



Exploring the Western Mediterranean through X-chromosome

J. F. Ferragut¹ · K. Bentayebi² · A. Barbaro³ · M. Ramírez¹ · A. Y. Saguillo¹ · C. Ramon¹ · A. Picornell¹

Received: 7 October 2020 / Accepted: 17 December 2020 / Published online: 3 January 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH, DE part of Springer Nature 2021

Abstract

In this study, we investigate the forensic and population genetics properties of 21 X-chromosome markers (9 X-Alu insertions and 12 X-STRs) in a dataset composed of 716 individuals from 11 Western Mediterranean populations. The high values of combined forensic parameters indicate that this 21 X-loci panel can complement autosomal or uniparental markers in kinship analysis and complex deficient paternity testing in the populations studied. Population analyses revealed a lower differentiation between Western Mediterranean human groups for X-STRs than for X-Alu insertion polymorphisms. Moreover, X-chromosome markers suggest a sex-biased migration rate, confirming the predominance of patrilocality in this area.

Keywords X-chromosome · Alu insertions · X-STR · Western Mediterranean · Calabria · Balearic Islands · Valencia · Morocco · Investigator Argus X-12 kit

The genetic history of human populations in the Mediterranean basin has been widely studied due to their complexity. The Mediterranean Sea is located between three continents (Europe, Africa, and Asia) and has historically been, and still is, an important migratory route, and its coasts a point of meeting of different people and cultures. This area can be described as a melting pot, as the interaction and eventual admixture over time of different peoples have contributed to shape the genetic pool of modern Mediterranean populations. However, the mode and extent to which different peoples and demographic events (recurrent colonizations, expansions, relative isolations, bottleneck processes, admixture...) have contributed to the genetic variation of the present-day Mediterranean inhabitants are not yet fully understood and therefore remains a matter of discussion (e.g. [1–4]). For this

reason, Mediterranean populations have been the target of a large amount of genetic investigations aimed at reconstructing their genetic structure and population history based on different molecular markers such as autosomal polymorphisms (see, e.g., [5]), Y-chromosome (see revision in [6]), mtDNA (e.g. [7]), or X-chromosome (e.g. [2, 8]). Evaluating the genetic sub-structure in the Western Mediterranean area is especially relevant due to the existence of geographical and/or cultural isolates, such as Balearic populations [9] and Calabrian populations [10]. At the X-chromosome level, this region had not been studied with the set of markers included in the present study.

The particular characteristics of the X-chromosome make the study of markers located on it a valuable complement of the variation in autosomes in the field of population genetics [11]. Due to its inheritance pattern, the X-chromosome compared with autosomes has a lower recombination rate, a lower mutation rate, and a smaller effective population size (N_e), resulting in faster genetic drift. Consequently, both linkage disequilibrium (LD) and population structure in the X-chromosome are expected to be stronger than those in the autosomes. On the other hand, the X-chromosome spends 2/3 of its history in women and, consequently, their genetic diversity compared with other markers can help reveal demographic histories involving sex-biased migration and breeding patterns [12]. The X-chromosome has also become an extremely useful tool in the field of forensic genetics [13], helping to solve complex cases where at least one female is involved, which

✉ A. Picornell
apicornell@uib.es

¹ Departament de Biologia, Institut Universitari d'Investigació en Ciències de la Salut (IUNICS) i Laboratori de Genètica, Universitat de les Illes Balears, Carretera de Valldemossa, km 7.5, 07122 Palma de Mallorca, Illes Balears, Spain

² Biochemistry and Immunology Laboratory-Forensic Unit, Faculty of Sciences, Mohammed V Agdal University, Avenue Ibn Batouta Agdal, 1014 Rabat, BP, Morocco

³ Sezione di Genetica Forense, Studio Indagini Mediche e Forensi (SIMEF), 89128 Reggio Calabria, Italy

can be applied to cases of missing persons, incest, immigration, deficiency paternity, and other questioned relationships. STRs are the most used X-markers in forensic genetics, especially those included in commercial panels, although Indels and Alu insertion polymorphisms are also a reliable set of polymorphisms that could assist in human forensic genetic investigations (e.g. [14, 15]). The purpose of the study was to analyse the diversity of 21 X-chromosome markers (12 STRs and 9 Alu insertion polymorphisms) in 11 human Western Mediterranean populations, aiming to gain better knowledge of the genetic landscape of this region and assess the forensic potential of these markers in these populations.

