Int J Legal Med (2000) 114 : 36–40

ORIGINAL ARTICLE

E. Bosch · F. Calafell · A. Pérez-Lezaun · D. Comas
H. Izaabel · O. Akhayat · A. Sefiani · G. Hariti
J. M. Dugoujon · J. Bertranpetit

Y chromosome STR haplotypes in four populations from northwest Africa

Received: 29 July 1999 / Accepted: 14 January 2000

Abstract The eight short tandem repeat (STR) polymorphic systems mapping on the male-specific region of the human Y chromosome, DYS19, DYS388, DYS389I, DYS389II, DYS390, DYS391, DYS392 and DYS393, were typed in four populations from northwest (NW) Africa (Moroccan Arabs, southern Moroccan Berbers, Saharawis and Mozabites). Allele frequency distributions showed statistically significant differences for all loci among all the populations except for DYS19. Complete typing was obtained for 185 chromosomes, which showed 74 different haplotypes. The two most frequent haplotypes were found in 16.2% and 15.1% of the individuals, although the latter was almost exclusively found in the Mozabites. Locus and haplotype informativeness were measured by means of the gene diversity (D). The haplotype diversity ranged from 0.856 (Mozabites) to 0.967 (southern Moroccan Berbers). For some loci, allele frequencies in NW Africans were clearly different from those in Europeans. The most common NW African haplotype was found only in one individual out of a total of 494 Europeans typed for the whole STR set. Thus, NW African and European Y chromosomes are clearly differentiated.

E. Bosch · F. Calafell · A. Pérez-Lezaun · D. Comas
J. Bertranpetit (✉)
Unitat de Biologia Evolutiva,
Facultat de Ciències de la Salut i de la Vida,
Universitat Pompeu Fabra, Doctor Aiguader 80,
08003 Barcelona, Spain
e-mail: jaume.bertranpetit@cexs.upf.es;
Tel.: +34-93-5422840; Fax: +34-93-5422802

H. Izaabel · O. Akhayat
Laboratoire de Biologie Cellulaire et Moléculaire,
Faculté des Sciences, Université Ibnou Zohr, Agadir, Morocco

A. Sefiani
Institut National d'Hygiène, Rabat, Morocco

G. Hariti
Hôpital Mustapha, CHU Alger Centre, Alger, Algeria

J. M. Dugoujon
CNRS ERS 1590, CHU Purpan, Toulouse, France

Key words Y chromosome · Short tandem repeats (STRs) · Haplotypes · Northwest African populations

Introduction

Because of its sex-determining function, the human Y chromosome is paternally inherited and does not recombine for most of its length. The genetic content is not homologous to any other part of the genome except for a few housekeeping genes that are homologous but highly divergent from genes in the X chromosome (Lahn and Page 1997). These special features make the Y chromosome an important genetic tool for the study of male aspects of human evolution and population diversity, as well as in forensic practice and paternity testing (Jobling and Tyler-Smith 1995; Jobling et al. 1997). Different types of genetic markers such as biallelic polymorphisms, short tandem repeat (STR) polymorphisms (also called microsatellites) and the MSY1 minisatellite (Jobling et al. 1998; Bouzekri et al. 1998) are now available on the non-recombining (or male-specific) region of the human Y chromosome. Amongst them, the PCR-amplifiable Y-STR loci have become some of the most useful markers to identify male individuals in forensic applications. In a multicentre study, a basic set of seven STRs (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393) has been recommended for standard Y-haplotyping in forensic and paternity casework (Kayser et al. 1997; de Knijff et al. 1997). Currently, one of the major goals of the forensic genetic community is the compilation of data for this selected basic set in a number of different reference populations. This will allow a well-established ethnic and geographic stratification of Y chromosome STR allele to be obtained as well as haplotype frequencies as needed for routine forensic casework.

