurred and the predominant mating must have been between slave African females and autochthonous males, due to social pressures and also for legal reasons: offspring of slave females would be slaves, whereas offspring of slave males would not. Therefore, breeding between slave African males and white females, besides being socially repressed, would not bring any economic profit. If the pattern of genetic admixture was markedly sex influenced, the signature of this recent African influence would be expected to be very

different in the maternally inherited gene pool and in the paternally inherited one. In a recent study based on Y chromosome biallelic markers (Pereira *et al*. 2000) we have reported the absence of typical sub-Saharan haplogroups in the Y chromosome Portuguese pool. This finding, and the detection of L sequences at 7.1% in the mitochondrial pool, both seem to support the above-mentioned pattern of admixture with African slaves.

CONCLUSIONS

Studies of large population samples, designed to characterise the molecular diversity in restricted geographical contexts, can produce valuable insights concerning specific demographic features that would remain undetectable in broader scale surveys. In this work we have studied mtDNA variability in Portugal, considering North, Central and South regions as micro-screening sample units.

The level of mtDNA diversity found, although characteristic of European populations, is high when the westernmost location of the country in Europe, and the reported European tendency for reduction of diversity toward Western Iberia (Corte-Real *et al*. 1996; Salas *et al*. 1998), are considered. The observed HVRI and HVRI + HVRII mismatch distributions were unimodal but smoother than others previously found in neighbouring populations.

This finding, as well as the high level of haplogroup diversity, suggests the influence of specific demographic factors acting in the Portuguese population, and led us to hypothesise that an important modulator of the present Portuguese mtDNA variability could have been the influx of distinct mtDNA lineages at historically quite different times.

Sharing the features of mtDNA diversity generally registered in Europeans (all European haplogroups were detected), Portugal has in addition received significant North and sub-Saharan African influences. Frequencies of haplogroups specific to these regions were higher than those reported for other European populations: 7% of North African sequences were detected (restricted to North Portugal and representing almost 3% of the total sample), and sub-Saharan African sequences were found to be spread throughout the country, with frequencies between 5% and 9.8%. Although statistically significant differences were not detected between the three sub-samples considered, the geographic distribution pattern observed for U6 and L sequences strongly suggest that different population movements were responsible for their introduction into the country, although none of them had enough demographic impact to induce regional differentiation.

The introduction of L sequences in Portugal was tentatively imputed mainly to the modern slave trade that occurred between the 15th and 19th centuries. Both the great number of slaves that entered Portugal and their very diverse African geographic origin are consistent with the data set now reported. However, we cannot exclude some North-African contribution to present-day Portuguese L lineages.

While the population movement associated with the slave trade may be responsible by some U6 inputs, we suggest that U6 sequences were predominantly introduced into Portugal during the Berber/Arab invasion of the Peninsula. However, the observation that haplogroup U6 is restricted to North Portugal is puzzling, considering the more pronounced impact of the Muslin rule in south Iberia and the widespread presence of African slaves throughout the country, and deserves further investigation.

We are deeply in debt to Vincent Macaulay, whose critical discussion of several aspects of this paper was very important for its improvement. We also thank Martin Richards for his assistance in the classification of some sequences. This work was partially supported by a grant (PRAXIS BD/13632/97) financed by Fundação para a Ciência e a Tecnologia.

REFERENCES

Anderson, S., Bankier, A. T., Barrell, B. G., de Bruijn, M. H. L., Coulson, A. R., Drouin, J., *et al*. (1981). Sequence and organisation of the human mitochondrial genome. *Nature* **290**, 457–465.

Andrews, R. M., Kubacka, I., Chinnery, P. F., Lightowlers, R. N., Turnbull, D. M. & Howell, N. (1999). Reanalysis and revision of the Cambridge

reference sequence for human mitochondrial DNA. *Nature Genet.* **23**, 147.

Arnaiz-Villena, A., Martínez-Laso, J., Gómez-Casado, E., Díaz-Campos, N., Santos, P., Martinho, A., Breda-Coimbra, H. (1997). Relatedness among Basques, Portuguese, Spaniards and Algerians studied by HLA allelic frequencies and haplotypes. *Immunogenetics* **47**, 37–43.

Avise, J., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E. *et al.* (1987). Intraspecific phylogeography: the molecular bridge between population genetics and systematics. *Ann. Ver. Ecol. Syst.* **18**, 489–522.

