Int J Legal Med (2007) 121:359–363
DOI 10.1007/s00414-006-0123-9

ORIGINAL ARTICLE

Y-chromosomal microsatellite mutation rates in a population sample from northwestern Germany

Carsten Hohoff · Koji Dewa · Ulla Sibbing · Karolin Hoppe · Peter Forster · Bernd Brinkmann

Received: 2 December 2005 / Accepted: 10 August 2006 / Published online: 26 October 2006

Abstract To estimate Y-chromosomal short tandem repeat (Y-STR) mutation rates, 15 loci (i.e., DYS19, DYS389 I/II, DYS390, and DYS393; DYS437, DYS438, DYS439, and DYS385; DYS391, DYS392, YCA II, and DXYS156) were analyzed in a sample of 1,029 father/son pairs from Westphalia, northwestern Germany. Among 15,435 meiotic allele transfers, 32 mutations were observed; thus, the mutation rate across all 15 Y-STR loci was 2.1×10^{-3} per locus (95% C.I.: 1.5–3.0×10^{-3}). With the exception of a three-repeat mutation at DYS385, all remaining mutations were single repeat mutations. Repeat losses were more frequent than gains (20:12), and the mutation rate appeared to increase with age. The Y haplogroups that were detected in the individuals showing a mutation reflect the haplogroup distribution in the Westphalian population. Additionally, the correlation of surnames and haplotypes was tested: Only 49 surnames occurred more than once, and only two men with the same rare surname shared the same haplotype. All other men with identical surnames carried different haplotypes.

Keywords Y-STRs · Mutation rates · Microsatellites

Electronic supplementary material Supplementary material is available in the online version of this article at http://dx.doi.org/10.1007/s00414-006-0123-9 and is accessible for authorized users.

C. Hohoff · K. Dewa · U. Sibbing · K. Hoppe · B. Brinkmann (✉)
Institut für Rechtsmedizin, Universitätsklinikum Münster,
Röntgenstrasse 23,
48149 Münster, Germany
e-mail: brinkma@uni-muenster.de

K. Dewa
Department of Legal Medicine, Graduate School of Medicine and Dental Science, Niigata University,
Niigata, Japan

P. Forster
McDonald Institute for Archaeological Research,
University of Cambridge,
Cambridge, UK

Introduction

The determination of the mutation rate of Y-chromosomal short tandem repeat (Y-STR) loci is very important for the correct interpretation of typing results in paternity and identity testing involving male relatives as well as for evolutionary biology serving, for example, as a molecular clock [1]. Studies of Y-STR mutation rates are few and have so far considered a restricted number of markers (e.g., [2–5], the latter containing the data of [6]). According to the novel recommendations of the DNA Commission of the International Society of Forensic Genetics (ISFG) [7], here we report not only the results of a population study from northwestern Germany but also the mutation data from more than 1,000 father/son pairs with confirmed paternity. We have investigated more than 15,000 meiotic transfers at 13 STRs (corresponding to 15 loci, since YCA II and DYS385 include 2 loci each), ranging from dinucleotide (i.e., YCA II [8]) to pentanucleotide loci (i.e., DYS438 and DXYS156-Y [9]).

Materials and methods

Included in this study were only father/son pairs where paternity had been confirmed with a probability of more than 99.9% (i.e., the probability values were required to exceed 99.9% on the basis of independent autosomal STRs and classical serological markers according to [10], even

