Int J Legal Med (2008) 122:529–533
DOI 10.1007/s00414-008-0265-z

SHORT COMMUNICATION

Population and segregation data on 17 Y-STRs: results of a GEP-ISFG collaborative study

Paula Sánchez-Diz · Cíntia Alves · Elizeu Carvalho · Mónica Carvalho · Rosa Espinheira · Oscar García · Mª Fátima Pinheiro · Lurdes Pontes · Mª João Porto · Oscar Santapa · Cláudia Silva · Denilce Sumita · Sérgio Valente · Martin Whittle · Iñaki Yurrebaso · Angel Carracedo · António Amorim · Leonor Gusmão · GEP-ISFG (The Spanish and Portuguese Working Group of the International Society for Forensic Genetics)

Received: 11 March 2008 / Accepted: 10 June 2008 / Published online: 24 July 2008

Abstract A collaborative work was carried out by the Spanish and Portuguese International Society for Forensic Genetics Working Group in order to extend the existing data on Y-short tandem repeat (STR) mutations at the 17 Y chromosome STR loci included in the AmpFlSTR YFiler kit (Applied Biosystems): DYS19, DYS385, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS448, DYS456, DYS458, DYS635, and GATA H4.1. In a sample of 701 father/son pairs, 26 mutations were observed among 11,917 allele transfers across the 17 loci. After summing previously reported mutation data with our sample, mutation rates varied between 4.25×10^{-4} (95% CI 0.05×10^{-3}–1.53×10^{-3}) at DYS438 and 6.36×10^{-3} (95% CI 2.75×10^{-3}–12.49×10^{-3}) at DYS458. All mutations were single step, and mutations in the same father/son pair were found twice.

Electronic supplementary material The online version of this article (doi:10.1007/s00414-008-0265-z) contains supplementary material, which is available to authorized users.

P. Sánchez-Diz (✉) · A. Carracedo
Institute of Legal Medicine, University of Santiago de Compostela,
C/San Francisco, s/n, 15782 Santiago de Compostela,
A Coruña, Spain
e-mail: paulaiml@usc.es

C. Alves · A. Amorim · L. Gusmão
IPATIMUP, Institute of Molecular Pathology and Immunology of the University of Porto,
4200-465 Porto, Portugal

E. Carvalho
Laboratorio de Diagnósticos por DNA, Instituto de Biologia,
Universidade do Estado do Rio de Janeiro,
Rio de Janeiro, Brazil

M. Carvalho · M. J. Porto
Institute of Legal Medicine,
Coimbra, Portugal

R. Espinheira · C. Silva
Serviço de Genética Forense Instituto Nacional de Medicina Legal, Faculdade de Medicina, Universidade de Lisboa,
Lisbon, Portugal

O. García · I. Yurrebaso
Sección de Genética Forense, Area de Laboratorio Ertzaintza,
Bizkaia, Spain

M. F. Pinheiro · L. Pontes
Delegação do Porto do Instituto Nacional de Medicina Legal,
Porto, Portugal

O. Santapa · S. Valente
Banco Nacional de Datos Genéticos, Hospital Dr. C.G. Durand,
Buenos Aires, Argentina

D. Sumita · M. Whittle
Genomic Engenharia Molecular Ltda,
São Paulo, Brazil

A. Amorim
Faculdade de Ciências, Universidade do Porto,
Porto, Portugal

Keywords Y-chromosome haplotypes · STRs · AmpFlSTR YFiler kit · Mutation rates

Introduction

In order to improve the discrimination power of Y chromosome haplotypes, an increasing number of short tandem repeats (STRs) are being used for forensic purposes and included in multiplex typing kits. When a large set of markers is used in kinship analysis, the probability of detecting mutations between closely related individuals in the same patrilineage is high, and therefore, estimation of mutation rates becomes important.

