Int J Legal Med (2007) 121: 24–27
DOI 10.1007/s00414-005-0075-5

ORIGINAL ARTICLE

Heike Rodig · Manja Grum · Hans-Dieter Grimmecke

Population study and evaluation of 20 Y-chromosome STR loci in Germans

Received: 12 August 2005 / Accepted: 6 December 2005 / Published online: 20 May 2006

Abstract The nine European minimal haplotype (EMH) loci, the two SWGDAM loci and five further single-copy Y-chromosomal short tandem repeats (Y-STRs) DYS446, DYS447, DYS448, DYS449, DYS463, and the multicopy loci DYS464 were evaluated in the German population groups Dresden, Hamburg, Rostock, Munich, and the Sorbs who are a Slavic-speaking minority in Lusatia. Highest gene diversities in all populations were shown for DYS464, DYS385, and DYS449 (D=0.8559–0.9486). The haplotype diversity for the European minimal haplotype loci ranged between 0.9852 for Sorbs and 0.9983 for the Hamburg population showing that there is a significant portion of haplotypes which could not be resolved. Advanced typing using DYS446, DYS447, DYS448, DYS449, DYS463, and DYS464 discriminated all non-related individuals of the Dresden, Hamburg, and Rostock populations. Evaluation of the Y-STRs was accomplished by sequence analysis of all allelic fragments of the allelic ladders and microvariant alleles of DYS385 and the determination of the amounts of stutter products of the loci.

Keywords Y-STR · European minimal haplotype · SWGDAM · Population study · Microvariants · DYS385 · DYS446 · DYS447 · DYS448 · DYS449 · DYS463 · DYS464

Electronic Supplementary Material Supplementary material is available for this article at http://dx.doi.org/10.1007/s00414-005-0075-5

H. Rodig (✉) · M. Grum · H.-D. Grimmecke
Biotype AG,
Moritzburger Weg 67,
01109 Dresden, Germany
e-mail: h.rodig@biotype.de
Tel.: +49-351-8838400
Fax: +49-351-8838403

Introduction

The analysis of Y-chromosomal short tandem repeats (STRs) is a powerful tool for analyzing mixed forensic stains and for paternity testing. In 1997, an evaluation of Y-chromosomal STRs was performed in a multicenter study [1]. It was shown that the analysis of a core set of Y-STRs can discriminate between most of the male individuals in a population. The analysis of this core set, the European minimal haplotype (EMH) standard, consisting of DYS19, DYS385, DYS389-I and -II, DYS390, DYS391, DYS392, and DYS393, was recommended for court use [2]. The haplotype diversity using the European minimal haplotype loci reaches between <0.95 for Albania or Finland and 1.00 for Vienna [3]. Hence, there are significant portions of haplotypes in several populations which cannot be resolved. Therefore, the European minimal haplotype loci, the SWGDAM loci, and a set of further Y-chromosomal STR loci including the multicopy marker DYS464 selected from the publication of Redd et al. [4] were evaluated in five German population groups.

Materials and methods

DNA samples

Blood samples and buccal swab samples from 235 healthy, unrelated male German volunteers from Dresden (n=89), Rostock (n=38), Hamburg (n=49), Munich (n=30), and Sorbs (n=29) were collected with written informed consent. DNA from the Dresden inhabitants and Sorbs was extracted from buccal swabs by NucleoSpin Tissue kit (Macherey-Nagel, Düren, Germany). DNA from the Hamburg inhabitants was extracted from EDTA blood with proteinase K digest, NaCl desalting, and final ethanol precipitation. DNA from Munich inhabitants was extracted from buccal swabs with the chelex protocol. The DNA from Rostock inhabitants was extracted from blood samples and buccal swabs with DNA IQ system (Promega GmbH, Mannheim, Germany) and NucleoMag 96 Blood

kit (Macherey-Nagel). Cell line DNA samples NA9948 and NA3657 were purchased from Coriell Cell Repositories (Camden, NJ, USA).