Previous genetic studies of the forensic applicability in northwest (NW) Africa included a set of 13 autosomal STRs in Moroccan Arabs and Berbers (Pérez-Lezaun et al. 2000), HLA class II in southern Moroccan Berbers (Izaabel et al. 1998), and mtDNA (mitochondrial DNA)

sequences in Mozabites (Côrte-Real et al. 1996) and in a number of NW African populations (Rando et al. 1998). Three autosomal STRs (Meyer et al. 1995) and four Y chromosome STRs (DYS19, DYS389I, DYS389II and DYS390) were typed in a sample of Moroccans living in Brussels, Belgium (Kayser et al. 1997); in the same sample a detailed study of DYS389 haplotypes was also performed (Rolf et al. 1998).

This study presents allele and haplotype frequencies for eight Y-specific STRs in four populations from NW Africa (Moroccan Arabs, southern Moroccan Berbers, Saharawis, and Mozabites), which represent four distinct ethnical, linguistic and cultural groups in the area. These populations are of interest not only to the local forensic geneticists, but to those working in Western Europe where many immigrants from NW Africa have settled recently.

Material and methods

Samples

Genetic analyses were performed in a sample of Y chromosomes from 185 unrelated healthy males from NW Africa. Appropriate informed consent was obtained from all participants in this study and, in most cases, information about geographic origin of the four grandparents and mother tongue was recorded. Samples include 44 Moroccan Arabs, 44 southern Moroccan Berbers, 29 Saharawis (Western Sahara), and 68 Mozabites from the town of Ghardaia (Algeria). DNA was extracted from fresh blood by standard phenol-chloroform protocols.

STR polymorphism typing

PCR amplifications for all loci (DYS19, DYS388, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393) were carried out in a 10 µl final reaction volume containing 100 ng of genomic DNA, 50 mM KCl, 10 mM tris-HCl (pH 8.3), 1.5 mM $MgCl_2$ (2.5 mM for DYS19), 250 µM dNTP, 0.2 µM each primer and 1 U Taq DNA polymerase (Gibco, BRL) using a Perkin Elmer 9600 thermal cycler. Forward primers were fluorescently labeled. PCR cycling conditions were as described in Pérez-Lezaun et al. (1999). PCR products were run in an ABI 377 sequencer. ABI GS500 TAMRA was used as internal lane standard. The GeneScan 672 and Genotyper 1.1 software packages were used to collect the data and analyse fragment sizes. Y chromosome STR alleles were named according to the number of repeat units they contain, which was established through the use of sequenced allele ladders and reference samples kindly provided by P. de Knijff (Leiden, The Netherlands).

Statistical analysis

Population differentiation was tested by an exact test (Raymond and Rousset 1995) as implemented in the Arlequin package (Schneider et al. 1997). Gene diversity (D, which corresponds to expected heterozygosity for autosomal loci) was computed for each locus as $D = 1 - \Sigma p_i^2$, where p_i is the allelic frequency. Haplotype diversity was computed with the same equation using haplotype frequencies instead of allele frequencies. In the Y chromosome, haplotype diversity, power of discrimination and chance of paternity exclusion are numerically identical and this parameter is sufficient to adequately describe the informativeness of Y chromosome haplotypes in a population.

Table 1 Allele frequencies for the eight Y chromosome STRs DYS19, DYS388, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393 in four NW African populations (*ARA* Moroccan Arabs *SOB* Southern Moroccan Berbers SAH Saharawis MZA Mozabites)

