Int J Legal Med (1999) 112 : 181–183

ORIGINAL ARTICLE

C. Brinkmann · P. Forster · M. Schürenkamp · J. Horst
B. Rolf · B. Brinkmann

Human Y-chromosomal STR haplotypes in a Kurdish population sample

Received: 25 September 1998 / Received in revised form: 4 November 1998 / Accepted: 5 November 1998

Abstract In an Iraqi Kurdish population sample (n = 101), seven polymorphic STR loci of the Y-chromosome (DYS19, 389, 390, 391, 392, 393, and DXYS156-Y) were typed, with DYS389 being subtyped for its four segments. The haplotype diversity was 97.83% and 82 different haplotypes were observed. The Kurds shared some Y-types with neighbouring south Turks but strikingly few with Germans: it is 20–30 times more likely to find a sequence match in a random pair of Kurds than in a random Kurd-German pair.

Key words Y-chromosome · Kurdistan · Microsatellite · Match · Probability

Introduction

Short tandem repeat (STR) polymorphisms on the non-recombining part of the Y-chromosome are inherited in the form of compound Y-haplotypes and can only evolve by mutation. The different Y-types in a given population can therefore be linked in an evolutionary tree (Zerjal et al. 1997; Forster et al. 1998), which generally links them by single repeat mutations, the most frequent type of mutation at STR loci (Weber and Wong 1993; Brinkmann et al. 1998). Since most crimes are committed by males, their Y-types can be illuminating for investigations aiming at identification especially in sexual assaults, where mixes of male and female DNA are frequently encountered. It is known that compound Y-types are strongly population-specific (Jobling et al. 1997), necessitating extensive population sampling for a future comparative Y database. To contribute to this international database, we have analysed the standard forensic set of seven Y-STRs (Kayser et al. 1997; de Knijff et al. 1997) in an Iraqi Kurdish sample of 101 males.

C. Brinkmann · J. Horst
Institut für Humangenetik, Westfälische Wilhelms-Universität, Vesaliusweg 12–14, D-48149 Münster, Germany

P. Forster · M. Schürenkamp · B. Rolf · B. Brinkmann (✉)
Institut für Rechtsmedizin, Westfälische Wilhelms-Universität, Von-Esmarch-Strasse 62, D-48149 Münster, Germany
email: brinkma@uni-muenster.de
Tel. +49-251-83-55161; Fax +49-251-83-55158

Materials and methods

DNA was extracted from saliva swabs of 101 unrelated Kurds from north Iraq using the chelex method (Walsh et al. 1991). The following loci were amplified and separated electrophoretically: DYS19, DYS390 (duplex), DXYS156-Y, DYS391, DYS392, DYS393 (quadruplex), four segments of DYS389 (nested PCR reactions 1 and 2). The primers are described in Kayser et al. (1997), de Knijff et al. (1997) and Rolf et al. (1998). The amplification protocol was as follows: 1.5 U AmpliTaqGold (DYS389: 1 U AmpliTaqGold), 0.1–0.8 µM each primer, 200 µM dNTPs, 25 mM $MgCl_2$, 100 mM TrisHCl pH 8.3, 500 mM KCl, 5 µg BSA (DYS389: 20 µg) in a total volume of 25 µl. The forward primers were labeled with the appropriate dyes (Gibco, ABI). The cycling conditions were as follows: for DYS19, 390, 392, 393, 389–2, DXYS156-Y, 94 °C – 10 min, 94 °C – 90 s, 55 °C – 30 s, 72 °C – 30 s, 30 cycles; for DYS389-1, 94 °C – 10 min, 94 °C – 90 s, 55 °C – 30 s, 72 °C – 60 s, 30 cycles. Electrophoresis was carried out by capillary electrophoresis (ABI310) with sequenced allelic ladders as an internal standard. Of the amplified PCR products, 1–3 µl were mixed with 12 µl formamide and 0.5 µl GeneScan500 (internal length standard) according to the manufacturer's instructions. PCR fragments were automatically analysed by the GeneScan software (ABI). The repeat nomenclature follows that of Kayser et al. (1997) and de Knijff et al. (1997), except for DXYS156-Y which accords with the repeat-based nomenclature of Karafet et al. (1998), and DYS389, which was typed at greater resolution: DYS389q corresponds to DYS389-I and the sum of m, n, and q corresponds to DYS389–II. Simple haplotype diversity dw_{min} within a population, i.e. the probability of obtaining different haplotypes when sampling two individuals, was calculated as

$$dw_{min} = 1 - \sum_{i=1}^{h} (x_i / N_x)^2,$$

where x_i is the absolute frequency of the i-th haplotype and h the total number of different haplotypes in the N_x samples (Nei 1987). When comparing between populations, the formula for the probability db_{max} of obtaining different haplotypes when sampling two individuals from two different populations simplifies to

$$db_{max} = 1 - \sum_{i=1}^{m} x_i y_i / (N_x N_y)$$

where only the m sequence matches between populations X and Y need to be considered. Both formulas assume that the absolute fre-