Table 1 Mutation rates of the 13 STRs studied

System	This paper				Total*			
	Allelic transfers	Mutations	Mutation rate ($\times10^{-3}$)	95% C.I. ($\times10^{-3}$)	Allelic transfers	Mutations	Mutation rate ($\times10^{-3}$)	95% C.I. ($\times10^{-3}$)
DYS19	1,027	6	5.84	2.7–12.6	8,156	18	2.21	1.4–3.5
DYS389 I	1,027	1	0.97	0.2–5.5	6,479	11	1.69	0.9–3.1
DYS389 II	1,027	5	4.87	2.1–11.4	6,679	19	2.84	1.8–4.4
DYS385	2,055	4	1.95	0.7–4.9	15,830	26	1.64	1.1–2.4
DYS390	1,027	2	1.95	0.5–7.0	7,638	17	2.23	1.4–3.6
DYS391	1,028	2	1.95	0.5–7.0	7,620	25	3.28	2.2–4.8
DYS392	1,026	0	0	0–3.7	7,551	1	0.13	0.3–0.7
DYS393	1,027	1	0.97	0.2–5.5	6,459	5	1.54	0.4–1.8
DYS437	1,027	0	0	0–3.7	3,237	5	1.54	0.7–3.6
DYS438	1,025	1	0.98	0.2–5.5	3,155	2	0.64	0.2–2.3
DYS439	1,027	7	6.82	3.3–14.0	3,132	18	5.75	3.7–9.1
YCA II	2,056	3	1.46	0.5–4.3	2,296	3	1.31	0.5–3.9
DXYS156-Y	1,027	0	0	0–3.7	1,027	0	0	0–3.7
Total	15,435	32	2.1	1.5–3.0	79,259	150	1.89	1.6–2.2

*Includes data from this work and others [1–5, 21–23]

with biostatistical inclusion of the mutation according to [11]). The data for this study were taken from routine parentage cases on caucasoids at the Institute of Legal Medicine Münster (northwest Germany). Surnames were recorded and the majority was found to be of German origin (approximately 90%). The remaining samples carried other surnames, such as Italian, Turkish, or British ones.

Genomic DNA from blood or saliva samples from 1,029 father/son pairs was isolated by the Chelex/proteinase K method. In three different multiplexes (I: DYS19, DYS389 I/II, DYS390, and DYS393; II: DYS437, DYS438, DYS439, and DYS385; III: DYS391, DYS392, YCA II, and DXYS156) as described in [12], 13 Y-STR markers were analyzed. DYS385 was also investigated by nested PCR to make an unambiguous assignment of the alleles to the loci DYS385a and DYS385b [13, 14]. In the event of a mismatch, the analysis was repeated starting with the DNA extraction and including a sequence analysis on both strands using BigDye terminator chemistry (v.2, ABI, Darmstadt, Germany).

Results and discussion

After investigating 1,029 father/son pairs (five fathers each had two sons in this study) from Westphalia with confirmed paternity covering a total of 15,435 allelic transfers, we found 32 mutations. No more than one mutation occurred within any single father–son germline transmission. Repeat gains were observed in 12 cases and losses in 20 cases. With the exception of a three-repeat mutation at DYS385, all remaining mutations involved single repeats.

A comparison of our observed mutation rates with other published mutation rates (e.g., [3, 5, 15], cumulated in Table 1) shows similar values with some differences in both directions.

When addressing the influence of the repeat type and the mutation rate, the dinucleotides and pentanucleotides were associated with lower mutation rates than the tetranucleotide repeats (Table 2), while until now we have not observed a mutation at the only trinucleotide repeat marker in the study, DYS392. No detailed trend should be extrapolated from these observations without first obtaining

Table 2 Comparison of mutation rates and repeat types

System	Allelic transfers	Mutation	Mean mutation rate ($\times10^{-3}$)	95% C.I. ($\times10^{-3}$)
2mer	2,056	3	1.46	0.5–4.3
3mer	1,026	0	0	0–3.7
4mer	10,272	28	2.73	1.9–3.9
5mer	2,055	1	0.5	0.1–2.8
all 13 STRs	15,435	32	2.1	1.5–3.0

Table 3 Correlation of mutation rate and age at the time of conception

Age at conception (years)	Total transfers	Transfers with mutation	Mutation rate ($\times10^{-3}$)	95% C.I. ($\times10^{-3}$)	
<24	5,115	8	1.56	0.81	3.08
25–29	3,915	7	1.8	0.88	3.68
30–39	3,585	8	2.23	1.15	4.39
40–64	1,290	4	3.10	1.26	7.91