PCR amplification

Three multiplex polymerase chain reaction (PCR) systems were used to amplify the Y-STR loci. Mentype Argus Y-MH PCR amplification kit (Biotype AG, Dresden, Germany) was used to amplify the European minimal haplotype loci DYS19, DYS385, DYS389-I and -II, DYS390, DYS391, DYS392, and DYS393. A second experimental screening multiplex system (RAR), which is not yet commercially available, was used to amplify the STR loci DYS446, DYS447, DYS448, DYS449, DYS463, DYS464, and DYS385 for confirmation of both test systems. For DYS447 (6-FAM), DYS449 (HEX), DYS463 (HEX), and DYS464 (6-FAM), the published primer sequences [4] were chosen. However, new primers were designed for DYS385, DYS446, and DYS448 to optimize PCR and integrate the STR loci in the RAR multiplex as follows:

DYS385:
5′-(6-FAM)-GAG CTA GAC ACC ATG CCA AAC AAC-3′,
5′- GGG CTG CTG ACC AGA TTT CTT T-3′
DYS446:
5′-(HEX)-TAT TTT CAG TCT TGT CCT GTC-3′,
5′-CTG TCT GAA GAG AGG AGA AGA GAA-3′
DYS448:
5′-(HEX)-TGT CAA AGA GCT TCA ATG GAG A-3′,
5′-TCT TCC TTA ACG TGA ATT TCC TC-3′

The third multiplex was used to amplify the SWGDAM loci DYS438 and DYS439 in one reaction. Primer sequences for DYS439 (6-FAM) were as published [5] and for DYS438 as follows:

DYS438:
5′-(6-FAM)-TGG GGA ATA GTT GAA CGG TAA-3′,
5′-GGA GGT TGT GGT GAG TCG AG-3′

PCR amplification for the three multiplexes was performed according to the manual of Mentype Argus Y-MH PCR amplification kit. The PCR products were analyzed by capillary electrophoresis using an ABI PRISM 310 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA) according to the manual of Mentype Argus Y-MH PCR amplification kit. The alleles were typed automatically using the Genotyper software (Applied Biosystems) and the appropriate allelic ladders.

Sequence analysis and allelic ladders

All allelic fragments of the allelic ladders and microvariants of DYS385 were sequenced. For this purpose, the fragments were amplified with non-labeled primers as described. Amplicons were cloned using the TOPO TA Cloning kit (Invitrogen, Karlsruhe, Germany). Sequencing of the DNA inserts was performed using the Big Dye Terminator Cycle Sequencing kit (Applied Biosystems). The allele nomenclature follows the guidelines of the DNA Commission of the International Society for Forensic Genetics [6–8]. The alleles of the multicopy marker DYS464 were called using the conservative typing method as published [9, 10].

The allelic ladder of the European minimal haplotype loci is provided in the Mentype Argus Y-MH PCR amplification kit. An allelic ladder for the screening multiplex RAR was created using the most frequent alleles amplified with non-labeled primers and cloned as described. For the allelic ladder, plasmids were isolated, alleles were amplified with labeled primers, and PCR products of DYS446 (alleles 8, 10–13, 15), DYS447 (alleles 22–24, 26), DYS448 (alleles 19–22), DYS449 (alleles 28–33), DYS463 (alleles 18, 21–22), DYS464 (alleles 12, 15–17), and DYS385 (alleles 9–21) were mixed. The typing of DYS438 and DYS439 was performed using reference DNAs NA9948 and NA3657.

Statistical analysis

Allele frequency (p_i) was estimated by gene counting. The observed gene diversity (D) was calculated using the formula $D=(n/n-1)\cdot\left(1-\sum p_i^2\right)$. The haplotype diversity (H) was based on the formula $H=(n/n-1)\cdot\left(1-\sum f_i^2\right)$, where f_i is the frequency of a haplotype [11]. The amounts of stutter artefacts were calculated for peaks with a height of between 1,000 and 4,500 relative flourescent units (RFU); peak areas were used for quantitative stutter calculation.

