DISSECTION OF THE DF27 PATERNAL LINEAGE IN SOUTH-WESTERN EUROPE

P. Villaescusa[1], N. Solé-Morata[2], A. Carracedo[3], K. Rouault[4], C. Férec[4], O. Hardiman[5], A. Santurtun[6], S. Jiménez[7], M. F. Pinheiro[8], B. M. Jarreta[9], F. Calafell[2], M. M. de Pancorbo[1]

[1] BIOMICs Research Group, Lascaray Research Center, University of Basque Country UPV/EHU, Vitoria-Gasteiz, Spain
[2] 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
[3] Forensic Genetics Unit, Institute of Legal Medicine, University of Santiago de Compostela, Spain; Galician Foundation of Genomic Medicine (SERGAS), CIBERER (University of Santiago de Compostela), Santiago de Compostela, Spain.
[4] Inserm UMR1078, Génétique, Génomique fonctionnelle et Biotechnologies, Brest 15 Cedex 2, France.
[5] National Neuroscience Centre, Beaumont Hospital, Dublin, Ireland.
[6] Unit of Legal Medicine, Department of Physiology and Pharmacology, University of Cantabria, Spain.
[7] Institute of Legal Medicine of Alicante, Spain.
[8] National Institute of Legal Medicine and Forensic Sciences, Portugal.
[9] Laboratory of Genetics and Genetic Identification, University of Zaragoza, Spain.

Introduction

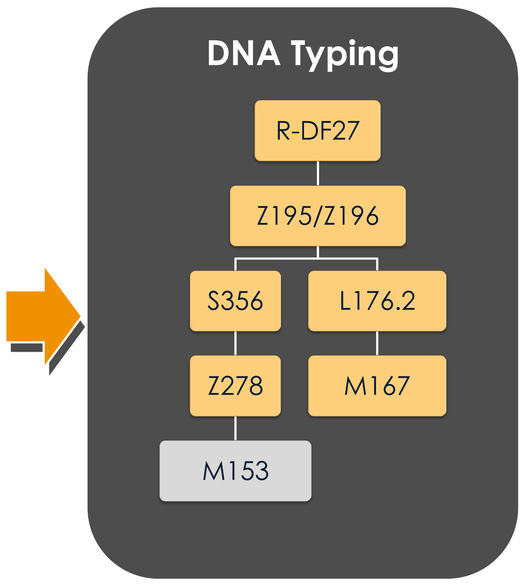
The analysis of Y chromosome SNPs (Y-SNPs) has revealed the existence of specific paternal lineages in human populations at continental and regional level [1], being the haplogroup R-M269 the most common in central and Western Europe [2,3,4]. Among its many subhaplogroups, DF27 (which branches from P312/S116) has revealed a striking frequency in the Iberian Peninsula, showing maximal frequencies in the Basque Country [4].

The aim of the present research was to investigate in detail the structure and variability of the DF27 haplogroup in populations from the Iberian Peninsula and South-Western Europe and to compare these distributions to other worldwide populations.

Materials and Methods

Populations
- Spain (N= 2153): ALI, AND, AST, ARA BCN, CANT, GAL, MAD, NV, NNV, TAR, CAS, CAT, GIR, LEI, MAL, PEN, PER, PIR, TER, and VAL.
- Portugal (N= 109)
- France (N= 585): ALS, AUV, BRT, PAR, TOU, LIL and MAR.
- Ireland (N= 146)

Data extracted from 1000 Genomes [5]
- Europe: GBR, TSI and FIN.
- America: CEU, CLM, MXL, PER and PUR.



