African Female Heritage in Iberia: A Reassessment of mtDNA Lineage Distribution in Present Times

LUÍSA PEREIRA,¹ CARLA CUNHA,^{1,2} CÍNTIA ALVES,¹ AND ANTÓNIO AMORIM^{1,2}

Abstract The Iberian peninsula is a peripheral region of Europe in close proximity to Africa. Its inhabitants have an overall mtDNA genetic landscape typical of European background, although with signs of some African influence, whose features we deemed to disclose by analyzing available mtDNA HVRI distributions and new data. We analyzed 1,045 sequences. The most relevant results are the following: (1) North African sequences (haplogroup U6) present an overall frequency of 2.39%, and sub-Saharan sequences reach 3.83%, values that are, in both cases, much higher than those generally observed in Europe; and (2) there is a substantial geographic heterogeneity in the distribution of these lineages (haplogroup L being the most frequent in the south, whereas haplogroup U6 is generally more common in the north). The analysis of the observed diversity within each haplogroup strongly suggests that both were recently introduced (in historical times). Although for haplogroup U6 the documented event that is demographically compatible is the Islamic period (beginning of the 8th century to the end of the 15th century), for haplogroup L the most probable origin is the modern slave trade (mid 15th century to the end of the 18th century). However, the observed geographic structuring for one of the haplogroups does not fit the expected distribution provided by simplistic historical considerations. In fact, although for haplogroup L the north-south increasing frequency is corroborated by historical data, the opposite trend, observed for haplogroup U6, is more difficult to reconcile with the magnitude and time span of the Islamic political and cultural influence, which lasted longer and was more intense in the south. To clarify this conundrum, we need not only a substantial increase in the amount of mtDNA data (particularly for North Africa) but also new historical data and interpretations.

The first reports on mtDNA distributions in the Iberian peninsula have placed the region on the edge of European westward-decreasing genetic diversity (Comas et al. 1997; Salas et al. 1998). This conclusion was based on the low values registered for general measures of genetic diversity, such as the mean number of

¹IPATIMUP (Instituto de Patologia e Imunologia Molecular da Universidade do Porto), R. Dr. Roberto Frias s/n, 4200-465 Porto, Portugal.

²Faculdade de Ciências da Universidade do Porto, Praça Gomes Teixeira, 4099-002 Porto, Portugal.

Human Biology, April 2005, v. 77, no. 2, pp. 213–229.

Copyright © 2005 Wayne State University Press, Detroit, Michigan 48201-1309

KEY WORDS: MTDNA, HVRI, HVRII, IBERIA, NORTH AFRICA, SUB-SAHARAN AFRICA.

pairwise differences and nucleotide diversity as well as the observed high frequency of haplogroup H. A very low genetic impact of Neolithic migrations was inferred, because Near Eastern lineages in Iberia accounted for a mere 10% (Richards et al. 2000). However, Iberia seems to have played a major role as a refuge during the Last Glacial Maximum, as detected by the expansion of haplogroup V from northeast Iberia (including France) toward the rest of Europe (Torroni et al. 1996, 2001a).

This description reflects a current bias in the analysis of mtDNA databases, especially in Europe, that restricts analyses to the migrations that have occurred in remote times, namely, in the Paleolithic and Neolithic periods. This bias results mainly from the fact that only those ancient migrations leave signs in present mtDNA pools amenable to some statistical analyses. For instance, the dating of an introduction depends on the expansion and diversification from a founder lineage, which requires a rather large time span. Therefore information about recent migrations is usually scarce, except in cases of severe founder effects followed by expansions (Helgason et al. 2001) or the occurrence of a mutation coincident with the migration (Rando et al. 1999).

In the European context, Iberia is a good starting point for the analysis of more recent migratory events. Indeed, because of its geographic location, it is in a privileged position for genetic flow with North Africa (the Strait of Gibraltar is only 15 km wide).

The significance of prehistoric North African influence is still controversial (Arnaiz-Villena et al. 1997; Comas et al. 1998), but there is no doubt that a major opportunity for contact occurred from A.D. 711 onward, after the Islamic invasion of Iberia. Indeed, the major proportion of the invading contingents was made up of Berbers. On the other hand, Islamic rule lasted until the end of the 13th century in southern Portugal and until the end of the 15th century in southern Spain. The Islamic influence (as judged by the criterion of political administration) has also been heterogeneous in terms of geography, because northern Iberia was never under permanent Islamic administration (Serrão and Marques 1993).

Although it is said that followers of the three religions of the Book—Muslims, Christians, and Jews—were allowed to live together under more liberal conditions during the Islamic period than afterward, little is known about the interbreeding between the members of these communities. However, it is expected that, precisely because of that tolerance and because at that time social organization was deeply based on religious affiliation, at least most of the officially recognized marriages would be made between members of the same religion. It has also been impossible to quantify the importance of religious conversions in this period (Torres 1992).

A symmetric influx of Iberians into North Africa occurred along with the Portuguese and Spanish maritime expansions, from the beginning of the 15th century onward. The influx was mainly male mediated with the conquest of coastal (mostly Moroccan) towns (Farinha 1999).

