The Maternal Legacy of Basques in Northern Navarre: New Insights Into the Mitochondrial DNA Diversity of the Franco-Cantabrian Area

Sergio Cardoso,1 Miguel A. Alfonso-Sánchez,1 Laura Valverde,1 Adrian Odriozola,1 Ana M. Pérez-Miranda,2 José A. Peña,2 and Marian M. de Pancorbo1*

1BIOMICs Research Group, Centro de Investigación y Estudios Avanzados “Lucio Lascaray”, Universidad del País Vasco, Miguel de Unamuno 3, 01006, Vitoria-Gasteiz, Spain
2Departamento de Genética, Antropología Física y Fisiología Animal, Facultad de Ciencia y Tecnología, Universidad del País Vasco, Apartado 644, 48080, Bilbao, Spain

KEY WORDS human isolate; genetic diversity; local microdifferentiation

ABSTRACT Autochthonous Basques are thought to be a trace from the human population contraction that occurred during the Last Glacial Maximum, based mainly on the salient frequencies and coalescence ages registered for haplogroups V, H1, and H3 of mitochondrial DNA in current Basque populations. However, variability of the maternal lineages still remains relatively unexplored in an important fraction of the Iberian Basque community. In this study, mitochondrial DNA diversity in Navarre (North Spain) was addressed for the first time. To that end, HVS-I and HVS-II sequences from 110 individuals were examined to identify the most relevant lineages, including analysis of coding region SNPs for the refinement of haplogroup assignment. We found a prominent frequency of subhaplogroup J1c (11.8%) in Navarre, coinciding with previous studies on Basques. Subhaplogroup H2a5, a putative autochthonous Basque lineage, was also observed in Navarre, pointing to a common origin of current Basque geographical groups. In contrast to other Basque subpopulations, comparative analyses at Iberian and European scales revealed a relevant frequency of subhaplogroup H3 (10.9%) and a frequency peak for U5b (15.5%) in Navarre. Furthermore, we observed low frequencies for maternal lineages HV0 and H1 in Navarre relative to other northern Iberian populations. All these findings might be indicative of intense genetic drift episodes generated by population fragmentation in the area of the Franco-Cantabrian refuge until recent times, which could have promoted genetic microdifferentiation between the different Basque subpopulations. Am J Phys Anthropol 145:480–488, 2011. ©2011 Wiley-Liss, Inc.

The Basque area currently encompasses around 20,750 km², distributed along the Biscay Gulf and to both sides of the border between Spain and France. Analyzing the genetic variability of present-day Basque population requires a thorough look at the spatial distribution of the Basque autochthonous language (Euskera), since language is a key cultural feature differentiating this singular anthropological group. The native Basque-speaking population of the Iberian Peninsula is concentrated in four northern Spanish provinces: Alava, Biscay, Guipuzcoa, and Navarre. Interestingly, the earliest references to the Basque people placed the ancient tribe of the Vascones (from which the word “Basque” is derived) in the current territory of Navarre (see Echaide, 1986; Fatás et al., 1993). The extant geo-linguistic distribution of the Euskera indicates that, in the specific case of Navarre, Basque-speakers are concentrated in its northernmost part, sharing with the adjacent province of Guipuzcoa the prevalence of the Euskera as the more habitual form of communication.

Despite being an important fraction of the Iberian Basque community, Navarrese Basques were the least scrutinized from the genetic viewpoint until recent times. The lack of genetic information has been partially solved throughout the last decade with several investigations focusing on the genetic diversity of this Basque geographical group (Pérez-Miranda et al., 2004, 2005; Calderón et al., 2006; García-Obregón et al., 2007). However, genetic variability of the maternal lineages still remains unexplored.

