Int J Legal Med (2001) 115 : 12–15

ORIGINAL ARTICLE

B. Rolf · W. Keil · B. Brinkmann · L. Roewer
R. Fimmers

Paternity testing using Y-STR haplotypes: assigning a probability for paternity in cases of mutations

Received: 10 April 2000 / Accepted: 2 December 2000

Abstract In parentage testing with male children, Y-chromosomal STR evidence is gaining more and more importance. In some cases, multilocus haplotypes of related persons can differ at a single locus due to a mutation. In this work, a likelihood approach is presented for the calculation of a probability for paternity under consideration of a single mutation event on the Y-chromosome. The new methodology is applied to two case examples.

Keywords Y-STR · Mutation · Likelihood · Paternity probability

Introduction

In parentage testing with male children and in identity testing with male relatives, Y-chromosomal STR evidence is gaining more and more importance (Chakraborty 1985; Santos et al. 1993; Kishida et al. 1996; Foster et al. 1998; Kayser et al. 1998). Due to the haploid nature of the Y-chromosome, the calculation of likelihood ratios is based on the estimation of the frequency of the observed haplotype in an appropriate reference database (de Knijff et al. 1997). To this end an increasing number of medium-sized European regional databases have been published (Caglia et al. 1998; Lessig and Edelmann 1998; Rossi et al. 1998; Pestoni et al. 1998; Brinkmann et al. 1999; Gene et al. 1999; Nata et al. 1999). More realistic frequency estimation for a large number of European Y-STR standard haplotypes (9 loci minimum) has been made possible by the recent introduction of a large Caucasian Y-STR haplotype reference database in the WorldWideWeb (http://ystr.charite.de). This database (n = 3,589 in April 2000) combines representative population samples from 26 European regions and actually provides frequency estimates for 1,964 different 9-locus minimal haplotypes (Roewer et al. 2000)

In some cases multilocus haplotypes of related persons could differ at a single locus due to mutation. The mutation rates of the Y-STR loci are similar to their autosomal counterparts (Weber and Wong 1993; Heyer et al. 1997; Brinkmann et al. 1998; Jobling et al. 1999; Kayser et al. 2000). In this work, we develop a general likelihood approach for the interpretation of those cases under consideration of a single mutation event. The new methodology is applied to two case examples.

Theory

Consider a paternity case where only an alleged father and a male child are typed with a set of Y-chromosomal markers and the observed haplotype shows a single exclusion/mutation at one locus and the two alleles differ by a single repeat unit. For the evaluation of the results, two hypotheses have to be compared, i.e. H_0 the man and child are unrelated and H_1 the alleged father is the true father. The respective likelihood under non-paternity is $Y = P(E|H_0)$:

$$Y = f(\text{af}) \times f(\text{child})$$

with f(af) being the frequency of the haplotype of the alleged father and f(child) the frequency of the haplotype of the child. The other hypothesis is that the alleged father is

B. Rolf (✉) · W. Keil
Institute of Legal Medicine, Ludwigs-Maximilians-Universität,
Frauenlobstrasse 7a, 80337 München, Germany
e-mail: burkhard.rolf@rechts.med.uni-muenchen.de

B. Brinkmann
Institute of Legal Medicine,
Westfälische-Wilhelms-Universität Münster,
Von-Esmarch-Strasse 62, 48149 Münster, Germany

L. Roewer
Institute of Legal Medicine, Humboldt-Universität,
Hannoversche Strasse 6, 10115 Berlin, Germany

R. Fimmers
Institut für Medizinischen Biometrie,
Informatik und Epidemiologie,
Rheinische Friedrich-Wilhelms-Universität,
Sigmund-Freud-Strasse 25, 53105 Bonn, Germany

the true father, and the respective likelihood under paternity is $X = P(E|H_1)$

$$X = f(\text{af}) \times \tilde{\mu}$$

where $\tilde{\mu}$ is the rate at which the paternal allele mutates to the offspring allele. Unfortunately, precise estimates of such allele-specific mutation rates are difficult to obtain. For microsatellites, however, most mutations are expansions or contractions by a single repeat unit so that, by first approximation, we may assume

$$\tilde{\mu} \approx \mu \times \frac{1}{2}$$

where μ is the overall mutation rate of the locus showing the single exclusion due to a one repeat difference. The resulting likelihood ratio is:

$$\frac{Y}{X} = \frac{f(\text{child})}{\frac{1}{2} \times \mu} \quad (1)$$

In cases without mutation, instead of $\mu/2$, the transmission probability 1 is used, and the resulting Y/X value for paternity is only the frequency of the haplotype. This formula ignores the fact that the other loci in the haplotype could mutate as well; however, since the mutation rate is small and the vast majority of loci always segregate unchanged, this error can be considered negligible.