Sub-Saharanans were initially brought to Iberia by Carthaginians, Romans, and Muslims, who were already trading them as slaves (Thomas 1998). However, those contributions were demographically insignificant compared with the huge numbers implied in the modern Atlantic slave trade, which was initiated in the mid 15th century by the Iberian kingdoms. It is believed that most slaves brought to Europe were introduced in Portugal, whereas Spanish slave trade was mainly devoted to its colonies in the Americas (Thomas 1998).

Previous analyses of Iberian mtDNA have already pinpointed the female African influence, by the detection of some characteristic lineages of North Africa (haplogroup U6) or sub-Saharan Africa (L haplogroups) (Côte-Real et al. 1996; Salas et al. 1998; Pereira et al. 2000a).

In this paper we provide a deeper and more global insight into the African female influence in Iberia, as judged from the maternally transmitted mtDNA. We analyzed a considerably enlarged Portuguese sample of 549 individuals, microgeographically structured, and all the Iberian data published so far, for a total of 1,045 individuals. The observed sequences whose origin is attributable to Africa were analyzed in the context of the available data from European, North African, and sub-Saharan populations.

Materials and Methods

Portuguese Samples. Five hundred forty-nine unrelated individuals born in Portugal were analyzed, and DNA was extracted from blood spots using the resin Chelex-100 method (Lareu et al. 1994). To avoid recent African admixture, we analyzed individuals of Portuguese ancestry only (excluding migrants after the independence of Portuguese ex-colonies, in 1975, even if they claimed exclusive European origin). We screened all Portuguese mainland administrative regions (districts). To obtain substantial sample sizes, we pooled the regions into larger geographic units (corresponding roughly to the longitudinal division of the country by the two main rivers, Douro and Tagus) as follows: northern Portugal (Viana do Castelo, 17 individuals; Braga, 41; Porto, 67; Vila Real, 20; and Bragança, 42; for a subtotal of 187); central Portugal (Aveiro, 25; Viseu, 38; Guarda, 37; Coimbra, 27; Leiria, 29; Castelo Branco, 23; Santarém, 4; and Lisbon, 56; for a subtotal of 239); and southern Portugal (Setúbal, 28; Portalegre, 25; Évora, 26; Beja, 15; and Faro, 29; for a subtotal of 123). This sample incorporates the one previously reported by Pereira et al. (2000a).

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'-CTGTTAAAAGTG CATACCGCCA-3') for HVRII. The temperature profile was 95°C for 10 s, 60°C for 30 s, and 72°C for 30 s for 35 cycles of amplification. The amplified samples were purified with Microspin S-300 HR columns (Amersham Biosciences), according to the manufacturer's specifications. The sequence reactions were carried

out using the kit Big-Dye Terminator Cycle Sequencing Ready Reaction (AB Applied Biosystems), with one of the listed 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 run in an automatic sequencer, ABI 377 or ABI 3100.

Genetic Analysis and Population Comparison. The nucleotide positions considered for the analysis of the sample were 16024–16383 for HVRI and 73–340 for HVRII. Length variation (often scored as transversions in HVRI) was not considered (Bendall and Sykes 1995). Sequence classification was done according to the nomenclature of Salas et al. (2002).

The Iberian database was constructed with the following data (see Figure 1 for sample location and coding): our Portuguese samples (northern Portugal, 187; central Portugal, 239; southern Portugal, 123); 92 individuals from Galicia (Salas et al. 1998); 118 individuals from Catalonia (Crespillo et al. 2000); 122 Basques [45 from Bertranpetit et al. (1995) and 77 from Richards et al. (1996)]; 61 individuals from Leon (Larruga et al. 2001); 38 individuals from Castile (Larruga et al. 2001); and 65 individuals from Andalusia [50 from Larruga et al. (2001) and 15 from Côté-Real et al. (1996)]; for a total of 1,045 individuals. The data from Côté-Real et al. (1996) and Pinto et al. (1996) were not considered because no microgeographic location of the samples was provided. Conversely, Crespillo et al.'s (2000) data were included (although their sampling was less stringent on individual's ancestry than for the other works) simply because it is still the only data available for northeastern Spain.

The North African databases used for haplogroup U6 comparison were Mozabites from Algeria (Côté-Real et al. 1996); Moroccan Berbers and non-Berbers, Saharans from West Sahara, and Mauritians from several different ethnic groups (Rando et al. 1998); Berbers from the Souss area, Morocco (Brakez et al. 2001); and south Arabian Bedouins of Berber origin (Di Rienzo and Wilson 1991).

Heterogeneity of haplogroup regional distributions was assayed through chi-square computations for the contingency tables. The computations were performed using the interactive software provided by the University of Hong Kong (available at http://department.obg.cuhk.edu.hk/researchsupport/RxC_contingency_table.asp).

The median-joining networks were calculated according to the method of Bandelt et al. (1995), by using the program Network 3.0 (available at <http://www.fluxus-engineering.com>), giving the same weight to all the substitutions.