Data from 716 individuals from 11 populations in the Western Mediterranean region (East Spain: Valencia, Ibiza, Minorca, and Majorca; Morocco: Arabs, Berbers, and Sahrawis; and South Italy: Sicily, Catanzaro, Cosenza, and Reggio Calabria) (Supplementary Figure 1) were analysed for two sets of X-chromosome genetic markers: (i) 9 X-chromosome Alu insertions (Ya5DP62, Yb8DP49, Yd3JX437, Yb8NBC634, Ya5DP77, Ya5NBC491, Yb8NBC578, Ya5DP4, and Ya5DP13) described by Callinan et al. [16] and (ii) 12 X-STRs included in the Investigator Argus X-12 kit (Qiagen GmbH, Hilden, Germany). This dataset is composed of the following: (a) new data on X-STRs from Sicily, Catanzaro, Cosenza, and Reggio Calabria populations and (b) data from genotyped samples as a part of other studies: X-STRs of Moroccan and Spanish individuals [17, 18] and X-Alu polymorphisms [19].

DNA samples were obtained after informed consent from unrelated healthy individuals with known ancestors until at least the third generation. Genotyping and statistical methods were performed as described in Ferragut et al. [14]. As a quality control, proficiency testing of the Spanish and Portuguese Speakers Working Group of the International Society for Forensic Genetics (GHEP-ISFG, <https://ghep-isfg.org>) was conducted. New variant alleles were sequenced by Sanger method and aligned using Geneious software version 7.1.3 (Biomatters, Ltd., Auckland, New Zealand). The relationship between distance matrices based on X- and Y-chromosome STRs and geographic locations of the 11 populations studies were estimated by performing standard Mantel correlation tests [20] using Arlequin v.3.5.1.2 software [21]. Statistical significance of the results was empirically tested over 100,000 permutations.

Allele frequencies for X-Alu polymorphisms and X-STRs of the eleven populations studied are presented in Supplementary Tables 1 and 2, respectively. Values of gene diversity (HET) are disclosed locus by locus in Supplementary Tables 3 and 4 and by population in Supplementary Table 5. Three Alu markers were polymorphic in all populations: YA5DP62, Yb8DP49, and Yd3JX437. The other six loci appeared as monomorphic in at least one of the

populations. The presence of the insertion was fixed: for Yb8NBC634 in Ibiza; for Ya5DP77 in Valencia and Moroccan Arab populations; for YA5DP13 in Reggio Calabria, Cosenza, and Catanzaro; and for Ya5NBC491 and Yb8NBC578 in 4 out of the 11 populations studied: Ibiza, Reggio Calabria, Cosenza, and Catanzaro (Supplementary Table 1). Absence of the Alu element was fixed for Ya5DPP4 in Minorca, Ibiza, and Cosenza. Most polymorphic Alu insertions showed diversities in the moderate to low range of the possible values for biallelic markers (≤ 0.5). Average values per marker ranged from 0.070 (Yb8NBC634) to 0.276 (Yb8DP49). Population average diversity (Supplementary Table 5) ranged from 0.057 (Cosenza) to 0.208 (Valencia). Values for Cosenza (0.057), Ibiza (0.080), and Catanzaro (0.088) were the lowest. As regards STRs, DXS10135 was found to be the most discriminating locus, with 32 alleles found throughout populations, whilst DXS10103 was the least polymorphic, with just 7 alleles. All populations showed a high average gene diversity for this set of X-STRs, with values ranging from 0.791 in Catanzaro to 0.831 in Moroccan Sahrawis. Regarding X-chromosome STR haplotypes, linkage groups (LG1 to LG4) presented 350, 250, 204, and 289 different haplotypes, respectively, in the 573 males studied (Supplementary Table 6). As can be seen in the summarised data of haplotypic diversity (Supplementary Table 7), the most polymorphic linkage group was LG1 and the least LG3. The most common haplotype, found in 20 individuals of 9 out of the 11 populations, was 19-12-30.2 in LG3. Ibiza and Sicilian populations revealed the lowest haplotype diversity (HD), with a low percentage of unique haplotypes compared to the other populations studied, while Moroccan Arabs and Valencia were the most diverse. The Moroccan population, altogether, showed higher proportions of unique haplotypes than Spanish and Italian ones. Pairwise linkage disequilibrium (LD) analysis, after the Bonferroni correction, showed only four significant p values, corresponding to the pairs Ya5dP62-Ya5NCB491 (Valencia), DXS10101-DXS10146 (Moroccans), DXS10103-DXS10101 (Sicily), and DXS10148-DXS10135 (Ibiza). Associations detected in Sicily and Ibiza involved STRs located in the same linkage group.

