tapraid5/c0-phan/c0-phan/c01202/c00619d02a | heckt | S=17 | 9/9/02 | 11:04 | Art: 00-1074

Mitochondrial DNA Affinities at the Atlantic Fringe of Europe

Ana M. González,[1]* Antonio Brehm,[2] José A. Pérez,[1] Nicole Maca-Meyer,[1] Carlos Flores,[1] and Vicente M. Cabrera[1]

[1]*Departamento de Genética, Universidad de La Laguna, 38271 La Laguna, Tenerife, Spain*
[2]*Human Genetics Laboratory, Center of Macaronesian Studies, University of Madeira, 9000 Funchal, Portugal*

KEY WORDS Western Europe; Portugal; Galicia; haplotypes; clines

ABSTRACT Mitochondrial DNA analysis of Atlantic European samples has detected significant latitudinal clines for several clusters with Paleolithic (H) and Neolithic (J, U4, U5a1, and U5a1a) coalescence ages in Europe. These gradients may be explained as the result of Neolithic influence on a rather homogeneous Paleolithic background. There is also evidence that some Neolithic clusters reached this border by a continental route (J, J1, J1a, U5a1, and U5a1a), whereas others (J2) did so through the Mediterranean coast. An important gene flow from Africa was detected in the Atlantic Iberia. Specific sub-Saharan lineages appeared mainly restricted to southern Portugal, and could be attributed to historic Black slave trade in the area and to a probable Saharan Neolithic influence. In fact, U6 haplotypes of specific North African origin have only been detected in the Iberian peninsula northwards from central Portugal. Based on this peculiar distribution and the high diversity π value (0.014 ± 0.001) in this area compared to North Africa (0.006 ± 0.001), we reject the proposal that only historic events such as the Moslem occupation are the main cause of this gene flow, and instead propose a pre-Neolithic origin for it. Am J Phys Anthropol 119:000–000, 2002.

Anthropological remains (Kozlowski, 1982) indicate that modern humans arrived in Eastern Europe around 40,000 years ago. They most probably came from the Near East, although a more or less simultaneous arrival from North Africa to the Iberian Peninsula has also been suggested (Ferembach, 1986). From the size and localization of their settlements, it has been deduced that European populations were sparse during the Paleolithic, suffering important constrictions and expansions due to climatic changes (Mellars, 1998). Notable densities were reached in the Mesolithic, mainly in southern areas, although population sizes did not increase greatly until the introduction of farming into Europe from the Near East in the Neolithic (Whittle, 1998). It has been proposed that these Neolithic farmers gradually displaced the indigenous Mesolithic hunters and gatherers (Ammerman and Cavalli-Sforza, 1973).

Recently, the molecular variation in mitochondrial DNA (mtDNA) has been used to give a maternal genetic perspective of this European demographic history. Two main approaches have been used to accomplish this goal. Global quantitative analyses, such as those based on population genetic distances (Comas et al., 1997), and spatial autocorrelation (Simoni et al., 2000), have been congruent in showing a certain degree of population homogeneity in Central Europe and a detectable east-west clinal variation for the Mediterranean area. All these data have been interpreted as in agreement with the Neolithic demic displacement hypothesis, although it has been pointed out that mtDNA does not offer the clear picture obtained by the use of autosomal and Y-chromosome markers. On the other hand, the establishment of phylogeographic relationships among mtDNA haplotypes has been a valuable tool to distinguish several major mtDNA lineages in west Eurasian populations (Macaulay et al., 1999).

When the divergence times for the European founders of these clusters were calculated, it was found that the bulk of them originated in pre-Neolithic times, whereas those dated in the Neolithic were only around 15%, considerably reducing the Near East demic impact into Europe (Richards et al., 1996; Torroni et al., 1998). Furthermore, preliminary studies on mtDNA diversity in the Iberian Peninsula compared with other European samples

Grant sponsor: DGES; Grant number: PB93-0558; Grant sponsor: Gobierno de Canarias; Grant number: COF1999/019.

*Correspondence to: Ana M. González, Departamento de Genética, Facultad de Biología, Universidad de La Laguna, 38271 Tenerife, Spain. E-mail: amglez@ull.es

Received 27 September 2000; accepted 17 July 2002.

DOI 10.1002/ajpa.10168
Published online in Wiley InterScience (www.interscience.wiley.com).

(Côrte-Real et al., 1996) showed greater diversity levels in Southern than in Northern European populations. In addition, diversity levels and the geographic distribution of some autochthonous mtDNA European lineages revealed that a Paleolithic population expansion from Southwestern to Central and Northern Europe probably occurred after the last glaciation (Torroni et al., 1998). Finally, moderate levels of mtDNA gene flow from North Africa to the Iberian Peninsula have also been detected (Côrte-Real et al., 1996; Rando et al., 1998; Rocha et al., 1999).

All the molecular data are congruent with the idea that the cultural and demic impact from the Near East in the Neolithic was very attenuated on the Atlantic edge of Europe, in particular the Iberian Peninsula. For this reason, it should be at this Atlantic fringe where the Paleolithic contribution to modern European populations is expected to be most conspicuous. In this paper, we studied the mtDNA variation in samples from the whole Western border of the Iberian Peninsula to determine the relative affinities and differences among them and other European populations. Using both population and molecular phylogenetic approaches, we further searched for molecular clues of the past demographic relationships in this Atlantic area.

MATERIALS AND METHODS

Samples

In this study, 299 individuals from Portugal and 43 from Galicia (Northwestern Spain) were analyzed. Blood samples from the Portuguese were taken from voluntary Army donors. Each one was assigned to a different geographic area, according to the origin of their most ancient known maternal ancestor, as follows: North Portugal (NP), 84; Center Portugal (CP), 78; and South Portugal (SP), 137. These numbers were respectively enlarged with 100, 84, and 59 individuals from another study (Pereira et al., 2000). Whenever Portugal was used as a total, 54 individuals from throughout Portugal were added (Côrte-Real et al., 1996). Galician samples were collected from voluntary donors resident in the Canary Islands but with all their known maternal ancestors from Galicia. To enlarge this sample, 92 unrelated individuals from the same region studied in Salas et al. (1998) were pooled (GA). When the Iberian Peninsula (IP) as a total is analyzed, it refers only to the pool of the above-mentioned Atlantic regions.

To determine the relative affinities of the IP population with other Atlantic areas, the following samples were included in the analysis: a pool of 255 already published northwest African (NA) mtDNA sequences, comprising samples from the following areas: 85 Berbers from northern Algeria analyzed for mtDNA sequences (Côrte-Real et al., 1996) and for cluster-diagnostic RFLP polymorphisms (Macaulay et al., 1999); 60 Moroccan Berbers, 32 non-Berber Moroccans, 25 west-Saharans, and 30 Mauritanians as detailed in Rando et al. (1998); and 23 Tuareg from Watson et al. (1997). England (EN) was represented by 92 individuals from Wales, 69 from Cornwall (Richards et al., 1996), and 100 UK Caucasians (Piercy et al., 1993). Scotland (SC) was represented by 874 sequences from the mainland (Helgason et al., 2001). Continental Europe representatives included: 180 French (FR), 20 from the mtDNA Concordance database (Miller et al., 1996), 50 from Rousselet and Mangin (1998), and 110 from Cali et al. (2001); 404 Germans were subdivided as: South Germany (SG), including 67 individuals from Hofmann et al. (1997) and 199 described in Lutz et al. (1998); and North Germany (NG), comprising 105 German and 33 Danish as detailed in Richards et al. (1996); also included were 50 non-Saami Finns (FI) from Sajantila et al. (1995) and 323 Norwegians (NO) from Helgason et al. (2001). Finally, for the analysis of shared sequence types between regions, 90 additional North African sequences, 33 French sequences, and one British unpublished sequence (J. Larruga, personal communication) were included.

