Forensic Science International: Genetics 27 (2017) 142–148

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Research paper

# Characterization of the Iberian Y chromosome haplogroup R-DF27 in Northern Spain

Patricia Villaescusa[a], María José Illescas[a], Laura Valverde[a], Miriam Baeta[a], Carolina Nuñez[a], Begoña Martínez-Jarreta[b], Maria Teresa Zarrabeitia[c,d], Francesc Calafell[e], Marian M. de Pancorbo[a,*]

[a] BIOMICs Research Group, Lascaray Research Center, University of the Basque Country UPV/EHU, Vitoria-Gasteiz, Spain
[b] Laboratory of Genetics and Genetic Identification, University of Zaragoza, Spain
[c] Unit of Legal Medicine, University of Cantabria, Av. Herrera Oria, s/n, 39011 Santander, Cantabria, Spain
[d] Instituto de Investigación Marqués de Valdecilla (IDIVAL). Avda. Cardenal Herrera Oria s/n, 39011 Santander, Cantabria, Spain
[e] Departament de Ciències Experimentals i de la Salut, Institute of Evolutionary Biology (CSIC-Universitat Pompeu Fabra), Universitat Pompeu Fabra, CEXS-UPF-PRBB, Barcelona, Catalonia, Spain

ARTICLE INFO

*Article history:*
Received 26 June 2016
Received in revised form 22 December 2016
Accepted 29 December 2016
Available online 29 December 2016

*Keywords:*
Y chromosome
R haplogroup
Iberia
DF27
Z195

ABSTRACT

The European paternal lineage R-DF27 has been proposed as a haplogroup of Iberian origin due to its maximum frequencies in the Iberian Peninsula. In this study, the distribution and structure of DF27 were characterized in 591 unrelated male individuals from four key populations of the north area of the Iberian Peninsula through the analysis of 12 Y-SNPs that define DF27 main sublineages. Additionally, Y-SNP allele frequencies were also gathered from the reference populations in the 1000 Genomes Project to compare and obtain a better landscape of the distribution of DF27. Our results reveal frequencies over 35% of DF27 haplogroup in the four North Iberian populations analyzed and high frequencies for its subhaplogroups. Considering the low frequency of DF27 and its sublineages in most populations outside of the Iberian Peninsula, this haplogroup seems to have geographical significance; thus, indicating a possible Iberian patrilineal origin of vestiges bearing this haplogroup. The dataset presented here contributes with new data to better understand the complex genetic variability of the Y chromosome in the Iberian Peninsula, that can be applied in Forensic Genetics.



## 1. Introduction

The inference of bio-geographical ancestry using markers with population-differentiated variation can provide investigative leads in forensic cases when eyewitness testimony or a database hit are not available [1]. A first step in forensic ancestry inference may be the analysis of the Y chromosome or the mitochondrial DNA (mtDNA), as they are uniparental markers differentiated geographically. Both lineages are not afected by recombination and correlate with continental regions [1,2]. Y chromosome and mtDNA, along with autosomal ancestry informative markers (AIMs) can be used for inferring the bio-geographical origin of a vestige, as they are tools that complement each other.

The analysis of Y chromosome SNPs (Y-SNPs) has revealed the existence of specific lineages in human populations at continental and regional levels [3]. The Y chromosome diversity analysis performed in multiple European populations disclosed the existence of significant frequency clines in the major patrilineal lineages [4]. The most frequent paternal lineage in Europe is R1b [5], being haplogroup R-M269 the most common in Central and Western Europe [6,7], with frequencies ≈0.4 in Italy and Germany, ≈0.6 in Britain, France, and the Iberian Peninsula; and up to >0.8 in Ireland and the Basque Country [7].

The origin of R-M269 has been the subject of great controversy [6–8], as it was originally believed to have originated in the Palaeolithic [9,10]. More recent analysis [11,12] suggested that this lineage had a Neolithic origin, but this claim was challenged [7] due to the Y-STR choice for computing the coalescence times and sample ascertainment. The last studies involving next-generation sequencing (NGS) of the Y-chromosome [8,13] and the analysis of ancient DNA [14] bring light to the debate, as they support more

* Corresponding author at: BIOMICs Research Group, Lascaray Research Center, University of the Basque Country UPV/EHU, Avda. Miguel de Unamuno, 3, 01006 Vitoria-Gasteiz, Spain.
*E-mail address:* marian.mdepancorbo@ehu.eus (M.M. de Pancorbo).

http://dx.doi.org/10.1016/j.fsigen.2016.12.013
1872-4973/

recent origin and continent-wide expansion of the main European patrilineages, including R-M269 (≈5 KYA, middle Neolithic).

R-M269 is split into geographically localized sublineages, being the main branches U106 (more frequent in Central-Northern Europe) and P312 (Western and South-Western Europe) [6,7]. The latter, in turn, trifurcates into U152, M529, and DF27. U152 is common in northern Italy and the Alpine region, while M529 is nearly restricted to the British Isles and Britanny [6,7,12]. DF27, instead, shows its maximum frequencies in the Basque Country, from where it decreases gradually into the rest of the Iberian Peninsula; elsewhere, it is much rarer. Therefore, it has been suggested that DF27 is a distinctive Iberian ancestry haplogroup, where it likely first originated [6].