Previous mitochondrial DNA (mtDNA) analyses centered in the distribution of maternal lineages across Europe have hypothesized that the autochthonous Basque people could be a trace from the human population contraction that occurred during the Last Glacial Maximum (LGM), based mainly on coalescence ages and the salient frequencies registered for mtDNA haplogroups V (~20%) and H (~52%)—specifically for subhaplogroups H1/C24 and H/C24. Other studies have suggested that the subhaplogroup H2a5, a putative autochthonous Basque lineage, was also observed in Navarre, pointing to a common origin of current Basque geographical groups. In contrast to other Basque subpopulations, comparative analyses at Iberian and European scales revealed a relevant frequency of subhaplogroup H3 (10.9%) and a frequency peak for U5b (15.5%) in Navarre. Furthermore, we observed low frequencies for maternal lineages HV0 and H1 in Navarre relative to other northern Iberian populations. All these findings might be indicative of intense genetic drift episodes generated by population fragmentation in the area of the Franco-Cantabrian refuge until recent times, which could have promoted genetic microdifferentiation between the different Basque subpopulations.

Additional Supporting Information may be found in the online version of this article.

S. Cardoso and M.A. Alfonso-Sánchez have contributed equally to this work.

Grant sponsor: Consolidación de Grupos de Investigación del Gobierno Vasco; Grant number: IT-424-07. Grant sponsors: Servicios Generales de Investigación (SGiker) of the University of the Basque Country (UPV/EHU, MICINN, GV/EJ, ERDF, and FSE), Basque Government (Dpto. de Educación, Universidades e Investigación).

*Correspondence to: Marian M. de Pancorbo, BIOMICs Research Group, Centro de Investigación y Estudios Avanzados “Lucio Lascaray”, Universidad del País Vasco UPV/EHU, Miguel de Unamuno 3, 01006 Vitoria-Gasteiz, Spain. E-mail: marianpancorbo@gmail.com

Received 16 September 2010; accepted 28 February 2011

DOI 10.1002/ajpa.21532
Published online 3 May 2011 in Wiley Online Library (wileyonlinelibrary.com).
and H3 (~14%)—in present-day Basques (Torrioni et al., 1998, 2001; Achilli et al., 2004). In a recent work, Alfonso-Sánchez et al. (2008) confirmed the relevant frequency of haplogroup H (50.9%) in the Iberian Basque area, but their findings did not corroborate the frequency peak for haplogroup V (5.5%). The same study unveiled outstanding frequencies of haplogroup J. These discrepancies were explained in terms of local genetic microdifferentiation promoted by genetic drift during the last glacial period.

This work was aimed at contributing new, pioneering data on the polymorphism of the mtDNA D-loop region in autochthonous Basques from northern Navarre, in an attempt to gradually fill the existing gaps in the knowledge of the mitochondrial genome diversity in Iberian (Spanish) Basques. We essentially focused on the analysis of the haplogroup composition to identify the most relevant mitochondrial lineages in the Franco-Cantabrian refuge area. Our study was also intended to provide evidence on local genetic differentiation among Iberian Basque geographical groups and to analyze the genetic position of Basques as a whole within the European context, bearing in mind the overall genetic diversity of maternal lineages of the different Basque subpopulations.

MATERIALS AND METHODS

Sample collection

Four rural villages representing the native Basque population settled in the northern fringe of Navarre (~28%) and H3 (~14%)—in present-day Basques (Torrioni et al., 1998, 2001; Achilli et al., 2004). In a recent work, Alfonso-Sánchez et al. (2008) confirmed the relevant frequency of haplogroup H (50.9%) in the Iberian Basque area, but their findings did not corroborate the frequency peak for haplogroup V (5.5%). The same study unveiled outstanding frequencies of haplogroup J. These discrepancies were explained in terms of local genetic microdifferentiation promoted by genetic drift during the last glacial period.

This work was aimed at contributing new, pioneering data on the polymorphism of the mtDNA D-loop region in autochthonous Basques from northern Navarre, in an attempt to gradually fill the existing gaps in the knowledge of the mitochondrial genome diversity in Iberian (Spanish) Basques. We essentially focused on the analysis of the haplogroup composition to identify the most relevant mitochondrial lineages in the Franco-Cantabrian refuge area. Our study was also intended to provide evidence on local genetic differentiation among Iberian Basque geographical groups and to analyze the genetic position of Basques as a whole within the European context, bearing in mind the overall genetic diversity of maternal lineages of the different Basque subpopulations.