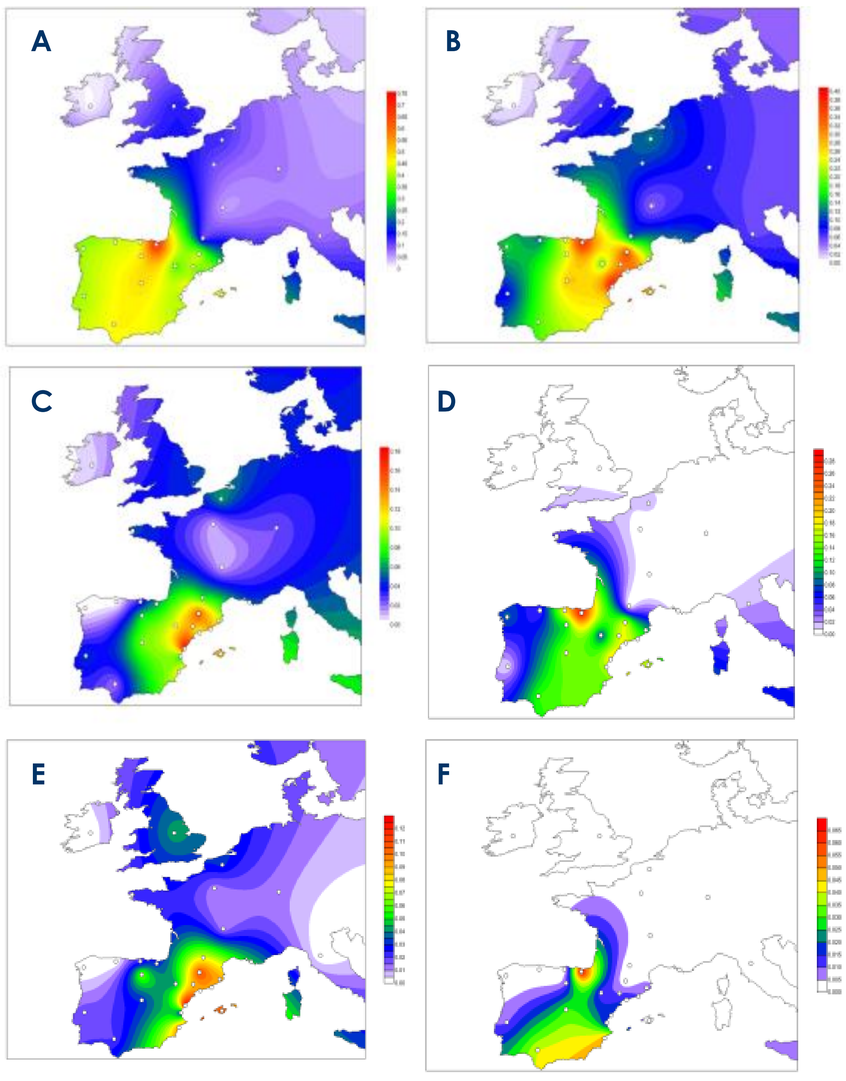
Statistical analyses
- Manual calculation of frequencies for each SNP
- Frequency maps were constructed with Golden Surfer Software v12
- Principal component analysis (PCA) implemented using Past3 software

Methodology details: In orange Y-SNPs analyzed by High resolution Melting (HRM); In light grey Y-SNPs analyzed by Sanger sequencing.

Population legend: Alicante (ALI), Andalucía (AND), Asturias (AST), Aragón (ARA), Barcelona (BCN), Cantabria (CANT), Galicia (GAL), Madrid (MAD), NV (Native Basques (NV), Non native Basques (NNV), TAR (Tarragona), Castellón (CAS), Central Cataluña (CAT), Girona (GIR), Lleida (LEI), Mallorca (MAL), Penedes (PEN), Peri (PER), Pyrenees (PIR), Terres de l'bre (TER), Valencia (VAL), Strasbourg (ALS), Clermont-Ferrand (AUV), Brest (BRT), Paris (PAR), Toulouse (TOU), Lille (LIL), Marseille (MAR). GBR (Britain), TSI (Tuscany), FIN (Finland), CEU (Utah residents), CLM (Colombia), Mexico (MXL), PER (Peru) and PUR (Puerto Rico).

Results and Discussion

Figure 1. Frequency distribution maps of the data compiled in this study and the data from Italy and Britain extracted from 1000 Genomes [5]. The white circles in the maps indicates the samples of populations analyzed. A) DF27; B) Z195; C) L176.2; D) S356; E) M167; F) M153 .



- Iberian Peninsula displays high frequencies (30-50%) of DF27, but it reaches striking frequencies in the native Basques(70%).
- Outside of Iberia the frequency decreases as it gets further away from the Peninsula, being below 15%.
- Z195 main sublineages, S356 and L176.2, display significant differences in their frequency distribution among the Iberian Peninsula.
- M167, Z278 and M153 maintain their maximum frequencies in the same geographical area as their predecessor haplogroups.
- The PCA displays Iberian populations grouped together and with Colombia and Puerto Rico, which show high frequencies for DF27 and its subhaplogroups.

