Int J Legal Med (2000) 114 : 125–129

SHORT COMMUNICATION

P. Grignani · G. Peloso · P. Fattorini · C. Previderè

Highly informative Y-chromosomal haplotypes by the addition of three new STRs DYS437, DYS438 and DYS439

Received: 29 November 1999 / Accepted: 17 March 2000

Abstract The Y chromosome STRs DYS437, DYS438 and DYS439 were selected from publicly available genome databases and used to analyse an Italian population sample. A tetraplex PCR reaction including the highly informative DYS385 locus, was set up and used for the analysis of 131 male samples to determine allele frequencies and STR diversity values. The number of different haplotypes and the haplotype diversity value found from the analysis of the STRs included in the tetraplex reaction were very similar to those found from the analysis of the basic set of 7 Y-STRs (DYS19, DYS389I/II, DYS390, DYS391, DYS392 and DYS393) previously carried out on the same population sample. By combining the allelic states of the 11 Y-chromosomal STRs we could construct highly informative haplotypes that allowed the discrimination of 93.8% (120 out of 128) of the samples tested. This approach represents a very powerful tool for individual identification and paternity testing in forensic medicine.

Key words Y-chromosome · STRs · Y-haplotypes

Introduction

The male-specific portion of the Y chromosome is paternally inherited and haploid. This special feature is particularly useful when Y-linked polymorphic loci are investigated for human evolutionary studies [1], for understanding population migrations and history [2] and in forensic medicine for male identification and paternity testing [3].

Y chromosome polymorphic markers can be divided into two categories: unique event polymorphisms (UEPs) and VNTR markers such as minisatellites or microsatellites (STRs). The former arise by unique mutational events as base substitutions or insertions/deletions. These markers are binary in nature and in combination, can define major groups of chromosomes in the population called haplogroups. The latter can be useful for elucidating the substructure of the main groups of Y chromosomes because of the higher heterogeneity. By combining the allelic states of several STR markers it is possible to define highly informative haplotypes that allow the discrimination of most unrelated males in a population sample. In a multicentre study [4] a panel of 13 STRs was used to analyse unrelated male samples from worldwide populations and a subset of 7 STR markers was proposed for standard analysis in forensic and paternity casework. Subsequently, the highly polymorphic DYS385 locus was validated for forensic practice and proposed as an additional marker to the basic 7 STR set [5].

The number of Y-chromosomal STRs seems to be quite limited and there is a major requirement for new Y microsatellites, in particular for applications in forensic medicine.

Recently, Ayub et al. [6] selected six new STRs from 1.33 Mb of Y chromosomal sequence available in databases and co-amplified them in a multiplex reaction. The loci DYS437, DYS438 and DYS439 seemed to have a high degree of polymorphism and to offer good individual discrimination power when a population sample from Pakistan was analysed.

Therefore, we decided to analyse an Italian population sample for these new STR markers by setting up a multiplex PCR reaction also including the highly variable DYS385 locus.

In agreement with the strategy suggested by Pascali et al. in a recent editorial [7], most of the individuals analysed in this study (102 out of 131) had been already typed for the 7 basic STRs and for 10 UEPs in a previous study [8].

P. Grignani
IRCCS Policlinico S. Matteo, Pavia, Servizio di Medicina Legale, Dipartimento di Medicina Legale e Sanità Pubblica, via Forlanini 12, 27100, Pavia, Italy

G. Peloso · C. Previderè (✉)
Dipartimento di Medicina Legale e Sanità Pubblica, via Forlanini 12, 27100 Pavia, Italy
e-mail: previde@unipv.it, Fax: +39-0382-528025

P. Fattorini
Istituto di Medicina Legale, via Molino a Vento 123, 34147 Trieste, Italy

Materials and methods

Blood samples were collected from 131 healthy donors and 18 father-child pairs from northwestern Italy with paternity already confirmed by autosomal STR analysis. DNA was isolated by organic extraction according to standard procedures. Primer sequences for the new Y-STRs were according to Ayub et al. [6] but with a different primer F labelling for DYS437 (TET labelled instead of HEX). Primer sequences for the DYS385 locus were according to Caglià et al.[5].

PCR conditions (tetraplex)

The PCR reactions (15 µl) contained 5–50 ng of template DNA, 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 1.5 mM $MgCl_2$, 0.6 µM (DYS437), 0.4 µM (DYS438), 0.2 µM (DYS439) and 0.2 µM (DYS385) each primer, 200 µM dNTPs and 2 U Taq polymerase (Roche).