Fig. 1. Map showing the location of the Basque territory in the northern fringe of the Iberian Peninsula and southern France. Navarre borders on the West with the Autonomous Community of the Basque Country, which includes the provinces of Alava, Biscay, and Guipuzcoa. On the north–east Navarre limits with the French Basque Country, which is traditionally subdivided into three provinces: Labourd, Lower Navarre, and Soule. The four rural villages of northern Navarre selected for the study are: (1) Bera, (2) Elizondo, (3) Lekaroz, and (4) Leitza. Also displayed are the Arratia valley (Biscay province) and the Goierri valley (Guipuzcoa province), the two Basque localities where the reference sample was collected (Alfonso-Sánchez et al., 2008).

Polymorphism of the mtDNA control region

DNA was extracted from peripheral blood, using the standard salting-out procedure (Miller et al., 1988). Hypervariable regions HVS-I and HVS-II of the mtDNA
D-loop region were amplified by PCR using the procedure described in Supporting Information Table 1. PCR products were sequenced using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA) on an ABI 3130 Genetic Analyzer (Applied Biosystems).

Regarding the analysis of DNA sequences, the region HVS-I was examined between nucleotide positions (np) 16,019 and 16,392, according to the Cambridge reference sequence, rCRS (Andrews et al., 1999), while analysis of HVS-II included np 40–390. Diversity parameters were computed as described in Alfonso-Sánchez et al. (2008). The HVS-I and HVS-II sequences reported in this article are available online at GenBank under accession nos. FJ667273–FJ667382 and FJ667383–FJ667492, respectively.

**Haplogroup classification**

Control region haplotypes were classified into mtDNA haplogroups following a recently released nomenclature revision (van Oven and Kayser, 2009). To confirm the classification of mtDNA sequences with a doubtful haplogroup assignment based on control region, direct sequencing of coding region SNPs was employed. This procedure (summarized in Supporting Information Table 1) mainly involved haplogroups H and U, whose classification through polymorphisms of control region may be relatively complex and often ambiguous.

**Phylogenetic network**

mtDNA genealogy of the Basque area was constructed using the median-joining network approach (Bandelt et al., 1995) available in the Network v4.5.0.0 software (Fluxus Technology). Polymorphisms were weighted depending on the mutability of the nucleotide position involved. By default, all polymorphisms were initially weighted by a value of 10. Bearing in mind that transitions are relatively infrequent, a value three times higher than transitions was assigned. Hotspots reported by Meyer et al. (1999) and those detected in this study were weighted by values between 0 and 5. Data on coding region polymorphisms were used to strengthen the topology of the network and they were given a comparatively high weight. HVS-I and HVS-II sequence data from an African individual [sample 4568 from Behar et al. (2008)] was used as outgroup for network construction.

**Phylogeographic analysis**

The genetic affinities of Navarrese Basques with 25 other western Eurasian collections based on mtDNA haplogroup composition were explored by correspondence analysis (CA). CA is an analog of principal component analysis, which is appropriate to discrete rather than to continuous variables (Hill, 1974). For the sake of comparability, from the mtDNA databases currently available for European populations only those collections with a well-defined geographic origin and, as far as possible, representative of relatively small areas were included in the analysis, to avoid bias in the results. Frequencies of mtDNA haplogroups and subhaplogroups in the 26 populations included in the CA and the corresponding sources are given in Supporting Information Table 2. CA was performed using PAST software (Hammer et al., 2001). Population differentiation from the mtDNA viewpoint was examined by computing $F_{st}$ values (Weir and Cockerham, 1984), which estimates the portion of genetic variation attributable to between-population divergence. Population pairwise $F_{st}$ values were calculated with the Arlequin v3.0 software (Excoffier et al., 2005).