Sequencing

Total DNA was isolated as described in Rudbeck and Dissing (1998). PCR was amplified as in Pinto et al. (1996), and directly sequenced for both complementary strands as detailed in Rando et al. (1998). The sequences of 403 bp of the first hypervariable segment (HVSI) of the control region of the mtDNA, from position 15997–16399, according to the Cambridge Reference Sequence (CRS; Anderson et al., 1981), were determined and aligned. To discriminate for ambiguous haplogroup classification, all individuals, sorted by sequence motif as H or U* (Torroni et al., 1996), were also analyzed for polymorphic site 73 (Wilkinson-Herbots et al., 1996). This analysis involved amplifying a fragment embraced by primers L29 and H408 as in Vigilant et al. (1989), using the following PCR conditions: 94°C, 15 sec; 52°C, 15 sec; and 72°C, 15 sec, repeated for 35 cycles, and restricted with *Alw*44I. We also screened samples for the 12308G mutation by using a mismatched primer that generates a *Hinf*I site when the 12308G mutation is present (Torroni et al., 1996). When ambiguity persisted, the following additional sites were tested: *Alu*I site at np 7025 (Torroni et al., 1996), *Mse*I site at np 14766 (Lamminen et al., 1997), and *Nla*III site at np 4577 (Torroni et al., 1996). To obtain more information for other haplogroup classification, several amplified segments were digested and restricted as detailed in the respective references. Sites tested were: 2349 *Mbo*I, 3592 *Hpa*I, 8994 *Hae*III, 10394 *Dde*I, 10397 *Alu*I (Chen et al., 1995; Torroni et al., 1996), 4216 *Nla*III (Macaulay et al., 1999), and 10871 *Mnl*I (Quintana-Murci et al., 1999).

tapraid5/c0-phan/c0-phan/c01202/c00619d02a | heckt | S=17 | 9/9/02 | 11:04 | Art: 00-1074

Population comparisons

AQ: 1 For the AMOVA analysis (Excoffier et al., 1992), the populations studied were grouped into the following geographic areas: Northern (Norway and Finland), Western (Scotland and England), Central (north Germany, south Germany, and France), Iberian (Galicia, and north, center, and south Portugal), and northwest Africa. The significance tests of the variance components were obtained by comparison with a distribution of 1,000 values calculated by randomization, using the computer program ARLEQUIN, version 1.1 (Schneider et al., 1997). Pairwise population comparisons were estimated as distance values. It is known that genetic distances between populations give us a measure of their present affinities, but give no clue about past interactions. For nonrecombinant systems such as mtDNA variation, the phylogeographic approach that joins phylogenetic clusterings of molecules with their geographic distribution, diversity, and coalescence times (Forster et al., 1996) aims to cover this gap at a global scale, but has no significant level of quantification for the present relatedness between close populations. Several mean distances were used to deal with this issue, but with inevitable shortcomings. Averaged distances between sequences do not take into account the fact that nucleotide differences between sequences belonging to different molecular lineages could be the same. On the other hand, distances based only on the relative frequency of different molecular lineages between populations miss the sequence heterogeneity existing within lineages. Distances based on the frequency of identical sequences between populations, although not appropriate to detect possible ancient connections, seem the most adequate to fairly well detect relatively recent gene flow between populations. We chose this distance for our population comparisons. In this case, mtDNA was considered as one locus with as many alleles as different haplotypes detected, and the linearized F_{ST} distance based on allele frequencies was used as implemented in the ARLEQUIN program. In order to include other published samples in the comparisons, only the variation between positions 069–365 was used. Nucleotide diversity measured as π values (Nei, 1987) was calculated using the DnaSP program, version 3.0 (Rozas and Rozas, 1999). The neighbor-joining (NJ) tree (Saitou and Nei, 1987) was used to depict graphical relationships among populations.

Phylogeographic analysis

All sequences used in this study were sorted into molecular clusters (haplogroups), as defined in Richards et al. (2000), but keeping their geographic origins. Relative frequencies of haplogroups, and matches within clusters, among and between areas, were used in conjunction with their previously calculated ages (Richards et al., 2000) and distributions in Europe to refine the global affinities obtained at population level. Due to the relatively low frequency of matches within haplogroups, geographical patterns of haplogroup variation were tested by spatial correlation, using haplogroup frequency and latitude (represented by that of the main town in each sampled area) as variables.

RESULTS

Haplotype classification and distribution

The Appendix lists our 342 sampled sequences according to their haplogroup status. Within haplotypes classified in superhaplogroup H* by their HVSI sequences, 4.0% belonged to the U* group, 2.9% to superhaplogroup R*, 1.7% to the pre-HV cluster, and 0.6% to the HV* cluster according to the RFLP analysis. However, in comparisons with samples where this sorting was not possible, we also grouped them as H*. The haplogroup frequencies and sample sizes for the populations analyzed are given in Table 1. The haplotype with the reference sequence (CRS, Anderson et al., 1981) is the most abundant haplotype in all samples, although values range from 11.7% in northwest Africa to 21.7% in north Portugal. T1

As expected, sub-Saharan African influence, represented by haplotypes classified in L and M1 clusters, is important in northwest Africa (26.1%) but negligible in Europe, with the exception of south Portugal (11.7%). On the other hand, subhaplogroup U6, of North African origin (Rando et al., 1998), has a local presence in Europe, being detected only in northwest Iberian Peninsula. The differential geographic distributions of these sub-Saharan African and northwest African haplogroups in the Iberian Peninsula are statistically significant: L and M1 clusters are more abundant in south Portugal ($\chi^2 = 9.81$; $P < 0.01$), and U6 in northern areas ($\chi^2 = 5.83$; $P < 0.05$). Cluster U5, with coalescence ages in the early Upper Paleolithic, and a probable European origin (Richards et al., 2000), reaches its highest frequencies for its ancestral motives in Britain ($\chi^2 = 11.74$; $P < 0.001$) when compared to other continental areas such as Scandinavia, Germany, France, and the Iberian Peninsula. On the other hand, haplogroup J and the subhaplogroups T1, T2, U3, U5a1a, and perhaps U5a1 have been considered of Neolithic influence (Richards et al., 2000). The basal subhaplogroup J ($\chi^2 = 11.22$; $P < 0.001$) also has its greatest frequency in Britain. In addition, different J derivatives have different geographic distributions. For example, J1a ($\chi^2 = 8.10$; $P < 0.01$) and J1b1 ($\chi^2 = 13.16$; $P < 0.001$) are more abundant in northern areas including Scandinavia, Britain, and northern Germany, whereas J2, with representatives in all southern populations, is only present in the northern population of Scotland. In a similar vein, the highest frequencies for U5a1 are in the north ($\chi^2 = 8.03$; $P < 0.01$). Although not statistically testable due to their low frequencies, similar trends are also detectable for some subclusters be-

TABLE 1. Frequency of haplogroups and the cambridge reference sequence (crs) in the different populations analyzed[1]