During STR profiling, a previously undescribed 19.3 allele was found in the DXS10135 marker in a Cosenza sample. Sanger sequencing was conducted to describe the structure of the new allele [22]. While the standard structure of a 20 allele is [AAGA]₃ GAAAG [GAAA]₁₇, the new allele displayed a single-base deletion (G), exhibiting the following structure: [AAGA]₃ [GAAA]₁₈.

In order to evaluate forensic efficiency, statistical parameters of forensic interest were calculated. Values for each marker and populations are presented in Supplementary Tables 3 and 4. Combined values are listed in Supplementary Table 5. Alu insertions were, as expected considering their nature of

biallelic polymorphisms, less informative than STRs. Nevertheless, biallelic markers can be useful in cases of inconclusive results using only STRs. In STRs, combined PD ranged from 1 in 3.336E+14 (Moroccan Sahrawis) to 1 in 9.007E+15 (Moroccan Arabs and Reggio Calabria) in females, and from 1 in 4.715E+08 (Catanzaro) to 1 in 2.771E+09 (Moroccan Arabs) in males. Combined MEC and MEC duo values indicated that these markers are highly informative in all populations (exceeding 0.99999999 and 0.99999, respectively). Catanzaro, Cosenza, and Ibiza displayed the lowest values in all forensic parameters for both sets of X-chromosome markers.

To test the genetic relationship between the Western Mediterranean populations studied and others from the literature, pairwise F_{ST} genetic distances were calculated and represented in an MDS plot (Supplementary Figure 2). Analyses were carried out separately for the two types of markers, since few populations in the literature have used the same 21 X-markers. In both MDS plots, it can be observed that the X-chromosome markers used were able to discriminate populations from different continents, with North African populations positioned closer to European than to sub-Saharan populations. Ibiza, for both sets of markers, showed a distanced position with respect to other geographically surrounding populations. This fact, together with lower genetic diversity, may be the result of the genetic drift experienced by this small, isolated population [9]. Calabria populations clustered separately from other European ones in the Alu markers' MDS plot, but not in the X-STR MDS. The F_{ST} values observed among Western Mediterranean populations (data not shown) were, on average, 10 times higher for Alu insertion markers than for STR markers, indicating that Alu polymorphisms are more discriminatory for population genetics studies in this region.

X-chromosome polymorphisms mostly behave as matrilineal markers, as two-thirds of the X-chromosomes descend from maternal origin. In order to compare the present results with the genetic variation in paternal lines in the same populations, a Mantel test was performed comparing genetic distances of Y-STRs vs X-STRs, yielding a significant correlation ($r_{Y1} = 0.648$, $p \leq 0.05$), in accordance with Tomas et al. [2] where X-chromosome SNPs were used. These results suggest that male and female movements have had a similar overall tendency around the Mediterranean Sea. Additionally, correlation between geographical vs Y-STR distances ($r_{Y1} = 0.954$) and vs X-STR variation ($r_{Y1} = 0.628$) was also both significant, contrasting with other studies [23, 24]. Nevertheless, Y-STR correlation index was much higher than X-STR value, indicating that genetic distances in males are more closely linked to geographical location. Therefore, the markers studied in the present study suggest that the migration rate in the Western Mediterranean has been higher in women than in men, as found with other markers in Europe and the

Mediterranean area [2, 25], except for particular human groups [14, 26].