Table 1 Y-STR types in Kurds

#/DYS	19	156Y	389m	389n	389p	389q	390	391	392	393	n	Germ	Turk
1	13	11	5	11	3	9	22	11	11	11	1	0	0
2	13	11	6	11	3	10	23	9	11	13	1	0	0
3	13	11	6	11	3	10	23	10	11	13	1	0	0
4	13	12	5	11	3	10	24	11	13	12	1	0	0
5	13	12	5	11	3	11	22	10	14	13	1	0	0
6	13	12	5	12	3	9	25	10	11	12	2	0	0
7	13	12	5	12	3	9	25	11	11	12	1	0	0
8	13	12	5	12	3	11	23	10	15	14	1	0	0
9	14	9	5	11	3	11	23	10	15	12	1	0	0
10	14	11	5	11	3	9	22	10	14	11	1	0	0
11	14	11	6	11	3	10	24	10	11	13	1	0	0
12	14	11	6	12	3	10	24	10	11	13	1	0	0
13	14	11	6	12	3	10	25	10	11	13	1	0	0
14	14	11	6	12	3	10	26	10	11	13	1	0	0
15	14	12	4	12	3	10	23	10	13	13	1	0	0
16	14	12	4	13	3	10	23	10	13	13	1	0	0
17	14	12	4	13	3	11	23	10	13	13	1	0	1
18	14	12	5	10	3	10	24	10	14	12	1	0	0
19	14	12	5	10	3	10	24	11	14	12	1	0	2
20	14	12	5	10	3	10	25	10	14	12	1	0	0
21	14	12	5	10	3	11	23	10	11	11	1	0	0
22	14	12	5	10	3	12	24	11	13	12	1	0	0
23	14	12	5	11	3	9	23	10	11	12	1	0	1
24	14	12	5	11	3	9	24	10	12	13	1	0	0
25	14	12	5	11	3	10	22	10	15	13	1	0	0
26	14	12	5	11	3	10	23	10	11	12	10	0	3
27	14	12	5	11	3	10	23	11	13	12	1	0	0
28	14	12	5	11	3	10	23	10	13	13	1	3	0
29	14	12	5	11	3	10	24	10	13	12	1	0	0
30	14	12	5	11	3	10	24	11	14	12	1	0	0
31	14	12	5	11	3	11	23	10	11	11	1	0	0
32	14	12	5	11	3	11	23	10	11	12	3	0	0
33	14	12	5	11	3	11	24	10	11	12	1	0	1
34	14	12	5	12	3	9	22	10	12	13	1	0	0
35	14	12	5	12	3	9	23	10	11	12	1	0	0
36	14	12	5	12	3	10	22	10	11	12	1	1	0
37	14	12	5	12	3	10	23	10	11	12	3	0	3
38	14	12	5	12	3	10	23	10	13	13	1	0	0
39	14	12	5	12	3	10	24	10	13	13	1	1	0
40	14	12	5	12	3	10	24	11	13	13	1	2	0
41	14	12	5	12	3	11	23	10	11	12	1	0	1
42	14	12	5	13	3	9	22	10	11	14	1	0	0
43	14	12	5	13	3	11	23	10	11	10	1	0	0
44	14	12	5	14	3	10	24	10	11	12	1	0	0
45	14	12	5	15	3	10	24	10	13	13	1	0	0
46	15	9	5	11	3	10	23	10	14	12	1	0	0
47	15	11	5	11	3	10	26	10	11	14	2	0	0
48	15	11	6	12	3	10	22	11	11	13	1	0	0
49	15	12	5	10	3	10	23	11	14	12	1	0	0
50	15	12	5	11	3	9	23	9	11	12	1	0	0
51	15	12	5	11	3	9	23	11	11	12	1	0	0
52	15	12	5	11	3	10	23	9	11	12	1	0	0
53	15	12	5	11	3	10	23	10	11	12	2	0	0
54	15	12	5	12	3	9	20	10	11	14	1	0	0
55	15	12	5	12	3	9	21	10	11	14	4	0	1
56	15	12	5	12	3	9	21	11	11	15	1	0	0
57	15	12	5	12	3	9	22	10	10	14	1	0	0
58	15	12	5	12	3	9	23	10	11	14	1	0	0
59	15	12	5	12	3	10	23	9	11	12	1	0	0
60	15	12	5	12	3	10	23	10	11	12	1	1	0
61	15	12	5	12	3	11	22	10	11	11	1	0	0
62	15	12	5	12	3	11	22	10	11	13	1	0	0
63	15	12	5	12	3	11	24	11	11	13	1	0	0
64	15	12	5	12	3	11	24	11	11	14	1	0	0
65	15	12	5	12	3	11	25	11	11	14	1	0	0
66	15	12	5	13	3	10	24	11	13	13	1	0	0
67	16	11	4	13	3	9	24	10	11	13	1	0	0
68	16	11	6	11	3	11	25	10	11	14	1	0	0
69	16	12	5	10	3	10	22	11	11	12	1	0	0
70	16	12	5	11	3	10	24	11	11	13	1	0	0
71	16	12	5	11	3	11	23	10	11	13	1	0	0
72	16	12	5	12	3	10	22	10	11	12	1	0	0
73	16	12	5	13	3	9	22	10	10	14	1	0	0
74	16	12	5	13	3	9	23	10	12	13	1	0	0
75	16	12	5	13	3	10	25	11	11	13	1	2	1
76	16	12	5	14	3	10	25	10	11	13	1	0	0
77	16	12	6	12	3	10	25	9	11	13	1	0	0
78	17	12	5	11	3	10	25	11	11	13	1	1	0
79	17	12	5	12	3	9	21	10	11	14	1	0	0
80	17	12	5	12	3	10	24	11	11	13	1	0	0
81	17	12	5	12	3	10	25	11	11	13	1	0	0
82	17	12	5	13	3	10	25	11	11	13	1	0	1

