Int J Legal Med (2002) 116 : 133–138

ORIGINAL ARTICLE

Francesco Calì · Peter Forster · Christian Kersting
Mario G. Mirisola · Rosalba D'Anna · Giacomo De Leo
Valentino Romano

DXYS156: a multi-purpose short tandem repeat locus for determination of sex, paternal and maternal geographic origins and DNA fingerprinting

Received: 27 June 2001 / Accepted: 2 October 2001

Abstract In forensic science and in legal medicine Y chromosomal typing is indispensable for sex determination, for paternity testing in the absence of the father and for distinguishing males in multiple rape cases. Another potential application is the estimation of paternal geographic origin or family name from a crime stain to narrow down the range of suspects and thus reduce costs of mass screenings. However, Y typing alone cannot provide a sufficiently resolved DNA fingerprint as required for court convictions. Thus, there is a dilemma whether or not to sacrifice valuable material for the sake of extensive Y chromosomal investigations when stain DNA is limited (typically allowing only few PCR amplifications). We here describe a Y-chromosome-specific nucleotide insertion in the duplicate short tandem repeat (STR) locus DXYS156 which allows us to distinguish males from females as does the commonly used amelogenin system, but with the advantage that this locus is multi-allelic, thus substantially contributing towards DNA fingerprinting of a sample and furthermore enabling the detection of sample contamination. Yet another bonus is that both the X and the Y copies of DXYS156 have alleles specific to different parts of the world, offering separate estimates of maternal and paternal descent of that sample. We therefore recommend the inclusion of DXYS156 in standard multiplexing kits for forensic, archaeological and genealogical applications.

Keywords Human · Ethnic · Population · Race · Phenotype · Evolution

F. Calì · V. Romano
Laboratorio di Genetica Molecolare,
Istituto OASI Maria Santissima (I.R.C.C.S.), Troina, Italy

P. Forster (✉)
The McDonald Institute for Archaeological Research,
University of Cambridge,
Downing Street, Cambridge CB2 3ER, England
e-mail: pf223@cam.ac.uk,
Fax: +44-1223-339285

C. Kersting
Institute of Legal Medicine, University of Münster, Germany

M. G. Mirisola · R. D'Anna · G. De Leo · V. Romano
Dip. Biopatologia e Metodologie Biomediche,
University of Palermo, Palermo, Italy

Introduction

New Y short tandem repeat (STR) loci are continuously being discovered with the progress of the Human Genome Project (e.g. White et al. 1999) and are gaining importance not only in studies of human evolution (Forster et al. 1998, 2000) but also in genealogical research, in forensic casework and in paternity testing (Jobling et al. 1997; Kayser et al. 1997; Honda et al. 1999). Due to their non-recombining nature, Y STRs are inherited uniparentally in male lineages and persist as Y haplotypes over generations of male descendency unless a mutation occurs. This feature of paternal inheritance makes Y STRs a powerful tool for family name studies (Sykes and Irven 2000). In medico-legal applications, Y typing is indispensable for sex testing (conventionally using the amelogenin system Mannucci et al. 1994; but see Santos et al. 1998; Brinkmann 2002; Steinlechner et al. 2002; Thangaraj et al. 2002), for paternity deficiency tests (i.e. when the putative father is dead or unavailable, the patrilineal relatives can be taken as substitutes, e.g. Foster et al. 1998; Thomas et al. 1998) and for multiple rape investigations (determination of the number and identity of male offenders in mixed semen stains). Another new and potentially very useful application is the determination of the paternal descent of a stain: in many police investigations, especially where expensive mass screenings are required, knowledge of the geographic origin and by extrapolation the phenotype, of an unknown stain donor could narrow down the list of suspects and substantially reduce time and cost of the screening.