Bertranpetit, J., Sala, I., Calafell, F., Underhill, P., Moral, P. & Comas, D. (1995). Human mitochondrial DNA variation and the origin of the Basques. *Ann. Hum. Genet.* **59**, 63–81.

Calafell, F., Underhill, P., Tolun, A., Aangelicheva, D. & Kaladjieva, L. (1996). From Asia to Europe: mitochondrial DNA sequence variability in Bulgarians and Turks. *Ann. Hum. Genet.* **60**, 35–49.

Comas, D., Calafell, F., Mateu, E., Pérez-Lezaun, A., Bosch, E., Martínez-Arias, R. *et al.* (1998). Trading genes along the Silk Road: mtDNA sequences and the origin of Central Asian populations. *Am. J. Hum. Genet.* **63**, 1824–1838.

Comas, D., Calafell, F., Mateu, E., Pérez-Lezaun, A. & Bertranpetit, J. (1996). Geographic variation in human mitochondrial DNA control region sequence: the population history of Turkey and its relationship to the European population. *Mol. Biol. Evol.* **13**, 1067–1077.

Côrte-Real, H., Macaulay, V. A., Richards, M. B., Hariti, G., Issad, M. S., Cambon-Thomsen, A. *et al.* (1996). Genetic diversity in the Iberian Peninsula determined from mitochondrial sequence analysis. *Ann. Hum. Genet.* **60**, 331–350.

Di Rienzo, A. & Wilson, A. C. (1991). Branching pattern in the evolutionary tree for the human mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* **88**, 1597–1601.

Forster, P., Harding, R., Torroni, A. & Bandelt, H. J. (1996). Origin and evolution of native American mtDNA variation: a reappraisal. *Am. J. Hum. Genet.* **59**, 935–945.

Francalacci, P., Bertranpetit, J., Calafell, F. & Underhill, P. A. (1996). Sequence diversity of the control region of mitochondrial DNA in Tuscany and its implications for the peopling of Europe. *Am. J. Phys. Anthrop.* **100**, 443–460.

Graven, L., Passarino, G., Semino, O., Boursot, P., Santachiara-Benerecetti, S., Langaney, A. & Excoffier, L. (1995). Evolutionary correlation between control region sequence and restriction polymorphisms in the mitochondrial genome of a large Sengalese Mandenka sample. Mol Biol Evol. **12**, 334–45.

Lareu, M. V., Phillips, C. P., Carracedo, A., Lincoln, A. J., Syndercombe-Court, D. & Thomson, J. A. (1994). Investigation of the STR locus HUMTH01 using PCR and two electrophoresis formats; UK and Galician Caucasian population surveys and usefulness in paternity investigations. *Forensic Sci. Int.* **66**, 41–52.

Macaulay, V., Richards, M., Hickey, E., Vega, E., Cruciani, F., Guida, V. *et al.* (1999). The emerging tree of west Eurasian mtDNAs: a synthesis of control-region sequences and RFLPs. *Am. J. Hum. Genet.* **64**, 232–249.

Marjoram, P. & Donnelly, P. (1994). Pairwise comparisons of mitochondrial DNA sequences in subdivided populations and implications for early human evolution. *Genetics.* **136**, 673–683.

Mateu, E., Comas, D., Calafell, F., Pérez-Lezaun, A., Abade, A. & Bertranpetit, J. (1997). A tale of two islands: population history and mitochondrial DNA sequence variation of Bioko and São Tomé, Gulf of Guinea. *Ann. Hum. Genet.* **61**, 507–518.

Mellars, P. (1998). The Upper Paleolithic Revolution. In *Prehistoric Europe: an illustrated history* (ed. B. W. Cunliffe). Oxford: Oxford University Press.

Meyer, S., Weiss, G. & von Haeseler, A. (1999). Pattern of nucleotide substitution and rate heterogeneity in the hypervariable regions I and II of human mtDNA. *Genetics* **152**, 1103–1110.

Pereira, L., Brión, M., Prata, M. J., Jobling, M. A., Carracedo, A. & Amorim, A. (2000). Gradient of Y chromosome haplogroup 21 across the Western Iberia. *Progress in Forensic Genetics* **8**, 281–283.

Piercy, R., Sullivan, K. M., Benson, N. & Gill, P. (1996). The application of mitochondrial DNA typing to the study of white Caucasian genetic identification. *Int. J. Leg. Med.* **106**, 85–90.

Pinto, F., González, A. M., Hernández, M., Larruga, J. M. & Cabrera, V. M. (1996). Genetic relationship between the Canary Islanders and their African and Spanish ancestors inferred from mitochondrial DNA sequences. *Ann. Hum. Genet.* **60**, 321–330.