A large number of Y-STR mutations have been reported [2–7, 11–21, 24–26], allowing a good estimation on the average rate, but the available data concern only a small set of markers that can be considered insufficient to infer reliable locus-specific mutation rates. For that reason, the DNA Commission of the International Society for Forensic Genetics (ISFG), in their last recommendations on the use of Y-STRs in forensic analysis, recognized the importance of extending the available data [9].

In the present work, we aimed to collect new mutation data on 17 Y-STR loci, frequently used in forensic investigation, by compiling haplotype information in confirmed father/son pairs from seven laboratories members of the Spanish and Portuguese Working Group of the International Society for Forensic Genetics (GEP-ISFG).

Materials and methods

The sample consisted of 701 father/son pairs from paternity cases in seven different laboratories from Argentina, Brazil, and Portugal and does not overlap the material previously reported by the GEP-ISFG working group [11]. All the individuals gave informed consent prior to inclusion in the study. Non-related father haplotype data from north Portugal partially overlap those from Alves et al. [1]. The biological relationship of all father/son pairs was previously confirmed by using autosomal STRs, with paternity indices above 10,000.

Each laboratory used its own routine DNA extraction method. Samples were amplified for the 17 Y-STRs using the AmpFlSTR Yfiler PCR Amplification Kit (Applied Biosystems [AB], Foster City, CA, USA) according to the manufacturer's instructions. Typing was achieved using ABI Genetic Analyzers (AB) and by comparison to reference sequenced ladders provided by the manufacturer (AB). Alleles for GATA H4.1 locus were named according to the ISFG recommendations [9] by adding nine repeats to the nomenclature included in the typing kit [20].

All the mutations were confirmed by a second laboratory.

Table 1 Total number of mutations and allele transmissions per locus

Marker	This work		Total[a]		Freq. ($\times 10^{-3}$)	95% CI ($\times 10^{-3}$)
Y-STR	No. mutations	Allele trans.	No. mutations	Allele trans.		
DYS19	2	701	23	9,658	2.381	1.510–3.571
DYS385	1	1402	31	14,896	2.081	1.414–2.953
DYS389 I	1	701	14	7,862	1.781	0.974–2.986
DYS389 II	1	701	22	7,849	2.803	1.757–4.241
DYS390	2	701	21	9,140	2.298	1.423–3.510
DYS391	3	701	28	9,089	3.081	2.048–4.449
DYS392	0	701	5	9,053	0.552	0.179–1.288
DYS393	1	701	7	7,842	0.893	0.359–1.838
DYS437	1	701	7	4,672	1.498	0.603–3.085
DYS438	0	701	2	4,709	0.425	0.051–1.533
DYS439	6	701	27	4,686	5.762	3.800–8.372
DYS448	1	701	2	1,258	1.590	0.193–5.731
DYS456	2	701	6	1,258	4.769	1.752–10.352
DYS458	2	701	8	1,258	6.359	2.749–12.492
DYS635	2	701	8	2,131	3.754	1.622–7.384
GATA H4.1	1	701	5	2,294	2.180	0.708–5.079
Total	26	11,917	216	97,655	2.212	1.929–2.530

Locus specific and overall mutation rate estimates and respective CI

[a] Includes data from this work and from Heyer et al. [12], Bianchi et al. [4], Kayser et al. [15], Tsai et al. [24], Dupuy et al. [7], Kurihara et al. [17], Budowle et al. [5], Ballard et al. [2], Berger et al. [3], Gusmão et al. [11], Hohoff et al. [13]; Turrina et al. [25], Mulero et al. [20], Domingues et al. [6], Lee et al. [18], and Pontes et al. [21]

Table 2 Mutations observed at STR loci with the respective paternity index (L) and father's age at the time the son was born