**RESULTS**

**Diversity parameters**

Polymorphic sites of the mtDNA control region in Navarre collection ($N = 110$), which defined 58 different haplotypes, are given as supplemental data (Supporting Information Table 3).

Concerning within-population diversity parameters estimated from mtDNA sequence data, gene diversity in Navarre ($0.9761 \pm 0.0056$) was slightly but not significantly higher than the estimated one for the Basque Country sample ($0.9650 \pm 0.0134$). As expected, gene diversity increased when total Basques were considered ($0.9792 \pm 0.0041$) but remains far from diversity levels registered for other Europeans (see for instance, Table 3 in Alfonso-Sánchez et al. (2008)). Other diversity parameters in Navarre, such as nucleotide diversity ($\pi_n = 0.0124 \pm 0.0065$) and mean number of pairwise differences ($\pi = 7.2417 \pm 3.4182$), fitted into the range of the European populations.

**Haplogroup composition**

The assignment of mtDNA haplogroups to haplogroups through polymorphism of HVS-I and HVS-II, as well as results of the analysis of coding region SNPs to refine haplogroup composition are summarized in Supporting Information Table 3.

mtDNA haplogroup composition in northern Navarre Basques compared with data reported for the Basque Country sample (Alfonso-Sánchez et al., 2008) is depicted in Table 1. Coinciding with most Eurasian populations, haplogroup H was the most common lineage in Navarre (34.5%), although far from the outstanding frequency registered in the Basque Country (54.6%). Bearing in mind that analysis of haplogroup H as a whole is not fully informative because of its complex phylogeny, we further analyzed SNPs from the coding region to genotype the most characteristic H sublineages in Basques, namely H1 (G3010A), H3 (T6776C), and the putative autochthonous lineage H2a5 (T4592C) (Achilli et al., 2004; Pereira et al., 2005; Álvarez-Iglesias et al., 2009). All of them appeared relatively well represented in Navarre. Thus, for instance, 16 individuals were found to belong to subhaplogroup H1 (~15%), although this frequency is clearly below the average in the Iberian Peninsula (~23.5%, Achilli et al., 2004; Pereira et al., 2005). At a microgeographic scale, both the frequency of haplogroup H ($G = 6.05, df = 1, P < 0.05$) and the frequency of subhaplogroup H1 ($G = 6.14, df = 1, P < 0.05$) proved to be significantly lower in northern Navarre than in the reference Basque sample, as indicated by results of the likelihood ratio test. The putative autochthonous Basque sublineage H2a5 described by Álvarez-Iglesias et al. (2009) was present at a frequency of 2.7% in Basques from northern Navarre. Lineages of subhaplogroup H3 tended to show a higher frequency in Navarre (~11%) than in the Basque Country sample (~5%), though this difference did not reach statistical significance ($G = 2.50, df = 1, P = 0.114$).
Subhaplogroup J1c (T14798C) showed a frequency of 11.8% in northern Navarre, even slightly higher than the frequency in the Basque Country (10.8%). Therefore, the frequency of subhaplogroup J1c in total Basques peaks at 11.3%. Considering published data on HVS-I and HVS-II plus data on complete mitochondrial genomes (see Supporting Information Table 4), we found that the frequency of J1c in total Basques is the highest value reported to date at a European scale. In fact, J1c frequency in Basques proved to be significantly higher than the estimated ones in other neighboring Iberian populations such as the Pas valley [G = 5.38, df = 1, P < 0.05 (Cardoso et al., in press)], Saragossa [G = 5.69, df = 1, P < 0.05 (Martinez-Jarreta et al., 2000)], or Madrid [G = 5.37, df = 1, P < 0.05 (Budowle et al., 1999)].

