International Congress Series 1239 (2003) 343–348

Y-chromosome STR haplotypes in a Swedish population

G. Holmlund *, H. Nilsson, A. Langö-Warensjö, B. Rosén, B. Lindblom

National Board of Forensic Medicine, Institute of Forensic Genetics, University Hospital, SE-581 85 Linköping, Sweden

Abstract

The Swedish Y-chromosome database of 350 individuals is based on an analysis of fathers and their sons. Individuals with Swedish names coming from different parts of Sweden are included. The Y-chromosomes were typed for the STR markers DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393 and DYS385. An evaluation of the 350 chromosomes revealed 241 different haplotypes of which the most common is found in only 15 chromosomes (4.3%). Nineteen haplotypes were found more than three times and 28 haplotypes were found twice. Six of the 12 most frequent haplotypes are present in frequencies that are significantly different from the frequencies among the 6124 Y-STR haplotypes reported to the European database. Seven of these can be assigned to a cluster of closely related haplotypes found in 50 individuals and another three to a different cluster with related haplotypes found in 19 individuals. The first cluster is significantly different from the corresponding cluster of haplotypes in the European database. Comparisons were done by a 2×2 table using Yates correction. The Y-STR polymorphism in the Swedish population with an estimated haplotype diversity of 0.994 and a discrimination capacity of 69% will be very useful for testing paternal lineages.

Keywords: Y-chromosome; STR; Population; Database

1. Introduction

Y-STR haplotyping has during the last few years become a valuable tool for paternity and kinship testing. Typing of closely linked loci that do not undergo recombination

* Corresponding author. Tel.: +46-13-22-30-36; fax: +46-13-13-60-05.
E-mail address: gunilla.holmlund@rmv.se (G. Holmlund).

0531-5131/03
PII: S0531-5131(02)00305-9

processes, as represented by the Y-STR loci, is of particular interest in deficiency cases and in testing paternal lineages over generations. Exclusion of unrelated haplotypes is usually straightforward. Inclusion, however, needs thorough statistical evaluation of the likelihood that an Y-chromosome comes from a related person by inheritance rather than that it is identical just by chance. The Y-STR loci are linked and the frequencies for the different haplotypes cannot be calculated via the product of the individual allele frequencies at the different loci. Therefore, the number of Y-chromosomes that are included in a database needs to be very large. Our aim has been to establish a Swedish Y-chromosome STR database including 350 Y-chromosomes.

2. Materials and methods

2.1. Samples and PCR reaction

Seven hundred Y-chromosomes from 350 father–son combinations were typed for the Y-STR markers DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393 and DYS385. Approximately 100 ng DNA from each individual was analysed in single 25-μl PCR reactions for each marker. For identification of the marker and detection of the PCR products, the forward primers were labelled with different fluorescent dyes. The PCR reactions were optimised by using different concentrations of $MgCl_2$ as follows: 2 mM for DYS385; 3.5 mM for DYS390, DYS391; 4 mM for DYS19, DYS389I, DYS389II, DYS393; and 4.5 mM for DYS392. Amplification was achieved with 30 cycles of 15 s denaturation at 94 °C, 15 s annealing at 57 °C and 30 s synthesis at 72 °C. The cycles were preceded by a denaturation for 11 min at 95 °C and ended with a hold of 45 min for full synthesis at 60 °C.

2.2. Definition of alleles

Alleles were defined by running a mixture of the PCR products from single amplifications of each marker per lane on an ABI 377 with filter D. The forward primers of DYS19, DYS392, DYS393 and DYS385 were labelled with Fam; DYS389I, DYS389II

Table 1
Alleles, with sizes in base pair, included in the ladders for each marker

DYS393	DYS19	DYS391	DYS389I	DYS389II	DYS390	DYS385	DYS392
13 (123)	13 (188)	10 (285)	11 (246)	27 (362)	22 (210)	11 (363)	11 (251)
14 (127)	14 (192)	11 (289)	12 (250)	28 (366)	23 (214)	12 (367)	12 (254)
15 (131)	15 (196)		13 (254)	30 (374)	25 (222)	13 (371)	13 (257)
	16 (200)		15 (162)	31 (378)		14 (374)	14 (250)
						15 (378)	
						16 (382)	
						17 (386)	
						21.2 (404)	

with Ned; and DYS390, DYS391 with Hex. GS500 was used as size marker and the allelic ladders were constructed using sequenced alleles (Table 1). The nomenclature is according to Kayser et al. 1997 [1]. Alleles outside the ladder were identified according to the number of repeats calculated from their sizes.