quency of each haplotype in a sample reflects its relative frequency in the population, which may be an overestimate in small, highly diverse samples. Thus the ratio $(1 - dw_{min})/(1 - db_{max})$ or its equivalent expression mw_{max}/mb_{min} is an overestimate of the probability ratio of obtaining a match when sampling a pair of individuals from the same population versus from two different populations. In order to obtain a lower bound for this interpopulation comparison, we assume that at the extreme all unique haplotypes in the sample reflect close to zero relative frequency. Then the within-population matching pair probability mw_{min} equals.

$$\sum_{i=1;x_i>1}^{h} \frac{\binom{x_i}{2}}{\binom{N_x}{2}}$$

The ratio mw_{min}/mb_{min} then yields a lower estimate of how much more probable it is to obtain a match within a population than between two populations.

Results and discussion

In 101 Iraqi Kurds we observed 82 different Y-types, of which 75 Y-types were unique. Seven Y-types were present in more than one person. Such "common" Y-types were shared by 2–4 persons with the exception of a frequent type that was found 10 times (Table 1). From a forensic point of view, it is important to evaluate matching

probabilities rather than genetic distances. We thus also analysed the Y-types shared with other Caucasoid populations (Table 1), i.e. west Germans around Münster (n = 179, 104 Y-types, Nata et al. 1999) and south Turks around Adana (n = 85, 77 Y-types). The haplotype diversities are high within all three populations, evidently because the rapid mutation rate of the STRs quickly eliminates identities: the haplotype diversity in the Kurds is 97.83%, in the Turks 98.55%, and in the Germans 98.06%. The maximum probability db_{max} of obtaining two different Y types when sampling a Kurd and a Turk is 99.41%, and when sampling a Kurd and a German it is 99.93%. In other words, in the maximum estimate it is about 3.7 times more likely to find a match between two Iraqi Kurds than between an Iraqi Kurd and a Turk, and 30.6 times more likely than finding a Kurd-German match. The lower estimates mw_{min}/mb_{min} are 2 and 17.9, respectively. These ethnic differences underline the necessity of microgeographic sampling for forensic applications.

References

Brinkmann B, Klintschar M, Neuhuber F, Hühne J, Rolf B (1998) Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. Am J Hum Genet 62: 1408–1415

Forster P, Kayser M, Meyer E, Roewer L, Pfeiffer K, Benkmann H, Brinkmann B (1998) Phylogenetic resolution of complex mutational features at Y-STR DYS390 in aboriginal Australians and Papuans. Mol Biol Evol 15: 1108–1115

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-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, Honda K, Jobling M, Krawczak M, Leim K, Meuser S, Meyer E, Oesterreich W, Pandya A, Parson W, Penacino G, Perez-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

de Knijff P, 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, Perez-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

Nata M, Brinkmann B, Rolf B (1999) Y-chromosomal STR haplotypes in a population from northwest Germany. Int J Legal Med (in press)

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York, p 178

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

Walsh PS, Metzger DA, Higuchi R (1991) Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. Biotechniques 4: 506–513

Weber JL, Wong C (1993) Mutation of human short tandem repeats. Hum Mol Genet 2: 1123–1128

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