Amplification conditions

PCR was carried out using a 2400 Perkin-Elmer thermocycler with a pre-denaturation step at 95 °C for 3 min followed by 30 cycles at 94 °C for 1 min, 56 °C for 1 min, 72 °C for 1 min and a final extension of 72 °C for 7 min.

Electrophoresis conditions and sequence analysis

Aliquots of 0.2 µl of PCR amplification products were mixed with TAMRA 500 internal size standard and processed by capillary electrophoresis in the ABI 310 capillary sequencer (Perkin-Elmer). PCR fragment lengths were determined using GeneScan software version 2.1 (Perkin-Elmer). For each locus (DYS437, DYS438 and DYS439) three alleles were sequenced on the ABI 310 sequencer using the Big Dye Terminator cycle sequencing kit (Perkin-Elmer) and the results were consistent with the sequencing data found by Ayub et al. [6]. The consensus repeat structure of alleles found for each locus in this study can be summarised as follows:

1. DYS437 (3 alleles)
   – $(GATA)_4(GACA)_2(GATA)_{8–10}$
2. DYS438 (6 alleles)
   – $(TTTTC)_{8–13}$
3. DYS439 (5 alleles)
   – $(GATA)_2N_4(GATA)_3N_{14}(GATA)N_3(GATA)N_7(GATA)_{10–14}$

The alleles were named according to the number of repeats as recommended by the DNA Commission of the International Society for Forensic Haemogenetics [9]. For the DYS437, DYS438 loci we adopted the nomenclature suggested by Ayub et al. [6] but modified it for the DYS439 locus to include all the repeat units of the consensus sequence GATA. For the DYS385 locus the nomenclature used was according to Caglià et al.[5].Allele scoring for each marker was obtained by comparisons to self-made allelic ladders.

The STR diversity value and the haplotype diversity value were calculated according to the same formula $1 – \Sigma p_i^2$, where p_i represents allele frequency for each allele found at a given locus or the haplotype frequency found by combining different loci, respectively.

Results and discussion

Allele frequencies and STR diversity values found in the Italian population sample for the three new Y-STRs DYS437, DYS438 and DYS439 are shown in Table 1.

Table 1 Allele frequencies for the three new STRs

Locus	Alleles		Frequency	STR diversity value
	bp	Units		
DYS437	186	14	0.313	0.6099
	190	15	0.511	
	194	16	0.176	
DYS438	213	8	0.008	0.7223
	218	9	0.206	
	223	10	0.298	
	228	11	0.061	
	233	12	0.374	
	238	13	0.053	
DYS439	242	17	0.038	0.6652
	246	18	0.328	
	250	19	0.435	
	254	20	0.191	
	258	21	0.008	

Table 2 Phenotype frequencies for the DYS385 locus

Phenotype	Frequency	Phenotype	Frequency	STR diversity value
10–13	0.008	13–19	0.008	0.8970
10–14	0.023	14–14	0.037	
10–15	0.015	14–15	0.037	
10–16	0.015	14–16	0.008	
10–18	0.008	14–17	0.037	
11–11	0.015	14–18	0.008	
11–13	0.037	14–19	0.008	
11–14	0.290	15–15	0.023	
11–16	0.023	15–16	0.023	
12–12	0.023	15–17	0.015	
12–13	0.023	15–18	0.015	
12–14	0.031	15–19	0.008	
12–15	0.015	16–16	0.008	
12–16	0.008	16–18	0.023	
12–17	0.008	16–19	0.008	
13–14	0.031	17–17	0.008	
13–15	0.031	17–18	0.015	
13–16	0.037	18–18	0.008	
13–17	0.015	18–19	0.008	
13–18	0.031	19–19	0.008	

Table 3 Number of different haplotypes and haplotype diversity values found for three combinations of Y-STR markers

	Haplotype A (DYS19, DYS389I/II, DYS390, DYS391, DYS392, DYS393)	Haplotype B (DYS437, DYS438, DYS439, DYS385)	Haplotype C (A + B)
Number of individuals	131	131	128
Different haplotypes	97	94	120
Haplotype diversity value	0.9753	0.9759	0.9907

Table 4 The 120 different haplotypes found in the Italian population sample with a combination of 11 Y-STRs