Rando, J. C., Cabrera, V. M. Larruga, J. M., Hernández, M., González, A. M., Pinto, F., & Bandelt, H.-J. (1999). Phylogeographic patterns of mtDNA reflecting the colonization of the Canary Islands. *Ann. Hum. Genet.* **63**, 413–428.

Rando, J. C., Pinto, F., González, A. M., Hernández, M., Larruga, J. M., Cabrera, V. M. & Bandelt, H.-J. (1998). Mitochondrial DNA analysis of Northwest African populations reveals genetic exchanges with European, Near-Eastern, and sub-Saharan populations. *Ann. Hum. Genet.* **62**, 531–550.

Richards, M. B., Macaulay, V. A., Bandelt, H.-J. & Sykes, B. C. (1998). Phylogeography of mitochondrial DNA in western Europe. *Ann. Hum. Genet.* **62**, 241–260.

Richards, M. B., Côrte-Real, H., Forster, P., Macaulay, V., Wilkinson-Herbots, H., Demaine, A. *et al.* (1996). Palaeolithic and Neolithic lineages in the European mitochondrial gene pool. *Am. J. Hum. Genet.* **59**, 185–203.

Rogers, A. R. & Harpending, H. (1992). Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.* **9**, 552–569.

Salas, A., Comas, D., Lareu, M. V., Bertranpetit, J. & Carracedo, A. (1998). mtDNA analysis of the Galician population: a genetic edge of European variation. *Eur. J. Hum. Genet.* **6**, 365–375.

Salas, A., Lareu, M. V., Sánchez-Diz, P., Calafell, F. & Carracedo, A. (2000). mtDNA hypervariable region II (HVII) sequences in human evolution studies: impact of mutation rate heterogeneity. *Progress in Forensic Genetics* **8**, 329–331.

Schneider, S., Kueffer, J. M., Roessli, D. & Excoffier, L. (1997). Arlequin ver.1.1: A software for population genetic data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Thomas, H. (1998). *The slave trade – the history of the Atlantic slave trade 1440–1870*. London: Macmillan Publishers Ltd.

Torroni, A., Bandelt, H.-J., D'Urbano, L., Lahermo, P., Moral, P., Sellitto, D., *et al*. (1998). MtDNA analysis reveals a major late Palaeolithic population expansion from southwestern to northeastern European populations. *Genetics* **62**, 1137–1152.

Torroni, A., Huoponen, K., Francalacci, P., Petrozzi, M., Morelli, L., Scozzari, R. *et al*. (1996). Classification of European mtDNAs from an analysis of three European populations. *Genetics* **144**, 1835–1850.

Vigilant, L., Stoneking , M., Harpending, H. Hawkes, K., Wilson, A. C. (1991) African populations and the evolution of human mitochondrial DNA. *Science* **253**, 1503–7.

Watson, E., Bauer, K., Aman, R., Weiss, G., von Haeseler, A. & Paabo, S. (1996). mtDNA sequence diversity in Africa. *Am. J. Hum. Genet.* **59**, 437–444.

APPENDIX

Haplotypes and their geographical distribution in North (N), Central (C) and South (S) Portugal.