Locus	Allele (repeat)	ARA $n = 44$	SOB $n = 44$	SAH $n = 29$	MZA $n = 68$
DYS19	13	0.592	0.682	0.760	0.822
	14	0.295	0.205	0.172	0.074
	15	0.045	0.045	0.034	0.059
	16	0.045	0.068	0	0.015
	17	0.023	0	0.034	0.015
	18	0	0	0	0.015
DYS388	12	0.795	0.841	0.828	0.926
	13	0.045	0.045	0	0.059
	14	0.023	0	0	0
	15	0	0.045	0	0
	16	0	0.023	0.034	0.015
	17	0.114	0.023	0.138	0
	18	0.023	0.023	0	0
DYS389I	9	0.205	0.114	0.103	0.044
	10	0.205	0.409	0.138	0.162
	11	0.590	0.431	0.759	0.794
	12	0	0.023	0	0
	13	0	0.023	0	0
DYS389II	23	0	0	0	0.015
	24	0.023	0	0	0
	25	0.045	0	0	0.015
	26	0.136	0.295	0.070	0.118
	27	0.728	0.523	0.448	0.764
	28	0.045	0.136	0.448	0.059
	29	0.023	0.023	0.034	0.029
	30	0	0.023	0	0
DYS390	20	0	0	0	0.015
	21	0.068	0.023	0.034	0.029
	22	0	0.068	0.034	0.029
	23	0.318	0.250	0.207	0.132
	24	0.546	0.545	0.310	0.604
	25	0.068	0.114	0.415	0.176
	26	0	0	0	0.015
DYS391	8	0	0.023	0	0
	9	0.409	0.613	0.794	0.147
	10	0.386	0.250	0.034	0.765
	11	0.182	0.114	0.172	0.088
	12	0.023	0	0	0
DYS392	10	0	0.023	0	0
	11	0.932	0.954	0.759	0.956
	12	0	0	0.241	0
	13	0.068	0.023	0	0.044
DYS393	10	0.023	0	0	0
	11	0	0	0	0
	12	0.159	0.114	0.172	0.015
	13	0.750	0.818	0.828	0.941
	14	0.068	0.068	0	0.015
	15	0	0	0	0.029

Results and discussion

Allele frequency distributions for loci DYS19, DYS388, DYS389I, DYS389II, DYS390, DYS391, DYS392 and DYS393 in the four NW African populations analysed are shown in Table 1. An exact test of population differentiation (Raymond and Rousset 1995) showed statistically significant differences ($p < 0.0001$ except $p = 0.008$ for DYS388 and $p = 0.020$ for DYS390) for all loci among all the populations except for DYS19 ($p = 0.15$).

Several loci showed allele frequencies that were very different from those found in European populations (Kayser et al. 1997). For instance, the most frequent allele for DYS19 is 13 in NW Africa but 14 in Europe. For DYS389I/II, the most frequent alleles in NW Africa are 11/27, but 10/26 in Europe. It is evident that European Y chromosome STR allele frequencies should not be used in forensic casework involving NW African subjects.

Table 2 Gene diversity (*D*) by population for the eight Y chromosome STRs analysed, as well as for the whole haplotype. The mean gene diversity for each locus is also given (Abbreviations as in Table 1)

Locus	ARA	SOB	SAH	MZA	Mean
DYS19	0.572	0.498	0.406	0.317	0.448
DYS388	0.359	0.294	0.305	0.140	0.275
DYS389I	0.580	0.647	0.409	0.346	0.496
DYS389II	0.458	0.634	0.613	0.403	0.527
DYS390	0.606	0.636	0.712	0.594	0.637
DYS391	0.665	0.560	0.352	0.392	0.492
DYS392	0.130	0.090	0.379	0.086	0.171
DYS393	0.417	0.320	0.296	0.115	0.287
Haplotype diversity	0.921	0.967	0.867	0.856	

Gene diversity (*D*) by population for each Y chromosome STR and for the whole haplotype are presented in Table 2. The most informative loci, as measured by the average gene diversity, were DYS390 ($D = 0.637$), DYS389II ($D = 0.527$), DYS389I ($D = 0.496$) and DYS391 ($D = 0.492$). The populations with the highest and lowest haplotype diversity were the southern Moroccan Berbers (0.967) and the Mozabites (0.856), respectively.

The Y haplotype frequency distributions constructed considering the eight Y STR loci studied (DYS19-DYS388-DYS389I-DYS389II-DYS390-DYS391-DYS392-DYS393) are given in Table 3. From a total of 185 chromosomes with complete typing, 74 distinct haplotypes were obtained. The two most frequent haplotypes, 19 (13–12–11–27–24–9–11–13) and 21 (13–12–11–27–24–10–11–13), were found in 30 (12 Moroccan Arabs, 4 Mozabites, 7 Saharawis and 7 southern Moroccan Berbers) and in 28 individuals (3 Moroccan Arabs, 24 Mozabites, 1 southern Moroccan Berber), respectively. Of the haplotypes 51 were observed in unique copies, and out of 74 different haplotypes, 61 (82.4%) were found only in one population.