Table 4 Sequences of the mutated alleles

System	Father	Son	Repeat gain/loss	Sequence
DYS439	12	11	Loss	...$(GATA)_{12\rightarrow11}$...
	14	13	Loss	...$(GATA)_{14\rightarrow13}$...
	12	11	Loss	...$(GATA)_{12\rightarrow11}$...
	12	11	Loss	...$(GATA)_{12\rightarrow11}$...
	13	14	Gain	...$(GATA)_{13\rightarrow14}$...
	11	10	Loss	...$(GATA)_{11\rightarrow10}$...
	12	13	Gain	...$(GATA)_{12\rightarrow13}$...
DYS19	14	15	Gain	...$(GATA)_3GGTA(GATA)_{11\rightarrow12}$...
	17	16	Loss	...$(GATA)_3GGTA(GATA)_{14\rightarrow13}$...
	18	17	Loss	...$(GATA)_3GGTA(GATA)_{15\rightarrow14}$...
	15	16	Gain	...$(GATA)_3GGTA(GATA)_{12\rightarrow13}$...
	18	17	Loss	...$(GATA)_3GGTA(GATA)_{15\rightarrow14}$...
	16	15	Loss	...$(GATA)_3GGTA(GATA)_{13\rightarrow12}$...
DYS389 II	32	31	Loss	...$(CTGT)_5(CTAT)_{14\rightarrow13}$...$(CTGT)_3(CTAT)_{10}$...
	30	29	Loss	...$(CTGT)_5(CTAT)_{12\rightarrow11}$...$(CTGT)_3(CTAT)_{10}$...
	29	30	Gain	...$(CTGT)_5(CTAT)_{11\rightarrow12}$...$(CTGT)_3(CTAT)_{10}$...
	30	31	Loss	...$(CTGT)_5(CTAT)_{12\rightarrow13}$...$(CTGT)_3(CTAT)_{10}$...
	29	30	Loss	...$(CTGT)_5(CTAT)_{11\rightarrow12}$...$(CTGT)_3(CTAT)_{10}$...
DYS385	14–14	14–15	Gain	...$(GAAA)_{14\rightarrow15}$...
	14–15	14–12	Loss	...$(GAAA)_{15\rightarrow12}$...
	14–17	14–16	Loss	...$(GAAA)_{17\rightarrow16}$...
	13–14	14–14	Gain	...$(GAAA)_{13\rightarrow14}$...
YCA II	19–23	19–22	Loss	...$(CA)_{23\rightarrow22}$...
	21	21–22	Gain	...$(CA)_{21\rightarrow22}$...
	19–21	19–20	Loss	...$(CA)_{21\rightarrow20}$...
DYS390	26	25	Loss	...$(CTGT)_8(CTAT)_{13\rightarrow12}$...$(CTGT)_1(CTAT)_4$...
	24	25	Gain	...$(CTGT)_8(CTAT)_{11\rightarrow12}$...$(CTGT)_1(CTAT)_4$...
DYS391	10	11	Gain	...$(CTAT)_{10\rightarrow11}$...
	11	10	Loss	...$(CTAT)_{11\rightarrow10}$...
DYS389 I	13	12	Loss	...$(CTGT)_6(CTAT)_{11}$...$(CTGT)_3(CTAT)_{10\rightarrow9}$...
DYS393	13	14	Gain	...$(GATA)_{13\rightarrow14}$...
DYS438	10	11	Gain	...$(TTTTC)_{10\rightarrow11}$...

more mutational data on non-tetrameric loci, as the tetramers represent 2/3 of all transfers in our data.

The relationship between the age at conception and the number of observed mutations (Table 3) revealed the following trend: With increasing paternal age, the mutation rate increases, that is, the mutation rate for older fathers (>40 years) is twice as high as for younger fathers (<24 years). This difference is not significant ($p=0.28$ according to a Fischer exact test).