Results and Discussion

Iberian Samples. Because this research is mainly centered on the analysis of African sequences, we present in Table 1 only the frequency distribution of observed haplogroups in our Portuguese sample (the full data can be provided upon

Figure 1. Frequencies of the North African (U6 haplogroup) and sub-Saharan (haplogroups L1, L2, and L3) mtDNA lineages in Iberia. NP, northern Portugal; CP, central Portugal; SP, southern Portugal; GA, Galicia; LE, Leon; CA, Castile; AN, Andalusia; BQ, Basque Country; CL, Catalonia. Values refer to frequencies (%) of the sequences that belong to the U6 (first value) and L haplogroups (L1, L2, and L3; second value).

request from L. Pereira). The description of the available data on North African and sub-Saharan sequences in Iberia is given in Table 2, and their distribution is summarized in Table 3.

North African Influence. The available information about the typical North African mtDNA haplogroup, U6, is still scarce, partly because it presents low frequencies in most North African populations, even in those that speak a Berber language (Figure 2). Furthermore, no clear gradient is observed, although frequencies seem to be higher in the west than in the east, reaching a maximum of 25% in Mozabites from Algeria (Côte-Real et al. 1996).

In Iberia, 25 U6 sequences were found in the total sample of 1,045 individuals, corresponding to a mean frequency of 2.39% ($\pm 0.47\%$). Regional variation

Table 1. Distribution of Observed mtDNA Haplogroups in Portugal

<i>Haplogroup</i>	<i>Northern Portugal</i>	<i>Central Portugal</i>	<i>Southern Portugal</i>
H	76	98	52
I	4	2	7
J	7	14	4
J1	2	0	0
J1b	0	1	2
J1b1	1	2	0
J2	3	1	0
K	9	17	6
L1b	0	2	7
L1c	0	0	2
L2a	3	2	0
L2b	0	1	0
L3*	2	4	3
L3*/N*	0	1	0
L3b	0	1	0
L3e1a	1	0	0
L3e2b	0	1	0
L3e3	0	0	2
M1	0	2	0
M1a	0	1	0
N1a	0	1	0
N1b	1	1	0
T	11	24	4
T1	8	3	2
U	8	5	5
U1a/M1?	0	0	1
U2	2	5	1
U3	3	0	0
U3?	0	1	0
U3 + 4	0	0	1
U4	2	2	3
U4/U5a	1	0	0
U5	4	3	0
U5a	3	2	0
U5a/b	2	0	1
U5a1	3	2	3
U5a1a	1	5	0
U5b	3	2	3
U5b2	1	1	1
U6a	6	6	2
U6b	4	0	0
U7	0	1	0
V	12	13	4
W	4	3	1
X	0	9	6
Total	187	239	123

Table 2. Regional Distribution of North African and Sub-Saharan Sequences in Iberia

Region	North African		Sub-Saharan		Other		Total N
	N	Frequency $\pm$ SD	N	Frequency $\pm$ SD	N	Frequency $\pm$ SD	
Galicia	2	0.022 $\pm$ 0.015	3	0.033 $\pm$ 0.019	87	0.946 $\pm$ 0.024	92
Northern Portugal	10	0.054 $\pm$ 0.017	6	0.032 $\pm$ 0.013	171	0.914 $\pm$ 0.021	187
Central Portugal	6	0.025 $\pm$ 0.010	12	0.050 $\pm$ 0.014	221	0.925 $\pm$ 0.017	239
Southern Portugal	2	0.016 $\pm$ 0.012	14	0.114 $\pm$ 0.029	107	0.870 $\pm$ 0.031	123
Castile	1	0.026 $\pm$ 0.026	0		37	0.974 $\pm$ 0.026	38
Leon	1	0.016 $\pm$ 0.016	1	0.016 $\pm$ 0.016	59	0.967 $\pm$ 0.023	61
Northeastern Spain	2	0.017 $\pm$ 0.012	3	0.025 $\pm$ 0.015	113	0.958 $\pm$ 0.019	118
Basque Country	0		0		122		122
Andalusia	1	0.015 $\pm$ 0.015	1	0.015 $\pm$ 0.015	63	0.969 $\pm$ 0.022	65
Total	25	0.024 $\pm$ 0.005	40	0.038 $\pm$ 0.006	980	0.938 $\pm$ 0.008	1,045