	Population											
Haplogroup	FI	NO	SC	EN	NG	SG	FR	GA	NP	CP	SP	NA
CRS	0.120	0.201	0.164	0.191	0.157	0.154	0.183	0.215	0.217	0.198	0.209	0.117
H*-CRS	0.240	0.263	0.268	0.340	0.364	0.361	0.385	0.422	0.239	0.315	0.265	0.201
V	0.120	0.043	0.035	0.031	0.057	0.045	0.033	0.052	0.065	0.031	0.041	0.049
U1		0.003				0.004						0.003
U1a		0.003	0.006			0.004						
U2	0.020		0.008		0.007	0.004	0.014	0.015	0.005	0.019	0.005	
U3		0.019	0.019		0.014	0.015	0.005		0.022	0.006	0.010	0.034
U4		0.043	0.026	0.023		0.041	0.019	0.007	0.022	0.012	0.015	
U5		0.006	0.011	0.019			0.005		0.005			0.006
U5a		0.009	0.010	0.008		0.008	0.005		0.016	0.006	0.005	
U5a1	0.020	0.043	0.025	0.027	0.043	0.019	0.014	0.015	0.011	0.006	0.015	0.009
U5a1a	0.020	0.012	0.023	0.004	0.014	0.015	0.005		0.005	0.019	0.010	
U5b	0.040	0.012	0.014	0.004	0.014	0.034	0.009	0.015	0.027	0.019	0.015	0.023
U5b1	0.020	0.009										
U6								0.022	0.043	0.019		0.149
U7											0.005	
K	0.040	0.056	0.080	0.073	0.079	0.056	0.056	0.052	0.038	0.080	0.061	0.034
J	0.040	0.087	0.076	0.118	0.057	0.056	0.038	0.052	0.016	0.056	0.046	0.032
J1		0.009	0.008	0.004		0.011	0.005		0.011			0.003
J1a		0.009	0.006	0.011	0.050	0.008						0.006
J1b											0.015	0.003
J1b1		0.025	0.023	0.023		0.004	0.005	0.007	0.005		0.005	
J2			0.009			0.004	0.009	0.015	0.011	0.019	0.015	
T		0.019	0.027	0.023	0.007	0.023	0.028		0.016	0.019	0.005	0.006
T1	0.020	0.009	0.023	0.023	0.036	0.019	0.019	0.007	0.065	0.012	0.020	0.023
T2	0.060	0.043	0.043	0.019	0.021	0.053	0.056	0.022	0.049	0.037	0.031	0.009
T3		0.003	0.005		0.007					0.012	0.010	0.011
T4		0.006	0.002		0.014						0.005	
T5		0.012	0.002				0.009	0.007	0.011	0.006	0.005	0.003
B							0.005					
N1a			0.001							0.006	0.005	
N1b		0.003			0.007	0.008			0.005			0.003
I	0.100	0.025	0.042	0.038	0.007	0.030	0.014		0.027		0.005	
W	0.100	0.019	0.013		0.014	0.011	0.028	0.022	0.033	0.012	0.020	0.006
X		0.003	0.021	0.015	0.007	0.011	0.009	0.007		0.019	0.036	0.011
C							0.005					
D					0.007							
M1			0.001							0.006	0.010	0.023
N/M/L3	0.040	0.003	0.007	0.004	0.007	0.004	0.023	0.015	0.011	0.025	0.020	0.020
L3b								0.007		0.006	0.005	0.023
L3d												0.009
L3e										0.006	0.015	0.011
L1a												0.006
L1b					0.007			0.015		0.012	0.020	0.066
L1c				0.004							0.005	0.006
L2			0.001				0.014	0.007	0.022	0.019	0.041	0.097
Sample size	50.0	323.0	874.0	262.0	140.0	266.0	213.0	135.0	184.0	162.0	196.0	349.0

[1] FI, Finland; NO, Norway; SC, Scotland; EN, England; NG, north Germany; SG, south Germany; FR, France; GA, Galicia; NP, north Portugal; CP, center Portugal; SP, south Portugal; NA, North Africa.

longing to haplogroup T. Thus, T4 shows greater frequencies in the north, whereas T5 is more prevalent in the south. These results show that the western fringe of Europe is not a homogeneous mitochondrial landscape; on the contrary, Paleolithic and Neolithic linages are heterogeneously distributed throughout the studied area.

Population affinities

AQ: 2 The AMOVA results showed that there is a very low level of mtDNA genetic structure in the geographic area analyzed. The bulk of the total variance (99.7%) is due to differences among individuals within populations, 0.1% attributable to differences among populations within areas, and 0.2% to heterogeneity among areas. In spite of this, the geographic partition among areas is statistically significant at the 0.01 level. Figure 1a shows the relationships among populations, based on F_{ST} values listed in Table 2. Northwest Africa is significantly different from all other populations; curiously, Finland appears as its most related pair. However, the cause for this peculiarity is that both have the lowest frequencies for the CRS sequence, which accounts for 87% of all matches used in the calculation of F_{ST} distances, masking the effect of other less frequent haplotypes.

F1 T2

Distances calculated without the CRS (Table 2) significantly increased the heterogeneity between populations, and showed somewhat different relationships among them in the generated NJ tree (Fig. 1b). As expected, northwest Africa continues to be the most divergent population, but congruently, now their closest relatives are the Iberian Peninsula

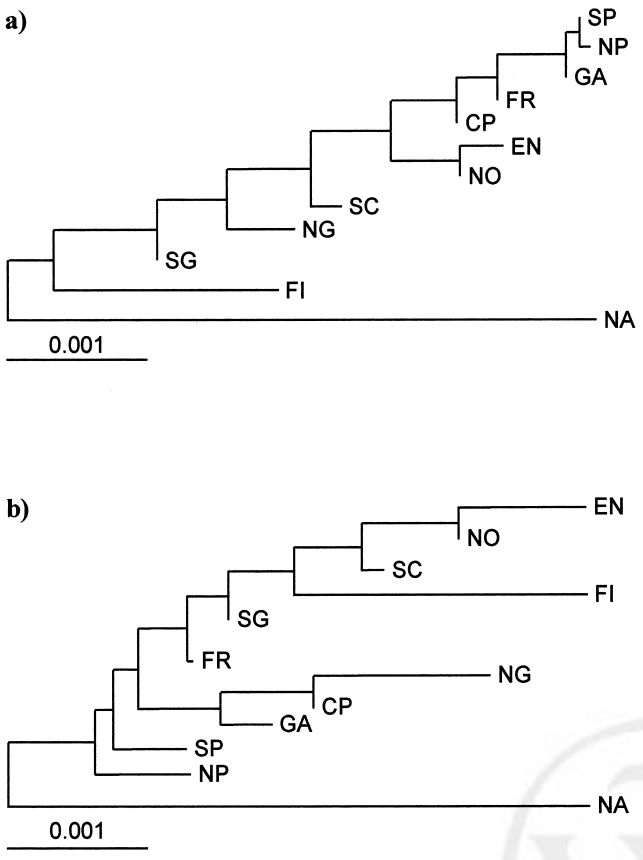

Fig. 1. Neighbor-joining trees constructed from F_{ST} distance matrix, (**a**) using all haplotypes and (**b**) excluding CRS. CP, center Portugal; EN, England; FI, Finland; FR, France; GA, Galicia; NA, North Africa; NG, North Germany; NO, Norway; NP, North Portugal; SC, Scotland; SG, South Germany; SP, South Portugal.

samples. In fact, the latter are, globally, more related to the Central European samples than to England and Scotland. However, center Portugal and north Germany cluster together in spite of their large geographic distance. Again, an inspection of the matches between them shows that this affinity is due to the high frequencies in both for common haplotypes, mainly in H* and K, and not to specific matches not shared with other areas.

Within-haplogroup population affinities

Sequence matches distribution within haplogroups. In order to keep adequate sample sizes within haplogroups, we analyzed matches only for the large geographic areas considered in the AMOVA analysis. Table 3 indicates that matches due to the CRS are the overwhelming majority of matches between areas (87.3 ± 3.0%). Furthermore, all areas share matches due to the basic and more frequent haplotypes for the following clusters: H*, J, K, T, T1, U5b, and V, that together summed up an additional 10.0 ± 2.5%. Only 2.7 ± 1.6% of the matches are specific between areas.

With rare exceptions such as the central motives of U5 and J1a, the Iberian Peninsula participates in

TABLE 2. Pairwise linearized F_{ST} for all haplotypes (above diagonal) and excluding Cambridge Reference Sequence (below diagonal)[1]

	FI	NO	SC	EN	NG	SG	FR	GA	NP	CP	SP	NA
FI		0.0042	0.0018	0.0056	0.0029	0.0007	0.0025	0.0064*	0.0071*	0.0043	0.0061*	0.0061*
NO	0.0024		0.0008	0.0000	0.0024	0.0012	0.0010	0.0011	0.0018	0.0002	0.0010	0.0088***
SC	0.0023	0.0007*		0.0012*	0.0018*	0.0003	0.0010	0.0020*	0.0031**	0.0008	0.0021*	0.0063***
EN	0.0057*	0.0007	0.0018**		0.0019	0.0016	0.0016	0.0013	0.0028*	0.0005	0.0024	0.0083***
NG	0.0049*	0.0030**	0.0032***	0.0029*		0.0003	0.0022	0.0024	0.0046*	0.0002	0.0029	0.0056**
SG	0.0016	0.0010	0.0008*	0.0022**	0.0014		0.0000	0.0019	0.0019	0.0003	0.0011	0.0044***
FR	0.0020	0.0022**	0.0017**	0.0033**	0.0037***	0.0000		0.0005	0.0007	0.0000	0.0005	0.0054**
GA	0.0040*	0.0029**	0.0020*	0.0030*	0.0023	0.0014*	0.0015*		0.0000	0.0000	0.0000	0.0083***
NP	0.0042*	0.0037***	0.0032***	0.0048***	0.0051***	0.0009	0.0014*	0.0016*		0.0000	0.0000	0.0086***
CP	0.0032	0.0015*	0.0012*	0.0020*	0.0003	0.0003	0.0009	0.0000	0.0012		0.0000	0.0067**
SP	0.0040*	0.0026*	0.0021**	0.0044***	0.0032**	0.0003	0.0014*	0.0017*	0.0012*	0.0006		0.0080***
NA	0.0094***	0.0077***	0.0071***	0.0084***	0.0073***	0.0053***	0.0048***	0.0057***	0.0054***	0.0055***	0.0055***	