All mutations were verified as being located in the repetitive regions by the sequencing of the affected alleles (Table 4). To estimate mutation rates for uninterrupted GATA repeats (equivalent to TAGA/TCTA), data for the tetranucleotide repeat systems (DYS19, DYS389I/II, DYS390, DYS391, DYS393, and DYS439) were pooled according to [5], that is, into short (<11 uninterrupted, homogeneous repeats) and long classes (≥12 repeats). The mutation rate in the long class was more than twofold higher than in the short class (Table 5). Interestingly, three times more repeat losses than gains were found in alleles from the long class, while, in contrast, alleles in the short class did not show a predominant mutational direction. It should be noted that the classification is somewhat arbitrary, since such a distinct threshold can hardly be

Table 5 Mutation rates of GATA repeats

GATA repeats	Transfers	Mutations	Loss	Gain	Mutation rate ($\times10^{-3}$)	95% C.I. ($\times10^{-3}$)
7–11	3,053	7	4	3	2.3	1.1–4.7
12–16	3,068	17	13	4	5.5	3.3–8.9

Table 6 Y-haplogroup distribution of the mutation cases

Haplogroup	Occurrence	Percent
R1b*	11	34.4
I*	9	28.1
R1a1*	4	12.5
E3* (xE3a*, E3b*)	2	6.3
E3a*	2	6.3
G*	1	3.1
N3*	1	3.1
R1*	1	3.1
K*	1	3.1

*Includes data from this work and others [1–5, 21–23]

established, and the mutation rate is expected to show a relationship to the repeat length that may be exponential [16].

Since different mutation rates can be expected in different haplogroup backgrounds, we typed 29 Y-chromosomal binary polymorphisms according to [17]. The two most frequent haplogroups in the Westphalian population [18] were also the most frequent haplogroups in the mutated samples (Table 6); thus we could not detect a haplogroup with particularly high or low mutation rate.

In the 1,024 independent father/son pairs, only 49 surnames occurred more than once, 33 of them twice, 10 names occurred three times, 4 names occurred four times and 2 names occurred five times (Table 7). Thus, the surname diversity according to [19] was 99.89%. Only two men with the same surname shared the same haplotype. This particular surname is rare in our population sample that is in agreement with the finding of King et al. [20], who found an increased likelihood for matching Y haplotypes in rare surnames but little likelihood among common surnames. All other men with identical surnames carried different haplotypes.

Table 7 Occurrence of the fathers' surnames

Surname	Occurrence
Beckmann, Bürger, Brinkmann, Budde, Busch, Fiedler, Fischer, Fritz, Hansen, Heitmann, Joldzic, Kettler, Lehmann, Meier, Merten, Neuhaus, Neumann, Nickel, Niehues, Ostmann, Peters, Pick, Saric, Schäfer, Schwede, Stahl, Surdyk, Tappe, Terheyden, Tiemann, Uhlenbusch, Will, Winter	2
Beyer, Brockmann, Drees, Jansen, Kaminski, Koch, Lange, Meyer, Schmitz, Schneider	3
Brandt, Hoffmann, Möller, Müller	4
Schmidt, Wolf	5

Table 8 Frequency of Y-STR haplotypes in the 1,024 fathers

Number of haplotypes	Occurence
772	1
49	2
14	3
7	4
2	5
3	6
1	7
1	8
1	11
1	13
1	17

On some Y chromosomes, an STR locus can occur in duplicate, and if the repeats at both homologous loci differ in length, electropherograms show biallelic patterns. In the 1,024 paternal haplotypes (see Electronic supplementary material), we observed six men and their sons with multiple alleles: at DYS19: 14,15 (twice); 15,16 (twice); and 16,17 (once) and at DYS385: 11-13-14. Due to pronounced peak height differences, the latter can be explained by the following mechanism: after a duplication of a Y-chromosomal segment (at least 40 kb long) that harbors the complete DYS385 locus (e.g., DYS385*11–13 and DYS385*11–**13**), a subsequent mutation as DYS385*11–13 and DYS385*11–**14** at one of the loci has taken place at two different times in the lineage that these males belong to.

In the 1,024 different fathers, we have observed 852 different and 772 unique haplotypes (Table 8). Thus, the haplotype diversity according to [19] was 99.8%. The most frequent haplotype occurred 17 times in our population sample and was found in the Y Chromosome Haplotype Reference Database (http://www.yhrd.org) 26 times and only in European populations (sample size, 4,445 persons).