Table 3. mtDNA Haplotypes of African Origin Observed in Iberia

<i>Haplotype Code</i>	<i>Haplogroup</i>	<i>n^a</i>	<i>HVRI^b</i>
1	U6a	1 CA	092 172 219 278
2	U6a	1 CP	145 172 219 235 278 362
3	U6a	1 NP, 2 CP, 1 LE	172 189 219 239 278
4	U6a	3 NP	172 189 219 278
5	U6a	1 CP, 1 GA	172 219 235 278
6	U6a	1 NP, 2 SP	172 219 235 278 355
7	U6a	1 NP	172 219 271 278
8	U6a	2 CP, 1 CL	172 219 278
9	U6b	2 NP	051 172 219 311
10	U6b	1 CL	129 172 219 311
11	U6b	1 NP	172 174 188 219 311
12	U6b	1 AN	172 174 219 311
13	U6b	1 NP	172 219 261 311
14	U6b1	1 GA	129 163 164 ^{A/T} 172 219 311
15	L1b	2 SP	104 187 189 223 270 278 289 293 311
16	L1b	1 SP	114 ^{C/A} 126 187 189 223 264 270 278 293 311 362
17	L1b	2 SP	126 145 187 189 223 264 270 278 293 311
18	L1b	1 SP	126 187 189 223 259 264 270 278 293 311
19	L1b	1 SP	126 187 189 223 261 264 270 278 293 311
20	L1b	1 CP	126 187 189 223 264 270 278 293 311 362
21	L1b	1 LE	126 187 189 223 264 270 278 293 311
22	L1b	1 CP, 1 CL	126 187 189 223 264 270 278 311
23	L1b	1 CL	126 187 189 223 264 278 293 311
24	L1b	1 CL	126 223 264 270 278 293 311
25	L1c	1 SP	093 129 187 189 223 265 ^{A/C} 278 286 ^{C/A} 292 294 311 360
26	L1c	1 SP	129 187 189 223 265 ^{A/C} 278 294 311 359 360
27	L2a	1 NP	037 223 278 294 309 357 390
28	L2a	1 CP	093 189 192 223 260 278 294 309
29	L2a	1 NP	093 189 192 223 278 294 309 390
30	L2a	1 NP	093 223 278 294 309 311 320 390
31	L2a	1 CP	223 278 294 309
32	L2a	1 AN	223 278 294 390
33	L2b	1 CP	093 114 ^{C/A} 129 213 223 271 278
34	L2b	1 GA	114 ^{C/A} 129 213 223 234 278
35	L3*	1 CP	041 223
36	L3*	1 CP	104 189 223
37	L3*	1 CP	167 209 223 311
38	L3*	1 GA	169 193 195 223 243 261
39	L3*	1 NP	169 193 195 223 243 261 311
40	L3*	1 NP	209 218 223 256 292 311
41	L3*	1 CP, 1 SP	209 223 311
42	L3*	1 CP	223 292 311 320
43	L3*	2 SP	223
44	L3b	1 GA	093 124 223 278 362
45	L3b	1 CP	223 278 362
46	L3e1a	1 NP	169 185 223 311 327
47	L3e2b	1 CP	172 189 223 248 320

Table 3. Continued.

Haplotype Code	Haplogroup	<i>n</i> ^a	HVRI ^b
48	L3e3	1 SP	223 265 ^{A/T}
49	L3e3	1 SP	093 223 265 ^{A/T} 316
50	M1	1 CP	129 185 189 ^{Del T} 223 311
51	M1	1 CP, 1 AN, 1 CL	129 189 223 249 311
52	M1a	1 CP, 1 AN	129 189 223 249 311 359

a. Number of individuals observed in each population (NP, northern Portugal; CP, central Portugal; SP, southern Portugal; GA, Galicia; LE, Leon; CA, Castile; AN, Andalusia; BQ, Basque Country; CL, Catalonia).

b. HVRI substitutions (minus 16000), which are transitions, unless specified by the base mutated.

is displayed in Figure 1, with frequencies varying between 0 (in Basques) and 5.35% in northern Portugal. The substantial increase of sampling effort for Iberia has led to the detection of U6 sequences also in southern Portugal, in contrast to a previous claim that U6 sequences were restricted to the north (Pereira et al. 2000a). However, in Portugal a decreasing north-south frequency of U6 sequences can still be observed.

Although the absolute value of observed U6 frequency in Iberia is low, it reveals a considerable North African female contribution, if we keep in mind that haplogroup U6 is not very common in North Africa itself and virtually absent in the rest of Europe. Indeed, because the range of variation in western North Africa is 4–28%, the estimated minimum input is 8.54%.

Figure 2. Frequencies of the North African (U6 haplogroup) (gray), sub-Saharan (haplogroups L1, L2, and L3) (black), and Eurasian (white) mtDNA lineages in several populations from North Africa and the Near East. B, Berber; A, Arabic.

Figure 3. Median-joining network for the U6 sequences observed in Iberia. White, northern Portugal; black, central Portugal; backward diagonal stripes, southern Portugal; horizontal stripes, Galicia; forward diagonal stripes, Leon; checkerboard pattern, Castile; vertical stripes, Andalusia; diagonal cross-hatching, Catalonia. Numbers on the branches refer to HVRI substitution (minus 16000; all are transitions except a transversion, which is represented by the base mutated), and numbers near the pie charts are the code for the sequence (as displayed in Table 3). Circle sizes are proportional to haplotype frequencies.