The problem is that Y haplotypes alone do not have a discrimination capacity high enough for court convictions, if only for the reason that many suspects will have

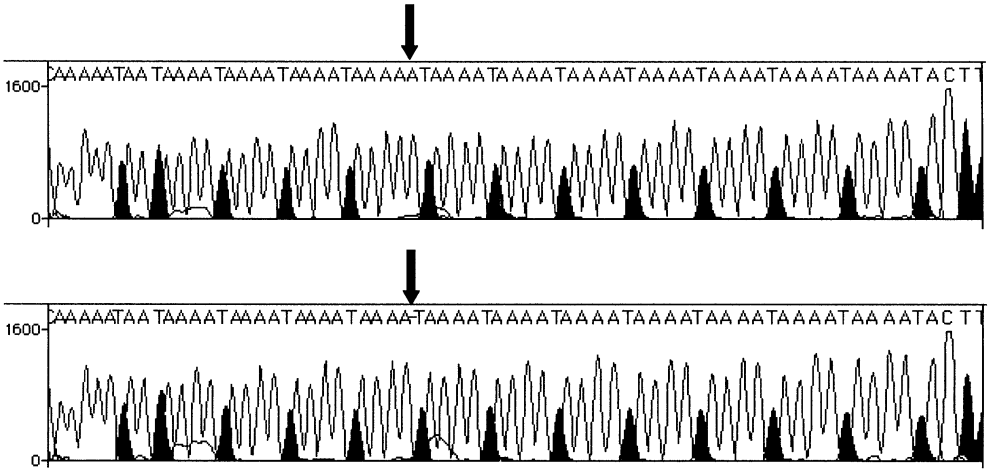

Fig. 1 Adenine insertion in the Y locus of DXYS156. The normal repeat motif is TAAAA which is extended to TAAAAA by the insertion in the fourth repeat

Table 1 DXYS156X allele frequencies in Sicily and Korea

Location	*n*	Allele							
		4	5	6	7	8	9	10	11
Sicily (Sciacca)	116	0.01	0	0	0.87	0.03	0.05	0.04	0
Sicily (Troina)	166	0	0	0	0.78	0.08	0.13	0.01	0.01
Korea (Kwangju)	48	0	0	0	0.96	0.04	0	0	0

male relatives who will have the same Y chromosome. It is therefore still indispensable to type autosomal loci which, due to their inherent feature of recombination in every generation, yield potentially unique genetic profiles (except in cases of identical twins). Thus, there is a dilemma whether or not to sacrifice valuable material for the sake of extensive Y chromosomal investigations when stain DNA is very limited (typically allowing only few PCR amplifications). Here, we resolve this conflict by presenting a multi-purpose XY-STR (DXYS156); its X and Y homologues can be distinguished from each other by virtue of an internal point insertion which we have discovered to be Y-specific.

Subjects and methods

Sicilian blood samples were obtained from healthy blood donors, selected for ancestry of all four grandparents from the Sicilian towns of Troina (59 women and 48 men) and Sciacca (33 women and 50 men). Korean samples (7 women and 34 men) were obtained from students in Kwangju, South Korea. DNA was extracted from peripheral blood leukocytes by standard procedures.

Extracted DNA was amplified by the polymerase chain reaction (PCR) using the published primers and thermocycling conditions (Chen et al. 1994). The PCR was performed in 50 µl containing 50 ng of genomic DNA, 1 U Taq DNA polymerase (Perkin Elmer, USA), 5 µl 10 × reaction buffer (20 mM Tris-HCl pH 8, 100 mM KCl, 0.1 mM EDTA, 1 mM DTT, 50% glycerol, 0.5% Tween 20, 0.5% Nonidet P40), 1.5 mM $MgCl_2$, 0.2 mM of each dNTP and 0.2 mM of each primer. Primer 1 was modified 5′ by addition of a dye label (TAMRA: N,N,N′,N′-tetramethyl-6-carboxyrhodamine). After PCR amplification, 1 µl of the products was diluted in 12 µl of deionised formamide and 1 µl of GeneScan 350 Rox (molecular weight DNA marker). The DNA was denatured at 95 °C for 3 min, cooled on melting ice and loaded on an ABI PRISM 310 genetic analyser (Perkin Elmer, USA) for amplicon length determination.

For those alleles that were chosen for sequencing, the alleles were first cloned using the TOPO-TA cloning kit (Invitrogen, Groningen, The Netherlands). The same primers were used to sequence both strands of DNA, using ABI PRISM BigDye terminator cycle sequencing ready reaction kit (PE, Applied Biosystems, Milano, Italy).