[1] Abbreviations as in Table 1.
* $P < 0.05$.
** $P < 0.01$.
*** $P < 0.001$.

tapraid5/c0-phan/c0-phan/c01202/c00619d02a | heckt | S=17 | 9/9/02 | 11:04 | Art: 00-1074

TABLE 3. Frequencies of matches (× 10^5) between areas for Cambridge Reference Sequence (CRS), common haplotypes (present in all populations), and rare haplotypes (shared by two or more populations), and percentage (%) of matches between areas for each class[1]

	Populations									
	NE × BR	NE × CE	NE × IP	NE × NA	BR × CE	BR × IP	BR × NA	CE × IP	CE × NA	IP × NA
CRS										
Total	3,234.0	3,137.0	4,062.0	2,236.0	2,800.0	3,626.0	1,996.0	3,517.0	1,936.0	2,507.0
%	81.8	84.9	90.5	88.0	84.3	89.9	87.1	90.3	86.3	90.2
Common										
H*-CRS	98.0	110.0	81.0	85.0	101.0	74.0	74.0	101.0	121.0	68.0
V	28.0	35.0	38.0	55.0	21.0	23.0	34.0	29.0	42.0	46.0
U5b	1.0	3.0	3.0	5.0	3.0	3.0	5.0	5.0	6.0	8.0
K	70.0	60.0	69.0	45.0	73.0	85.0	61.0	71.0	49.0	61.0
J	307.0	164.0	100.0	81.0	159.0	97.0	78.0	52.0	42.0	25.0
T	7.0	1.0	2.0	2.0	2.0	5.0	4.0	1.0	0.0	1.0
T1	12.0	12.0	10.0	18.0	13.0	11.0	20.0	11.0	19.0	16.0
Total	524.0	384.0	303.0	290.0	371.0	298.0	275.0	269.0	279.0	226.0
%	13.2	10.4	6.8	11.4	11.2	7.4	12.0	6.9.0	12.5	8.1
Rare										
H*-CRS	37.8	32.1	16.5	3.1	42.7	33.5	9.3	27.0	8.3	9.0
V	0.9	2.6	1.5	7.7	0.4			0.9	4.6	2.7
U1a		0.4								
U2					0.3	1.2		1.1		
U3	0.7				4.3	3.9	1.5	4.4	2.3	1.6
U4	14.9	0.9	7.3		0.3	4.8		4.4		
U5	1.2	0.4		0.8			1.3			
U5a	0.2				0.1	0.1		0.4		
U5a1	7.3	9.1	3.3		9.1	3.3	1.0	3.1		2.4
U5a1a	7.3	9.1	3.3		9.1	3.3	1.0	3.1		2.4
U5b	0.2							0.2		
U6										13.7
K	3.3	0.9	1.5		5.1	2.9		0.7		
J	0.9	0.9	1.8		0.6	1.8		2.9		0.4
J1	0.9				1.4	1.2		0.9		
J1a	2.4	6.9		3.1	5.8		2.5		7.4	
J1b1	20.1		3.3			1.8				
J2					1.1	2.0		1.8		
T	2.4	3.9	0.7	0.8	4.1	0.2	1.3	0.2	2.3	
T1						0.8		0.4		0.4
T2	56.4	76.2	72.3		37.7	34.0	0.3	46.0		
T3			0.4		0.3	0.5		0.4		
T4					0.3					
T5			2.9							
N1b								0.2	0.5	0.4
I	36.1	19.1	3.7		18.8	4.6		2.9		
W	3.8	11.3	2.2		3.7	1.4		2.4		0.8
X	1.7	0.9	1.8		2.6	4.9	0.3	2.9		
M1										0.8
N/M/L3					2.1	2.4	1.3	2.7	1.4	2.4
L3b										0.4
L3e										2.4
L1b										0.4
L2						0.4	1.5	0.4		7.1
Total	198.0	175.0	122.0	15.0	150.0	109.0	21.0	109.0	27.0	47.0
%	5.0	4.7	2.7	0.6	4.5	2.7	0.9	2.8	1.2	1.7

[1] NE, North Europe; BR, Britain; CE, Central Europe; IP, Iberian Peninsula; NA, North Africa.

all the matches between northwest Africa and Europe. Quantitatively, the Iberian Peninsula shares more matches with Britain than with the Continent, in particular for H*, I, J1b1, T1, and X, whereas matches in clusters J, T2, U5a, and V approximate the Iberian Peninsula to the Continent. In fact, Britain seems more related to Central Europe for H*, J1a, J2, T, U3, and X, and to the North for I, J, J1b1, U4, U5, U5a1a, and V.

The match connection between northwest Africa and the Atlantic Iberian Peninsula deserves special attention. At least two different African influences can be detected on the basis of the geographic areas and haplogroups implied. On the one hand, there are matches due to haplotype sharing for sub-Saharan Africa lineages L and M1, affecting mainly southern Portugal. On the other, there are matches within the specific northwest African sub-haplogrup U6 (Rando et al., 1998) that occur only in the northern areas. Furthermore, U6 haplotypes were not detected in southern Portugal (Table 1).

Latitudinal correlations. Significant latitudinal correlations were found for several haplogroups (Table 4 and Fig. 2). The H* and J2 haplogroups show strong negative correlations. However, the extensions of the clines are different. The highest frequencies for H* are found in Galicia, decreasing to the north but also to the south, whereas the J2 gradient is from south Portugal to the north. On the other hand, significant positive correlations were found for subcluster U4, all the Js excluding J1b and J2, and the recent subclusters U5a1 and U5a1a. Although gradients for clusters belonging to Js and U5 run

T4-F2

tapraid5/c0-phan/c0-phan/c01202/c00619d02a | heckt | S=17 | 9/9/02 | 11:04 | Art: 00-1074

parallel, there is an evident difference in their behavior. Finally, some less frequent Caucasian subhaplogroups were only detected in localized areas such as U1a and U5b1 in the north, or U7 and J1b in the south.

DISCUSSION

Global analysis shows European populations rather homogeneous at the mtDNA level, differing significantly only from northwest Africa. Reanalysis without the CRS, which is responsible for the majority of matches between populations, significantly increases the spatial differentiation. This shows that all European populations shared an ancestral common background, but that more recent movements had a different impact on different areas.

The haplogroup distribution of matches shows that the only young subcluster with a widespread representation is T1, to which coalescence has been dated at recent Neolithic times (Macaulay et al., 1999). On the contrary, there was no overall sharing for U5 haplotypes, i.e., the most ancient subhaplogroup in Europe (Richards et al., 1996, 1998). The basic haplotypes, 12308G and 16270T, have only been detected in Northern Europe, Britain, and northwest Africa. Since the earliest human settlements in Norway and Finland were dated around 13,000–9,000 years ago (Mellars, 1998), their presence in this area is most probably the result of an upward migration from more southern places. Although of Neolithic introduction in Europe, the northern pattern for basal J can be explained in the same way.

The significant latitudinal gradients found here for several mtDNA clusters with very different coalescence and expansion ages in Europe (Richards et al., 2000) were also detected with synthetic maps

TABLE 4. Correlation coefficients (R) and two-tailed significance values (P) obtained for widespread mtDNA clusters

Cluster	N[1]	R	*P*
H* (without CRS)[2]	8	−0.97	<0.001
J2	11	−0.85	0.001
Js (without J2 and J1b)	11	0.78	0.005
U5a1 plus U5a1a	11	0.81	0.001
U4	11	0.64	0.036

[1] Number of pairs.
[2] Cambridge Reference Sequence.

