Int J Legal Med (2009) 123:137–141
DOI 10.1007/s00414-008-0302-y

ORIGINAL ARTICLE

Genetic sub-structure in western Mediterranean populations revealed by 12 Y-chromosome STR loci

V. Rodríguez · C. Tomàs · J. J. Sánchez · J. A. Castro ·
M. M. Ramon · A. Barbaro · N. Morling · A. Picornell

Received: 10 June 2008 / Accepted: 18 November 2008 / Published online: 6 December 2008

Abstract Haplotype and allele frequencies of 12 Y-chromosome short tandem repeat (Y-STR) loci (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS385 a/b, DYS437, DYS438 and DYS439) included in the Powerplex® Y System were determined in seven western Mediterranean populations from Valencia, Majorca, Ibiza (eastern Spain), Sicily and three Calabrian provinces (southern Italy). Amongst the 554 males included in the study, 443 different haplotypes were observed of which 372 were only observed once. The other haplotypes were shared by two to seven men. The overall haplotype diversity was 0.9988±0.0002. These Y-STRs markers showed a low capacity of discrimination (56.3%) in the Ibiza population probably due to genetic drift. Comparisons between the populations studied and other neighbouring populations showed a clear genetic sub-structure in the western Mediterranean area.

Keywords Y-STRs · Eastern Spain · Ibiza · Sicily · Calabria

Introduction

The analysis of Y-chromosome short tandem repeats (Y-STRs) has become a very useful tool, both in evolutionary studies and forensic casework. Two sets of Y-STR systems called minimal and extended haplotypes have been widely accepted by the forensic community and are currently used in the worldwide database: “Y-STR haplotype reference database” (YHRD; http://www.ystr.org) [38]. Here, we present a population study carried out using the Powerplex® Y System (Promega, Madison, WI, USA) containing 12 Y-STR loci (the nine Y-loci of the minimal haplotype, as well as the DYS437, DYS438 and DYS439 loci).

The International Society of Forensic Genetics (ISFG) recommended the use of regional Y-STR haplotype databases and the verification that no population sub-structure exists before pooling data from different regions [16]. It is especially important to evaluate genetic sub-structure in the western Mediterranean area, due to the existence of some geographical and/or cultural isolates such as Balearic populations [24, 37] and Calabrian populations [36]. We studied the genetic constitution of seven western Mediterranean populations (Valencia and two Balearic Islands, from eastern Spain, Sicily and three provinces in the Calabria region, southern Italy) with the aim of establishing a Y-STR

Electronic supplementary material The online version of this article (doi:10.1007/s00414-008-0302-y) contains supplementary material, which is available to authorized users.

V. Rodríguez · J. A. Castro · M. M. Ramon · A. Picornell (✉)
Institut Universitari d’Investigacions en Ciències de la Salut (IUNICS) i Laboratori de Genètica, Departament de Biologia, Universitat de les Illes Balears,
Carretera de Valldemossa, km 7.5,
07122 Palma de Mallorca, Illes Balears, Spain
e-mail: apicornell@uib.es

C. Tomàs · J. J. Sánchez · N. Morling
Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health Sciences, University of Copenhagen,
11 Frederik V’s Vej,
2100 Copenhagen, Denmark

J. J. Sánchez
Delegación de Canarias,
Instituto Nacional de Toxicología y Ciencias Forenses,
Campus de Ciencias de la Salud, La Cuesta,
38320 Tenerife, Spain

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

haplotype regional database and to evaluate sub-structuring of geographical sub-sets in this Mediterranean region.

Materials and methods

Population samples

Blood samples were obtained from 554 unrelated males belonging to seven populations in the western Mediterranean area (Figure shown in ESM 1)—three Spanish populations: Valencia (eastern coast of the Iberian Peninsula; *n*=59), Majorca (*n*=91) and Ibiza (*n*=96) (Balearic Islands) and four Italian populations: Sicily (*n*=115) and three populations from the region of Calabria (southern Italy), Reggio Calabria (*n*=97), Cosenza (*n*=37) and Catanzaro (*n*=59). Protocols were approved by the Danish local ethical committee (KF-01-037/03). DNA was extracted by using QIAamp spin columns (Qiagen, Hilden, Germany) following the manufacturer's recommendations.

Y-STR typing

Amplification of the 12 Y-chromosome STRs loci DYS19, DYS389 I/II, DYS390, DYS391, DYS392, DYS393, DYS385 a/b, DYS437, DYS438 and DYS439 was carried out according to the kit Powerplex® Y System protocol, from 2 ng DNA template in a 10-µl final reaction volume, using a GeneAmp PCR system 2400 Thermal Cycler (Perkin-Elmer, Waltham, MA, USA).