Haplotype 19, which is the most frequent in NW Africa (16.2%), was found in one of 56 Basques (Pérez-Lezaun et al. 1997), but was not found in 33 Catalans (Pérez-Lezaun et al. 1997), 125 central and southern Italians (Caglià et al. 1998) or 280 Finns (Kittles et al. 1998). In a sample of 93 Galicians (Pestoni et al. 1988) typed for DYS19, DYS389I, DYS389II, DYS390 and DYS393, two individuals partially matched haplotype 19. Altogether, this is evidence for the specificity of NW African Y chromosome haplotypes.

Since the Y chromosome has one-quarter of the effective population size of any autosome, genetic drift acts more strongly on the Y chromosome than on the autosomes. As demontrated by Pérez-Lezaun et al. 1997 this means that:

Table 3 Y chromosome STR haplotypes for the four NW African populations analysed (abbreviations as in Table 1)

Haplotype no.	DYS19	DYS388	DYS389I	DYS389II	DYS390	DYS391	DYS392	DYS393	ARA	SOB	SAH	MZA	Total
1	13	12	9	25	24	10	11	13				1	1
2	13	12	9	26	23	10	11	10	1				1
3	13	12	9	26	23	10	11	13	2				2
4	13	12	9	26	23	10	11	14		1			1
5	13	12	9	26	24	10	10	13		1			1
6	13	12	9	26	24	10	11	13	1				1
7	13	12	9	26	25	10	11	13		1			1
8	13	12	10	26	23	9	11	13		2			2
9	13	12	10	26	24	9	11	13		2	1	3	6
10	13	12	10	26	24	10	11	13				3	3
11	13	12	10	27	24	9	11	13	1	1			2
12	13	12	10	28	22	9	11	13		2			2
13	13	12	10	28	24	10	11	13		1			1
14	13	12	10	30	22	9	11	13		1			1
15	13	12	11	27	23	8	11	13		1			1
16	13	12	11	27	23	9	11	13		2	1		3
17	13	12	11	27	23	10	11	13				5	5
18	13	12	11	27	23	11	11	13	1				1

Table 3 (continued)

Haplo-type no.	DYS19	DYS388	DYS389I	DYS389II	DYS390	DYS391	DYS392	DYS393	ARA	SOB	SAH	MZA	Total
19	13	12	11	27	24	9	11	13	12	7	7	4	30
20	13	12	11	27	24	9	11	14		1			1
21	13	12	11	27	24	10	11	13	3	1		24	28
22	13	12	11	27	24	11	11	13	1	1			2
23	13	12	11	27	25	9	11	13	2	1			3
24	13	12	11	27	25	10	11	13				7	7
25	13	12	11	27	25	10	11	15				1	1
26	13	12	11	28	24	9	11	13	1		1		2
27	13	12	11	28	24	10	11	13				4	4
28	13	12	11	28	25	9	11	13			6		6
29	13	12	11	28	25	9	12	13			6		6
30	13	12	12	28	23	9	11	13		1			1
31	13	12	13	29	23	9	11	13		1			1
32	13	13	10	26	24	9	11	13		1			1
33	13	13	10	28	25	10	11	13		1			1
34	13	13	11	27	23	9	11	13	1			3	4
35	13	13	11	27	25	10	11	13				1	1
36	14	12	9	24	23	10	11	14	1				1
37	14	12	9	25	24	10	13	13	1				1
38	14	12	9	26	23	10	11	13	1				1
39	14	12	9	26	25	10	11	13				1	1
40	14	12	10	27	24	11	13	13	1				1
41	14	12	10	29	25	10	11	15				1	1
42	14	12	11	27	24	9	11	13	1	4			5
43	14	12	11	27	24	10	11	13	1			1	2
44	14	12	11	29	20	10	11	14				1	1
45	14	14	11	29	25	10	11	12	1				1
46	14	15	9	26	23	10	11	12		1			1
47	14	15	10	26	24	10	11	12		1			1
48	14	16	10	27	22	11	11	12				1	1
49	14	16	10	27	23	11	11	12		1			1
50	14	16	11	29	23	11	11	12			1		1
51	14	17	9	26	23	11	11	12			1		1
52	14	17	10	26	23	11	11	12	1	1			2
53	14	17	10	27	23	10	11	12	1				1
54	14	17	10	27	23	11	11	12	3		3		6
55	14	18	10	27	23	12	11	12	1				1
56	14	18	10	27	24	11	11	12		1			1
57	15	12	9	26	22	11	11	13				1	1
58	15	12	9	27	21	10	11	13			1		1
59	15	12	9	27	21	10	11	14	2				2
60	15	12	10	23	23	10	13	13				1	1
61	15	12	10	27	21	11	11	13				1	1
62	15	12	10	27	25	10	11	13		1			1
63	15	12	10	28	24	10	11	13		1			1
64	15	12	11	27	26	11	13	13				1	1
65	16	12	9	26	21	10	11	14		1			1
66	16	12	10	26	25	11	13	13		1			1
67	16	12	10	27	21	11	11	13				1	1
68	16	12	11	27	24	9	11	13		1			1
69	16	12	11	27	24	11	13	13	1				1
70	16	12	11	28	21	10	11	13	1				1
71	17	12	9	27	22	9	12	13			1		1
72	17	12	11	27	24	11	13	13				1	1
73	17	13	10	25	23	10	11	13	1				1
74	18	12	11	27	25	10	11	13				1	1