The median-joining network of the Iberian U6 sequences (Figure 3) shows the high level of diversity presented by these sequences, which are also geographically scattered in the Iberian peninsula. The distribution is not significantly heterogeneous, as judged by the chi-square analysis of a 2 (U6 vs. all other sequences) $\times$ 9 (geographic samples) contingency table ($p = 0.2055$). This pattern in itself is strong evidence against the hypothesis of a major single founding event as an explanation for the presence of this haplogroup in Iberia. Whether the introduction resulted from a source population with a diverse set of U6 lineages or proceeded over a long period from several sources is not yet clear. Nevertheless, because no local lineage expansions were detected, it seems that the introduction (unique or not) was recent and made by the sampling of a set of already diverse sequences.

By analyzing jointly the U6 sequence diversity observed in Iberia and in North Africa (Figure 4), we confirm that haplotypes more frequent in Iberia are also the most frequent in North Africa (haplotypes 16172-16219-16278, 16172-16189-16219-16278, and 16172-16189-16219-16239-16278). Interestingly, in the branch linking haplotype 8 to haplotypes 2 and 18, the alternating Iberian and North African haplotypes occupy what would be empty nodes if both sets

Figure 4. Median-joining network for the U6 sequences observed in Iberia (white) and North Africa (black). Numbers on the branches refer to HVRI substitution (minus 16000; all are transitions except a transversion, which is represented by the base mutated), and numbers near the pie charts are the code for the sequence (as displayed in Table 3). Circle sizes are proportional to haplotype frequencies.

were analyzed separately, showing once again how patchy the U6 haplogroup evolutionary picture still is, even for the putative region of origin (North Africa).

Another striking aspect to consider is that combining both Iberian and North African sequences increases subhaplogroup U6a diversity considerably, in contrast with what happens for subhaplogroup U6b, for which just a single haplotype has been described in North Africa (in the nomadic population of Bedouins; Di Rienzo and Wilson 1991). It must be said, however, that some U6b sequences have been observed in sub-Saharan Africa, namely, one Fulbe from Nigeria (16172-16173-16219-16311-16320; Watson et al. 1997) and one Wolof from Senegal (16172-16193-16219-16311; Rando et al. 1998). Rando et al. (1999) further defined a subclade of subhaplogroup U6b: U6b1, characterized by a transition at np 16163 and typical of the Canary Islands. U6b1's starlike phylogeny suggests a unique introduction from the hypothetical ancestral lineages 16163-16172-16219-16311 and 16092-16163-16172-16219-16311. This pattern is quite different from the one observed for Iberia. Two instances of the Canarian type were observed in Iberia: one in Galicia (sequence 14 in Figures 3 and 4; Salas et al. 1998) and one in Portugal (16163-16164^{A/T}-16172-16219-16311; Crte-Real et al. 1996), which most probably represent introductions from those islands.

Figure 5. Median-joining network for the sub-Saharan sequences observed in Iberia. White, northern Portugal; black, central Portugal; backward diagonal stripes, southern Portugal; horizontal stripes, Galicia; forward diagonal stripes, Leon; cross-hatching, Castile; vertical stripes, Andalusia; diagonal cross-hatching, Catalonia. Numbers near the pie charts are the code for the sequence (as displayed in Table 3). Circle sizes are proportional to haplotype frequencies, and mutations were omitted from the branches for legibility sake.

Sub-Saharan African Influence. The mean frequency for the sequences belonging to superhaplogroup L, typical for sub-Saharan populations, reaches 3.83% ($\pm 0.59\%$) in Iberia, representing 40 sequences in the total sample of 1,045 individuals. The frequency is clearly higher (Figure 1) in Portugal (32 sequences in 549 individuals; 5.83%) than in Spain (8 out of 496; 1.61%) and without parallel in the rest of Europe. Furthermore, in western Iberia, increasing frequencies are observed for Galicia and northern Portugal (3.26% and 3.21%, respectively—a value similar to the one found in the rest of Spain) through the center (5.02%) and to the south (11.38%). The overall geographic Iberian heterogeneity is highly significant ($p = 0.0003$ in the chi-square for the 2×9 contingency table), as is the heterogeneity for the four western populations ($p = 0.0110$). This pattern is also consistent with historical reports referring to a predominant introduction of slaves in southern Portugal (Lahon 1999).

All major sub-Saharan haplogroups are represented, and they contribute to a very high observed diversity and a nonstarlike network (Figure 5). Again, as in

the case of haplogroup U6, the absence of signs of local expansions seems to point to historically recent introduction(s) of these lineages in Iberia.

Furthermore, the geographic origin of these lineages is also heterogeneous. Indeed, a high proportion (32.5%) is almost undoubtedly ascribable to the west coast (L1b). The second most frequent group of L sequences in Iberia (L2a; 15.0%) is widespread in Africa (although also more diverse on the west coast and potentially involved in the Bantu expansion to the south; Pereira et al. 2001). However, a smaller proportion (10.0%; L3e) is attributable to the Bantu southern migration (Bandelt et al. 2001), and one case found in Portugal (2.5%; L3e1a) is typical of the southern east coast [one step away from the sequence 16185-16223-26311-26327 observed as a unique instance in Mozambique (Pereira et al. 2001) and in Kikuyu from Kenya (Watson et al. 1997)]. The same proportion (5%) is observed for haplogroups L1c (widespread in low frequencies; see, e.g., Brehm et al. 2002), L2b (common in Senegalese; Torroni et al. 2001b), and L3b (also Bantu; Watson et al. 1997), but 11 sequences could not be assigned further than paragroup L3*.