For genetic typing, an ABI PRISM® 3100 Genetic Analyser along with GeneScan® 3.7 and Genotyper® v. 3.7 software (Applied Biosystems, Foster City, CA, USA) were used. All samples were tested twice. Allelic designation was based on comparison to the Powerplex® Y System allelic ladder. Allele nomenclature was according to the ISFG guidelines [16].

Sequence analysis

A new variant allele was sequenced on both strands. Briefly, the samples were amplified using unlabelled primers [3], amplicons were purified with a QIAquick PCR purification kit (Qiagen) and the sequence was determined using the Big Dye® Terminator Cycle Sequencing kit v. 3.1 (Applied Biosystems) and an ABI PRISM® 3130 Genetic Analyser (Applied Biosystems). Sequences were aligned using the Bioedit program v. 7.0.5.3 [17].

Quality control

Proficiency testing of the Spanish and Portuguese Working Group of the International Society for Forensic Genetics (GEP-ISFG, http://www.gep-isfg.org/) was carried out as quality control.

Statistical analysis

Allele and haplotype frequencies were estimated by gene counting. Haplotype and gene diversities, population differentiation parameters (F_{ST} and R_{ST}) and analysis of molecular variance (AMOVA) were calculated using ARLEQUIN v. 3.01 [11]. Discrimination capacity was calculated as the percentage of different haplotypes and haplotype match probability as 1-haplotype diversity. All statistical parameters were calculated for both minimal and extended haplotypes.

In order to examine the relationship of the populations studied with other neighbouring populations, Reynolds' genetic distances [30], calculated using PHYLIP v. 3.67 [12], were used to generate the multi-dimensional scaling (MDS) plot performed using the SPSS v. 15.0 (SPSS, Inc., Chicago, IL, USA).

Results and discussion

Allele frequencies and gene diversities of each Y-STR of the populations under study are shown in the table in ESM 2. DYS392 and DYS438 showed bimodal distribution of allele frequencies. In DYS392, modality was shared by DYS392-11 and DYS392-13 alleles, with DYS392-13 the most common allele amongst the Spanish populations and DYS392-11 the most frequent in southern Italy. These results are consistent with previous studies showing a longitudinal decrease of frequencies from the west to the east of the European landscape for DYS392-13 and, conversely, a decrease in the opposite direction for DYS392-11. The Neolithic demic diffusion could explain these two opposite patterns, with the DYS392-13 allele present in the proto-European gene pool [27, 28]. The clinal frequency pattern observed in the DYS438 system, with DYS438-12 as the most frequent in the Spanish and Sicilian populations and DYS438-10 the most frequent in Calabria, could also be due to the same Neolithic effect.

Gene diversities ranged from around 0.85 (in DYS385) to approximately 0.50 (in DYS392). Generally, Italian populations had higher gene diversities than the Spanish populations, following the same pattern found in bi-allelic Y-chromosome markers, with an increasing diversity trend from Spain to Greece, maybe due to the impact of the arrival of haplotypes in Europe from the Middle East [13]. Ibiza showed especially low gene diversities for DYS389II, DYS390, DYS391, DYS385, DYS438 and DYS439 loci.

Extra peaks were reproducibly obtained at DYS19 and DYS385, representing the presence of duplicated regions in

Table 1 Forensic parameters for the seven western Mediterranean populations studied using the minimal and the Powerplex® Y haplotypes

	Ibiza (n=96)	Majorca (n=91)	Valencia (n=59)	Sicily (n=115)	Catanzaro (n=59)	Cosenza (n=37)	R. Calabria (n=97)	Total (n=554)
Minimal 9 Y-STR haplotype								
Number of haplotypes	43	74	49	98	52	33	83	379
Unique haplotypes	19	53	37	72	37	23	58	299
Haplotype diversity ± SD	0.9695± 0.0061	0.9927± 0.0035	0.9930± 0.0047	0.9968± 0.0018	0.9959± 0.0039	0.9925± 0.0088	0.9968± 0.0020	0.9968± 0.0006
Discrimination capacity (%)	44.79	81.32	83.05	85.22	88.14	89.19	85.57	68.41
Match probability (%)	3.05	0.73	0.70	0.32	0.41	0.75	0.32	0.32
Powerplex 12 Y-STR haplotype								
Number of haplotypes	54	82	56	103	54	33	85	443
Unique haplotypes	30	70	49	86	43	26	68	372
Haplotype diversity ± SD	0.9807± 0.0050	0.9968± 0.0025	0.9982± 0.0035	0.9976± 0.0017	0.9971± 0.0037	0.9925± 0.0088	0.9972± 0.0020	0.9988± 0.0002
Discrimination capacity (%)	56.25	90.11	94.92	89.57	91.53	89.19	87.63	79.96
Match probability (%)	1.93	0.32	0.18	0.24	0.29	0.75	0.28	0.12

the Y-chromosome (e.g. [2, 6, 19, 20, 32]). Duplications were observed in six individuals: at locus DYS19, alleles 13 and 14 (once) and at loci DYS385, alleles 13–14–15 (once), 13–17–18 (twice) and 13–18–19 (twice).