Table 2
The allele frequencies for each locus given per 1000 individuals

DYS393		DYS19		DYS391		DYS389I		DYS389II		DYS390		DYS385		DYS392	
11	3	**11**	14	**8**	3	**11**	3	**27**	9	**21**	14	**9,13**	3	**10**	3
12	97	**13**	69	**9**	17	**12**	311	**28**	257	**22**	186	**10,13**	3	**11**	606
13	743	**14**	589	**10**	569	**13**	503	**29**	317	**23**	283	**10,14**	9	**12**	37
14	149	**15**	209	**11**	377	**14**	180	**30**	291	**24**	266	**10,16**	3	**13**	249
15	9	**16**	94	**12**	34	**15**	3	**31**	91	**25**	143	**11,11**	20	**14**	103
		17	26					**32**	31	26	6	**11,12**	9	**15**	3
								33	3	**27**	3	**11,13**	91		
												11,14	238		
												11,15	31		
												11,16	3		
												12,12	6		
												12,14	40		
												12,15	9		
												12,16	3		
												12,17	6		
												13,13	26		
												13,14	111		
												13,15	20		
												13,16	11		
												13,17	3		
												13,18	9		
												13,19	11		
												13,22	3		
												14,14	100		
												14,15	63		
												14,16	11		
												14,17	11		
												14,19	3		
												15,15	9		
												15,16	14		
												15,17	11		
												16,16	9		
												16,17	17		
												16,18	14		
												16,19	3		
												16,20	6		
												17,18	3		
												17,19	6		
												18,18	3		
												19,20	3		

The names of the alleles are in boldface.

3. Results and discussion

3.1. Confirmation of typing results

For results from a father to be included in the database, they had to be in agreement with the results from the son. In one case, we found a mutation in DYS390, a loss of one repeat from allele 26 to 25. This deletion was confirmed by multiple analysis.

3.2. Allele frequencies

Allele frequencies, given as the numbers found in 1000 individuals, at the individual loci are shown in Table 2. The largest variability is found for the DYS385 locus that contains a duplication of the repetitive sequence. For DYS19, we found an allele 11 that was not present in the European database. This allele was present in three different haplotypes.

3.3. Haplotype diversity

Haplotypes are given in the order the STR loci are found on the Y-chromosome [2,3]. We found 241 different haplotypes among the 350 chromosomes analysed. When contributed to the European Y-chromosome STR database, 47% of these haplotypes were new to the database. The most common haplotype was found on 15 chromosomes and 194 different haplotypes were found only once. The 12 most frequent haplotypes ($n > 3$) listed in Table 3 have been compared with the European database. Six of these haplotypes are present in frequencies that are significantly different from the frequencies in the European database (6124 Y-STR haplotypes, Swedish population excluded). When contributing the Swedish data to the Y-STR database, 141 haplotypes (53%) were completely new and 2 haplotypes contained new DYS385 alleles (Lutz Roewer, personal communication).

Table 3
The 12 most frequent Y-STR haplotypes in the Swedish population (*S*, $n = 350$) compared with the European Y-chromosome database (*E*, $n = 6124$)

No.	DYS393	DYS19	DYS391	DYS389I	DYS389II	DYS390	DYS385	DYS392	*S*	E	P>
1	13	14	10	12	28	23	14,14	11	15	29	99.95
2	13	14	10	12	28	23	14,15	11	9	13	99.95
3	13	14	10	12	28	23	13,14	11	5	37	n.s.
4	13	14	10	12	28	22	13,14	11	8	73	n.s.
5	13	14	10	12	28	22	14,14	11	4	26	n.s.
6	13	15	10	12	28	22	13,14	11	5	12	99.95
7	13	15	10	12	28	22	13,13	11	4	1	–
8	13	14	10	13	29	24	12,14	13	4	13	99.00
9	13	14	11	13	29	24	11,14	13	11	200	n.s.
10	13	14	11	13	30	24	11,14	13	4	30	n.s.
11	14	14	11	14	30	24	11,13	14	6	8	99.95
12	13	16	11	13	29	25	11,14	11	4	14	99.00

The comparison was done by a 2×2 table using Yates correction. Haplotypes 1–7 and 8–10 form groups of closely related Y-chromosomes (see Fig. 1).