HVRI	HVRII	N	C	S	Hap
051 162	73 263 311.1	—	1	—	H
051 257	152 263 303.1 311.1	1	—	—	H
075 183$^{A/C}$ 189 249 356	263 303.3 311.1	—	—	1	H
092 129 239	73 263 311.1	—	—	1	H ?
093 126	199 263 303.1 311.1	1	1	—	H ?
093 213 215 263$^{T/A}$	263 303.1 311.1	1	—	—	H
093 263	263 311.1	—	1	—	H
124	146 263 303.1 311.1	—	1	—	H
124 256	146 263 303.1 311.1	—	1	—	H
129	146 263 311.1	1	—	—	H
129	152 263 303.1 311.1	—	1	—	H
162 209	73 263 311.1	—	—	1	H
162 209 293	73 263 303.1 311.1	—	—	1	H
163	263 311.1	—	—	1	H
172	263 311.1	1	—	—	H
176	195 204 263 311.1	1	—	—	H
176 218	200 251 263 303.1 311.1	1	—	—	H
180 278	263 303.1 311.1	1	—	—	H
183$^{A/C}$ 189 356 362	263 311.1	1	1	—	H
183$^{A/C}$ 189	263 303.2 311	—	—	2	H
184	146 263 303.2 311.1	—	1	—	H
189	263 303.1 311.1	—	1	—	H
192 274 362	239 263 303.1 311.1	1	—	—	H
192 274 362	152 239 263 303.1 311.1	1	—	—	H
192	263 303.1 311.1	—	—	1	H
209	263 303.1 311.1	1	1	—	H
209 304	263 311.1	—	1	—	H
218 299	263 303.1 311.1	1	—	—	H
248	257 263 303.1 311.1	1	—	—	H
248	257 263 303.2 311.1	—	1	—	H
259	195 263 303.1 311.1	1	—	—	H
260	263 303.1 311.1	—	1	—	H
261	93 263 311.1	—	1	—	H
265$^{A/C}$	263 303.1 311.1	1	1	—	H
269 270	263 311.1	—	—	1	H
272 304	263 303.1 311.1	1	—	—	H
274 294$^{C/G}$	152 263 311.1	1	1	—	H
274	263 303.1 311.1	—	1	—	H
293 311	195 263 303.1 311.1	—1	—	H	
304	263 311.1	1	—	—	H
304 327	263 311.1	1	—	—	H
311	146 195 263 311.1	—	1	—	H
320$^{C/A}$	263 311.1	3	—	—	H
335	263 311.1	—	—	1	H
344	93 263 303.2 311	—	—	1	H
362	150 239 263 311.1	—	1	—	H
362	263 303.1 311.1	—	—	1	H
CRS	263 311.1	11	3	2	H
CRS	263 303.1 311.1	1	3	3	H

HVRI	HVRII	N	C	S	Hap
CRS	152 263 303.1 311.1	1	1	1	H
CRS	151 152 263 311.1	1	—	—	H
CRS	152 263 311.1	1	2	—	H
CRS	263 269$^{C/A}$ 303.2 311.1	1	—	—	H
CRS	185 263 303.1 311.1	1	—	—	H
CRS	195 257 263 303.2 311.1	1	—	—	H
CRS	263 303.2 311.1	—	1	2	H
CRS	146 263 303.1 311.1	—	1	—	H
CRS	150 263 311.1	—	1	1	H
CRS	151 262 263 303.2 311.1	—	—	1	H
CRS	151 263 303.2 311.1	—	—	1	H
CRS	150 263 303.1 311.1	—	—	1	H
CRS	195 257 263 303.1 311.1	—	—	1	H
CRS	263 303.2 311.1 338	—	—	1	H
129 223 278 311 (391)	73 199 204 250 263 303.2 311.1	1	—	—	I
129 172 223 311	73 199 203 204 250 263 311.1	—	—	1	I
063 069 126	73 228 263 295 311.1	—	1	—	J*
069 126 172	73 228 263 295 311.1	1	—	—	J*
069 126 286	73 185 228 263 295 303.1 311.1	1	—	—	J*
069 126	73 185 263 295 303.1 311.1	—	1	—	J*
069 126 311	73 185 263 295 303.1 311.1	—	1	—	J*
069 126 311	73 228 263 295 303.1 311.1	—	1	—	J*
069 126	73 185 207 228 263 295 311.1	—	—	1	J*
069 126	73 146 185 188 222 228 263 295 311.1	—	—	1	J*
069 126 319	73 185 228 263 295 303.1 311.1	—	—	1	J*
069 126 261	73 146 185 228 263 295 303.1 311.1	1	—	—	J1
069 126 261	73 146 185 228 263 295 311.1	1	—	—	J1
069 126 145 222 235 261 271	73 263 295 311.1	—	—	1	J1b
069 126 145 222 256 261 278	73 199 263 295 311.1	—	—	1	J1b
069 126 193 319 360	73 150 152 263 295 303.1 311.1	1	1	—	J2
069 126 193 319 360 362	73 150 152 263 295 303.1 311.1	1	1	—	J2
093 224 311	73 195 263 303.1 311.1	1	—	—	K
093 189 224 311	73 195 263 311.1	—	1	—	K
093 224 311	73 152 263 311.1	—	1	—	K
093 224 290 311	73 263 303.1 311.1	—	—	1	K
093 224 311	73 150 195 263 303.1 311.1	—	—	1	K
224 311	73 263 303.1 311.1	1	1	—	K
224 311	73 263 303.2 311.1	1	—	1	K
224 256 311	73 263 303.1 311.1	—	1	—	K
224 311	73 146 152 263 311.1	—	1	—	K
224 311	73 146 263 311.1	—	1	—	K
224 311	73 195 263 303.1 311.1	—	—	1	K
126 187 189 223 264 270 278 293 311 362	73 152 182 185$^{G/T}$ 195 247 263 303.1 311.1	—	1	—	L1b
126 145 187 189 223 264 270 278 293 311	73 152 182 185$^{G/T}$ 195 247 263 311.1 357	—	—	1	L1b
037 223 278 294 309 357 (390)	73 143 146 152 195 263 311.1	1	—	—	L2
093 189 192 223 278 294 309 (390)	73 143 146 152 195 263 303.1 311.1	1	—	—	L2
093 223 278 294 309 311 320 (390)	73 143 146 152 195 263 311.1	1	—	—	L2
093 114$^{C/A}$ 129 213 223 271 278	73 146 150 152 182 195 198 207 263 311.1	—	1	—	L2
093 189 192 223 260 278 294 309	73 143 146 152 195 207 263 303.1 311.1	—	1	—	L2
223 278 294 309	73 143 146 152 195 263 311.1	—	—	1	L2
041 223	73 150 263 311.1	—	1	—	L3*
086 147$^{C/A}$ 164 172 223 248 320 355	73 152 199 204 207 263 311.1	—	1	—	L3*
104 183$^{A/C}$ 189 223	73 263 311.1	—	1	—	L3*
167 209 223 311	73 189 200 263 311.1	—	1	—	L3*
169 185 223 311 327	73 150 185 189 200 263 311.1	1	—	—	L3*
169 193 195 223 243 261 311	73 150 200 235 249$^{del\ A}$ 263 303.1 311.1	1	—	—	L3*
172 182$^{A/C}$ 183$^{A/C}$ 189 223 248 320	73 150 195 263 311.1	—	—	1	L3*
209 223 311	73 189 200 263 311.1	—	1	—	L3*