In summary, our results indicate that this set of 21 X-chromosome markers is highly informative in the Western Mediterranean populations included in this study, and as such is therefore a useful complementary tool for kinship analysis, human identification, and resolution of historical and anthropological cases. Genetic distances between populations indicate that X-Alu insertion polymorphisms could be more discriminatory for population genetics studies than X-STR markers in this region, since they clustered human isolates such as Ibiza and Calabria separately. The Mantel test suggests a sex-biased migration rate in the Western Mediterranean, confirming the predominance of patrilocality in this area.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00414-020-02498-4>.

Acknowledgements We are grateful to Maria Trinidad Garcia (from the Serveis Científicotècnics of the Universitat de les Illes Balears) for her assistance and help with capillary electrophoresis procedures and Meryl Jones for the English language corrections. We would also like to sincerely thank all the people who volunteered to participate in this study.

Funding This work was partially supported by grant AAEE133/09 from the Direcció General de R + D + I (Comunitat Autònoma de les Illes Balears) and European Regional Development Fund (ERDF). KB was a recipient of an AVERROES grant (*Erasmus MUNDUS, European Commission*).

Compliance with ethical standards All participants provided written informed consent statements in accordance with the approval of the respective local committees: Italian Law D.lgs. 196/2003 (and subsequent supplement) and SIMEF internal procedures, Ethical Committee of the University of Mohammed V-Agdal (Morocco), and the *Direcció General de R + D + I* (Government of Balearic Islands, Spain) (procedure AAEE133/2009). The anonymity of the recruited samples was preserved throughout the study.

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Pimenta J, Lopes AM, Comas D, Amorim A, Arenas M (2017) Evaluating the Neolithic expansion at both shores of the Mediterranean Sea. *Mol Biol Evol* 34(12):3232–3242
2. Tomas C, Sanchez JJ, Barbaro A, Brandt-Casadevall C, Hernandez A, Ben Dhiab M, Ramon M, Morling N (2008) X-chromosome SNP analyses in 11 human Mediterranean populations show a high overall genetic homogeneity except in North-west Africans (Moroccans). *BMC Evol Biol* 8(1):75
3. Hernández CL, Pita G, Cavadas B, López S, Sánchez-Martínez LJ, Dugoujon JM, Novelletto A, Cuesta P, Pereira L, Calderón R (2020) Human genomic diversity where the Mediterranean joins the Atlantic. *Mol Biol Evol* 37(4):1041–1055