Three alleles not included in the Powerplex® Y allelic ladder were observed. DYS438-7 and DYS438-13 alleles have been reported in other populations (e.g. [8, 23]), but to our knowledge, DYS19-9 has not been reported before. Sequence analysis (GenBank: FJ196286) confirmed the number of repeats attributed: $(TAGA)_3tagg(TAGA)_6$.

Amongst the 554 western Mediterranean males analysed, 443 different haplotypes were observed (Table in ESM 3), of which 372 were only observed once. The other haplotypes were shared by two to seven men. The most frequent haplotypes were h314 (15–12–29–22–10–11–14–14,14–16–10–12) and h356 (15–13–29–24–11–13–13–11,14–14–12–13), both found in seven men from the Ibiza population but absent from the other studied populations.

The haplotypes (without locus DYS437, not included in the YHR database) were searched against the haplotypes in the YHRD (release 18), and 149 haplotypes were matched to at least one YHRD sample. The most frequent haplotypes h314 and h356 matched with three and six samples of European origin, respectively, in a worldwide database of 38,761 haplotypes. It is noteworthy that 46 of the other haplotypes (almost all from Valencia, the Balearic Islands and Sicily) matched with north African or African samples. This result is concordant with other studies showing African influences in these Mediterranean populations (e.g. [14, 15, 24, 35]).

Table 1 shows the forensic parameters for the 12-loci Powerplex® Y System haplotypes compared with the diversity values of haplotypes based on the nine-loci minimal haplotype. The overall haplotype diversity only increased by 0.20% (ranging from 0% for Cosenza to 1.15% for Ibiza) by using the 12-loci Powerplex® Y System instead of the minimal haplotype.

The discrimination capacity ranged from 87.63% (Reggio Calabria) to 94.92% (Valencia) except in the Ibizan population (56.25%). Ibiza also showed a reduced genetic diversity in previous genetic studies [24, 25, 37]. These results are in accordance with the historical and demographic data of the island population (an isolated, consanguineous population with a reduced effective population size) [1, 22]. Therefore, the high haplotype match probability in Ibiza (1.93%) must be taken into account in forensic practice.

AMOVA analysis of the seven Mediterranean populations showed a significant value (F_{ST}=0.0499, P<0.0001). Pairwise analyses (Table in ESM 4) evidenced two

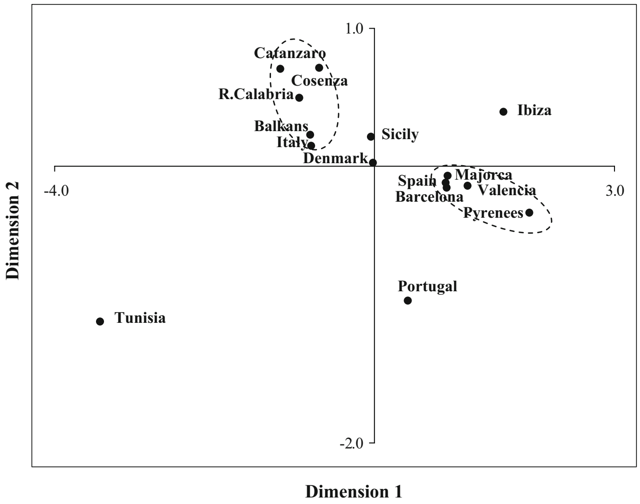

Fig. 1 MDS plot based on Reynolds' distances (Spain [23], Pyrenees [21], Barcelona [33], Portugal [26], Italy [29], the Balkans [5], Tunisia [9], Denmark [18], Majorca, Valencia, Ibiza, Catanzaro, Cosenza, Reggio Calabria and Sicily (this study))

significantly different sub-sets: one with the Spanish populations (Valencia, Ibiza and Majorca) and another with the Calabrian samples (Catanzaro, Cosenza and Reggio Calabria). Sicily was closer to the Italian than to the Spanish populations, although presented significant differences with Reggio Calabria ($P<0.005$). Significant differences between groups ($F_{CT}=0.0603$, $P<0.0001$) were found with a three-hierarchical AMOVA performed grouping the populations according to pairwise analysis results (Spanish populations, Sicily and Calabrian populations).