Results and discussion

Fourteen single-copy and two multicopy Y-chromosomal STRs were evaluated in five German populations from Dresden, Hamburg, Rostock, Munich, and the ethnic group Sorbs who are a Slavic-speaking minority in East Germany (Lusatia). Allele frequencies, gene diversities, haplotype distributions, and haplotype diversities for each population group are listed in supplementary Tables 1, 2, 3, and 4. Some of the Y-STRs like DYS389-II revealed similar gene diversities in all populations. Different diversities were shown for DYS464, especially for the Sorbs (D=0.7931), in comparison to the four other populations (D=0.9423–0.9715). Especially low gene diversities, in comparison to the other populations, were calculated for DYS19 and the Hamburg population or DYS448, DYS392, and DYS393 for the Sorbs. In addition, cumulative values were calculated for allele frequencies and gene diversities. The loci with the highest cumulative gene diversity and, therefore, with the highest ranking position were DYS464 (D=0.9486), DYS385 (D=0.8675), and DYS449 (D=0.8559). According to our data, the European minimal haplotype loci DYS391, DYS392, and DYS393 were

displaced to the end of the ranking list (supplementary Table 3).

Considering the European minimal haplotype standard alone, haplotype diversities between 0.9852 for the Sorbs in Lusatia and 0.9983 for the Hamburg population were calculated. Using the EMH loci together with the SWGDAM loci, the haplotype diversity was increased only slightly. In the Rostock population, complete individualization was reached by typing EMH loci plus SWGDAM loci. The use of the EMH loci together with the RAR markers DYS446, DYS447, DYS448, DYS449, DYS463, and DYS464 increased the haplotype diversities for the populations of Dresden, Hamburg, and Rostock to 1.0000. Diversities of Munich and Sorbs were increased to 0.9977 and 0.9901, respectively. However, no complete discrimination of all non-related male individuals was possible in these two populations even with the typing of all 20 Y-STR loci. These findings indicate that the Munich population and especially the ethnic group Sorbs are separate populations [23].

In our study, it became obvious that 100% individualization of the DNA samples of the Dresden, Hamburg, and Rostock populations was reached by the use of nine EHM loci plus five highly discriminating single-copy marker and DYS464 as a multicopy marker. De Souza Goés et al. [12] individualized 100% of the samples in a Brazil population using 18 Y-STR including the EMH, too. In the study of Berger et al. [13], a haplotype diversity of 0.9994 was achieved in an Austrian population by the use of EMH loci plus eight additional loci. Comparing our data with these two studies indicates that minimum haplotype diversity for EMH loci of at least 0.9958 was needed to completely discriminate all DNA samples with further seven to nine Y-STR loci. In the case of lower haplotype diversities for the EHM panel, even a higher number of Y-STR loci or more diverse Y-STR loci, as published by Kayser et al. [14], would be needed to completely resolve all haplotypes in separated populations like Munich, Sorbs, or Austria.

For the use of Y-STRs in forensic practice, it is important to know about the amount of stutter products which are PCR artefacts due to strand slippage [15, 16]. The level of stutter products was calculated from peak area (peak heights 1,000–4,500 RFU) amplified with JumpStart Taq DNA polymerase (Sigma, Taufkirchen). The minimum, maximum, and median values of stutter artefacts are listed in supplementary Table 5. Most Y-STR loci revealed small median stutter products between 3.08 and 7.80%. Two Y-STR loci had median stutter products which were higher than 10%: DYS389-II with 11.88% and DYS449 with 15.89%, respectively. High amounts of stutter products for DYS389-II have also been published by Daniels et al. [17]. The smallest stutter artefacts were analyzed for DYS447 and DYS448 which have pentanucleotide and hexanucleotide repeats. The highest single stutter artefacts were detected for the pentanucleotide repeat DYS446 (29.52%), but the median of stutter peaks for this locus is low (6.42%).