Haplotype	*n*	DYS19	DY389-I	DY389-II	DY390	DY391	DY392	DY393	DYS437	DYS438	DYS439	DYS385
1	1	13	10	26	22	10	13	13	15	9	19	13–15
2	1	13	10	26	23	9	11	13	14	10	18	15–15
3	1	13	10	26	24	10	13	13	14	10	19	16–18
4	1	13	10	27	24	9	11	13	14	10	19	15–18
5	1	13	10	27	24	10	11	13	14	10	19	15–18
6	1	13	10	27	24	10	11	13	14	10	19	16–18
7	1	13	10	27	25	10	11	15	14	10	19	14–18
8	1	13	10	28	23	10	11	13	14	10	19	17–18
9	1	13	10	28	24	10	11	14	14	10	19	16–19
10	1	13	10	28	24	10	11	14	14	10	19	17–18
11	1	13	11	26	23	10	13	13	15	9	18	13–17
12	1	13	11	28	24	10	11	13	14	10	19	15–19
13	1	13	11	28	24	10	11	13	15	10	20	10–13
14	1	14	8	24	23	11	11	13	16	10	18	13–15
15	1	14	9	25	22	10	11	13	16	10	18	13–14
16	1	14	9	25	23	10	11	12	16	10	18	14–15
17	1	14	9	25	24	10	13	13	15	12	20	11–14
18	1	14	9	25	25	11	13	13	14	12	21	11–14
19	1	14	9	26	23	11	13	13	15	12	19	11–14
20	1	14	9	26	24	9	11	13	14	10	17	18–18
21	1	14	9	27	22	10	11	14	14	10	19	18–19
22	1	14	9	27	24	11	13	13	14	13	18	11–14
23	1	14	10	25	24	11	13	13	15	12	20	11–13
24	1	14	10	25	25	10	11	12	15	9	18	13–15
25	1	14	10	26	22	10	11	12	16	9	19	14–17
26	1	14	10	26	23	10	11	12	15	9	18	13–18
27	1	14	10	26	23	10	11	13	15	9	19	14–17
28	1	14	10	26	23	10	13	13	15	12	19	11–11
29	1	14	10	26	23	11	12	12	15	9	18	13–18
30	1	14	10	26	23	11	13	13	15	12	19	11–14
31	2	14	10	26	23	11	13	13	15	12	20	11–14
32	1	14	10	26	23	11	13	14	15	12	19	11–14
33	1	14	10	26	24	10	13	13	15	12	19	11–13
34	1	14	10	26	24	10	13	13	15	13	19	11–14
35	1	14	10	26	24	10	13	13	16	11	19	11–14
36	1	14	10	26	24	10	14	12	15	12	20	11–14
37	1	14	10	26	24	10	15	13	16	13	19	13–14
38	1	14	10	26	24	11	13	13	14	12	19	10–14
39	1	14	10	26	24	11	13	13	14	13	20	11–14
40	1	14	10	26	24	11	13	13	15	11	19	11–16
41	1	14	10	26	24	11	13	13	15	12	18	11–14
42	1	14	10	26	24	11	13	13	15	12	18	14–14
43	1	14	10	26	24	11	13	13	15	12	18	10–14
44	4	14	10	26	24	11	13	13	15	12	19	11–14
45	2	14	10	26	24	11	13	13	15	12	20	11–14
46	1	14	10	26	24	11	13	13	15	12	20	12–13
47	1	14	10	26	24	11	13	13	15	13	19	11–14
48	1	14	10	26	24	11	13	13	15	13	19	12–17
49	1	14	10	26	24	11	13	13	16	12	19	11–14
50	1	14	10	26	24	11	13	15	15	12	20	11–14
51	1	14	10	26	24	12	14	13	15	12	20	11–14
52	1	14	10	26	25	10	13	13	15	12	18	11–11
53	1	14	10	26	25	11	13	12	14	12	17	12–15
54	1	14	10	26	25	11	13	13	14	12	19	11–13
55	1	14	10	27	22	11	11	12	14	9	20	12–12
56	1	14	10	27	23	10	11	12	15	9	20	14–15
57	1	14	10	27	23	10	12	11	16	10	19	17–17

Table 4 (continued)