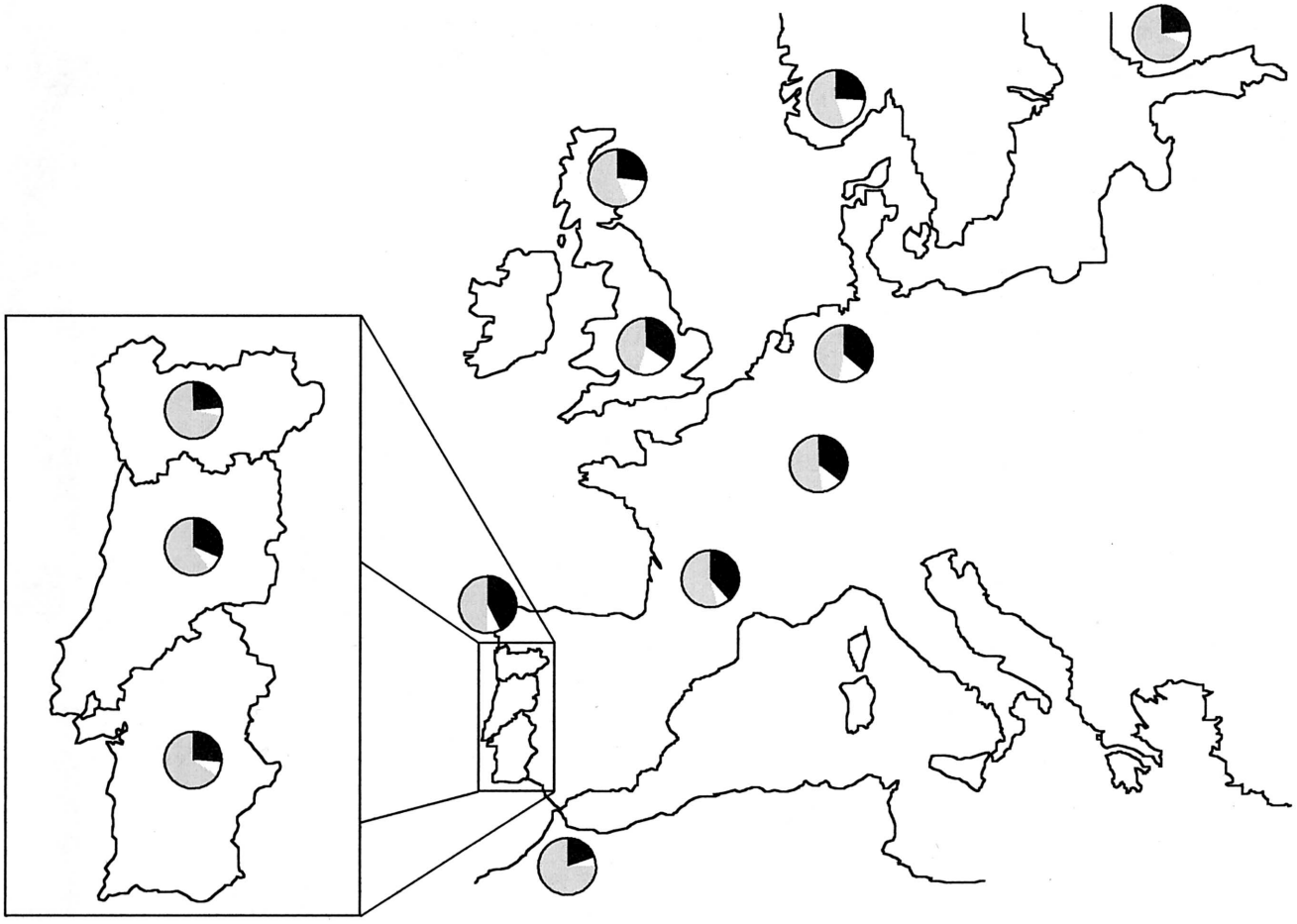

Fig. 2. Frequency distribution for some mtDNA haplogroups that showed significant latitudinal clines. Pie charts represent populations from Table 1. Black indicates frequency of haplogroup H* without CRS. White includes frequencies of Js (excluding J2 and J1b) and U5a1 plus U5a1a. Gray represents frequency of rest of haplogroups.

obtained from classical markers (Cavalli-Sforza et al., 1994) and for Y chromosome haplogroups (Lucotte and Loirat, 1999; Rosser et al., 2000), which have been interpreted as a result of the Neolithic expansion in Europe. A high frequency of H* is the main characteristic of the Basques, considered a European Paleolithic isolate (Bertranpetit et al., 1995), and possibly it was characteristic of all Paleolithic European settlers. The geographic gradient detected for this main haplogroup may be the result of a differential impact of posterior migrations on a common Paleolithic substrate in Europe. As southwestern areas were comparatively more populated than the north (Mellars, 1998), Neolithic or more recent waves from the East, where H* has lower frequencies, could have comparatively more influence in northern regions. On the basis of archaeological evidence, it has been proposed that the Neolithic in Europe had both a continental and a Mediterranean coastal introduction (Renfrew, 1987). This might explain why Portugal and northwest Africa do not follow the H* gradient, since a strong Neolithic influence on Portugal, in contrast with northern Iberia, is well-documented (Martí-Oliver, 1998). For the same reason, the negative gradients for J2 and the exclusive southern presence of J1b may be explained as signatures of the Mediterranean Neolithic wave into the Atlantic border. On the other hand, the significant positive correlations found for the rest of the Js, and the younger clusters of U5a1 and U5a1a, are better explained as the result of the continental penetration of the Neolithic. Finally, Portugal deserves some particular comments. In comparisons within Europe, north Portugal has the highest frequency for T1 ($\chi^2 = 16.40$; $P < 0.001$), and is the only region in the Iberian Peninsula where J1 has been detected. At the other end, the main peculiarity of south Portugal is that this area harbors the highest frequency of X in Europe ($\chi^2 = 6.90$; $P < 0.01$). With respect to northwest Africa, the geographically localized distribution of matches and haplotypes of sub-Saharan African and northwest African origin in the Iberian Peninsula is noteworthy. This distribution cannot be totally explained by a historic genetic influence from the Moslem occupation (Pereira et al., 2000). During that time, the haplotype composition of northwest Africa had to be similar to that of the present, and for this reason, sub-Saharan African L and northwest African U6 haplotypes should be uniformly distributed in the Iberian Peninsula. It has been pointed out that one of the most important demic influences on northwest African populations from the Sahara occurred around 7,000 years ago, with the expansion of the Neolithic culture that flourished at that time in the Sahara (Dutour et al., 1994). If we admit that this expansion brought, for the first time, sub-Saharan African haplotypes into the Maghreb, it seems plausible that they could also have reached the Iberian Peninsula across the Gibraltar Strait, meaning that the L haplotypes detected in southern Portugal could be the result of African Neolithic influence in this region. The presence of U6 only in the northern Iberian Peninsula could represent the remnants of a pre-Neolithic expansion from northwest Africa to the Iberian Peninsula.

Some molecular data are congruent with this picture. The diversity value for U6 is significantly higher in the Iberian Peninsula (0.0135 ± 0.0014) than in northwest Africa (0.0055 ± 0.0007), a finding that cannot be explained by a recent and limited gene flow. Most probably, other undetected haplogroups also participated in this pre-Neolithic northwest Africa gene flow. It seems, however, that these African waves did not reach farther than Galicia, as haplotypes belonging to the U6 lineage have not been found in other European samples (Richards et al., 2000). However, with respect to the sub-Saharan Africa lineages, the recent history of the Black slave trade carried out by the Portuguese (mainly in the 15th and 16th centuries), with a well-documented import in southern Portugal (Godinho, 1983), could also be a plausible alternative to explain the presence of these African haplotypes in this region (Pereira et al., 2000). To test this possibility, we compared the proportion of sub-Saharan Africa haplotype matches between the Iberian Peninsula and northwest Africa (0.75%) with those of the Iberian Peninsula and a sample of sub-Saharan Africans from the Gulf of Guinea. The sample includes 45 Bubis from Bioko and 49 individuals from Sâo Tomé (Mateu et al., 1997), 32 Yoruba (Watson et al., 1997), and 72 Equatorial Guineans (Pinto et al., 1996; J. Larruga, personal communication). The percentage obtained (0.35%) is roughly half of the former, and in addition, the majority of them (97%) are also shared with northwest Africa, although matches between sub-Saharan and North African samples are only 0.95%. These results suggest that, although both prehistoric and historical influences likely contributed to the sub-Saharan African haplotype pool present in the Iberian Peninsula, the former seems to be more important. Our results are in agreement with the gene flow (19.5%) from northwest Africa to the Iberian Peninsula estimated in a recent study of variation in the autosomic CD4 locus (Flores et al., 2000b), and with the evidence of northwest African male input in Iberia calculated at around 20%, using the relative frequency of northwest African Y-chromosome-specific markers in Iberian samples (Flores et al., 2000a). Furthermore, our results clearly reinforce, extend, and clarify the preliminary clues of an important mtDNA contribution from northwest Africa into the Iberian Peninsula (Côrte-Real et al., 1996; Rando et al., 1998; Flores et al., 2000a; Rocha et al., 1999). On the basis of the L1b frequencies detected in Spanish and Portuguese samples (2–3%) and those found in western Africa (10–30%), a significant influence (at least 10%) of North Africans in the Iberian gene pool has also been admitted (Rocha et al., 1999).