In order to exclude the possibility of sample confusion or contamination, the recorded sex was confirmed for each sequenced sample by the amelogenin test (Sullivan et al. 1993; Mannucci et al. 1994) using the AmpF1 STR Green kit (PE Applied Biosystems, Milano, Italy).

Results

Adenine insertion in the Y homologue of DXYS156

Our initial sequencing in Sicilian and Korean males and females of the whole range (4–14 repeats) of DXYS156 alleles consistently revealed an inserted adenine in the fourth repeat motif from the 5′ end (Fig. 1) in males. To avoid the inconvenience of DNA sequencing for routine applications, we then generated an allelic ladder which contains the most common alleles on the X and Y chromosomes (Fig. 2). Length analysis using this ladder confirmed the presence of the insertion in all males (98 out of

Table 2 DXYS156Y allele frequencies in Sicily and Korea

Location	*n*	Allele[a]				
		10.1	11.1	12.1	13.1	14.1
Sicily (Sciacca)	50	0.02	0.14	0.84	0	0
Sicily (Troina)	48	0.02	0.21	0.77	0	0
Korea (Kwangju)	34	0	0.12	0.50	0.32	0.06

[a] Elsewhere in the tables and text, we omit the insertion label ".1" to simplify comparisons with the published undifferentiated alleles

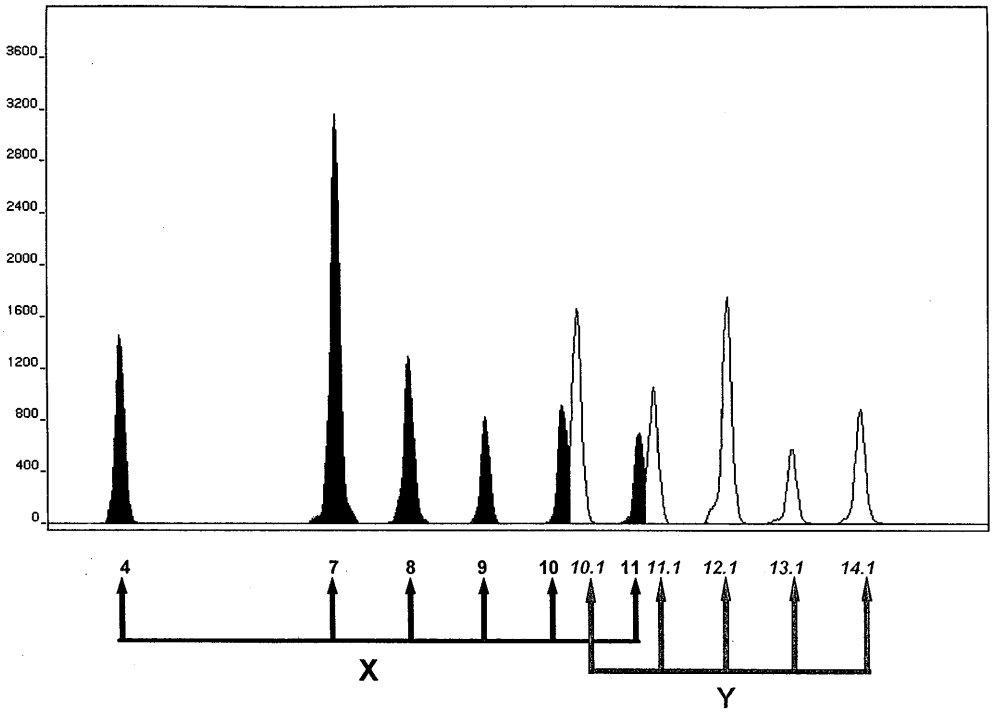

Fig. 2 Electrophoretic allele ladder for distinguishing adenine insertion alleles from uniform alleles at DXYS156