HVRI	HVRII	N	C	S	Hap
223	73 150 195 263 303.1 311.1	—	—	1	L3*
129 183$^{A/C}$ 189 223 249 311	73 195 263 303.1 311.1	—	1	—	M1
037 126 186 189 222	73 152 263 303.1 311.1	1	—	—	T*?
051 126 294 296 304	73 151 204 263 303.1 311.1	1	—	—	T*
093 126 271 294 296 304	73 151 263 303.1 311.1	—	1	—	T*
114 126 153 192 294	73 150 263 303.1 311.1	—	1	—	T*
126 153 294	73 150 263 311.1	1	—	—	T*
126 192 294 296 304	73 151 263 311.1	1	—	—	T*
126 256 294 296	73 152 263 303.1 311.1	—	1	—	T*
126 260 294 296 319	73 263 311.1	—	1	—	T*
126 292 294	73 263 303.1 311.1	—	1	1	T*
126 294 296 304	73 263 311.1	—	1	1	T*
126 294 296 304	73 195 263 311.1	—	1	—	T*
126 294 304	73 152 263 311.1	—	1	—	T*
126 153 189 294 296	73 150 263 303.1 311.1	—	—	1	T*
126 218 294 296 324	73 263 311.1	—	—	1	T*
126 294 296 304	73 151 263 311.1	—	—	1	T*
126 294 296 304	73 151 260 263 303.1 311.1	—	—	1	T*
037 126 163 186 189	73 152 263 303.1 311.1	1	—	—	T1?
126 163 186 187 189 294	73 152 195 263 303.1 311.1	1	—	—	T1
126 163 186 189 294	73 195 263 303.1 311.1	3	—	—	T1
126 163 186 189 249 294 311	73 263 303.1 311.1	2	1	—	T1
CRS	73 263 311.1	1	—	—	U*?
CRS	73 263 303.1 311.1	1	—	—	U*?
CRS	73 152 263 303.1 311.1	1	—	—	U*?
142$^{C/A}$ 311	73 263 311.1	—	1	—	U*
179	73 195 263 303.1 311.1	—	1	—	U*
184 264 291	73 263 303.1 311.1 316	1	—	—	U*
189	73 263 303.1 311.1	—	1	—	U*?
189	73 263 311.1	—	—	1	U*?
051 092 129$^{G/C}$ 174 183$^{A/C}$ 189 362	73 152 217 263 311.1	—	1	—	U2
051 129$^{G/C}$ 189 256 311	73 152 217 263 311.1 340	—	1	—	U2
343	73 150 263 311.1	1	—	—	U3
343 356 (390)	73 150 263 311.1	1	—	—	U3
343 356	73 150 195 263 303.1 311.1	—	—	—	U3
134 288 356	73 152 195 263 311.1	—	1	—	U4
179 356	73 150 195 263 311.1	1	—	—	U4
179 356	73 195 263 303.1 311.1	1	1	—	U4
224 270	73 150 199 263 279 311.1	1	—	—	U5
167 192 270 311 318 356	73 150 263 303.1 311.1	1	—	—	U5a
167 192 270 311 356	73 150 263 303.1 311.1	—	1	—	U5a
192 235 270 304	73 146 150 263 311.1	1	—	—	U5a
192 270	73 150 151 228 263 303.2 311.1	—	—	1	U5a
114$^{C/A}$ 192 270 294	73 263 303.1 311.1	1	—	—	U5a1
189 192 256 270 362	73 195 263 303.1 311.1	—	1	—	U5a1
192 256 270 291 (399)	73 263 303.1 311.1	1	—	—	U5a1
192 256 270	73 263 303.2 311.1	—	—	1	U5a1
189 256 270 362	73 185 204 263 303.1 311.1	1	—	—	U5a1a
256 270	73 263 311.1	—	2	—	U5a1a
256 270	73 146 263 284$^{A/C}$ 311.1	—	—	1	U5a1a
183$^{A/C}$ 187 189 192 270	73 150 195 200 263 311.1	1	—	—	U5a/b
093 111 189 270	73 150 263 311.1	—	—	1	U5b
114 189 192 270	73 140 150 263 311.1	—	1	—	U5b
189 270	73 150 263 311.1	1	—	—	U5b
051 172 219 311	73 263 311.1	1	—	—	U6
172 182$^{A/C}$ 183$^{A/C}$ 189 219 278	73 263 303.1 311.1	1	—	—	U6a
172 183$^{A/C}$ 189 219 278	73 263 303.1 311.1	1	—	—	U6a
172 219 235 278 355	73 146 263 303.1 311.1	1	—	—	U6a
172 219 271 278	73 152 263 303.1 311.1	1	—	—	U6a
172 174 188 219 311	73 263 311.1	1	—	—	U6b
172 219 261 311	73 263 303.1 311.1	1	—	—	U6b
309 318$^{A/T}$	73 152 263 303.2 311.1	—	—	1	U7
124 298 311 319	(72) 263 311.1	—	—	1	V
129 298	195 263 311.1	1	—	—	V
183$^{A/C}$ 189 259$^{C/G}$ 298	263 303.2 311.1	—	—	1	V