tapraid5/c0-phan/c0-phan/c01202/c00619d02a | heckt | S=17 | 9/9/02 | 11:04 | Art: 00-1074

In a similar way, and discarding possible genetic drift effects, our own data allow us to make minimal estimates of the maternal African pre-Neolithic, Neolithic, and/or recent slave trade input into Iberia. For the former, we consider only the mean value of the U6 frequency in northern African populations, excluding Saharans, Tuareg, and Mauritanians (16%), as the pre-Neolithic frequency in that area, and the present frequency in the whole Iberian Peninsula (2.3%) as the result of the northwest African gene flow at that time. The value obtained (14%) could be as high as 35% using the data of Côrte-Real et al. (1996), or 27% with our north Portugal sample. In the same vein, the Saharan Neolithic gene flow can be estimated as 13%, taking the actual frequencies for the sub-Saharan African haplogroups (51%) in southern northwest African samples (Tuareg, Saharans, and Mauritanians) as the frequency of the African Neolithic, and that of the Iberian Peninsula (6.8%) as the result of the putative Neolithic maternal gene flow. This value could rise to 23% when only south Portugal is taken into account. However, if we admit a recent 10% of slave trade input into this region, as historically documented (Godinho, 1983), 13% would be left for the putative Saharan Neolithic gene flow. In spite of the simplicity of the method, different estimates give congruent results, that in turn give molecular support to the earlier hypothesis of prehistoric northwest African influences in the Iberian Peninsula, proposed on the basis of archaeological and anthropological affinities (Barandiarán, 1998).

ACKNOWLEDGMENTS

We thank H. Bandelt for his help and comments. This study was supported by grants of the DGES (PB93-0558) and Gobierno de Canarias (COF1999/019) to V.M.C. The authors are grateful to the Por-

APPENDIX. Hypervariable segment I (HVSI) haplotypes and restriction fragment length polymorphisms (RFLPs) in Portuguese and Galician samples[1]

	RFLPs														
	7 3	2 3 4 9	3 5 9 2	4 2 1 6	4 5 7 7	7 0 2 5	8 9 9 4	10 3 9 4	10 3 9 7	10 8 7 3	11 2 0 8	4 7 6 6	Samples			
HVSI type	w	j	h	q	q	a	e	c	a	y	g	u	GA	NP	CP	SP
H*																
CRS																1
CRS	−										−		4	14	19	24
CRS	+										−	−		4		
CRS	+			−							−	+				2
CRS	+										+			1	1	
093	−										−				1	
114	−										−		1			
126	+			+		+		−	−	+	−	+				1
129	−										−		1	1		1
141	−										−					1
147	−										−				1	
148	−										−		1			
153	−										−		1			
172	−										−				1	
179	+										+					1
189	−										−	−	2		2	2
189	+										−	−		1		
189	+										+					1
221	−										−	−				1
222	−										−					1
240	−										−				1	
261	−										−					1
262	−										−					1
264	−										−				1	
265C	−										−			1		
274	−										−					1
278	+										−	−				1
278	+										+				1	
289	−										−					1
291	−										−					1
292	+			−							−	+				1
293	−										−		2	1		
304	−										−		1	2	4	1
311	−										−	−		1	2	1
311	+										+				1	
320A	−										−					2

APPENDIX. (Continued)

	RFLPs												Samples			
	7 3	2 3 4 9	3 5 9 2	4 2 1 6	4 5 7 7	7 0 2 5	8 9 9 4	1 0 9 4	1 0 3 9 7	1 0 8 9 3	11 2 3 0 8	4 7 6 6				
HVSI type	w	j	h	q	q	a	e	c	a	y	g	u	GA	NP	CP	SP
320	–										–				2	1
335	–										–			1		
344	–										–			1		1
355	–										–					1
362	–										–	–		2		2
093 189	–										–			2		
093 263	–										–		1			
093 270	–					–					–					1
093 319	–										–				1	
111 311	–										–		1			
124 220	–										–				1	
124 362	–										–			1		
129 242	–										–			1		
129 362	–										–			1		1
188A 189	–										–					1
189 262	–										–				1	
189 356	–										–				1	1
192 262	–										–					1
213 304	–										–		1			
220T 311	–										–		1			
221 291	–										–		1			
235 291	–										–		1	2		
269 270	–					–					–					1
278 311	–										–	–				1
278 311	–										–	+				1
292 318	–			–							–	+		1		
293 311	–										–					1
311 319	+										+					1
093 189 293	–										–				1	
093 189 311	–										–					1
114 311 362	–										–				1	
126 264 362	–										–				1	
126 361 362	–										–					1
129 344 362	–										–		1			
172 192 311	–										–			1		
172 233 362	–										–				1	
180 219 221	–										–				1	
189 356 362	–										–					1
213 254 304	–										–		1			
217 218 311	–										–					1
235 291 293	–										–		1			
126 291A 355 362	–										–					1
													22	39	47	66
V (298)																
298					–											2
153 298					–								1			
189 298					+											1
239 298					–									1		
240 298					–								1	1		
256 298					+											1
258 298					–									1		
298 311					–									1		
298 319					+								1			
240 298 311					–								1			
													4	4	0	4
U2 (051 129C)																
051 129C 189 362													2	1	1	1
													2	1	1	1
U3 (343)																
343 390															1	1
343 356 390														2		
													0	2	1	1
U4 (356)																
356														1		2
179 356														1		1
													0	2	0	3

APPENDIX. (Continued)

	RFLPs												Samples			
	7 3	2 3 4 9	3 5 9 2	4 2 1 6	4 5 7 7	7 0 2 5	8 9 9 4	10 3 9 4	10 3 7 7	11 8 7 3	2 3 0 8	4 7 6 6				
HVSI type	w	j	h	q	q	a	e	c	a	y	g	u	GA	NP	CP	SP
U5a (192 270)																
192 270 304 311														1		
													0	1	0	0
U5a1 (192 256 270)																
192 256 270																1
192 256 270 362																1
													0	0	0	2
U5a1a (256 270 399)																
256 270																1
188 256 270 399															1	
													0	0	1	1
U5b (189 270)																
189 270																1
189 270						+								1		
189 192 270															2	
189 270 390														1		
093 189 192 270														1		
093 189 192 270 311																1
													0	3	2	2
U6 (172 219)																
051 172 219 311													1			
172 219 235 278 355														1	1	
172 189 219 239 278 362															2	
													1	1	3	0
K (224 311)																
224 311													2	1	4	2
093 224 311														1	1	
192 224 311													1			
224 234 311															2	
224 235 311																1
224 304 311														1		
224 311 320																3
093 110 224 311														1		
218 224 311 320																1
222 224 311 360																1
093 156 224 240C 311													1			
													4	4	7	8
J (069 126)																
069 126														1	3	1
069 126 150																1
069 126 241																1
069 126 256													1			1
069 126 311															1	
069 126 324 366 390																2
069 126 163 266 311															1	
													1	1	5	6
J1b (069 126 145 222 261)																
069 126 145 222 256 261 278																1
													0	0	0	1
J1b1 (069 126 145 172 222 261)																
069 126 145 172 222 261														1		1
													0	1	0	1
J2 (069 126 193)																
069 126 193 278															1	3
069 126 193 319 360													1			
													1	0	1	3
T (126 294)																
126 294														1	1	1
126 256 294 296														1		
093 126 189 294 296														1		
													0	3	1	1
T1 (126 163 186 189 294)																
037 126 163 186 189													1	1		
126 163 186 189 294														2		
126 163 171 186 189 294																1
126 163 186 189 261 294																2
126 163 186 189 249 294 311															1	1

tapraid5/c0-phan/c0-phan/c01202/c00619d02a | heckt | S=17 | 9/9/02 | 11:04 | Art: 00-1074