It is interesting to note that in our study, the surname diversity was higher than the genetic diversity based on Y haplotypes that are comprised of 15 Y-STR loci.

This study has increased the number of well-documented Y-STR mutations and the respective allele transmissions, which were previously published, by approximately 25%, and these pooled data now permit a reliable estimate of, for example, locus-specific mutation rates. Nevertheless, there are still some topics (e.g., the influences of paternal age and repeat type on the mutation rate) that demand the publication of additional Y-STR mutation data, as recommended by the DNA Commission of the ISFG [7], to draw statistically significant conclusions on these issues.

Acknowledgment The authors thank M. Heinrich (present address: Institute of Legal Medicine, Freiburg, Germany) for Y-SNP typing.

References

1. 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 63:1862–1871
2. 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 66:1580–1588
3. Dupuy BM, Stenersen M, Egeland T, Olaisen B (2004) Y-chromosomal microsatellite mutation rates: differences in mutation rate between and within loci. Human Mutat 23:117–124
4. Kurihara R, Yamamoto T, Uchihi R et al (2004) Mutations in 14 Y-STR loci among Japanese father–son haplotypes. Int J Leg Med 118:125–131
5. Gusmao L, Sanchez-Diz P, Calafell F et al (2005) Mutation rates at Y chromosome specific microsatellites. Human Mutat 26:520–528
6. de Souza Goes AC, de Carvalho EF, Gomes I et al (2005) Population and mutation analysis of 17 Y-STR loci from Rio de Janeiro (Brazil). Int J Leg Med 119:70–76
7. Gusmao 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 Leg Med 120:191–200
8. Schmidt U, Meier N, Lutz S (2003) Y-chromosomal STR haplotypes in a population sample from southwest Germany (Freiburg area). Int J Leg Med 117:211–217
9. Cali F, Forster P, Kersting C et al (2002) DXYS156: a multi-purpose short tandem repeat locus for determination of sex, paternal and maternal geographic origins and DNA fingerprinting. Int J Leg Med 116:133–138
10. Brinkmann B, Pfeiffer H, Schürenkamp M, Hohoff C (2001) The evidential value of STRs. An analysis of exclusion cases. Int J Leg Med 114:173–177
11. Rolf B, Keil W, Brinkmann B, Roewer L, Fimmers R (2001) Paternity testing using Y-STR haplotypes: assigning a probability for paternity in cases of mutations. Int J Leg Med 115:12–15
12. Dewa K, Tuyen NQ, Rand S, Hohoff C, Brinkmann B (2003) 13 Y-chromosomal STRs in a Vietnamese population. Prog Forensic Genet 9:315–318
13. Kittler R, Erler A, Brauer S, Stoneking M, Kayser M (2003) Apparent intrachromosomal exchange on the human Y chromosome explained by population history. Eur J Hum Genet 11:304–314
14. Niederstätter H, Berger B, Oberacher H, Brandstätter A, Huber CG, Parson W (2005) Separate analysis of DYS385a and b versus conventional DYS385 typing: is there forensic relevance? Int J Leg Med 119:1–9
15. 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 10:629–633
16. 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
17. Brion M, Sanchez JJ, Balogh K et al (2005) Introduction of an single nucleodite polymorphism-based "Major Y-chromosome haplogroup typing kit" suitable for predicting the geographical origin of male lineages. Electrophoresis 26:4411–4420
18. Kayser M, Lao O, Anslinger K et al (2005) Significant genetic differentiation between Poland and Germany follows present-day political borders, as revealed by Y-chromosome analysis. Hum Genet 117:428–443
19. Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York, p 178
20. King TE, Ballereau SJ, Schürer KE, Jobling MA (2006) Genetic signatures of coancestry within surnames. Curr Biol 16:384–388
21. Heyer E, Puymirat J, Dieltjes P, Bakker E, de Knijff P (1997) Estimating Y chromosome specific microsatellite mutation frequencies using deep rooting pedigrees. Hum Mol Genet 6:799–803
22. 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 150:1–15
23. 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 155:65–70