Haplotype	*n*	DYS19	DY389-I	DY389-II	DY390	DY391	DY392	DY393	DYS437	DYS438	DYS439	DYS385
58	1	14	10	27	23	10	12	14	14	10	18	14–16
59	1	14	10	27	23	11	13	13	14	12	19	11–16
60	1	14	10	27	24	10	11	12	15	9	17	13–16
61	1	14	10	27	24	10	11	13	14	10	19	19–19
62	1	14	10	27	24	10	13	13	15	12	18	12–15
63	1	14	10	27	24	10	13	13	15	12	19	11–16
64	1	14	10	27	24	10	13	13	15	12	19	12–14
65	1	14	10	27	24	10	13	14	15	12	18	11–14
66	3	14	10	27	24	11	13	13	15	12	19	11–14
67	1	14	10	27	25	10	13	13	15	12	18	11–14
68	1	14	10	27	25	11	13	13	14	12	20	11–14
69	1	14	10	28	23	10	11	13	15	9	19	13–16
70	1	14	10	28	24	11	11	13	14	10	20	15–16
71	1	14	11	26	23	10	13	14	15	9	19	14–19
72	1	14	11	27	23	11	13	13	15	12	19	11–13
73	1	14	11	27	23	11	13	13	15	12	19	11–14
74	1	14	11	27	23	11	14	13	15	12	20	11–14
75	1	14	11	27	24	10	13	13	15	12	20	11–14
76	1	14	11	27	24	11	13	13	14	12	18	11–14
77	1	14	11	27	25	11	13	13	14	11	17	11–14
78	1	14	11	28	23	10	13	13	14	9	19	14–17
79	1	14	11	29	24	10	11	12	14	10	18	16–16
80	1	15	9	25	21	10	11	15	16	10	18	13–16
81	1	15	9	25	22	10	11	13	16	10	19	13–15
82	1	15	9	25	24	10	11	12	15	9	19	15–17
83	1	15	9	25	24	10	11	12	16	9	18	13–19
84	1	15	9	25	25	10	11	12	16	9	18	13–16
85	1	15	9	25	25	10	11	12	16	10	18	14–17
86	1	15	9	26	22	10	11	14	16	8	20	13–14
87	2	15	9	26	22	10	11	14	16	10	18	14–14
88	1	15	9	26	22	10	11	14	16	10	19	14–14
89	1	15	9	26	23	10	11	12	15	9	18	13–18
90	1	15	9	26	24	10	11	13	15	9	19	10–18
91	1	15	9	27	22	10	10	14	14	11	18	15–15
92	1	15	9	27	22	10	11	13	15	10	18	14–15
93	1	15	10	26	23	9	11	12	14	9	19	10–16
94	1	15	10	26	23	9	13	14	14	9	19	15–16
95	1	15	10	26	23	10	11	13	15	10	18	12–13
96	1	15	10	26	23	11	13	13	14	12	19	10–15
97	1	15	10	26	24	10	13	13	14	12	20	10–15
98	1	15	10	26	25	11	13	13	15	12	18	11–13
99	1	15	10	27	23	11	13	13	15	11	18	12–12
100	1	15	11	27	22	11	11	12	16	9	18	13–18
101	1	15	11	27	22	11	11	14	16	10	18	13–15
102	1	15	11	27	23	10	11	12	15	9	19	10–16
103	1	15	11	27	23	10	13	14	14	9	18	14–17
104	1	15	11	28	22	10	11	12	14	9	18	15–17
105	1	15	11	28	23	10	11	11	16	10	20	14–14
106	1	15	11	28	25	11	11	13	15	13	20	11–14
107	1	15	11	28	25	11	13	13	15	12	20	12–14
108	1	16	10	26	23	10	11	12	14	9	18	12–16
109	1	16	10	26	23	10	12	14	14	10	18	16–18
110	1	16	10	26	25	10	11	13	14	11	18	12–14
111	1	16	10	28	24	11	11	13	15	10	20	14–15
112	1	16	10	28	25	11	11	13	15	10	20	14–15
113	1	16	11	26	24	9	13	13	15	9	19	12–14
114	1	16	11	27	25	10	11	13	14	11	17	11–14

Table 4 (continued)

Haplotype	*n*	DYS19	DY389-I	DY389-II	DY390	DY391	DY392	DY393	DYS437	DYS438	DYS439	DYS385
115	1	16	11	29	23	10	12	14	14	10	18	15–15
116	1	17	9	25	24	10	11	12	16	9	19	13–16
117	1	17	9	27	23	10	12	14	16	10	18	15–16
118	1	17	10	25	23	10	11	13	15	10	18	12–12
119	1	17	10	26	23	11	11	13	15	10	19	12–13
120	1	17	10	27	25	10	11	13	14	11	18	10–14

The DYS437 and DYS439 loci showed a unimodal allelic distribution (modal number 15 and 19, respectively) and the DYS438 locus showed a bimodal distribution (modal numbers 10 and 12). By combining the allelic state of these 3 STRs it was possible to define 39 different haplotypes of which 20 were shared and 19 were "unique". The haplotype diversity value was 0.9410.