Table 3 DXYS156X allele frequencies worldwide

Location	*n*	Reference	Allele								
			4	5	6	7	8	9	10	11	12
Sicily (Sciacca)	116	This study	0.01	0	0	0.87	0.03	0.05	0.04	0	0
Sicily (Troina)	166	This study	0	0	0	0.78	0.08	0.13	0.01	0.01	0
Germany	116	Kersting et al. (2001)	0	0	0	0.81	0.08	0.07	0.03	0	0
Italy	100	Kersting et al. (2001)	0	0	0	0.81	0.06	0.12	0.01	0	0
Turkey	194	Kersting et al. (2001)	0.02	0	0.01	0.83	0.02	0.11	0.02	0	0
Morocco	129	Kersting et al. (2001)	0.01	0.01	0	0.76	0.04	0.17	0.02	0.01	0
E. Africa	21	Kersting et al. (2001)	0.14	0	0.04	0.33	0.28	0.14	0.04	0	0
W. Africa	361	Kersting et al. (2001)	0.33	0.01	0	0.20	0.26	0.14	0.03	0.01	0
Namibia (Ovambo)	179	Kersting et al. (2001)	0.31	0	0	0.21	0.30	0.13	0.02	0.01	0
Mongolia	20	Kersting et al. (2001)	0	0	0	1.00	0	0	0	0	0
China	94	Kersting et al. (2001)	0	0.01	0	0.93	0.05	0	0	0	0
Korea	48	This study	0	0	0	0.96	0.04	0	0	0	0
Japan	138	Kersting et al. (2001)	0	0.01	0.01	0.92	0.04	0.02	0	0	0
PNG highlands	114	Kersting et al. (2001)	0	0.01	0	0.96	0.03	0	0	0	0
Africa (incl. Egypt)	223	Karafet et al. (1998)	0.19	0	0.01	0.24	0.27	0.18	0.08	0.01	0.02
Europe	312	Karafet et al. (1998)	0	0	0	0.80	0.07	0.10	0.03	0	0
N. Asia	657	Karafet et al. (1998)	0	0	0	0.91	0.06	0.02	0	0	0
E. Asia	552	Karafet et al. (1998)	0	0	0	0.93	0.05	0.02	0	0	0
Australasia	104	Karafet et al. (1998)	0	0	0	0.96	0.03	0.01	0	0	0
Americas	442	Karafet et al. (1998)	0	0	0	0.91	0.02	0.07	0	0	0

98 Sicilians and 34 out of 34 Koreans) but never in females (0 out of 92 Sicilians and 0 out of 7 Koreans). Sex testing by amelogenin (Sullivan et al. 1993) confirmed the sex recorded on the sample labels. We conclude that the point insertion is specific to the Y chromosome (see Tables 1 and 2). Worldwide surveys have shown that DXYS156 generally has allele lengths ranging from 4 to 12 repeats on the X chromosome (Table 3) and 8 to 15 repeats on the Y chromosome (Table 4). Although the X locus generally has allele lengths of 10 repeats or shorter and the Y locus generally has 11 repeats or longer, there is an overlap of X and Y allele ranges amounting to several percent in many populations (Tables 3 and 4). This overlap has previously hindered the secure chromosomal assignment of alleles despite efforts at statistical separation (Karafet et al. 1998). By typing the adenine insertion, which we have shown to be Y-specific, the Y alleles can be unambiguously distinguished from the X alleles.