Another point worth mentioning is that subhaplogroup U5b showed an outstanding frequency in Navarre (15.5%), whereas it barely reached 2.4% in the Basque Country sample (G = 9.07, df = 1, P < 0.003). Other European populations have been reported to exhibit even higher U5b frequencies, such as Finns (Hedman et al., 2007), Vienna Karelians (Lappalainen et al., 2008), and Saami (Tambets et al., 2004), with 18.0%, 21.8%, and 47.6%, respectively. Yet, frequency of U5b in these northern European populations results mainly from the profusion of the “Saami motif” (U5b1b1). This specific lineage was not detected among Basques from Navarre. In total, maternal lineages H, J1c, and U5b accounted for more than 61.0% of the haplogroup diversity in the population examined.

Finally, haplogroup cluster HV0, which includes haplogroup V (Torrioni et al., 2006) showed a rather low frequency in northern Navarre (2.7%), in accord with findings for the Basque Country reference sample (5.5%), but in contrast to earlier studies on Basque mtDNA (~20%, Torroni et al., 2001). Overall, seven haplogroups and subhaplogroups were common to both Basque subpopulations, whereas 11 haplogroups were exclusive from Navarre and eight were found only in the Basque Country sample.

**Phylogenetic network**

Figure 2 shows the phylogenetic network for total Basques (Navarre + Basque Country reference sample). The topology of the network showed three major mtDNA haplogroup clusters: R0, JT, and UK. Cluster R0 included haplogroups R0a1, HV*, HV0, and H. Within this cluster, haplogroup H formed a star-like cluster mainly represented by subhaplogroups H1 (G3010A) and H3 (T6776C). Subhaplogroup H1 grouped most of the Basque Country individuals, whereas subhaplogroup H3 was mainly represented by Navarre. In both cases, we observed haplotypes shared between the two Basque subpopulations. Common haplotypes were also found for the autochthonous lineage H2a5.

Within the cluster JT, haplogroup T was exclusively represented in Navarre, except for an individual from the Basque Country classified into subhaplogroup T2. On the other hand, haplogroup J was subdivided into subhaplogroups J1b1a, J1c, J2a, and J2b1a, the lineage J1c being the most frequent in Navarre. Both the Navarrese and the Basque Country collections showed J1c haplotypes characterized by a specific transition at np 16,366. The J1c haplotype harboring polymorphisms at positions 16,069, 16,126, 16,278, and 16,366 has been previously reported in a number of Iberian populations, such as Liebana in Cantabria province (Maca-Meyer et al., 2003), Castile, Andalusia (Larruga et al., 2001), and a historic Basque population from Alava (Alzuale et al., 2006).

In the resultant phylogeny, a number of samples were assigned to cluster UK, and more specifically to subhaplogroups R1, U, U2e, U4, U5a, U5b, U5b1c, and U6a. Subhaplogroup U5b was clearly the most represented in total Basques, with a specific transition at np 16,366 detected in both subpopulations. This polymorphism had been previously identified in U5b individuals from the Basque Country (Bertranpetit et al., 1995), as well as in individuals from Saragossa (Martinez-Jarreta et al., 2000), Northeastern Spain (Crespo et al., 2000), and the French Basque Country (GenBank A.N. EU566779 and EU566787). The haplogroup K was present in both Basque subpopulations, but there were no common haplotypes.

We found homoplasy cases in different haplogroups, mainly associated to fast positions in the D-loop region. Some of them occurred at haplogroup and subhaplogroup defining sites: 16,296 (subhaplogroup T2), 16,311 (haplogroup K), 72 (haplogroup HV0), and 73 (haplogroup H).
Results of the CA are shown in Figure 3. The two-dimensional representation accounted for 32.9% of the total inertia (variation), and it proved to be discriminant of the population samples from Portugal, Italy, the Basque area, and the rest of Spain, which are the best represented regions in the analysis. Populations from central and southeastern Europe all clustered close to the centroid of the distribution. As can be noticed, the position of Navarre in the CA was mostly determined by the comparatively high frequency of haplogroups J1c and U5b within the European context.