Given the forensic interest that derives from the near-specificity of DF27 in Iberia, we have characterized the structure and distribution of the DF27 lineage through the dissection in its sublineages [15–17]. For that purpose, we have analyzed the population that showed the highest frequency for DF27 [6], along with other four surrounding populations from the north of the Iberian Peninsula. In order to reach this aim, the Y-SNPs defining DF27 and its sublineages, Z195, Z196, L617, L881, Z220, Z278, M153, L176.2, S68, M167, and DF17 were genotyped. Additionally, Y-STR data from the Northern Iberia populations here analyzed were compiled from other studies [6,18–20].

## 2. Material and methods

### 2.1. Sample collection

A total of 591 healthy unrelated males from four different populations from Northern Iberia (from West to East: Asturias, Cantabria, Basque Country, and Aragon) were obtained. Informed consent was obtained from all individuals participating in the study. Human DNA samples were extracted from saliva (Asturias and Basque Country), peripheral blood (Cantabria and Basque Country) or blood stains collected on FTA paper (Aragon). Samples from Cantabria and Aragon were provided by the collection of the University of Cantabria and from the University of Zaragoza, respectively. Samples from Asturias were provided by the Spanish National DNA Bank-Carlos III (BNADN). Favorable ethical reports were obtained (Faculty of Pharmacy UPV/EHU, September 26th 2008; CEISH/119/2012, BNADN Ref. 12/0031). Fig. 1 shows the geographic location of the selected populations.

The Basque population was treated as two separate groups, native and resident Basques. Significant genetic differences were found between the two Basque groups, so it is not recommended to treat these populations as a single sample [21]. The inclusion criteria used in order to define natives was the Basque origin of surnames and birthplaces of the individuals and their ancestors going back at least three generations. The resident group corresponds to those individuals that live in the Basque Country but whose paternal ancestors are not native Basques, coming from elsewhere in Spain.

### 2.2. Y-SNP analysis

The samples were genotyped in a hierarchical manner for the following Y-SNPs within the R-DF27 haplogroup: DF27, Z195, Z196, Z220, Z278, M153, L176.2, M167 (also known as SRY2627 [16]), S68, DF17, L617, and L881 [15,22]. More details about the phylogeny are represented in Fig. 2. The Y-SNPs selected in this study correspond to the diagnostic positions that determine the main sublineages of DF27 haplogroup. These positions were chosen following the minimal reference phylogeny for the human Y chromosome [22], supplemented when necessary with the

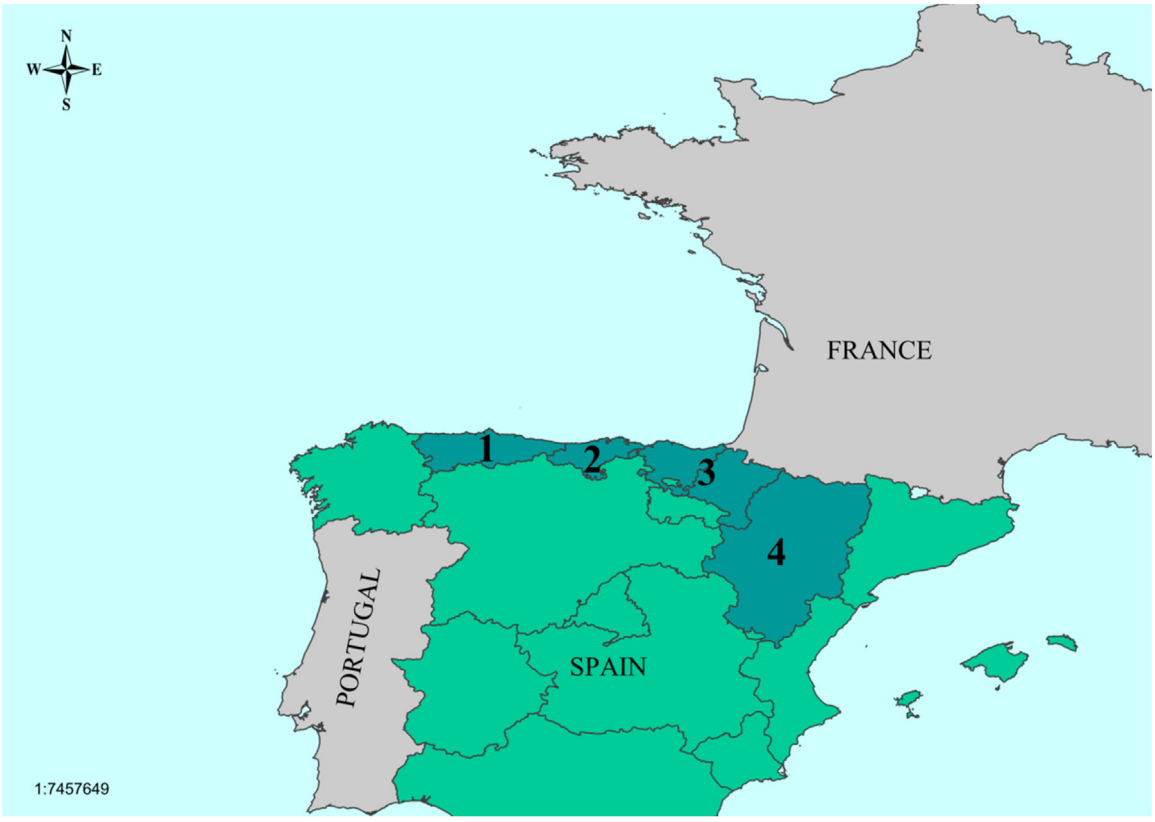


**Fig. 1.** Geographic location of the studied populations in Northern Spain. 1: Asturias; 2: Cantabria; 3: Basque Country; 4: Aragon.