Table 4 DXYS156Y allele frequencies worldwide

Location	*n*	Reference	Allele							
			8	9	10	11	12	13	14	15
Sicily (Sciacca)	50	This study	0	0	0.02	0.14	0.84	0	0	0
Sicily (Troina)	48	This study	0	0	0.02	0.20	0.77	0	0	0
Italy (Modena)	99	Rossi et al. (1999)	0	0	0	0	0.99	0.01	0	0
Germany	179	Nata et al. (1999)	0	0	0.01	0.03	0.97	0	0	0
Finland	54	Sajantila et al. (1996)	0	0	0	0	0.98	0.02	0	0
Estonia	20	Sajantila et al. (1996)	0	0	0	0	1.00	0	0	0
Saami	28	Sajantila et al. (1996)	0	0	0	0	1.00	0	0	0
Sweden	40	Sajantila et al. (1996)	0	0	0	0	0.98	0.02	0	0
Basques	25	Sajantila et al. (1996)	0	0	0	0.08	0.88	0.04	0	0
Switzerland	51	Sajantila et al. (1996)	0	0	0	0	1.00	0	0	0
Switzerland	99	Kayser et al. (1997); De Knijff et al. (1997)	0	0	0	0	0.97	0.02	0	0
Netherlands	89	Kayser et al. (1997); De Knijff et al. (1997)	0	0	0	0.05	0.94	0.01	0	0
Turkey	39	Forster et al. (2000)	0	0	0	0.13	0.82	0.05	0	0
Kurdistan	101	Brinkmann et al. (1999)	0	0.02	0	0.13	0.86	0	0	0
Sinai	67	Salem et al. (1996)	0	0	0	0	1.00	0	0	0
Egypt	153	Salem et al. (1996)	0	0	0	0.46	0.54	0	0	0
Morocco	44	Kersting et al. (2001)	0	0	0	0.66	0.32	0.02	0	0
E. Africa	21	Kersting et al. (2001)	0.05	0.05	0	0.90	0	0	0	0
W. Africa	181	Kersting et al. (2001)	0	0	0.01	0.90	0.09	0.01	0	0
Namibia (Ovambo)	28	Forster et al. (2000)	0	0	0	0.96	0.04	0	0	0
Thailand	50	Horst (1999)	0	0	0	0.14	0.56	0.14	0.16	0
Mongolia	20	Forster et al. (2000)	0	0	0	0.45	0.40	0.10	0.05	0
China	35	Forster et al. (2000)	0	0	0	0.31	0.23	0.46	0	0
Korea	34	This study	0	0	0	0.12	0.50	0.32	0.06	0
Japan	44	Forster et al. (2000)	0	0	0	0.43	0.34	0.20	0.02	0
Australia	32	Forster et al. (1998)	0	0	0	0.56	0.44	0	0	0
PNG highlands	47	Forster et al. (1998)	0	0	0	0.09	0.91	0	0	0
Africa (incl. Egypt)	204	Karafet et al. (1998)	0.02	0.03	0	0.72	0.22	0.01	0	0
Europe	231	Karafet et al. (1998)	0	0	0	0.04	0.94	0.02	0	0
N. Asia	644	Karafet et al. (1998)	0	0	0	0.34	0.64	0.02	0	0
E. Asia	497	Karafet et al. (1998)	0	0	0	0.42	0.30	0.17	0.10	0.01
Australasia	68	Karafet et al. (1998)	0	0	0	0.27	0.69	0.04	0	0
Americas	362	Karafet et al. (1998)	0	0	0.01	0.06	0.92	0.01	0	0