The relatively isolated position of northern Navarre in the CA was statistically corroborated, since $F_{st}$ estimates revealed significant heterogeneity in all but one of the pairwise comparisons between Basques from Navarre and the rest of European populations (including Iberian Basques), suggesting a weak influence of the gene flow in the shaping of the mitochondrial genome of Navarrese Basques. The only comparison where genetic heterogene-
ity was not detected was the one between Navarre and French Basques ($P = 0.472 \pm 0.014$). It is not completely clear whether failure to find genetic heterogeneity here reflects gene flow between both regions or merely lack of statistical robustness of the test arising from the small size of the French Basque sample ($N = 24$).

Interestingly, those Iberian populations that have been related to the Franco-Cantabrian refuge such as Navarre (NAVA), Pas valley (PASV), and the Basque Country samples (BASG: Basques from the Goierri valley), BASF (French Basque Country), SARG (Saragossa), MADR (Madrid), NESP (Northeast Spain), GALC (Galicia), NPOR (North Portugal), CPOR (Central Portugal), SPOR (South Portugal), WAUT (West Austria), MECK (Mecklenburg, northeastern Germany), MACE (Macedonia), WSLK (West Slovakia), ESLK (East Slovakia), WBOH (West Bohemia, Czech Republic), ITBO (Bologna, Italy), ITMO (Modena, Italy), ITTFI (Firenze, Italy), ITAN (Ancona, Italy), ITRO (Roma, Italy), ITPV (Pavia, Italy), ITTO (Torino, Italy), and ITTR (Terni, Italy). Solid circles, Basque samples; open squares, Iberian Peninsula, excluding Basques; solid squares, Central and Southeastern Europe; and open circles, Italian Peninsula.

As for the rest of the European populations involved, all the Iberian samples plotted in the positive segment of axis 1 and were separated by the second axis in two (Spanish and Portuguese) relatively well-differentiated clusters (excepting PASV). Likewise, all Italian samples jointly with the central European collections segregated in the negative segment of axis 1 and relatively separated between them by axis 2. It is worth mentioning that, contrasting with the results obtained for the Basque subgroups, pairwise $F_{st}$ values revealed no significant genetic heterogeneity between the eight Italian samples included in the CA, except for the comparison between Bologna and Firenze ($P = 0.0137 \pm 0.004$). Similar results were obtained for the different subgroups from Portugal and Slovakia. In all, results of the CA were revealing of a certain degree of genetic heterogeneity among the Basque subpopulations and, at the same
time, of their distinctiveness within the European context from the mtDNA viewpoint (see Supporting Information Table 5).

**DISCUSSION**

mtDNA diversity data on the autochthonous Basque population settled in northern Navarre are presented herein for the first time. The most remarkable findings emerging from this work point toward two essential issues. On the one hand, similarities concerning the low frequency of haplogroup V, the prominent frequency of subhaplogroup J1c, and above all, the presence of the putative Basque lineage H2a5 (Alvarez-Iglesias et al., 2009) in the native Basque collections from northern Navarre and the Basque Country, are all supportive of a common ancestry for the different Iberian Basque subpopulations. On the other hand, substantial differences in the frequency distribution of some specific mtDNA lineages (e.g., subhaplogroups H1 and U5b) provide noticeable signs of local genetic microdifferentiation among Basque geographical groups, supported by the findings of the pairwise \( F_{st} \) comparisons. Genetic drift processes associated with local founding events and/or population bottlenecks would be the causative agents of the microevolutionary changes detected in our study.

Torroni et al. (1998, 2001) proposed that haplogroup V originated in the easternmost area of the Cantabrian Cornice soon after the LGM, based on both the coalescence age and the spatial distribution pattern of this lineage across Eurasia. However, haplogroup HV0 (the cluster including haplogroup V) proved to have a low frequency in northern Navarre (2.7%), coinciding with the results obtained by Alfonso-Sánchez et al. (2008) in a present-day Basque collection and also with findings derived from ancient DNA studies on prehistoric and historic populations from the Basque Country (Izagirre et al., 1999; Alzuale et al., 2005). Likewise, in a recent mtDNA-based phylogenetic study at a European scale (Garcia et al., 2010), the authors found no evidence for the expansion of haplogroup HV0 from the Cantabrian Cornice.

