

Commentary

Recent Radiation of R-M269 and High Y-STR Haplotype Resemblance Confirmed

Neus Solé-Morata, Jaume Bertranpetit, David Comas and Francesc Calafell*

Institut de Biologia Evolutiva (CSIC-Universitat Pompeu Fabra), Barcelona, Catalonia, Spain

Keywords: Y-chromosome STRs, Y-chromosome SNPs, haplogroups, R-M269

Dear Editor,

In the March 2014 issue of the *Annals of Human Genetics*, Larmuseau et al. (2014) published a paper on the frequent Y-chromosome haplogroup R-M269. In it, they describe the genetic variation in over 1000 males from Flanders (Belgium) and part of the Netherlands, and find an extreme homogeneity of Y-STR haplotypes across R-M269 subhaplogroups, which they attribute to the recent, rapid radiation of R-M269. However, Larmuseau et al. also wonder whether their results, obtained from a geographically restricted sample, can be generalized to the rest of Europe.

We have collected over 2500 male samples from Catalonia, the Balearic Islands, and Valencia (Spain), from volunteers carrying one of 50 different surnames. These samples have been typed for 68 Y-chromosome SNPs (Table S1) with the OpenArray technology (Martinez-Cruz et al., 2011) and 17 Y-STRs (from the AmpFISTR Yfiler kit), and we can attempt to replicate the findings by Larmuseau et al. The SNPs we genotyped within haplogroup R are different from those in Larmuseau et al., but they widely overlap and permit a similar degree of phylogenetic resolution (Fig. S1). For the analyses described later, for the sake of independence among samples, we have selected groups of individuals of different surnames, or within the same surname but with an expected coalescence among groups that would predate the establishing of surnames by setting a minimum difference of two STR mutations (Martínez-González et al., 2012). The frequency of R-M269 (or, in our case, of the phylogenetically close R-L23) is similar in both populations (61.0% in Flanders vs. 66.0% in Catalonia, Valencia, and the Balearic Islands), although the subhaplogroup composition is markedly different, with R-P312* rather than R-U106 prevailing in Catalonia, Valencia, and the Balearic Islands (Table S2). In our sample, out of 1341 different Y-chromosome 17-STR haplotypes, 53 were found in more than one subhaplogroup: all of them

were shared across R-M269 subhaplogroups, and none in any other branch of the Y-chromosome phylogeny ($\chi^2 = 26.91$, $P \cong 2 \times 10^{-7}$). Forty-two combinations of STR and SNP alleles were found to be shared across men bearing different surnames; 36 of those carried R-M269 subhaplogroups ($\chi^2 = 7.08$, $P = 0.0078$). Both observations were also found by Larmuseau et al., although with a wider set of STRs testing to the shallowness of the R-M269/R-L23 phylogeny, and contradicting the general expectation that genetic variation at the Y chromosome is deeply structured by haplogroup (Bosch et al., 1999). It should be noted that we typed fewer Y-STRs than did Larmuseau et al. (17 and 38, respectively); although we may expect to observe less haplotype diversity, the relative proportions of haplotypes shared within and across haplogroups are not expected to be different among both studies.

Next, we constructed 17-YSTR median networks with the haplotypes of six randomly chosen men for each R-M269 subhaplogroup, as Larmuseau et al. did (Fig. S2). We find more extensive clustering by subhaplogroup, with the haplotypes in the R-Z220 branch (R-Z220*, R-Z278*, and R-M153) grouping together in the network; this was confirmed in a network with the 10 STRs with the slowest mutation rates (Fig. S3). Larmuseau et al. found some clustering only for R-U198, which we did not type, but which, given its N. European distribution, is unlikely to be found at non-negligible frequencies in our sample. SNPs Z220 and Z278 were recently discovered from 1000 genomes data (Rocca et al., 2012; The 1000 Genomes Project Consortium, 2012); they were not typed by Larmuseau et al., and, symmetrically to U198, they are not bound to be present in Flanders in meaningful frequencies. In the 1000 genome data set (Rocca et al., 2012), they are restricted to the Iberian population or to populations of Iberian descent (Mexicans, Puerto Ricans, and Colombians); only two out of 1200 Sardinian samples would fall in this branch (Francalacci et al., 2013). In our sample, their joint frequency is 15.0%, and, if any is found in the sample in Larmuseau et al., they would fall within R-Z195* (3.5% in Flanders).