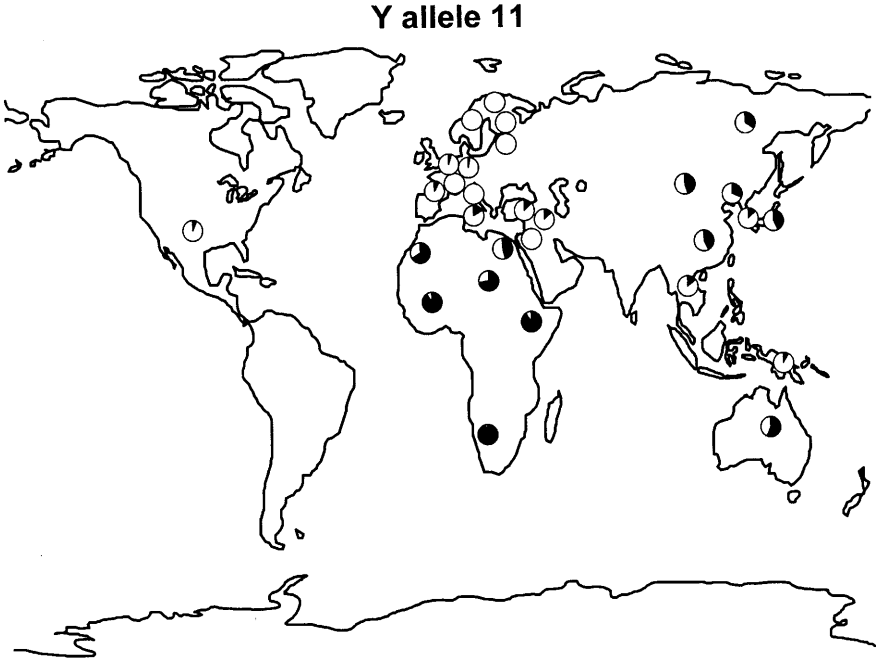

Fig. 3 Geographical distribution of DXYS156Y allele length 11 repeats

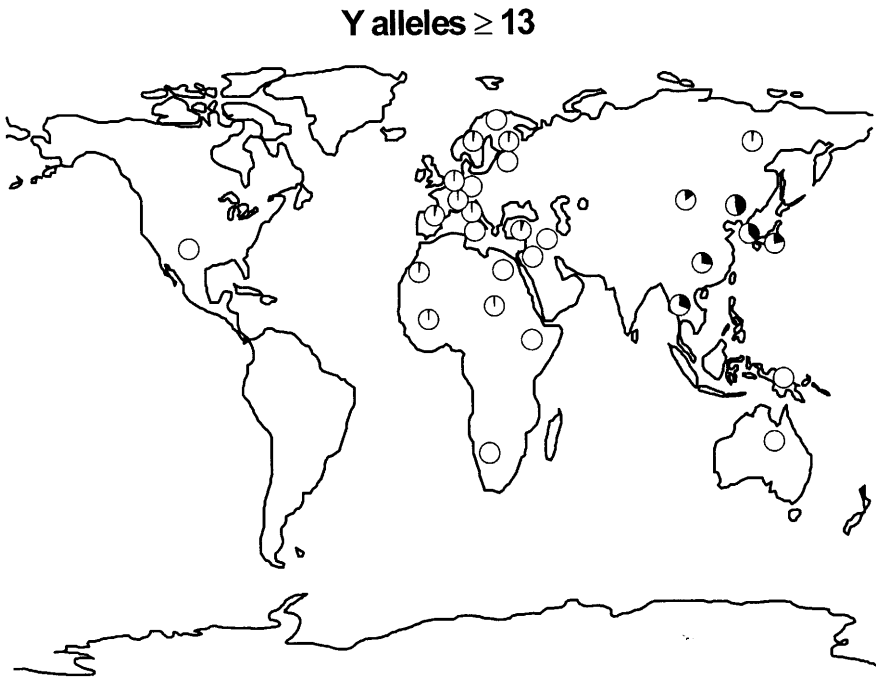

Fig. 4 Geographical distribution of DXYS156Y alleles ≥ 13 repeats

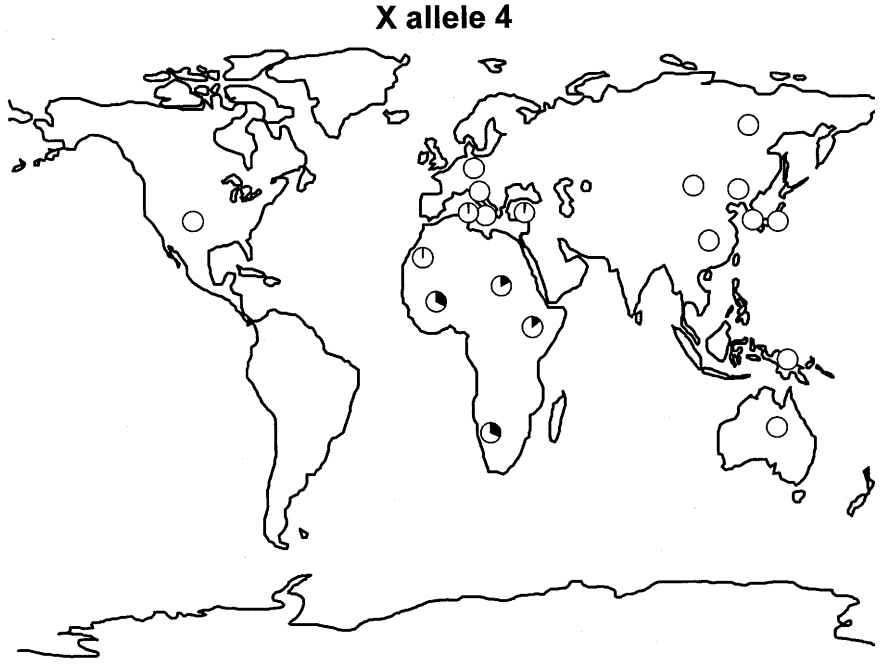

Fig. 5 Geographical distribution of DXYS156X allele length 4 repeats

Geographic specificities of alleles

The predominant Y allele in the Sicilian samples (inhabitants of Troina and Sciacca) is allele length 12, which is by far the most common European allele (Table 4). However, the allele 11 in the Troina sample is unusually common (10 out of 48 males) for a European population (Fig. 3). We therefore investigated the family names of the Troina donors and found that 6 of the 10 males with allele 11 share either of 2 family names. Evidently some of the increased percentage of allele 11 in Troina is due to these two paternal founders who must have lived some time after the introduction of family names in Italy (thirteenth to fourteenth century AD), but before three generations ago, which was our limit for tracing family relationships between donors.

While European males predominantly have Y allele 12 and Africans generally have Y allele 11, east Asians are distinctive in having the longest Y alleles at high frequency, namely Y alleles 13, 14 and 15 (Fig. 4). On the X chromosome, X allele 4 is found at high frequency only in Africans (Fig. 5). Hence, there are several X and Y alleles which are diagnostic for different regions of the world, which can assist the geographical assignment of the mater-

nal and paternal line of a male sample of unknown origin. For female samples, in ideal cases paternal and maternal X alleles may be distinguished by inference if the abundant, maternally inherited, mitochondrial DNA (mtDNA) of a sample is additionally typed and located geographically (mtDNA localisation is possible with a geographic precision of 0–2000 km in two-thirds of cases, see Röhl et al. 2001 and Forster et al. 2002): if the mtDNA type is specific to the same part of the world as one of the X alleles, then it may be reasonable to equate the geographic specificity of the other X allele with the paternal origin of the sample.