Figure 1 shows a multi-dimensional scaling plot where 15 populations have been included. The Tunisia population showed a displaced position, in accordance with other studies that have suggested that the Mediterranean Sea may have acted as a relative north-to-south geographic barrier to gene flow [4, 10, 28]. With the exception of Sicily and Ibiza, the Spanish and Italian populations under study grouped together with other Spanish and Italian samples, respectively. On the one hand, Sicily presented an intermediate position between Italian and Spanish populations. No consensus on the genetic landscape of the Sicilian population has been established to date. Whilst some authors claim the differentiation of the Sicilian population from Italy and from the western Mediterranean basin [7, 31], other studies indicate Sicily is closely related to other Italian populations [13, 15, 34]. On the other hand, Ibiza showed a large distance from the Spanish group, in accordance with the fact that Ibiza has important historical and genetic differences from other insular and continental populations in the western Mediterranean area [24, 25, 37].

In conclusion, the results of the present study provide a useful Y-STR haplotype dataset, for the western Mediterranean region, where some geographical and/or cultural isolates exist based on demographic, historical and genetic data. A clear genetic sub-structure between population groups (Spanish, Sicilian and Calabrian populations) was observed. Therefore, local databases must be used in the forensic field to correctly weigh the value of the evidence of a Y profile match. Special care should be taken in male identification in the Ibizan population, due to the very low discrimination capacity found for the 12 Y-STR loci included in the Powerplex® Y System.

Acknowledgements This work was partially supported by grants PRDT-2003-12099 and PRDIB-2006-687872 from the Direcció General de R+D+I (Comunitat Autònoma de les Illes Balears) and by Ellen and Aage Andersen's Foundation.

References

1. Alarco Von Perfall C (1981) Cultura y personalidad en Ibiza. Editora Nacional, Madrid
2. Balaresque P, Parkin EJ, Roewer L et al (2008) Genomic complexity of the Y-STR DYS19: inversions, deletions and founder lineages carrying duplications. Int J Legal Med (in press)
3. Beleza S, Alves C, González-Neira A, Lareu M, Amorim A, Carracedo A, Gusmão L (2003) Extending STR markers in Y chromosome haplotypes. Int J Legal Med 117:27–33
4. Bosch E, Calafell F, Comas D, Oefner PJ, Underhill PA, Bertranpetit J (2001) High-resolution analysis of human Y-chromosome variation shows a sharp discontinuity and limited gene flow between northwestern Africa and the Iberian Peninsula. Am J Hum Genet 68:1019–1029
5. Bosch E, Calafell F, González-Neira A et al (2006) Paternal and maternal lineages in the Balkans show a homogeneous landscape over linguistic barriers, except for the isolated Aromuns. Ann Hum Genet 70:459–487
6. Butler JM, Decker AE, Kline MC, Vallone PM (2005) Chromosomal duplications along the Y-chromosome and their potential impact on Y-STR interpretation. J Forensic Sci 50:853–859
7. Calò CM, Garofano L, Mameli A, Pizzamiglio M, Vona G (2003) Genetic analysis of a Sicilian population using 15 short tandem repeats. Hum Biol 75:163–178
8. Chang YM, Perumal R, Keat PY, Kuehn DL (2007) Haplotype diversity of 16 Y-chromosomal STRs in three main ethnic populations (Malays, Chinese and Indians) in Malaysia. Forensic Sci Int 167:70–76
9. Cherni L, Pereira L, Goios A et al (2005) Y-chromosomal STR haplotypes in three ethnic groups and one cosmopolitan population from Tunisia. Forensic Sci Int 152:95–99
10. Comas D, Calafell F, Benchemsi N, Helal A, Lefranc G, Stoneking M, Batzer MA, Bertranpetit J, Sajantilla A (2000) Alu insertion polymorphisms in NW Africa and the Iberian peninsula: evidence for a strong genetic boundary through the Gibraltar straits. Hum Genet 107:312–329
11. Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. Evol Bioinf Online 1:47–50
12. Felsenstein J (2007) PHYLIP (Phylogeny Inference Package) version 3.67. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle
13. Francalacci P, Morelli L, Underhill PA et al (2003) Peopling of three Mediterranean islands (Corsica, Sardinia, and Sicily) inferred by Y-chromosome biallelic variability. Am J Phys Anthropol 121:270–279
14. Gérard N, Berriche S, Aouizerate A, Dieterlen F, Lucotte G (2006) North African Berber and Arab influences in the western Mediterranean revealed by Y-chromosome DNA haplotypes. Hum Biol 78:307–316
15. Ghiani ME, Piras IS, Mitchell RJ, Vona G (2004) Y-chromosome 10 locus short tandem repeat haplotypes in a population sample from Sicily Italy. Legal Med 6:89–96
16. Gusmão L, Butler JM, Carracedo A et al (2006) DNA Commission of the International Society of Forensic Genetics (ISFG): an update of the recommendation on the use of Y-STRs in forensic analysis. Int J Legal Med 120:191–200
17. Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98
18. Hallenberg C, Nielsen K, Simonsen B, Sánchez J, Morling N (2005) Y-chromosome STR haplotypes in Danes. Forensic Sci Int 155:205–210
19. Hohoff C, Dewa K, Sibbing U, Hoppe K, Forster P, Brinkmann B (2007) Y-chromosomal microsatellite mutation rates in a population sample from northwestern Germany. Int J Legal Med 121:359–363