4. Zalloua PA, Platt DE, El Sibai M et al (2008) Identifying genetic traces of historical expansions: Phoenician footprints in the Mediterranean. *Am J Hum Genet* 83(5):633–642
5. Botigué LR, Henn BM, Gravel S et al (2013) Gene flow from North Africa contributes to differential human genetic diversity in southern Europe. *Proc Natl Acad Sci U S A* 110(29):11791–11796
6. Larmuseau MH, Ottoni C (2018) Mediterranean Y-chromosome 2.0—why the Y in the Mediterranean is still relevant in the postgenomic era. *Ann Hum Biol* 45(1):20–33
7. De Angelis F, Scorrano G, Martínez-Labarga C, Scano G, Macciardi F, Rickards O (2018) Mitochondrial variability in the Mediterranean area: a complex stage for human migrations. *Ann Hum Biol* 45(1):5–19
8. Santamaría J, Álvarez-Álvarez MM, Esteban ME, Ramon-Gurrea E, Moral P (2018) Dinucleotide (CA)_n tandem repeats on the human X-chromosome and the history of the Mediterranean populations. *Ann Hum Biol* 45(1):72–76
9. Biagini SA, Solé-Morata N, Matisoo-Smith E, Zalloua P, Comas D, Calafell F (2019) People from Ibiza: an unexpected isolate in the Western Mediterranean. *Eur J Hum Genet* 27(6):941–951
10. Tagarelli A, Piro A, Tagarelli G, Zinno F (2000) Color-blindness in Calabria (Southern Italy): a north–south decreasing trend. *Am J Hum Biol* 12(1):17–24
11. Schaffner SF (2004) The X chromosome in population genetics. *Nat Rev Genet* 5(1):43–51
12. Bustamante CD, Ramachandran S (2009) Evaluating signatures of sex-specific processes in the human genome. *Nat Genet* 41(1):8–10
13. Gomes I, Pinto N, Antão-Sousa S, Gusmão L, Gomes V, Amorim A (2020) Twenty years later: a comprehensive review of the X chromosome use in forensic genetics. *Front Genet* 11:926
14. Ferragut JF, Bentayebi K, Pereira R, Castro JA, Amorim A, Ramon C, Picornell A (2017) Genetic portrait of Jewish populations based on three sets of X-chromosome markers: Indels, Alu insertions and STRs. *Forensic Sci Int Genet* 31:e5–e11
15. Singh G, Sandhu HS, Srinivas Y, Matharoo K, Bhanwer AJS (2019) Genetic portrait of North-West Indian population based on X chromosome Alu insertion markers. *Int J Legal Med* 134(5):1655–1657
16. Callinan PA, Hedges DJ, Salem AH, Xing J, Walker JA, Garber RK, Watkins WS, Bamshad MJ, Jorde LB, Batzer MA (2003) Comprehensive analysis of Alu-associated diversity on the human sex chromosomes. *Gene* 317:103–110
17. Bentayebi K, Picornell A, Bouabdeallah M, Castro JA, Aboukhalid R, Squalli D, Misericòrdia M, Amzazi S (2012) Genetic diversity of 12 X-chromosomal short tandem repeats in the Moroccan population. *Forensic Sci Int Genet* 6(1):e48–e49
18. Ferragut JF, Bentayebi K, Castro JA, Ramon C, Picornell A (2015) Genetic analysis of 12 X-chromosome STRs in Western Mediterranean populations. *Int J Legal Med* 129(2):253–255
19. Bentayebi K, Ramon MM, Castro JA, Barbaro A, Bouabdeallah M, Amzazi S, Picornell A (2011) Inferring ethnicity from the X-chromosome ALU insertions: data from Western Mediterranean human groups. *Forensic Sci Int Genet Suppl Ser* 3(1):e27–e28
20. Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Res* 27(2 Part 1):209–220
21. Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* 10(3):564–567
22. Sumita DR, Whittle MR (2009) Updated allelic structures of the DXS10135 and DXS10078 STR loci. *Forensic Sci Int Genet Suppl Ser* 2(1):51–52
23. Bosch E, Calafell F, Pérez-Lezaun A, Clarimón J, Comas D, Mateu E, Martínez-Arias R, Morera B, Brakez Z, Akhayat O, Sefiani A, Hariti G, Cambon-Thomsen A, Bertranpetit J (2000) Genetic structure of north-west Africa revealed by STR analysis. *Eur J Hum Genet* 8(5):360–366
24. Quintana-Murci L, Veitia R, Fellous M, Semino O, Poloni ES (2003) Genetic structure of Mediterranean populations revealed by Y-chromosome haplotype analysis. *Am J Phys Anthropol* 121(2):157–171
25. Rasteiro R, Bouttier PA, Sousa VC, Chikhi L (2012) Investigating sex-biased migration during the Neolithic transition in Europe, using an explicit spatial simulation framework. *Proc Biol Sci* 279(1737):2409–2416
26. Robino C, Lacerenza D, Aneli S, Di Gaetano C, Matullo G, Robledo R, Calò C (2018) Allele and haplotype diversity of 12 X-STRs in Sardinia. *Forensic Sci Int Genet* 33:e1–e3

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.