Population	Haplotype[a]	Locus	Father	Son	L	Age (years)
Portugal_North	H498	DYS390	24	23	2.5×10^{9}	50
	H496	DYS439	12	13	2.3×10^{8}	64
	H553	DYS458	18	19	1.9×10^{7}	32
	H502	DYS456	17	18	9.6×10^{8}	21
	H567	DYS385	13–15	14–15	3.6×10^{4}	30
	H353	DYS389II	29	30	2.9×10^{7}	44
	H390	DYS439	13	14	6.7×10^{8}	29
	H430	DYS635	21	20	5.7×10^{9}	20
	H435	DYS390	24	23	4.8×10^{8}	19
Portugal_Centre	H33	DYS635	23	24	4×10^{8}	30
	H44	GATA H4.1	21	20	9×10^{7}	33
	H44	DYS393	14	13	9×10^{7}	33
	H58	DYS391	10	11	6.7×10^{7}	31
	H58	DYS437	17	16	6.7×10^{7}	31
	H97	DYS456	16	15	1.2×10^{7}	27
Portugal_South	H609	DYS439	13	14	4.2×10^{9}	22
	H596	DYS448	19	20	7.7×10^{9}	50
Argentina_Buenos Aires	H336	DYS391	11	10	1.5×10^{6}	33
	H347	DYS439	12	13	2.8×10^{7}	39
	H286	DYS439	14	13	9.4×10^{4}	30
Brasil_general population	H170	DYS19	17	18	1.4×10^{9}	21
	H172	DYS439	13	12	1.2×10^{13}	17
	H213	DYS19	17	16	1.2×10^{6}	31
	H234	DYS389 I	14	13	1.3×10^{10}	–[b]
	H235	DYS391	10	11	6.4×10^{7}	20
Brasil_Rio de Janeiro	H82	DYS458	17	18	3.3×10^{8}	38

[a] Haplotypes coded as in Supplementary material Table S1
[b] No information available

Population pairwise genetic distances (R_{ST}) were calculated using software ARLEQUIN ver 3.0 [8]. In population comparisons, DYS385 was not considered, and the number of repeats in DYS389I was subtracted from DYS389II.

Table 3 Fathers' age distribution in 5 year intervals

Age group	Number of fathers
15–19	31
20–24	147
25–29	155
30–34	143
35–39	72
40–44	52
45–49	36
50–54	17
55–59	11
60–64	7
65–69	4
70–73	3
Mean age	31.6

Father's age could not be recorded in 19 out of the 697 cases included in the present study

Confidence intervals (CI) for mutation rates were estimated from the binomial standard deviation.

Results and discussion

Results were compiled from seven participating laboratories that typed 17 Y-STRs in a total of 701 father/son pairs.

Haplotypes from non-related individuals included in father samples from Brazil (Rio de Janeiro, n=50, and general population, n=103), Argentina (Buenos Aires, n=100), and Portugal (north, n=244; center, n=100; and south, n=100) are listed in supplementary material Table S1. A total of 674 haplotypes were observed, 653 being unique.

Low genetic distance values were obtained in the comparisons between the population samples analyzed ($R_{ST}\leq0.016$; $P\geq0.036$). After applying the Bonferroni correction for multiple tests, no significant R_{ST} P values were obtained, which is expected bearing in mind that the Y-chromosome haplotypic background of all the studied populations is mainly of European origin [10, 22, 23].

Null alleles were found in three cases: at DYS438 (H196), at DYS448 (H188), and at DYS456 (H44),

respectively. Duplications were also detected in DYS19 (H371, H382, H407, and H449), DYS439 (H605 and H639), DYS448 (H434), and DYS635 (H666). The observed duplications and silent alleles deserve some considerations. Concerning the latter, each of them was found at different loci, and therefore, it is parsimonious to assume that they constitute rare events resulting from recent mutations in distinct haplotype backgrounds. The same was observed for the duplications at DYS19, which were detected in backgrounds molecularly wide apart, implying that either the duplication is recurrent or very ancient. On the contrary, the duplications at DYS439 were detected in haplotypes just one mutational step apart, so it seems plausible to assume that both have a common recent origin.

Transmission

Among 11,917 allele transfers, 26 mutations were observed: one at DYS385, DYS389I, DYS389II, DYS393, DYS437, DYS448, and GATA H4.1, two at DYS19, DYS390, DYS456, DYS458, and DYS635, three at DYS391, and six at DYS439 (Table 1).