The application of the formula to deficiency cases with pedigrees spanning several generations requires some modifications. The hypotheses are now: H_0, the two males are unrelated; and H_1, the alleged relative is a true relative. The likelihood under non-relationship in such cases is again $Y = P(E|H_0)$:

$$Y = f(\text{pmr}) \times f(\text{child})$$

with f(pmr) being the frequency of the haplotype of the putative male relative. In cases without mutation, the likelihood under relationship is identical to the "one generation" case and the resulting likelihood ratio in these cases without mutation is simply the frequency of the observed haplotype. For this "no mutation" likelihood, a transmission probability of 1 is assumed. The true figure will be smaller, however, the error is small if the number of generations is not to large (< 10).

If the haplotype of the putative male relative and child differ only at one locus by a single-step difference, the following considerations have to be made:

The probability for the segregation of a haplotype showing one mutation over k generations reads:

$$p = \binom{k}{1} \tilde{\mu} (1 - \tilde{\mu})^{k-1}$$

and the likelihood under relationship is $X = P(E|H_1)$:

$$X = f(\text{pmr}) \times k\tilde{\mu} (1 - \tilde{\mu})^{k-1} \approx \frac{1}{2} k\mu$$

assuming again that $\mu/2$ is the rate at which a mutation event either decreases or increases the size of the paternal allele. The resulting likelihood ratio is:

$$\frac{Y}{X} = \frac{f(\text{child})}{k\mu (1 - \mu)^{k-1} \times \frac{1}{2}} \approx \frac{f(\text{child})}{\frac{1}{2} k\mu} \quad (2)$$

This formula ignores the individual mutation rates of the loci showing no mutation and we again assume a transmission probability of 1 for the other loci. However, the resulting figure is a good estimate for the likelihood quotient in deficiency case spanning several generations.

Case examples

Case 1

We were commissioned to find out whether three males, all sons of a deceased woman, were full brothers. For two of the children the paternity of the deceased husband was not in question, but for brother 1, this was doubtful. The results of the Y-STR typing are shown in Table 1. Except for a mutation/exclusion pattern at the DYS390 locus, the haplotypes were identical. Assuming that the haplotype of brothers 2 and 3 was also present in the father, these results can be used to calculate a paternity probability according to Eq. 1: the haplotype of brother 1 was found 97 times in 3,589 Caucasian haplotypes with the same 9-locus format. The mutation rate for DYS390 is 8.58×10^{-3} (Kayser et al. 2000). The paternity index, based on the Y-STR results, was 0.159 and the paternity probability for the deceased husband was 13.7% (based on the assumption of 0.5 prior probability). In this case, the paternity could be successfully established by additional typing of 10 autosomal STR loci in the three brothers and in a sister (data not shown).

Table 1 Y-STR haplotypes of the three males involved in case 1 (the mutation is indicated in bold type)

DNA locus	Brother 1	Brother 2	Brother 3
DYS19	14	14	14
DYS385 1/2	11/14	11/14	11/14
DYS389-1	10	10	10
DYS389-2	26	26	26
DYS390	**24**	**23**	**23**
DYS391	11	11	11
DYS392	13	13	13
DYS393	13	13	13

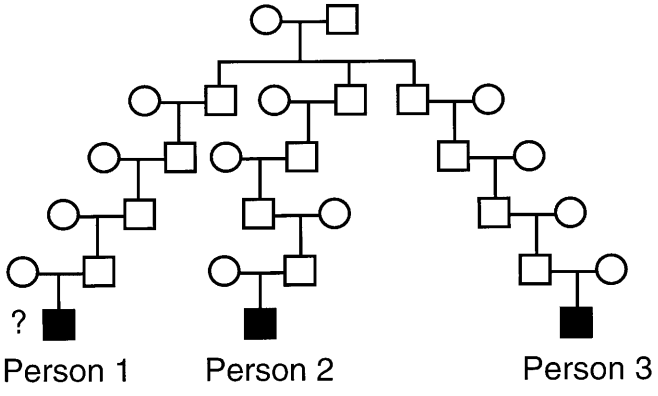

Fig. 1 Pedigree of the family involved in case 2

Table 2 Y-STR haplotypes of the three males involved in case 2 (the mutation is indicated in bold type)