The control DNA XY1 and all fragments used in the allelic ladders of Mentype Argus Y-MH and the experimental screening multiplex RAR were completely sequenced (supplementary Table 6). For the reference DNAs NA9948 and NA3657, alleles for all loci are listed in supplementary Table 6. Most sequences are in concordance to the published ones [1, 4] and the NIST standard reference material 2395 (SRM 2395; http://www.cstl.nist.gov/biotech/strbase/srm2395.htm). We found small differences (allelic variants) for DYS389-I, DYS389-II, and DYS390 as compared to published data. The repeat motifs $[TCTG]_4$ $[TCTA]_8$ and $[TCTG]_5$ $[TCTA]_{11}$ $[TCTG]_4$ $[TCTA]_8$ were found for DYS389-I allele 12 and DYS389-II allele 28, but only $[TCTG]_3$ and not $[TCTG]_4$ in the repeat motif are known from the literature [18]. DYS390 had the repeat motif $[TCTG]_7$ $[TCTA]_{12}$ $[TCTG]_1$ $[TCTA]_4$ for allele 24 in our samples, but only $[TCTG]_8$ motifs were known for SRM 2395.

For the locus DYS448, the allele nomenclature was in accordance with the published ones [4, 8, 19]. This means that in the repeat motif $[AGAGAT]_n$ N_{10} $[AGAGAT]_3$ N_{14} $[AGAGAT]_n$, the $[AGAGAT]_3$ motif was not counted.

Two amplicons of different lengths were generated for DYS446 with the primers used. According to Schoske [20], the shorter fragment represents a DYS446 homologous sequence on the X chromosome. The X-chromosomal sequence from GenBank accession number AL33512 is 42 nucleotides shorter than the respective Y-chromosomal sequence from GenBank accession number AC006152. In our experiments, two fragments were found for the male DNA samples: one from the X-chromosome and the other from the Y-chromosome. The difference of the two fragments was one to ten repeats. Following the rule of Schoske [20], only the longer fragments were Y-chromosomal fragments and, therefore, they were used for allele calling.

Allelic microvariants were typed for DYS385 and sequenced, and allele sequences were submitted to GenBank (supplementary Table 7; accession numbers BV679981-BV679987). At the locus DYS385, typical intermediate alleles like 10.1 with an A insert or 15.2 and 16.2 with a GA insert in the repeat region were analyzed which is in accordance with Füredi et al. [21] and Bhoopat et al. [22]. Allele 13.2 contains 14 repeats and a 2 bp deletion in the 3′ repeat flanking region which was shown for two individual DNA samples. Alleles 15.3, 16.3, and 17.3 represent the alleles 16, 17, and 18, respectively, with a T deletion in the 3′ repeat flanking region which was also published by Füredi et al. [21]. Furthermore, we found complete sequence concordance for all other fragments tested so far.

The data of the European minimal haplotype loci are available at the YHRD database. All haplotype data are available at the Biotype homepage (http://www.biotype.de) or via e-mail on request.

Acknowledgements This work was supported by the German Federal Ministry of Education and Research (BMBF, project number 03i4008). The authors would like to thank Lutz Roewer for kind cooperation. We also thank all the DNA donors and Christa Augustin, Volker Weirich, Burkhard Rolf, and Ivana Furac for kindly provided DNA samples. We acknowledge Stefanie Weidlich, Susann Gerber, and Mareen Müller for excellent technical assistance.