Y-STR haplotypes in the R-Z220 branch are distinct: their consensus haplotypes (constructed with the modal allele at

*Corresponding author: Francesc Calafell, Institut de Biologia Evolutiva (CSIC-Universitat Pompeu Fabra), Doctor Aiguader 88, 08003 Barcelona, Catalonia, Spain. Tel: 34-93-316-0842; E-mail: francesc.calafell@upf.edu

each locus) differ from the consensus haplotype for R-M269 at 2–4 loci (Table S3); no such difference is found for any other subhaplogroup.

Thus, we have been able to replicate most of the findings by Larmuseau et al.: R-M269, which is carried by two thirds of Western European men, is a recent haplogroup which rose quite rapidly in frequency, to the point that Y-STR haplotypes do not carry relevant phylogenetic information (with the R-U198 and R-Z220 exceptions). Their use in some genealogical and forensic applications may not be advisable unless SNPs are also typed.

Acknowledgements

This work has been made possible by the 2560 volunteers who contributed their samples. Sample collection was brilliantly organized by Ms. Cristina de Vasconcelos. Funding was provided by the Institut d'Estudis Catalans.

References

- Bosch, E., Calafell, F., Santos, F. R., Pérez-Lezaun, A., Comas, D., Benchemsi, N., Tyler-Smith, C. & Bertranpetit, J. (1999) Variation in short tandem repeats is deeply structured by genetic background on the human Y chromosome. *Am J Hum Genet* **65**, 1623–1638.
- Francalacci, P., Morelli, L., Angius, A., Berutti, R., Reinier, F., Atzeni, R., Pilu, R., Busonero, F., Maschio, A., Zara, I., Sanna, D., Useli, A., Urru, M. F., Marcelli, M., Cusano, R., Oppo, M., Zoledziewska, M., Pitzalis, M., Deidda, F., Porcu, E., Poddie, F., Kang, H. M., Lyons, R., Tarrier, B., Gresham, J. B., Li, B., To-fanelli, S., Alonso, S., Dei, M., Lai, S., Mulas, A., Whalen, M. B., Uzzau, S., Jones, C., Schlessinger, D., Abecasis, G. R., Sanna, S., Sidore, C. & Cucca, F. (2013) Low-pass DNA sequencing of 1200 Sardinians reconstructs European Y-chromosome phylogeny. *Science* **341**, 565–569.
- Larmuseau, M. H. D., Vanderheyden, N., Van Geystelen, A., van Oven, M., de Knijff, P. & Decorte, R. (2014) Recent radiation within Y-chromosomal haplogroup R-M269 resulted in high Y-STR haplotype resemblance. *Ann Hum Genet* **78**, 92–103.
- Martinez-Cruz, B., Ziegler, J., Sanz, P., Sotelo, G., Anglada, R., Plaza, S. & Comas, D. (2011) Multiplex single-nucleotide polymorphism typing of the human Y chromosome using TaqMan probes. *Investig Genet* **2**, 13.
- Martínez-González, L. J., Martínez-Espín, E., Álvarez, J. C., Albardaner, F., Rickards, O., Martínez-Labarga, C., Calafell, F. & Lorente, J. A. (2012) Surname and Y chromosome in Southern Europe: A case study with Colom/Colombo. *Eur J Hum Genet* **20**, 211–216.
- Rocca, R. A., Magoon, G., Reynolds, D. F., Krahn, T., Tilroe, V. O., Op den Velde Boots, P. M. & Grierson, A. J. (2012) Discovery of Western European R1b1a2 Y chromosome variants in 1000 genomes project data: An online community approach. *PLoS One* **7**, e41634.
- The 1000 Genomes Project Consortium (2012) An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56–65.

Supporting Information

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

Table S1 SNPs typed in our samples.

Table S2 Frequencies of subhaplogroups within R-L23 on a sample of 1148 independent chromosomes from Catalonia, Valencia, and the Balearic Islands (Spain).

Table S3 Modal allele, average, and variance of the allele size for each STR in each subhaplogroup.

Figure S1 Phylogenetic tree defined by the R-M269 SNPs genotyped in our samples.

Figure S2 Median-joining network based on 17 AmpFISTR loci.

Figure S3 Median-joining network based on 10 STRs with slow mutation rates: DYS389I, DYS390, DYS19, DYS393, DYS391, DYS392, GATAH4, DYS437, DYS438, and DYS448.

Received: 6 March 2014

Accepted: 8 April 2014