DNA locus	Person 1	Person 2	Person 3
DYS19	**15**	**14**	**14**
DYS385 1/2	11/15	11/15	11/15
DYS389-1	10	10	10
DYS389-2	26	26	26
DYS390	24	24	24
DYS391	11	11	11
DYS392	13	13	13
DYS393	13	13	13

Case 2

This case was a paternity case in an aristocratic family with ancestral records back to the seventeenth century (Fig. 1). The paternity at the end of the left branch in the pedigree in Fig. 1 was questioned and the alleged father had died 30 years previously. The results of the Y-STR typing of the three persons (filled squares in Fig. 1) are shown in Table 2. Except for a mutation/exclusion pattern at the DYS19 locus, the haplotypes of the three persons were identical. Assuming that the haplotype of persons 2 and 3 was present in the common male ancestor too, there would be five generations between person 1 and the male ancestor. The frequency of the haplotype of person 1 in the Caucasian Y-STR reference database is 2 in 3,589, the mutation rate for DYS19 is 2.01×10^{-3} (Kayser et al. 2000). Application of Eq. 2 results in a paternity index of 9.02 and a paternity probability of 90.0% (based on the assumption of 0.5 prior probability). This case was solved after a half-brother of the deceased alleged father was found and could be typed. This man showed the same "mutated" haplotype present in person 1.

Discussion

It is obvious from the presented equations as well as from the case examples, that the size of the reference database is essential for the evaluation of the Y-chromosome evidence. Especially in cases of mutation, where the hypothesis "unrelated" becomes more probable, a rare haplotype in a large database can result in a likelihood quotient that favours the hypothesis "related". The database used here for the evaluation of the two cases is by far the largest Y-STR database available, maintained and quality controlled at the Institute of Legal Medicine of Humboldt University in Berlin and available on the internet for use in interactive haplotype queries (http://ystr.charite.de). This database uses a highly discriminative haplotype format comprising at least 9 Y-STR loci (minimal haplotype; Pascali et al. 1999). The haplotype diversity has been calculated previously as $h = 0.993$ for $n = 2{,}439$, 7-locus haplotypes excluding the duplicated locus DYS385 (Roewer et al. 2000).

A precise knowledge of the mutation rates of the loci involved is essential for the interpretation of these kinds of cases. In a recent paper (Kayser et al. 2000), the mutation rates for the loci of the Yh1-format are compiled. For the two loci showing a single mutation/exclusion pattern in our case examples, the figures are 2.01×10^{-3} and 8.58×10^{-3} for DYS19 and DYS390, respectively. Xu et al. (2000) examined 236 mutations at 122 autosomal tetranucleotide repeat markers. They found that the vast majority of mutations were one-step mutations and that the overall rate of expansion mutations does not differ from that of contractions. Although larger alleles seem to mutate more often than smaller alleles (Brinkmann et al. 1998), and the direction of mutations seems to be dependent on allele length (Xu et al. 2000), the present knowledge of Y-STR mutations does not allow allele-specific mutation rates to be defined. Therefore, we suggest that the locus-specific mutation rate should be used for all alleles combined and half of this figure for mutations that either increase or decrease the size of an allele by one step. The application of the formula to cases showing more than one repeat unit difference between the two haplotypes would require estimation for "double-step" and "three-step", etc. mutations. However, the currently available data on Y-STR mutations does not allow a clearer definition of these figures; therefore, further data on Y-STR mutations would be necessary.

The formulae derived here could, in principle, be applied to mtDNA evidence as well. Large databases for mtDNA sequences are available (Budowle et al. 1999; Pfeiffer et al. 2001, D. Krause, Institut für Rechtsmedizin Magdeburg, http://www.d-loop-base.de, personal communication). However, in contrast to Y-STR haplotypes, a typical mtDNA haplotype consists of 400–800 individual loci, i.e. base pairs. Although still a matter of debate (Parsons et al. 1997; Jazin et al. 1998), the mutation rates of the individual bases in the hypervariable regions are much lower than the respective Y-STR mutation rate. Furthermore, intermediate, i.e. heteroplasmic haplotypes complicate the estimation of the mutation rate.

In summary, the new method presented enables the biostatistical treatment of single-step mutation events in Y-STR haplotypes found during paternity testing. Using the new formula, the evidence derived from a mutated haplotype could still favour paternity versus non-paternity in some cases.

Acknowledgement The authors would like to thank Charles Brenner and Michael Krawczak for helpful comments.