References

1. Kayser M, Caglia A, Corach D et al (1997) Evaluation of Y-chromosomal STRs: a multicenter study. Int J Legal Med 110:125–133, 141–149
2. Pascali VL, Dobosz M, Brinkmann B (1998) Coordinating Y-chromosomal STR research for the courts. Int J Legal Med 112:1
3. Roewer L, Croucher PJ, Willuweit S et al (2005) Signature of recent historical events in the European Y-chromosomal STR haplotype distribution. Hum Genet 116:279–291
4. Redd AJ, Agellon AB, Kearney VA et al (2002) Forensic value of 14 novel STRs on the human Y chromosome. Forensic Sci Int 130:97–111
5. Ayub Q, Mohyuddin A, Qamar R et al (2000) Identification and characterisation of novel human Y-chromosomal microsatellites from sequence database information. Nucleic Acids Res 28: E8–E8
6. Gill P, Brenner C, Brinkmann B et al (2001) DNA Commission of the International Society of Forensic Genetics: recommendations on forensic analysis using Y-chromosome STRs. Forensic Sci Int 124:5–10
7. Gill P, Brenner C, Brinkmann B et al (2001) DNA Commission of the International Society of Forensic Genetics: recommendations on forensic analysis using Y-chromosome STRs. Int J Legal Med 114:305–309
8. Gusmão L, Butler JM, Carracedo A et al (2005) 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 Legal Med 26:1–10
9. Butler JM, Schoske R (2004) Forensic value of the multicopy Y-STR marker DYS464. In: Doutremepuich C (ed) Progress in forensic genetics 10—proceedings of the 20th International ISFG Congress, Arcachon, France, 9–13 September 2003. Elsevier Science, Amsterdam, The Netherlands, pp 278–280
10. Schoske R, Vallone PM, Kline MC, Redman JW, Butler JM (2004) High-throughput Y-STR typing of U.S. populations with 27 regions of the Y chromosome using two multiplex PCR assays. Forensic Sci Int 139:107–121
11. Nei M, Tajima F (1981) DNA polymorphism detectable by restriction endonucleases. Genetics 97:145–163
12. de Souza Goés 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 Legal Med 119:70–76
13. Berger B, Lindinger A, Niederstätter H, Grubwieser P, Parson W (2005) Y-STR typing of an Austrian population sample using a 17-loci multiplex PCR assay. Int J Legal Med 119:241–246
14. Kayser M, Kittler R, Erler A et al (2004) A comprehensive survey of human Y-chromosomal microsatellites. Am J Hum Genet 74:1183–1197
15. Hauge XY, Litt M (1993) A study of the origin of 'shadow bands' seen when typing dinucleotide repeat polymorphisms by the PCR. Hum Mol Genet 2:411–415
16. Walsh PS, Fildes NJ, Reynolds R (1996) Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA. Nucleic Acids Res 24:2807–2812
17. Daniels DL, Hall AM, Ballantyne J (2004) SWGDAM developmental validation of a 19-locus Y-STR system for forensic casework. J Forensic Sci 49:668–683
18. Komuro T, Tsutsumi H, Mukoyama R, Nakamura M (1998) Repeat structure of DYS389 locus. Nippon Hoigaku Zasshi 52:227–232
19. Butler JM, Schoske R, Vallone PM, Kline MC, Redd AJ, Hammer MF (2002) A novel multiplex for simultaneous amplification of 20 Y chromosome STR markers. Forensic Sci Int 129:10–24
20. Schoske R (2003) The design, optimization and testing of Y chromosome short tandem repeat megaplexes. Thesis, Faculty of the College of Arts and Sciences of the American University, Washington
21. Füredi S, Woller J, Padar Z, Angyal M (1999) Y-STR haplotyping in two Hungarian populations. Int J Legal Med 113:38–42
22. Bhoopat T, Hohoff C, Steger HF (2003) Identification of DYS385 allele variants by using shorter amplicons and northern Thai haplotype data. J Forensic Sci 48:1108–1112
23. Immel UD, Krawczak M, Udolph J, Richter A, Rodig H, Kleiber M, Klintschar M (2006) Y-chromosomal STR haplotype analysis reveals surname-associated strata in the East-German population. Eur J Hum Genet 14:577–582