Conclusions

The ability to unambiguously distinguish the X and Y homologues via the point mutation greatly enhances the discrimination capacity and sex testing capacity of DXYS156, thus combining into a single PCR amplification experiment what hitherto needs to be typed in three unrelated PCRs (i.e. DNA fingerprinting through a multi-allelic autosomal STR, Y typing through a Y STR, and sex testing with the amelogenin system). Similar to the amelogenin system, DXYS156 offers a positive control for sex testing (the X homologue should always appear), except that DXYS156 is multi-allelic and therefore additionally may warn the user of the presence of contaminant alleles in the sample. As a further bonus, the unambiguously distinguishable X and Y alleles display geographic specificities, allowing separate estimates of the maternal and paternal geographic origins of a given sample. We therefore recommend that DXYS156 should replace less informative STRs in standard multiplexing kits.

Acknowledgements This work was partially supported by Progetto Finalizzato C.N.R., Sottoprogetto 4 Beni Culturali: "Cultural heritage", and by Progetto di Ricerca Finalizzata "Eterogeneità genetica della PKU in Italia: Strategie e tecniche di analisi per la ricerca delle mutazioni del gene PAH e per la valutazione della loro severità", the Italian Ministery of Health.

References

Brinkmann B (2002) Editorial. Int J Legal Med 116:63

Brinkmann C, Forster P, Schürenkamp M, Horst J, Rolf B, Brinkmann B (1999) Human Y-chromosomal STR haplotypes in a Kurdish population sample. Int J Legal Med 112:181–183

Chen H, Lowther W, Avramopoulos D, Antonarakis SE (1994) Homologous loci DXYS156X and DXYS156Y contain a polymorphic pentanucleotide repeat $(TAAAA)_n$ and map to human X and Y chromosomes. Hum Mut 4:208–211

Forster P, Kayser M, Meyer E, Roewer L, Pfeiffer H, Benkmann H, Brinkmann B (1998) Phylogenetic resolution of complex mutational features at Y-STR DYS390 in Aboriginal Australian and Papuans. Mol Biol Evol 15:1108–1114