The prominent frequency of subclade J1c in northern Navarre (11.8%) is the highest value observed to date at a European scale. Overall, the frequency of haplogroup J is substantially higher than those reported in previous mtDNA studies for other Basque samples (Bertranpetit et al., 1995; Corte-Real et al., 1996; Garcia et al., 2011), whereas it is similar to data reported for different prehistoric and historic Basque populations (Alzuale et al., 2005). Considering also the high frequency of lineage J1c in the Basque Country reference sample (10.9%), the pooled frequency of J1c in total Basques represents a maximum (11.4%) as well. The frequencies of J* and J1* have also been reported to show relevant values in the British Islands (Helgason et al., 2001), but these values have been obtained from data solely on HVS-I. Consequently, the lack of HVS-II and/or coding region data does not allow assessing the specific contribution of subhaplogroup J1c in these European populations. The frequency distribution pattern of this lineage within the European context still requires in-depth studies of the phylogeny of haplogroup J, especially in populations from the Franco-Cantabrian region.

Interestingly, subhaplogroup H2a5, recently proposed as an autochthonous clade in the Basque Country (Alvarez-Iglesias et al., 2009), was also observed in Navarrese Basques. Such a finding gives further support to the notion of both Basque subpopulations having a common origin, although partly masked as a result of local microdifferentiation processes promoted by genetic drift.

Concerning the differentiating features between the Basque geographical groups, the exceptionally high frequency of subhaplogroup U5b in northern Navarre relative to other European collections is another major finding. The relative abundance of U5b in Navarre contrasts notably with the low frequency of this subhaplogroup in the Basque Country (Alfonso-Sánchez et al., 2008). Based on the high diversity of the U5b1b cluster in western and southern Europe, Tambets et al. (2004) suggested that these regions, rather than Eastern Europe, could be the place of origin of U5b1b. Likewise, Achilli et al. (2005) postulated that this lineage could have been connected with groups of hunter-gatherers that inhabited the Franco-Cantabrian refuge during the Upper Paleolithic, on account of the existence of individuals classified as U5b1b sharing specific polymorphisms at the coding region in distant populations such as Saami and Berbers. The high frequency of U5b in northern Navarre seems to support this hypothesis, even though the analysis of the control region does not allow a more refined classification of this lineage. It is important to note that the specific transition (G16319A) found in U5b individuals from northern Navarre was not present in the phylogeny constructed by Achilli et al. (2005), while we have found that it is distributed in present-day populations from the Basque Country to Catalonia (Crespollo et al., 2000; Martinez-Jarreta et al., 2000; Alfonso-Sánchez et al., 2008), albeit usually at low frequencies (except in northern Navarre). This specific transition might therefore be directly related to the area of the glacial refuge in northern Spain; yet, analyses of complete mitochondrial genomes are crucial to tentatively subclassify this lineage as a new clade within subhaplogroup U5b.

Another conspicuous difference between Navarre and the Basque Country reference sample regarding haplogroup composition is the frequency of haplogroup H, and specifically subhaplogroups H1 and H3, whose implication in postglacial expansion events from the Franco-Cantabrian refuge is under discussion (Torroni et al., 2006; Garcia et al., 2011). Navarre collection exhibited a high frequency of subhaplogroup H3 in contrast to the predominance of H1 in the Basque Country, which provides a new line of evidence for local genetic differentiation among Basque geographical groups. Along these lines, Achilli et al. (2004) stated that the scenario proposed to explain the evolutionary history and present-day distribution of haplogroup V in Europe could be directly transposed to subhaplogroups H1 and H3, because of similarities in coalescence ages and frequency distribution maps. Thus, the high frequencies of H1 and H3 in Iberian Basque populations (and thereby in the Franco-Cantabrian refuge area) seem to give further support to this hypothesis.