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

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

Budowle B, Wilson MR, DiZinno JA, Stauffer C, Fasano MA, Holland MM, Monson KL (1999) Mitochondrial DNA regions HVI and HVII population data. Forensic Sci Int 103:23–35

Caglia 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

Chakraborty R (1985) Paternity testing with genetic markers: are Y-linked genes more efficient than autosomal ones? Am J Med Genet 21:297–305

De Knijff P, Kayser M, Caglia 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, Roewer L (1997) Chromosome Y microsatellites: population genetic and evolutionary aspects. Int J Legal Med 110:134–149

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

Gene M, Borrego N, Xifro A, Pique E, Moreno P, Huguet E (1999) Haplotype frequencies of eight Y-chromosome STR loci in Barcelona (North-East Spain). Int J Legal Med 112:403–405

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

Jazin E, Soodyall H, Jalonen P, Lindholm E, Stoneking M, Gyllensten U (1998) Mitochondrial mutation rate revisited: hot spots and polymorphism. Nat Genet 18:109–110

Jobling MA, Heyer E, Dieltjes P, de Knijff P (1999) Y-chromosome-specific microsatellite mutation rates re-examined using a minisatellite, MSY1. Hum Mol Genet 8:2117–2120

Kayser M, Krüger C, Nagy M, Geserick G, de Knijff P, Roewer L (1998) Y-chromosomal DNA analysis in paternity testing: experiences and recommendations. In: Olaisen B, Brinkmann B, Lincoln PJ (eds) Progress in forensic genetics 7. Elsevier Science, Amsterdam, pp 494–496

Kayser M, Roewer L, Hedman M, Henke L, Henke J, Brauer S, Krüger C, Krawczak M, Nagy M, Dobosz T, Szibor R, de Knijff P, Stoneking M, Sajantila A (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

Kishida T, Wang W, Fukuda M, Tamaki Y (1996) Two deficiency cases of disputed paternity: sibling or half-sibling? Nippon Hoigaku Zasshi 50:78–81

Lessig R, Edelmann J (1998) Y chromosome polymorphisms and haplotypes in West Saxonia (Germany). Int J Legal Med 111: 215–218

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

Parsons TJ, Muniec DS, Sullivan K, Woodyatt N, Alliston-Greiner R, Wilson MR, Berry DL, Holland KA, Weedn VW, Gill P, Holland MM (1997) A high observed substitution rate in the human mitochondrial DNA control region. Nat Genet 15:363–368

Pascali VL, Dobosz M, Brinkmann B (1999) Coordinating Y-chromosomal STR research for the Courts. Int J Legal Med 112:1

Pestoni C, Cal ML, Lareu MV, Rodriguez-Calvo MS, Carracedo A (1998) Y chromosome STR haplotypes: genetic and sequencing data of the Galician population (NW Spain). Int J Legal Med 112:15–21

Pfeiffer H, Forster P, Ortmann C, Brinkmann B (2001) The results of an mtDNA study of 1200 inhabitants of a German village in comparison to other Caucasian databases and its relevance for forensic casework. Int J Leg Med 114:169–172

Roewer L, Kayser M, Anslinger K, Augustin C, Cagliá A, Corach D, Füredi S, Geserick G, Henke L, Hidding M, Kärgel HJ, de Knijff P, Lessig R, Pascali V, Parson W, Prinz M, Rolf B, Schmitt C, Schneider P, Szibor R, Teifel-Greding J, Krawczak M (2000) Caucasian Y-STR haplotype reference database for forensic application. In: Sensabaugh GF, Lincoln PJ, Olaisen B (eds) Progress in forensic genetics 8. Elsevier Science, Amsterdam, pp 613–615

Roewer L, Kayser M, de Knijff P, Anslinger K, Betz A, Cagliá A, Corach D, Füredi S, Henke L, Hidding M, Kärgel HJ, Lessig R, Nagy M, Pascali VL, Parson W, Rolf B, Schmitt C, Szibor R, Teifel-Greding J, Krawczak M (2000) A new method for the evaluation of matches in non-recombining genomes: application to Y-chromosomal short tandem repeat (STR) haplotypes in European males. Forensic Sci Int 114:31–43

Rossi E, Rolf B, Schürenkamp M, Brinkmann B (1998) Y-chromosome STR haplotypes in an Italian population sample. Int J Legal Med 112:78–81

Santos FR, Epplen JT, Pena SDJ (1993) Testing deficiency paternity cases with a Y-linked tetranucleotide repeat polymorphism. In: Pena SDJ, Chakraborty R, Epplen JT, Jeffreys AJ (eds) DNA fingerprinting: the state of the science. Birkhäuser, Basel, pp 221–230

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

Xu X, Peng M, Fang Z (2000) The direction of microsatellite mutations is dependent upon allele length. Nat Genet 24:396–399