1. Gene diversities in the Y chromosome STRs are lower than those found in autosomal STRs. Recently, White et al. (1999) selected, out of 185 possible Y chromosome $(GATA)_n$ STRs, six loci with an average $D = 0.663$ in an unidentified sample, which is still far from the usual levels of polymorphism in the autosomal STRs used in forensic casework (> 0.75, Pérez-Lezaun et al. 2000).

2. More differentiation is found among populations for Y chromosome STRs than for autosomal STRs, and thus more subpopulations need to be typed to obtain reliable population data for Y chromosome markers.

These basic genetic properties explain the high levels of between-population differentiation observed in NW Africa for the Y chromosome markers and the reduction of the internal diversity in one of the populations, the Mozabites.

Acknowledgements We thank all the blood donors that made this study possible. This research was supported by Dirección General de Investigación Científica y Técnica in Spain (PB95-0267-CO2-01) and by Direcció General de Recerca, Generalitat de Catalunya (1998SGR00009). Comissionat per a Universitats i Recerca, Generalitat de Catalunya supported E.B. (FI/96–1153). The Mozabite samples from Ghardaia were obtained in the framework of an INSERM network: 490NS1 under the coordination of Dr. A. Cambon-Thomsen (Toulouse) and with the help of the Medical team of Ghardaia and Pr. M.S. Isaad (Alger). The help of A. Cambon-Thomsen and A. Sevin (CNRS Toulouse) is gratefully acknowledged. We also thank all persons involved in reaching the Saharawi donors as well as Elisabeth Pintado (Sevilla), Josep Lluís Fernández Roure and Alba Bosch (Mataró) for their help in contacting Moroccan donors.

References

Bouzekri N, Taylor PG, Hammer MF, Jobling MA (1998) Novel mutation processes in the evolution of a haploid minisatellite, MSY1: array homogenization without homogenization. Hum Mol Genet 7: 655–659

Caglià A, Dobosz M, Boschi I, d'Aloja E, Pascali VL (1998) Increased forensic efficiency of a STR-based Y-specific haplotype by addition of the highly polymorphic DYS385 locus. Int J Legal Med 111: 142–146

Côrte-Real HBSM, Macaulay V, Richards MB, Hariti G, Issad MS, Cambon-Thomsen A, Papiha SS, Bertranpetit J, Sykes B (1996) Genetic diversity in the Iberian Peninsula determined from mitochondrial sequence analysis. Ann Hum Genet 60: 331–350

Izaabel H, Garchon HJ, Caillat-Zucman S, Akhayat O, Bach J-F, Sanchez-Mazas A (1998) HLA class II DNA polymorphism in Moroccan population from Souss, Agadir area. Tissue Antigens 51: 106–110

Jobling MA, Tyler-Smith C (1995) Fathers and sons: the Y chromosome and the human evolution. Trends Genet 11: 449–456

Jobling MA, Pandya A, Tyler-Smith C (1997) The Y chromosome in forensic analysis and paternity testing. Int J Legal Med 110: 118–124