This pattern strongly supports the explanation for the introduction of these lineages in Iberia by means of the modern slave trade, for which the main sources of slaves were widespread on the western coast, from Senegal to Angola (Thomas 1998). Moreover, particularly in the Portuguese case, a substantial Mozambican (east coast) contribution is known to have also been involved by the end of this period (17th century onward), representing 8% of the total slave trade. Although especially directed to the Americas (Thomas 1998; Pereira et al. 2001), some of those slaves traveled the route Mozambique-Brazil-Portugal (Lahon 1999).

In the network shown in Figure 5, the lineages belonging to haplogroup M1 are also depicted. This haplogroup is said to have originated in northeast Africa, in the region of Ethiopia (Quintana-Murci et al. 1999), but it is also present in North Africa and in the Middle East, so it is not directly ascertained as having a sub-Saharan introduction. For that reason, it was not included in the frequency calculations.

The pattern of L sequences in Iberia is different from the pattern observed in the rest of Europe, where just a few instances were observed, even in countries also involved in the slave trade. Indeed, a general survey of the literature on Europeans provided the following cases of sub-Saharan sequences: one L1b (16126-16175-16189-16223-16264-16270-16278-16311) in 161 German-Danish (Richards et al. 1996); one L1c (16129-16189-16223-16278-16284-16294-16311-16360) in 100 British (Piercy et al. 1993); one L1a (16086-16129-16148-16166-16168-16172-16187-16188^{C/G}-16189-16223-16230-16311-16320) in 42 Albanians (Belledi et al. 2000); one L3 (16189-16223-16291^{C/A}) in 50 Finnish (Sajantila et al. 1995); one L1a1 (16129-16148-16168-16172-16187-16188^{C/G}-16189-16223-16230-16311-16320) and one L2a (16093-16189-16192-16223-16278-16294) in 69 Sardinians (Di Rienzo and Wilson 1991); and one L2a (16145-16189-16192-16223-16278-16294-16309-16390) in 106 Sicilians (Cali et al. 2001).

Global Analysis of the African Influence in Iberia. It has been shown that inferring North African U6 founder types in Iberia is impossible, as is dating the introduction event(s). However, the absence of local expansions is consistent with their occurrence in recent times. The most probable associated historical event is obviously the Islamic invasion of Iberia (A.D. 711), because it was carried out mainly by a North African contingent, most of it composed of recently converted Berbers. Nevertheless, this explanation meets some difficulties. In fact, Islamic rule lasted until almost the end of the 13th century in southern Portugal and the 15th century in southern Spain, whereas some regions in northern Iberia, such as northern Portugal, Galicia, and Basque Country, studied here, remained untouched (or were contacted just slightly by sporadic raids). Thus it appears paradoxical that the highest frequencies for this haplogroup are to be found in northern Iberia and in northern Portugal, in particular. The reasons for that distribution are probably to remain historically obscure because of the absence of quantified records. We can at least speculate that during the first period of the “Reconquista” (involving only the north), interbreeding between Christians and (female) Muslims (not necessarily slaves) was easier than in the final period. Indeed, during the first period, not only were many “Moorish” slaves (particularly females) recorded, but also special privileges were given to them in the charters of new (or refounded) towns, whereas during the second period (in the south) the Islamic communities were, at least in urban areas, isolated into ghettos (*mourarias*) and the colonization and/or repopling of rural areas was performed by the already vast demographic surpluses from the north (Gomes 1999). It must also be recalled that male and female gene pool introgressions can show opposite behaviors. Classically, the reproduction of high-status men with women under manumission or even slaves was quite tolerated or even encouraged, but a few high-ranking males (particularly during the Islamic period, when polygyny but not polyandry was the rule in the ruling class) could have had a disproportionate contribution to the next generation (e.g., Richards et al. 2003; Pereira et al. 2000b).

The patchy pattern of U6 sequences observed today in North Africa is consistent with a picture of endogamous small groups of Berbers resulting from the Arabic invasion in the 7th century. It would be essential to obtain data from ancient DNA in the original Berber populations before Islamization, because the increasing description of present-day population diversity in North Africa (including some Berber-speaking populations) has not shed any light on the subject.

Concerning the sub-Saharan lineages found in Iberia, our analysis on their observed diversity points to the slave trade initiated in the 15th century along with the maritime expansion of Iberian countries as the most probable population movement for their introduction. Estimates for the number of slaves entering Europe in this period (and for the total European population at that time) are uncertain but seem to amount to a maximum of around a half-million individuals [200,000 according to Thomas (1998) and 300,000–500,000 according to Lahon

(1999)]. Whatever the exact figure, it seems clear that, at least on the female side, Iberian countries not only preceded other European states in this trade but also greatly exceeded them in the number of imported slaves. This difference is particularly evident when we look at the high frequency (11%) of sub-Saharan lineages in southern Portugal.