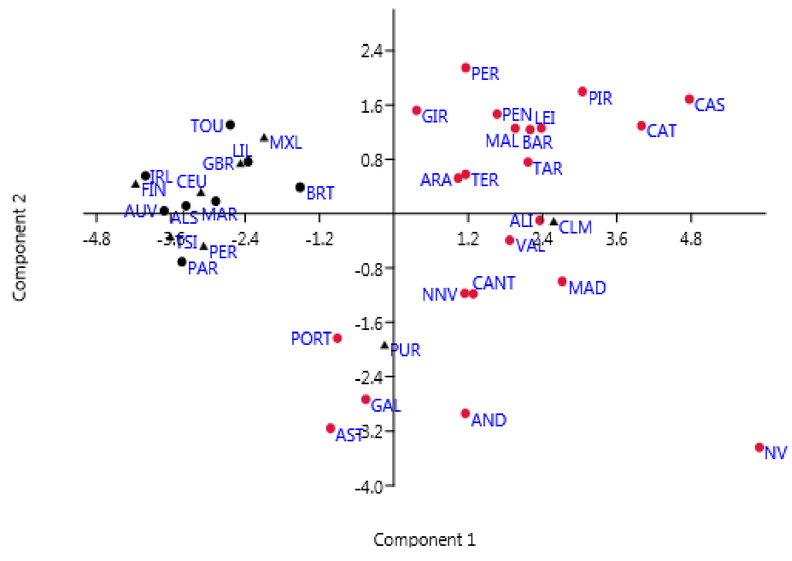


Figure 2. Principal component analysis (PCA) based on haplogroup frequencies of the populations analyzed and data compiled from 1000 Genomes [5]. Component 1: 57.71%; Component 2: 15.73%. Red dots correspond to populations from the Iberian Peninsula, black dots correspond to other European populations and black triangles correspond to populations extracted from 1000 Genomes data [5]. Population legend: Alicante (ALI), Andalucía (AND), Asturias (AST), Aragón (ARA), Barcelona (BCN), Cantabria (CANT), Galicia (GAL), Madrid (MAD), NV (Native Basques (NV), Non native Basques (NNV), TAR (Tarragona), Castellon (CAS), Central Cataluña (CAT), Girona (GIR), Lleida (LEI), Mallorca (MAL), Penedes (PEN), Peri (PER), Pyrenees (PIR), Terres de l'bre (TER), Valencia (VAL), Strasbourg (ALS), Clermont-Ferrand (AUV), Brest (BRT), Paris (PAR), Toulouse (TOU), Lille (LIL), Marseille (MAR), Ireland (IRL), Porto (PORT), Britain (GBR), Tuscany (TSI), Finland (FIN), Utah residents (CEU), Colombia (CLM), Mexico (MXL), Peru (PER) and Puerto Rico (PUR).

Conclusion

- DF27 displays the highest frequencies in populations from the Iberian peninsula.
- The frequency for DF27 and its sublineages decreases as it gets further away from Iberia.
- Colombia and Puerto Rico show high frequencies for DF27 and its subhapogroups, as expected since these two areas have received historically known Spanish migrations.

References

[1] P.A. Underhill, P. Shen, A.A. Lin, et al., Y chromosome sequence variation and the history of human populations., Nat. Genet. 26 (2000) 358–361.
[2] N.M. Myres, S. Rootsi, A.A. Lin, et al., A major Y-chromosome haplogroup R1b Holocene era founder effect in Central and Western Europe., Eur. J. Hum. Genet. 19 (2011) 95–101.
[3] G.B.J. Busby, F. Brisighelli, P. Sánchez-Diz, et al., The peopling of Europe and the cautionary tale of Y chromosome lineage R-M269., Proc. Biol. Sci. 279 (2012) 884–92.
[4] L. Valverde, M.J. Illescas, P. Villaescusa, et al., New clues to the evolutionary history of the main European paternal lineage M269: dissection of the Y-SNP S116 in Atlantic Europe and Iberia., Eur. J. Hum. Genet. 24(2016) 437-41.
[5] 1000 Genomes Project Consortium, A. Auton, L. Brooks, et al., A global reference for human genetic variation., Nature 526 (2015) 68-74.

BIOMICs Research Group
Universidad del País Vasco UPV/EHU
Lascaray Research Center
Avda. Miguel de Unamuno, 3
01006 Vitoria-Gasteiz (Spain)

Tel: +34-945014528
Fax: +34-945014458

Web:
www.biomics.es

ACKNOWLEDGEMENTS

FUNDING