20. Lim SK, Xue Y, Parkin EJ, Tyler-Smith C (2007) Variation of 52 new Y-STR loci in the Y Chromosome Consortium worldwide panel of 76 diverse individuals. Int J Legal Med 121:124–127
21. López AM, Alvarez S, Gusmão L et al (2004) Population data for 16 Y-chromosome STRs in four populations from Pyrenees (Spain). Forensic Sci Int 140:125–129
22. Macabich I (1966) Historia de Ibiza. Volumen I. Daedalus, Palma de Mallorca
23. Martín P, García-Hirschfeld J, García O et al (2004) A Spanish population study of 17 Y-chromosome STR loci. Forensic Sci Int 139(1):231–235
24. Picornell A, Gómez-Barbeito L, Tomàs C, Castro JA, Ramon MM (2005) Mitochondrial DNA HVRI variation in Balearic populations. Am J Phys Anthropol 128:119–130
25. Picornell A, Miguel A, Castro JA, Ramon MM, Arya R, Crawford MH (1996) Genetic variation in the population of Ibiza (Spain): genetic structure, geography and language. Hum Biol 68:899–913
26. Pontes ML, Cainé L, Abrantes D, Lima G, Pinheiro MF (2007) Allele frequencies and population data for 17 Y-STR loci (AmpFlSTR Y-filer) in a northern Portuguese population sample. Forensic Sci Int 170:62–67
27. Quintana-Murci L, Semino O, Minch E, Passarino G, Brega A, Santachiara-Benerecetti AS (1999) Further characteristics of proto-European Y chromosomes. Eur J Hum Genet 7:603–608
28. 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 121:157–171
29. Rapone C, Geraci A, Capelli C et al (2007) Y chromosome haplotypes in central-south Italy: implication for reference database. Forensic Sci Int 172:67–71
30. Reynolds J, Weir BS, Cockerham CC (1983) Estimation of the coancestry coefficient: basis for a short-term genetic distance. Genetics 105:767–779
31. Robino C, Inturri S, Gino S et al (2006) Y-chromosomal STR haplotypes in Sicily. Forensic Sci Int 159:235–240
32. Roewer L, Willuweit S, Krüger C et al (2008) Analysis of Y chromosome STR haplotypes in the European part of Russia reveals high diversities but non-significant genetic distances between populations. Int J Legal Med 122:219–223
33. Sánchez C, Barrot C, Xifró A et al (2007) Haplotype frequencies of 16 Y-chromosome STR loci in the Barcelona metropolitan area population using Y-Filer™ kit. Forensic Sci Int 172:211–217
34. Scozzari R, Cruciani F, Pangrazio A et al (2001) Human Y-chromosome variation in the western Mediterranean area: implications for the peopling of the region. Hum Immunol 62:871–884
35. Semino O, Torroni A, Scozzari R, Brega A, De Benedictis G, Santachiara Benerecetti AS (1989) Mitochondrial DNA polymorphisms in Italy. III. Population data from Sicily: a possible quantitation of maternal African ancestry. Ann Hum Genet 53:193–202
36. Tagarelli A, Piro A, Tagarelli G, Zinno F (2000) Color-blindness in Calabria (southern Italy): a north–south decreasing trend. Am J Hum Biol 12:17–24
37. Tomàs C, Jiménez G, Picornell A, Castro JA, Ramon MM (2006) Differential maternal and paternal contributions to the genetic pool of Ibiza Island, Balearic Archipelago. Am J Phys Anthropol 129:268–278
38. Willuweit S, Roewer L, on behalf of the International Forensic Y Chromosome User Group (2007) Y chromosome haplotype reference database (YHRD): update. Forensic Sci Int: Genetics 1:83–87