APPENDIX. (Continued)

	RFLPs												Samples			
	7 3	2 3 4 9	3 5 9 2	4 2 1 6	4 5 7 7	7 0 2 5	8 9 9 4	1 0 3 4	1 0 3 7	1 0 8 3	11 2 3 0	4 7 6 6				
HVSI type	w	j	h	q	q	a	e	c	a	y	g	u	GA	NP	CP	SP
126 163 186 189 294 304 311														1		
													1	4	1	4
T2 (126 294 304)																
126 294 304														1		
126 172 294 304														1		
126 294 296 304														2	2	1
051 126 294 296 304																1
093 126 294 296 304														1		
126 147 294 297 304														1		
126 294 296 304 344																1
126 294 296 304 365													1			
093 126 271 294 296 304														1		
													1	7	2	3
T3 (126 292 294)																
126 292 294 296 324															1	
126 171 292 294 296 324																1
													0	0	1	1
T5 (126 153 294)																
126 153 294														1		
126 153 274 293 294 296													1			
													1	1	0	0
N1a (147A/G 172 223 248 355)																
086 147A 172 223 248 320 355		−						+	−	+						1
													0	0	0	1
N1b (145 176G 223)																
145 176G 223 390														1		
													0	1	0	0
I (129 223 391)																
129 223 391														1		
129 172 223 311 391														2		
129 172 223 293 311 391														1		
													0	4	0	0
W (223 292)																
223 292																1
223 292 311															1	
223 292 311								−	−	+			1			
223 292 318														1		
223 292 320														1		
223 292 362														1		2
192 223 292 325														1		
129 209 223 292 295 311 343T		−						+	+	−						1
													1	4	1	4
X (189 223 278)																
189 223 278																4
189 223 278			−		+			−	−	+			1			
189 209 223 278			−		+			+	−	−						1
													1	0	0	5
M1 (129 189 223 249 311)																
129 189 223 249 311																1
129 186 189 218 223 249 311																1
													0	0	0	2
L3 (223)																
223							+	−	−	+						1
185 223																1
111 169 193 195 223 243 261 311													1			
													1	0	0	2
L3b (124 223 278 362)																
223 278 362																1
124 189 223 234 278															1	
													0	0	1	1
L3e (172 223 320)																
172 189 223 320																2
172 189 223 320			+													1
172 189 223 248 320															1	
													0	0	1	3
L1b (126 187 189 223 264 270 278 311)																
187 189 223 264 270 278 311																1

APPENDIX. (Continued)

	RFLPs															
								1	1	1	11					
		2	3	4	4	7	8	0	0	0	2	4				
		3	5	2	5	0	9	3	3	8	3	7				
	7	4	9	1	7	2	9	9	9	7	0	6				
	3	9	2	6	7	5	4	4	7	3	8	6	Samples			
HVSI type	w	j	h	q	q	a	e	c	a	y	g	u	GA	NP	CP	SP
104 187 189 223 270 278 289 293 311																1
104 187 189 215 223 270 278 289 293 311																1
126 187 189 223 264 270 278 293 311 357			+						+	−	−		1			
126 187 189 223 264 270 278 293 311 362									+	−	−		1			
126 187 189 223 264 270 278 293 311 362															1	
													2	0	1	3
L1c (129 187 189 223 278 294 311 360)																
129 187 189 223 265C 278 292 294 311 360																1
													0	0	0	1
L2 (223 278 390)																
189 223 278 390																1
223 278 294 309 390																2
148 223 278 294 355 390																1
093 223 260 278 294 309 390																1
093 223 278 294 309 311 320 390														1		
189 192 223 278 294 309 316 390															1	
189 223 234 249 278 294 295 390															1	
086 129 148 189 223 278 300 354 390															1	
												0	1	1	7	
Total samples												43	83	78	137	

[1] Position numbers as in Anderson et al. (1981). Those of HVSI less 16,000.

tuguese Army Chief of Staff for allowing the collection of blood samples covering the whole territory

LITERATURE CITED

Ammerman AJ, Cavalli-Sforza LL. 1973. A population model for the diffusion of early farming in Europe. In: Renfrew C, editor. The explanation of culture change: models in prehistory. London: Duckworth. p 348–358.

Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJH, Staden R, Young IG. 1981. Sequence and organization of the human mitochondrial genome. Nature 290:457–465.

Barandiarán I. 1998. El Paleolítico y el Mesolítico. In: Barandiarán I, Martí B, del Rincón MA, Maya JL, editors. Prehistoria de la Península Ibérica. Barcelona: Ariel, S.A. p 1–120.

Bertranpetit J, Sala J, 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.

Cali F, Le Roux MG, D'Anna R, Flugy A, DeLeo G, Chiavetta V, Ayala GF, Romano V. 2001. MtDNA control region and RFLP data for Sicily and France. Int J Legal Med 114:229–231.

Cavalli-Sforza LL, Menozzi P, Piazza A. 1994. The history and geography of human genes. Princeton: Princeton University Press.

Chen Y-S, Torroni A, Excoffier L, Santachiara-Benerecetti AS, Wallace DC. 1995. Analysis of mtDNA variation in African populations reveals the most ancient of all human continent-specific haplogroups. Am J Hum Genet 57:133–149.

Comas D, Calafell F, Mateu E, Pérez-Lezaun A, Bosch E, Bertranpetit J. 1997. Mitochondrial DNA variation and the origin of the Europeans. Hum Genet 99:443–449.

Côrte-Real HB, Macaulay VA, Richards MB, Hariti G, Issad MS, Cambon-Thomsen A, Papiha S, Bertranpetit J, Sykes BC. 1996. Genetic diversity in the Iberian Peninsula determined from mitochondrial sequence analysis. Ann Hum Genet 60:331–350.

del Rincón MA. 1998. El Calcolítico y la Edad del Bronce. In: Barandiarán I, Martí B, del Rincón MA, Maya JL, editors. Prehistoria de la Península Ibérica. Barcelona: Ariel, S.A. p 197–315.

Dutour O, Vernet R, Aumassip G. 1994. Le peuplement prehistorique du Sahara. In: Aumassip G, Ferhat N, Heddouche A, Vernet R, Thinon M, Dutour O, Onrubia-Pintado J, Grebenart D, Ould Khattar M, Tauveron M, Striedter KH, Dupuy C, Amblard S, Quechon G, Gaussen J, Bedeaux R, Bathily M, Mori F, editors. Milieux, hommes et techniques du Sahara prehistorique. Problèmes actuels. Paris: L'Harmattan. p 39–52.

Excoffier L, Smouse PE, Quattro JM. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA data. Genetics 131:479–491.

Ferembach D. 1986. Homo sapiens sapiens en Afrique: des origines au néolithique. In: Ferembach D, Susanne C, Chamla MC, editors. L'homme, son évolution, sa diversité. Paris: CNRS, Douin. p 245–256.

Flores C, Hernández M, González AM, Cabrera VM. 2000a. Genetic affinities among human populations inhabiting the Subsaharan area, northwest Africa, and the Iberian Peninsula. In: Arnaiz-Villena A, editor. Prehistoric Iberia: genetics, anthropology, and linguistics. New York: Kluwer Academic/Plenum Publishers. p 33–50.

Flores C, Maca-Meyer N, González AM, Cabrera VM. 2000b. Northwest African distribution of the CD4/Alu microsatellite haplotypes. Ann Hum Genet 64:321–327.

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.

Godinho VM. 1983. Os descobrimentos e a economia mundial, volume 4, 2nd ed. Lisbon: Presença Editors.

Helgason A, Hikey E, Goodacre S, Bosnes V, Stefánsson K, Ward R, Sykes B. 2001. MtDNA and the islands of the North Atlantic: estimating the proportions of Norse and Gaelic ancestry. Am J Hum Genet 68:723–737.