Forster P, Röhl A, Lünnemann P, Brinkmann C, Zerjal T, Tyler-Smith C, Brinkmann B (2000) A short tandem repeat-based phylogeny for the human Y chromosome. Am J Hum Genet 67: 182–196

Forster P, Calì F, Röhl A, Metspalu E, D'Anna R, Mirisola M, De Leo G, Flugy A, Salerno A, Ayala G, Kouvatsi A, Villems R, Romano V (2002) Continental and subcontinental distributions of mtDNA control region types. Int J Legal Med 116:99–108

Foster EA, Jobling MA, Taylor PG, Donnelly P, de Knijff P, Mieremet R, Zerjal T, Tyler-Smith C (1998) Jefferson fathered slave's last child. Nature 396:27–28

Honda K, Roewer L, de Knijff P (1999) Male typing from 25-year-old vaginal swabs using Y-STR polymorphisms in retrial request case. J Forensic Sci 44:868–872

Horst B, Eigel A, Sanguasermsri T, Brinkmann B (1999) Human Y-chromosomal STR types in north Thailand. Int J Legal Med 112:211–212

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

Karafet T, de Knijff P, Wood E, Ragland J, Clark A, Hammer MF (1998) Different patterns of variation at the X- and Y-chromosome-linked microsatellite loci DXYS156X and DXYS156Y in human populations. Hum Biol 70:979–992

Kayser M, Caglià A, Corach D, Fretwell N, Gehrig C, Graziosi G, Heidorn F, Herrmann S, Herzog B, Hidding M, et al, (1997) Evaluation of Y-chromosomal STRs: a multicenter study. Int J Legal Med 110:125–133, 141–149

Kersting C, Hohoff C, Rolf B, Brinkmann B (2001) Pentanucleotide short tandem repeat locus dxys156 displays different patterns of variations in human populations. Croat Med J 42: 310–314

Knijff De P, Kayser M, Caglià A, Corach D, Fretwell N, Gehrig C, Graziosi G, et al, (1997) Chromosome Y microsatellites: population genetic and evolutionary aspects. Int J Legal Med 110: 134–149

Mannucci A, Sullivan KM, Ivanov PL, Gill P (1994) Forensic application of a rapid and quantitative DNA sex test by amplification of the X-Y homologous gene amelogenin. Int J Legal Med 106:190–193

Nata M, Brinkmann B, Rolf B (1999) Y-Chromosomal STR haplotypes in a population from north west Germany. Int J Legal Med 112:406–408

Röhl A, Brinkmann B, Forster L, Forster P (2001) An annotated mtDNA database. Int J Legal Med 115:29–39

Rossi E, Rolf B, Schürenkamp M, Brinkmann B (1999) Y-chromosome STR haplotypes in an Italian population sample. Int J Legal Med 112:78–81

Sajantila A, Salem A-H, Savolainen P, Bauer K, Gierig C, Pääbo S (1996) Paternal and maternal DNA lineages reveal a bottleneck in the founding of the Finnish population. Proc Natl Acad Sci USA 93:12035–12039

Salem A-H, Badr FM, Gaballah MF, Pääbo S (1996) The genetics of traditional living: Y-chromosomal and mitochondrial lineages in the Sinai peninsula. Am J Hum Genet 59:741–743

Santos FR, Pandya A, Tyler-Smith C (1998) Reliability of DNA-based sex tests. Nat Genet 18:103

Steinlechner M, Berger B, Niederstätter H, Parson W (2002) Rare failures in the amelogenin sex test. Int J Legal Med (in press)

Sullivan KM, Mannucci A, Kimpton CP, Gill P (1993) A rapid and quantitative DNA sex test: fluorescence-based PCR analysis of X-Y homologous gene amelogenin. Biotechniques 15:636–641

Sykes B, Irven C (2000) Surnames and the Y chromosome. Am J Hum Genet 66:1417–1419

Thangaraj K, Reddy AG, Singh L (2002) Is the amelogenin gene reliable for gender identification in forensic casework and prenatal diagnosis? Int J Legal Med 116:121–123

Thomas MG, Skorecki K, Ben-Ami H, Parfitt T, Bradman N, Goldstein DB (1998) Origins of Old Testament priests. Nature 394:138–140

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