The allelic combinations from the two co-amplified PCR products from the DYS385 locus are shown in Table 2 together with the STR diversity value. A total of 40 different phenotype combinations were found, of which 25 were shared and 15 were "unique". Our frequency data for the DYS385 locus were compared to those from another Italian population sample [5] by means of a R × C contingency table; no statistically significant differences were found ($p > 0.07$). By combining the allelic states of the STR markers used in the tetraplex PCR reaction, we could find 94 different haplotypes with a haplotype diversity value of 0.9759. These values are very similar to those found by combining the allelic state of the basic set of 7 STRs (Table 3) for the same population sample. The combination of the complete set of 11 STR markers allowed the discrimination of 93.8% (120 out of 128) of the males from our population sample with a haplotype diversity value of 0.9907. The distribution of the haplotypes for the set of 11 STRs in the Italian population sample analysed in this study is shown in Table 4.

The forensic usefulness of the STR markers used in the tetraplex PCR reaction was tested by amplifying female DNA samples and mixed samples with a female/male ratio of 100:1. No PCR products in the range of the ladders were seen for the female DNA amplifications and only male specific profiles were seen in the mixed samples. In addition, 18 father-child pairs were analysed for the 11 Y-chromosomal STRs and no mutations were found.

The complete set of 11 Y-chromosomal STRs can be amplified in 3 different multiplex PCR reactions and this approach represents a very powerful tool in forensic medicine for male identification, mixed sample analysis and paternity testing.

Acknowledgements The authors wish to thank Chris Tyler-Smith, University of Oxford, for providing them with the primer sequences and for helpful comments on the manuscript.

References

1. Jobling MA, Tyler-Smith C (1995) Fathers and sons: the Y chromosome and human evolution. Trends Genet 11: 449–456
2. Zerjal T, Dashnyam B, Pandya A, Kayser M, Roewer L, Santos FR, Schiefenhövel W, Fretwell N, Jobling MA, Harihara S, Shimizu K, Semjidmaa D, Sajantila A, Salo P, Crawford MH, Ginter EK, Evgrafov OV, Tyler-Smith C (1997) Genetic relationships of Asians and Northern Europeans revealed by Y-chromosomal DNA analysis. Am J Hum Genet 60: 1174–1183
3. Jobling MA, Pandya A, Tyler-Smith C (1997) The Y chromosome in forensic analysis and paternity testing. Int J Legal Med 110: 118–124
4. Kayser M, Caglià A, Corach D, Fretwell N, Geherig 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, Perez-Leuzaun A, Piccinini A, Prinz M, Schmitt C, Schneider PM, Szibor R, Teifel-Greding J, Weichhold D, de Knijff P, Roewer L (1997) Evaluation of Y-chromosomal STRs: a multicenter study. Int J Legal Med 110: 125–133
5. 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
6. Ayub Q, Mohyuddin A, Qamar R, Mazhar K, Zerjal T, Medhi Q, Tyler-Smith C (2000) Identification and characterisation of novel human Y-chromosomal microsatellites from sequence database information. Nucleic Acids Res 28(2): e8
7. Pascali VL, Dobosz M, Brinkmann B (1998) Coordinating Y-chromosome STR research for the Courts. Int J Legal Med 112: 1
8. Previderè C, Tyler-Smith C, Grignani P, Peloso G (2000) Y-chromosomal haplotypes and haplogroups in male identification. In: Sensabaugh GF, Lincoln PJ, Olaisen B (eds) Progress in forensic genetics 8. Elsevier Science, Amsterdam, pp 257–259
9. DNA Commission of the ISFH (1994) Report concerning further recommendations of the DNA Commission of the ISFH regarding PCR-based polymorphisms in STR (short tandem repeats). Int J Legal Med 107: 159–160