Hofmann S, Jaksch M, Bezold R, Mertens S, Aholt S, Paprotta A, Gerbitz KD. 1997. Population genetics and disease susceptibility: characterization of Central European haplogroups by mtDNA gene mutations, correlations with D loop variants and association with disease. Hum Mol Genet 6:1835–1846.

AQ: 4

Kozlowski JK, editor. 1982. Excavation in the Bacho Kiro Cave (Bulgaria). Final report. Warsaw: Panstwowe Wydawnietwo Naukowe.

Lamminen T, Huoponen K, Sistonen P, Juvonen V, Lahermo P, Aula P, Nikoskelainen E, Savontaus ML. 1997. MtDNA haplotype analysis in Finnish families with Leber hereditary optic neuropathy (LHON). Eur J Hum Genet 5:271–279.

Lucotte G, Loirat F. 1999. Y-chromosome DNA haplotype 15 in Europe. Hum Biol 71:431–437.

Lutz S, Weisser HJ, Heizmann JP. 1998. Location and frequency of polymorphic positions in the mtDNA control region of individuals from Germany. Int J Legal Med 111:67–77.

Macaulay V, Richards M, Hickey E, Vega E, Cruciani F, Guida V, Scozzari R, Bonné-Tamir B, Sykes B, Torroni A. 1999. The emerging tree of West Eurasian mtDNAs: a synthesis of control-region sequences and RFLPs. Am J Hum Genet 64:232–249.

Martí-Oliver B. 1998. El Neolítico. In: Barandiarán, I, Martí B, del Rincón MA, Maya JL, editors. Prehistoria de la Península Ibérica. Barcelona: Ariel S.A. p 121–195.

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

Mellars P. 1998. La revolución del Paleolítico Superior. In: Cunliffe B, editor. Prehistoria de Europa Oxford. Barcelona: Crítica, Grijalbo Mondadori, S.A. p 47–81.

Miller KWP, Dawson JL, Hagelberg E. 1996. A concordance of nucleotide substitutions in the first and second hypervariable segments of the human mtDNA control region. Int J Legal Med 109:107–113.

Nei M. 1987. Molecular evolutionary genetics. New York: Columbia University Press.

Pereira L, Prata MJ, Amorim A. 2000. Diversity of mtDNA lineages in Portugal: not a genetic edge of European variation. Ann Hum Genet 64:491–506.

Piercy R, Sullivan KM, Benson N, Gill P. 1993. The application of mitochondrial DNA typing to the study of white Caucasian genetic identification. Int J Legal Med 106:85–90.

Pinto F, González AM, Hernández M, Larruga JM, Cabrera VM. 1996. Genetic relationship between the Canary Islanders and their African and Spanish ancestors inferred from mitochondrial DNA sequences. Ann Hum Genet 60:321–330.

Quintana-Murci L, Semino O, Bandelt HJ, Passarino G, McElreavey K, Santachiara-Benerecetti AS. 1999. Genetic evidence of an early exit of *Homo sapiens sapiens* from Africa through eastern Africa. Nat Genet 23:437–441.

Rando JC, Pinto F, González AM, Hernández M, Larruga JM, Cabrera VM, 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.

Renfrew C. 1987. Archaeology and language. London: Penguin.

Richards M, Côrte-Real H, Forster P, Macaulay V, Wilkinson-Herbots H, Demaine A, Papiha S, Hedges R, Bandelt H-J, Sykes B. 1996. Paleolithic and Neolithic lineages in the European mitochondrial gene pool. Am J Hum Genet 59:185–203.

Richards M, Macaulay V, Bandelt H-J, Sykes B. 1998. Phylogeography of mitochondrial DNA in Western Europe. Ann Hum Genet 62:241–260.

Richards M, Macaulay V, Hickey E, Vega E, Sykes B, Guida V, Rengo C, Sellitto D, Cruciani F, Kivisild T, Villems R, Thomas M, Rychkov S, Rychkov O, Rychkov Y, Golge M, Dimitrov D, Hill E, Bradley D, Romano V, Cali F, Vona G, Demaine A, Papiha S, Triantaphyllidis C, Stefanescu G. 2000. Tracing European founder lineages in the Near Eastern mtDNA pool. Am J Hum Genet 67:1251–1276.

Rocha H, Flores C, Campos Y, Arenas J, Vilarinho L, Santorelli FM, Torroni A. 1999. About the "pathological" role of the mtDNA T3308C mutation. Am J Hum Genet 65:1457–1459.

Rosser ZH, Zerjal T, Hurles ME, Adojaan M, Alavantic D, Amorim A, Amos W, Armenteros M, Arroyo E, Barbujani G, Beckman G, Beckman L, Bertranpetit J, Bosch E, Bradley DG, Brede G, Cooper G, Côrte-Real HB, de Knijff P, Decorte R, Dubrova YE, Evgrafov O, Gilissen A, Glisic S, Golge M, Hill EW, Jeziorowska A, Kalaydjieva L, Kayser M, Kivisild T, Kravchenko SA, Krumina A, Kucinskas V, Lavinha J, Livshits LA, Malaspina P, Maria S, McElreavey K, Meitinger TA, Mikelsaar AV, Mitchell RJ, Nafa K, Nicholson J, Norby S, Pandya A, Parik J, Patsalis PC, Pereira L, Peterlin B, Pielberg G, Prata MJ, Previdere C, Roewer L, Rootsi S, Rubinsztein DC, Saillard J, Santos FR, Stefanescu G, Sykes BC, Tolun A, Villems R, Tyler-Smith C, Jobling MA. 2000. Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. Am J Hum Genet 67:1526–1543.

Rousselet F, Mangin P. 1998. Mitochondrial DNA polymorphisms: a study of 50 French Caucasian individuals and application to forensic casework. Int J Legal Med 111:299–304.

Rozas J, Rozas R. 1999. DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics 15:174–175.

Rudbeck L, Dissing J. 1998. Rapid, simple alkaline extraction of human genomic DNA from whole blood, buccal epithelial cells, semen and forensic stains for PCR. Biotechniques 25:588–592.

Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425.

Sajantila A, Lahermo P, Anttinen T, Lukka M, Sistonen P, Savontaus ML, Aula P, Beckman L, Tranebjaerg L, Gedde-Dahl T, Issel-Tarver L, DiRienzo A, Pääbo S. 1995. Gene and lenguages in Europe: an analysis of mitochondrial lineages. Genome Res 5:42–52.

Salas A, Comas D, Lareu MV, Bertranpetit J, Carracedo A. 1998. MtDNA analysis of the Galician population: a genetic edge of European variation. Eur J Hum Genet 6:365–375.

Schneider S, Kueffer J-M, Roessli D, Excoffier L. 1997. Arlequin version 1.1: a software for population genetic data analysis. Geneva: Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Simoni L, Calafell F, Pettener D, Bertranpetit J, Barbujani G. 2000. Geographic patterns of mtDNA diversity in Europe. Am J Hum Genet 66:262–278.

Torroni A, Huoponen K, Francalacci P, Petrozzi M, Morelli L, Scozzari R, Obinu D, Savontaus M-L, Wallace D. 1996. Classification of European mtDNAs from an analysis of three European populations. Genetics 144:1835–1850.

Torroni A, Bandelt H-J, D'Urbano L, Lahermo P, Moral P, Sellitto D, Rengo C, Forster P, Savontaus M-L, Bonné-Tamir B, Scozzari R. 1998. MtDNA analysis reveals a major late Palaeolithic population expansion from southwestern to northeastern Europe. Am J Hum Genet 62:1137–1152.

Vigilant L, Pennington R, Harpending H, Kocher TD, Wilson AC. 1989. Mitochondrial DNA sequences in single hairs from a southern African population. Proc Natl Acad Sci USA 86:9350–9354.

Watson E, Foster P, Richards M, Bandelt H-J. 1997. Mitochondrial footprints of human expansions in Africa. Am J Hum Genet 61:691–704.

Whittle A. 1998. Los primeros agricultores-ganaderos. In: Cunliffe B, editor. Prehistoria de Europa Oxford. Barcelona: Crítica, Grijalbo Mondadori, S.A. p 138–168.

Wilkinson-Herbots HM, Richards MB, Forster P, Sykes BC. 1996. Site 73 in hypervariable region II of the human mitochondrial genome and the origin of European populations. Ann Hum Genet 60:499–508.