As discussed above, various findings of this study provide evidence for a certain degree of genetic heterogeneity among Basque subpopulations. The inclusion of the Basque Country sample (Alfonso-Sánchez et al., 2008) in the analyses allowed detecting significant frequency differences for subhaplogroups H1 and U5b, among other signals of local genetic differentiation. The variation in haplogroup composition among samples from the Iberian Basque area had been already noted in previous mtDNA
studies on prehistoric and historic populations of the Basque Country (Izagirre et al., 1999; Alzuñá et al., 2005). These authors claimed post-Neolithic restructuring of the population settled in Basque lands during the Late Antiquity, between 5,000 and 1,500 years BP, to account for the differences in haplogroup distribution (especially for haplogroups H and V) between extinct and extant Basque populations (Alzuñá et al., 2005). However, this argument fails to explain the variability of maternal lineages among present-day Basque samples such as, for instance, variation in subhaplogroups H1 and U5b between northern Navarre and the Basque Country reference sample, or haplogroup V and subhaplogroup J1c between these two Basque collections and those analyzed in earlier studies (Bertranpetit et al., 1995; Corte-Real et al., 1996; Torroni et al., 1998). In our view, genetic microdifferentiation processes caused by genetic drift episodes could account for the spatial structuring of the mitochondrial genome detected in present-day native Basque subpopulations. Genetic drift would have been promoted by founder effects due to small demographic sizes of the human groups arriving in the Franco-Cantabrian refuge, and/or by population bottlenecks that could have occurred later, in a region historically isolated and characterized by low population densities. The conspicuously low diversity of mtDNA haplotypes observed in Navarre and in the Basque Country reference sample (see Alfonso-Sánchez et al., 2008) relative to other European populations seems to support the substantial impact of the genetic drift in this region of the Iberian Peninsula. Population subdivision, and accordingly, the spatial genetic heterogeneity among autochthonous Basque geographical groups have been reported in several previous studies (de Pancorbo et al., 2001; Pérez-Miranda et al., 2004, 2005). Genetic drift could be also a plausible explanation for the great genetic heterogeneity in terms of haplogroup composition between the two Basque subgroups and the Pas valley population (see Fig. 3), which is characterized by relevant frequencies of lineages HV0 and T2 (Cardoso et al., 2010). This population, also located in the Franco-Cantabrian refuge area, is settled in the mountainous region of Cantabria province (northern Spain), and has remained genetically isolated from neighboring populations (including Basques) until very recent times.

Results of this work provide a different insight, from the maternal legacy viewpoint, to the question of whether the Basque geographical groups are genetically homogeneous or they present a certain genetic structuring, a debate that has been recently stimulated by conflicting findings from studies based on high-density SNP genotyping (Layouni et al., 2010; Rodríguez-Ezpeleta et al., 2010). Our analysis of the mtDNA haplogroup composition in northern Navarre strongly suggests the existence of local microdifferentiation among populations from the Franco-Cantabrian refuge area, and more specifically, among Basque subpopulations. In addition, our findings seem to give further support to the major role of the Franco-Cantabrian refuge in the postglacial resettlement of Europe. Bearing in mind that long-term population fragmentation in the Franco-Cantabrian refuge area could have generated extreme genetic drift events, additional studies aimed at the molecular dissection of major mtDNA haplogroups into clades of younger ages and more restricted geographic distribution could be highly advantageous in exploring all the diversity of female genetic lineages that coexisted in the glacial refuges. To that end, analyses at the highest resolving power (complete mtDNA sequences) in different native populations from the Cantabrian Cornice region would constitute a reliable approach to construct a more complete and refined mitochondrial phylogeny of this interesting southwestern European region.

ACKNOWLEDGMENTS

The authors are grateful to all voluntary donors who cooperated generously for the development of this study.

LITERATURE CITED


American Journal of Physical Anthropology