All mutations were single step, and double mutations in the same father/son pair were observed twice (H44 and H58 in Table 2).

The average age of the fathers involved in mutation events (31.8 years; Table 2) was similar to the fathers of the whole studied sample (31.6 years; Table 3).

Single locus mutation rates were estimated for the 17 Y-STRs under study as the frequency of mutations in the total number of allele transfers, adding our results to previously reported in Heyer et al. [12], Bianchi et al. [4], Kayser et al. [15], Tsai et al. [24], Dupuy et al. [7], Kurihara et al. [17], Ballard et al. [2], Berger et al. [3], Budowle et al. [5], Gusmão et al. [11], Mulero et al. [20], Turrina et al. [25], Domingues et al. [6], Hohoff et al. [13], Lee et al. [18], and Pontes et al. [21] (Table 1). The obtained locus-specific mutation rates varied between 4.25×10^{-4} (95% CI 0.05×10^{-3}–1.53×10^{-3}), at DYS438, and 6.36×10^{-3} (95% CI 2.75×10^{-3}–12.49×10^{-3}), at DYS458, with an average of 2.21×10^{-3} (95% CI 1.93×10^{-3}–2.53×10^{-3}) in a total of 97,655 allele transmissions. It is worth highlighting that a subset of these markers (DYS438, DYS392, and DYS393) displays mutation rates below the CI limit of the average; on the contrary, both DYS439 and DYS458 present a much higher mutation rate, above the CI of the average value. The higher values obtained for these two markers, as well as DYS456, are not unexpected, bearing in mind that the most frequent alleles at both loci display a high number of perfect repeats.

In conclusion, our work confirms that average mutation rates at Y-STRs can vary up to a factor of ten across loci and that much more data are needed to confirm the suggestion that allele heterogeneity inside each locus is still more relevant.

Acknowledgments IPATIMUP is partially supported by Fundação para a Ciência e a Tecnologia, through POCI (Programa Operacional Ciência e Inovação 2010). P.S.D. is supported by the Isidro Parga Pondal program (Plan Galego de Investigación, Desenvolvemento e Innovación Tecnolóxica-INCITE (2006–2010) from Xunta de Galicia).

References

1. Alves C, Gomes V, Prata MJ, Amorim A, Gusmão L (2007) Population data for Y-chromosome haplotypes defined by 17 STRs (AmpFlSTR YFiler) in Portugal. Forensic Sci Int 171(2–3):250–255
2. Ballard DJ, Phillips C, Wright G, Thacker CR, Robson C, Revoir AP, Syndercombe Court D (2005) A study of mutation rates and the characterisation of intermediate, null and duplicated alleles for 13 Y chromosome STRs. Forensic Sci Int 155(1):65–70
3. Berger B, Lindinger A, Niederstätter H, Grubwieser P, Parson W (2005) Y-STR typing of an Austrian population sample using a 17-loci multiplex PCR assay. Int J Legal Med 119(4):241–246
4. Bianchi NO, Catanesi CI, Bailliet G et al (1998) Characterization of ancestral and derived Y-chromosome haplotypes of New World native populations. Am J Hum Genet 63(6):1862–1871
5. Budowle B, Adamowicz M, Aranda XG et al (2005) Twelve short tandem repeat loci Y chromosome haplotypes: Genetic analysis on populations residing in North America. Forensic Sci Int 150(1):1–15
6. Domingues PM, Gusmão L, da Silva DA, Amorim A, Pereira RW, de Carvalho EF (2007) Sub-Saharan Africa descendents in Rio de Janeiro (Brazil): population and mutational data for 12 Y-STR loci. Int J Legal Med 121(3):238–241
7. Dupuy BM, Stenersen M, Egeland T, Olaisen B (2004) Y-chromosomal microsatellite mutation rates: differences in mutation rate between and within loci. Hum Mutat 23(2):117–124
8. Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evol Bioinform Online 1:47–50
9. Gusmão L, Butler JM, Carracedo A et al (2006) DNA Commission of the International Society of Forensic Genetics (ISFG): an update of the recommendations on the use of Y-STRs in forensic analysis. Int J Legal Med 120(4):191–200
10. Gusmão L, Sánchez-Diz P, Alves C et al (2003) Results of the GEP-ISFG collaborative study on the Y chromosome STRs GATA A10, GATA C4, GATA H4, DYS437, DYS438, DYS439, DYS460 and DYS461: Population data. Forensic Sci Int 135 (2):150–157
11. Gusmão L, Sánchez-Diz P, Calafell F et al (2005) Mutation rates at Y chromosome specific microsatellites. Hum Mutat 26(6):520–528
12. Heyer E, Puymirat J, Dieltjes P, Bakker E, Knijff P (1997) Estimating Y-chromosome specific microsatellite mutation frequencies using deep rooting pedigrees. Hum Mol Genet 6(5):799–803
13. 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 (5):359–363
14. Holtkemper U, Rolf B, Hohoff C, Forster P, Brinkmann B (2001) Mutation rates at two human Y-chromosomal microsatellite loci using small pool PCR techniques. Hum Mol Genet 10(6):629–633
15. Kayser M, Roewer L, Hedman M et al (2000) Characteristics and frequency of germline mutations at microsatellite loci from the