HVRI	HVRII	N	C	S	Hap
189 298	(72) 263 311.1	—	—	1	V
242	(72) 152 263 303.1 311.1	—	1	—	V ?
254 298	(72) 263 303.1 311.1	—	1	—	V
264 298	195 263 303.1 311.1	o	—	—	V
264 298	(72) 195 227 263 311.1	—	—	1	V
298	(72) 131 263 303.2 311.1	1	—	—	V
298	(72) 263 303.1 311.1	2	1	—	V
298	(72) 263 311.1	1	—	—	V
298	(72) 195 263 303.1 311.1	1	1	—	V
298 311	(72) 263 311.1	1	—	—	V
298	73 263 311.1	—	1	—	V ?
298 344	73 263 311.1	—	1	—	V ?
192 223 292 325	73 189 194 195 204 207 263 303.1 311.1	—	1	—	W
223 292 311	73 189 195 204 207 263 311.1	1	—	—	W
223 292 362	73 189 194 195 204 207 263 303.1 311.1	1	—	—	W
048 189 223 255 278	73 146 153 195 225 263 303.1 311.1	—	—	1	X
172 183$^{A/C}$ 189 278	73 146 152 185 263 303.1 311.1	—	—	1	X ?
183$^{A/C}$ 189 223 260 278	73 153 195 225 226 263 303.1 311.1	—	1	—	X
189 223 255 278	73 146 153 172$^{T/G}$ 195 225 226 263 303.1 311.1	—	1	—	X
189 223 255 278	73 146 153 195 225 226 263 311.1	—	1	—	X

Variant positions from the Cambridge Reference Sequence (CRS) of Anderson *et al*. (1981) are shown (minus 16000 in HVRI). Transversions are further specified by the appropriate base change. Haplogroups where the classification presents ambiguity are assigned with a question mark (?). In the case of positions 303 and 311, the presence of one, two or three Cs is referred by .1, .2 and .3, respectively, following the base position. In some cases, additional positions outside the referred regions are indicated inside brackets.