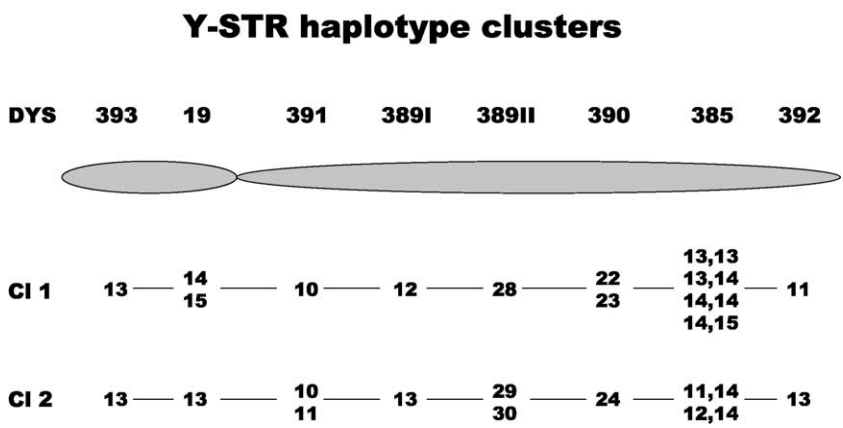

Fig. 1. Y-STR loci in the order they are located on the chromosome. Two clusters of closely related haplotypes were found. The first cluster (Cl 1) consists of seven different haplotypes of which three are significantly different for the Swedish population (P>99.95). Fifty of the 350 haplotypes studied belonged to this cluster which is significantly different from the corresponding cluster of European haplotypes (P>99.95). Cluster 2 (Cl 2) consists of three different haplotypes shared among 19 individuals. One of the haplotypes in this group is significant for the Swedish population, but the cluster is not (see Table 3 for more details).

3.4. Population comparison

Haplotype nos. 1, 2 and 6 (Table 3) that are significant for the Swedish population are also found in high frequencies in the Norwegian population. Haplotype nos. 5 and 11 are also relatively frequent in the Estonian population and no. 12 in the Latvian population although the number of observations is small. The high-frequency haplotype nos. 9 and 10 are found in many populations and seem to be frequent in Europeans in general.

3.5. Haplotype clusters

Two major clusters of haplotypes could be found if we assume that the diversity of the haplotypes has evolved by a single-repeat-change of one allele at only one Y-STR locus at a time (Fig. 1). That is, each haplotype can be derived from another haplotype in the same cluster. The stepwise-mutation model for the evolution of microsatellite diversity in general [4] has also been shown to be valid for the Y-chromosome STR sequences [5]. Haplotype nos. 11 and 12 (Table 3) are both significant for the Swedish population, but could not be included in these clusters.

4. Conclusion

The estimated haplotype diversity is 0.994 and the discrimination capacity is 69% [6]. Both values are in good agreement with the data found for other European populations [7] and the Swedish Y-chromosome STR data will be valuable for analysis of paternal lineages for forensic purposes in Sweden.

References

[1] M. Kayser, A. Cagliá, D. Corach, N. Fretwell, C. Gehrig, G. Graziosi, F. Heidorn, S. Herrmann, B. Herzog, M. Hidding, K. Honda, M.A. Jobling, M. Krawczak, K. Leim, S. Meuser, E. Meyer, W. Oesterreich, A. Pandya, W. Parson, A. Piccinini, A. Perez-Lezaum, M. Prinz, C. Schmitt, P.M. Schneider, R. Szibor, J. Teifel-Greding, G. Weichhold, P. de Knijff, L. Roewer, Evaluation of Y-chromosomal STRs: a multicenter study, Int. J. Legal Med. 101 (3) (1997) 125–33, 141–9.

[2] Q. Ayub, A. Mohyuddin, R. Qamar, K. Mazhar, T. Zerjal, S.Q. Mehdi, C. Tyler-Smith, Identification and characterisation of novel human Y-chromosomal microsatellites from sequence database information, Nucleic Acids Res. 28 (2) (2000) E8–e8.

[3] D.R. Carvalho-Silva, F.R. Santos, M.H. Hutz, F.M. Salzano, S.D.J. Pena, Divergent human Y-chromosome microsatellite evolution rates, J. Mol. Evol. 49 (2) (1999) 202–214.

[4] A.M. Valdes, M. Slatkin, N.B. Freimer, Allele frequencies at microsatellite loci: the stepwise mutation model revisited, Genetics 133 (3) (1993) 737–749.

[5] M. Kayser, L. Roewer, M. Hedman, L. Henke, J. Henke, S. Brauer, C. Krüger, M. Krawczak, M. Nagy, T. Dobosz, R. Szibor, P. de Knijff, M. Stoneking, A. Sajantila, 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 (2000) 1580–1588.

[6] T. Melton, R. Peterson, A.J. Redd, N. Saha, A.S.M. Sofro, J. Martinson, M. Stoneking, Polynesian genetic affinities with southeast Asian populations as identified by mtDNA analysis, Am. J. Hum. Genet. 57 (1995) 403–414.

[7] L. Roewer, The Y chromosome: forensic application and evolutionary aspects, in: B. Olaisen, B. Brinkmann, P.J. Linkoln (Eds.), Prog. Forensic Genet. 7, Elsevier, Amsterdam, 1998, pp. 407–412.