Alternative scenarios, either invoking much older and demographically more significant introductions (at least for North African sequences; e.g., González et al. 2003) or claiming a substantial role of the Roman and/or Islamic periods on the introduction of sub-Saharan lineages, seem difficult to reconcile with the available data. Against the first case, we can point out the absence of signs of local expansions and the patchy geography and diversity of observed U6 distribution. The acceptance of the second case is seriously diminished by the fact that sub-Saharan and North African lineages show a considerably different geographic distribution, which should be identical if both were introduced by the same migratory movement.

As final remarks, we would like to state that (1) African female-mediated genetic input in Iberia is much greater than the overall European levels, (2) most of this contribution seems to have taken place in recent times (Islamic period, 711–1500 for North African lineages; modern slave trade, 1500–1800 for sub-Saharan lineages), and (3) the patterns of microgeographic heterogeneity of their distribution in Iberia are, in some cases, easily reconciled with historical data (for sub-Saharan lineages), whereas in other cases (North African) they are harder to fit into a simplistic historical frame. Therefore, to clarify these issues, we urgently need not only a substantial increase in mtDNA data (particularly for North Africa) but also new historical data and interpretations, considering the genetic results available so far.

Acknowledgments This research was partially supported by a research grant awarded to L. Pereira (SFRH/BPD/7121/2001) from Fundação para a Ciência e a Tecnologia and Instituto de Patologia e Imunologia Molecular da Universidade do Porto (IPATIMUP) through the Programa Operacional Ciência, Tecnologia e Inovação (POCTI), Quadro Comunitário de Apoio III.

Received 16 January 2003; revision received 4 November 2003.

Literature Cited

- Arnaiz-Villena, A., J. Martínez-Laso, E. Gómez-Casado et al. 1997. Relatedness among Basques, Portuguese, Spaniards, and Algerians studied by HLA allelic frequencies and haplotypes. *Immunogenetics* 47:37–43.
- Bandelt, H.-J., J. Alves-Silva, P. E. M. Guimarães et al. 2001. Phylogeography of the human mitochondrial haplogroup L3e: A snapshot of African prehistory and Atlantic slave trade. *Ann. Hum. Genet.* 65:549–563.

- Bandelt, H.-J., P. Forster, B. C. Sykes et al. 1995. Mitochondrial portraits of human populations using median networks. *Genetics* 141:743–753.
- Belledi, M., E. S. Poloni, R. Casalotti et al. 2000. Maternal and paternal lineages in Albania and the genetic structure of Indo-European populations. *Eur. J. Hum. Genet.* 8:480–486.
- Bendall, K. E., and B. C. Sykes. 1995. Length heteroplasmy in the first hypervariable segment of the human mtDNA control region. *Am. J. Hum. Genet.* 57:248–256.
- Bertranpetit, J., J. Sala, F. Calafell et al. 1995. Human mitochondrial DNA variation and the origin of Basques. *Ann. Hum. Genet.* 59:63–81.
- Brakez, Z., E. Bosch, H. Izaabel et al. 2001. Human mitochondrial DNA sequence variation in the Moroccan population of the Souss area. *Ann. Hum. Biol.* 28:295–307.
- Brehm, A., L. Pereira, H.-J. Bandelt et al. 2002. Mitochondrial portrait of the Cabo Verde archipelago: The Senegambian outpost of Atlantic slave trade. *Ann. Hum. Genet.* 66:49–60.
- Cali, F., M. G. Le Roux, R. D'Anna et al. 2001. mtDNA control region and RFLP data for Sicily and France. *Int. J. Legal Med.* 114:229–231.
- Comas, D., E. Calafell, E. Mateu et al. 1997. Mitochondrial DNA variation and the origin of the Europeans. *Hum. Genet.* 99:443–449.
- Comas, D., E. Mateu, F. Calafell et al. 1998. HLA class I and class II DNA typing and the origin of Basques. *Tissue Antigens* 51:30–40.
- Côrte-Real, H., V. A. Macaulay, M. B. Richards et al. 1996. Genetic diversity in the Iberian peninsula determined from mitochondrial sequence analysis. *Ann. Hum. Genet.* 60:331–350.
- Crespillo, M., J. A. Luque, M. Paredes et al. 2000. Mitochondrial DNA sequences for 118 individuals from northeastern Spain. *Int. J. Legal Med.* 114:130–132.
- Di Rienzo, A., and A. C. Wilson. 1991. Branching pattern in the evolutionary tree for human mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* 88:1597–1601.
- Farinha, A. D. 1999. *Os Portugueses em Marrocos*. Lisbon: Instituto Camões, Coleção Lazúli.
- Gomes, S. A. 1999. Grupos étnico-religiosos e estrangeiros. In *Nova História de Portugal*, v. 3, J. Serrão and A. O. H. Marques, coord. Lisbon: Editorial Presença, 309–383.
- González, A. M., A. Brehm, J. A. Perez et al. 2003. Mitochondrial DNA affinities at the Atlantic fringe of Europe. *Am. J. Phys. Anthropol.* 120:391–404.
- Helgason, A., E. Hickey, S. Goodacre et al. 2001. mtDNA and the islands of the North Atlantic: Estimating the proportions of Norse and Gaelic ancestry. *Am. J. Hum. Genet.* 68:723–737.
- Lahon, D. 1999. *O negro no coração do império: Uma memória a resgatar—séculos XV–XIX*. Lisbon: Coleção Entreculturas, História e histórias.
- Lareu, M. V., C. P. Phillips, A. Carracedo et al. 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.
- Larruga, J. M., F. Díez, F. M. Pinto et al. 2001. Mitochondrial DNA characterization of European isolates: The Maragatos from Spain. *Eur. J. Hum. Genet.* 9:708–716.
- Pereira, L., V. Macaulay, A. Torroni et al. 2001. Prehistoric and historic traces in the mtDNA of Mozambique: Insights into the Bantu expansions and the slave trade. *Ann. Hum. Genet.* 65:439–458.
- Pereira, L., M. J. Prata, and A. Amorim. 2000a. Diversity of mtDNA lineages in Portugal: Not a genetic edge of European variation. *Ann. Hum. Genet.* 64:491–506.
- Pereira, L., M. J. Prata, M. Brion et al. 2000b. Clinal variation of YAP + Y-chromosome frequencies in western Iberia. *Hum. Biol.* 72:937–944.
- Piercy, R., K. M. Sullivan, N. Benson et al. 1993. The application of mitochondrial DNA typing to the study of white Caucasian genetic identification. *Int. J. Legal Med.* 106:85–90.
- Pinto, F., A. M. González, M. Hernández et al. 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., O. Semino, H.-J. Bandelt et al. 1999. Genetic evidence of an early exit of *Homo sapiens sapiens* from Africa through eastern Africa. *Nat. Genet.* 23:437–441.