human Y chromosome, as revealed by direct observation in father/son pairs. Am J Hum Genet 66(5):1580–1588
16. Kayser M, Sajantila A (2001) Mutations at Y-STR loci: implications for paternity testing and forensic analysis. Forensic Sci Int 118(2–3):116–121
17. Kurihara R, Yamamoto T, Uchihi R et al (2004) Mutations in 14 Y-STR loci among Japanese father–son haplotypes. Int J Legal Med 118(3):125–131
18. Lee HY, Park MJ, Chung U, Lee HY, Yang WI, Cho SH, Shin KJ (2007) Haplotypes and mutation analysis of 22 Y-chromosomal STRs in Korean father–son pairs. Int J Legal Med 121(2):128–135
19. Luca F, Basile M, Di Giacomo F, Novelletto A (2005) Independent methods for evolutionary genetic dating provide insights into Y-chromosomal STR mutation rates confirming data from direct father–son transmissions. Hum Genet 118(2):153–165
20. Mulero JJ, Chang CW, Calandro LM, Green RL, Li Y, Johnson CL, Hennessy LK (2006) Development and validation of the AmpFlSTRYfiler trade mark PCR amplification kit: a male specific, single amplification 17 YSTR multiplex system. J Forensic Sci 51(1):64–75
21. Pontes ML, Caine L, Abrantes D, Lima G, Pinheiro MF (2007) Allele frequencies and population data for 17 Y-STR loci (AmpFlSTR (R) Yfiler trade mark) in a northern Portuguese population sample. Forensic Sci Int 170(1):62–67
22. Silva DA, Carvalho E, Costa G, Tavares L, Amorim A, Gusmão L (2006) Y-chromosome genetic variation in Rio de Janeiro population. Am J Hum Biol 18(6):829–837
23. Toscanini U, Gusmão L, Berardi G, Amorim A, Carracedo A, Salas A, Raimondi E (2007) Testing for genetic structure in different urban Argentinean populations. Forensic Sci Int 165(1):35–40
24. Tsai LC, Yuen TY, Hsieh HM et al (2002) Haplotype frequencies of nine Y-chromosome STR loci in the Taiwanese Han population. Int J Legal Med 116(3):179–183
25. Turrina S, Atzei R, De Leo D (2006) Y-chromosomal STR haplotypes in a Northeast Italian population sample using 17plex loci PCR assay. Int J Legal Med 120(1):56–59
26. Zhivotovsky LA, Underhill PA, Cinnioglu C et al (2004) The effective mutation rate at Y chromosome short tandem repeats, with application to human population-divergence time. Am J Hum Genet 74(1):50–61