Jobling MA, Bouzekri N, Taylor PG (1998) Hypervariable digital DNA codes for human paternal lineages: MVR-PCR at the Y-specific minisatellite, MSY1 (DYF155S1). Hum Mol Genet 7: 643–653

Kayser M, Caglià A, Corach D, Fretwell N, Gehrig C, Graziosi G, Heidorn F, Herrmann S, Herzog B, Hidding M, Honda K, Jobling M, Krawczak M, Leim K, Meuser S, Meyer E, Oesterreich W, Pandya A, Parson W, Penacino G, Pérez-Lezaun A, Piccinini A, Prinz M, Schmitt C, Schneider PM, Szibor R, Teifel-Greding J, Weichhold G, de Knijff P, Roewer L (1997) Evaluation of Y-chromosomal STRs: a multicenter study. Int J Legal Med 110: 125–133

Kittles RA, Perola M, Peltonen L, Bergen AW, Aragon RA, Virkkunen M, Linnoila M, Goldman D, Long JC (1998) Dual origins of Finns revealed by Y chromosome haplotype variation. Am J Hum Genet 62: 1171–1179

Knijff P de, Kayser M, Caglià A, Corach D, Fretwell N, Gehrig C, Graziosi G, Heidorn F, Herrmann S, Herzog B, Hidding M, Honda K, Jobling MA, Krawczak M, Leim K, Meuser S, Meyer E, Oesterreich W, Pandya A, Parson W, Penacino G, Pérez-Lezaun A, Piccinini A, Prinz M, Schmitt C, Schneider PM, Szibor R, Teifel-Greding J, Weichhold G, Roewer L (1997) Chromosome Y microsatellites: population genetic and evolutionary aspects. Int J Legal Med 110: 134–140

Lahn BT, Page DC (1997) Functional coherence of the human Y chromosome. Science 278: 675–680

Meyer E, Wiegand P, Brinkmann B (1995) Phenotype differences of STRs in 7 human populations. Int J Legal Med 107: 314–322

Pérez-Lezaun A, Calafell F, Seielstad MT, Mateu E, Comas D, Bosch E, Bertranpetit J (1997) Population genetics of Y chromosome short tandem repeats in humans. J Mol Evol 45: 265–270

Pérez-Lezaun A, Calafell F, Comas D, Mateu E, Bosch E, Martínez-Arias R, Clarimón J, Fiori G, Luiselli D, Facchini F, Pettener D, Bertranpetit J (1999) Gender-specific migration in Central Asian populations revealed by the analysis of Y-chromosome STRs and mtDNA. Am J Hum Genet 65: 208–219

Pérez-Lezaun A, Calafell F, Clarimón J, Bosch E, Mateu E, Gusmao L, Amorim A, Benchemsi N, Bertranpetit J (2000) Allele frequencies of 13 short tandem repeats in population samples of the Iberian Peninsula and Northern Africa. Int J Legal Med 113: 208–214

Pestoni C, Cal ML, Lareu MV, Rodríguez-Calvo MS, Carracedo A (1998) Y chromosome STR haplotypes: genetic and sequencing data of the Galician population (NW Spain). Int J Legal Med 112: 15–21

Rando JC, Pinto F, Gonzalez AM, Hernández M, Larruga JM, Cabrera VM, Bandelt HJ (1998) Mitochondrial DNA analysis of Northwest African populations reveals genetic exchanges with European, near-Eastern, and sub-Saharan populations. Ann Hum Genet 62: 531–550

Raymond M, Rousset F (1995) An exact test of differentiation. Evolution 49: 1280–1283

Rolf B, Meyer E, Brinkmann B, Knijff P de (1998) Polymorphism at the tetranucleotide repeat locus DYS389 in 10 populations reveals strong geographic clustering. Eur J Hum Genet 6: 583–588

Schneider S, Kueffer J-M, Roessli D, Excoffier L (1997) Arlequin ver 1.1: A software for population genetic data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland

White PS, Tatum OL, Deaven LL, Longmire JL (1999) New, male-specific microsatellite markers from the human Y chromosome. Genomics 57: 433–437