- Rando, J. C., V. M. Cabrera, J. M. Larruga et al. 1999. Phylogeographic patterns of mtDNA reflecting the colonization of the Canary Islands. *Ann. Hum. Genet.* 63:413–428.
- Rando, J. C., F. Pinto, A. M. González et al. 1998. Mitochondrial DNA analysis in northwestern African populations reveals genetic exchanges with European, Near Eastern, and sub-Saharan populations. *Ann. Hum. Genet.* 62:531–550.
- Richards, M. B., H. Côrte-Real, P. Forster et al. 1996. Paleolithic and Neolithic lineages in the European mitochondrial gene pool. *Am. J. Hum. Genet.* 59:185–203.
- Richards, M., V. Macaulay, E. Hickey et al. 2000. Tracing European founder lineages in the Near Eastern mtDNA pool. *Am. J. Hum. Genet.* 67:1251–1276.
- Richards, M., C. Rengo, F. Cruciani et al. 2003. Extensive female-mediated gene flow from sub-Saharan Africa into Near Eastern Arab populations. *Am. J. Hum. Genet.* 72:1058–1064.
- Sajantila, A., P. Lahermo, T. Anttinen et al. 1995. Genes and languages in Europe: An analysis of mitochondrial lineages. *Genome Res.* 5:42–52.
- Salas, A., D. Comas, M. V. Lareu et al. 1998. mtDNA analysis of the Galician population: A genetic edge of European variation. *Eur. J. Hum. Genet.* 6:365–375.
- Salas, A., M. Richards, T. De la Fe et al. 2002. The making of the African mtDNA landscape. *Am. J. Hum. Genet.* 71:1082–1111.
- Serrão, J., and A. H. O. Marques. 1993. *Nova História de Portugal, v. 2, Portugal das Invasões Germânicas à "Reconquista."* Lisbon: Editorial Presença.
- Thomas, H. 1998. *The Slave Trade: The History of the Atlantic Slave Trade 1440–1870.* London: Macmillan.
- Torres, C. 1992. O Garb-Al-Andaluz. In *História de Portugal*, J. Mattoso, dir. Lisbon: Círculo de Leitores.
- Torroni, A., H.-J. Bandelt, L. D'Urbano et al. 1998. mtDNA analysis reveals a major late Paleolithic population expansion from southwestern to northeastern European populations. *Am. J. Hum. Genet.* 62:1137–1152.
- Torroni, A., H.-J. Bandelt, M. Vincent et al. 2001a. A signal, from human mtDNA, of postglacial recolonization in Europe. *Am. J. Hum. Genet.* 69:844–852.
- Torroni, A., C. Rengo, V. Guida et al. 2001b. Do the four clades of the mtDNA haplogroup L2 evolve at different rates? *Am. J. Hum. Genet.* 69:1348–1356.
- Watson, E., P. Forster, M. Richards et al. 1997. Mitochondrial footprints of human expansions in Africa. *Am. J. Hum. Genet.* 61:691–704.