SHORT COMMUNICATION



Multiplex assay development and mutation rate analysis for 13 RM Y-STRs in Chinese Han population

Wenqiong Zhang^{1,2} · Chao Xiao¹ · Jin Yu³ · Tian Wei¹ · Fei Liao¹ · Wei Wei¹ · Daixin Huang¹

Received: 8 September 2016 / Accepted: 25 October 2016 / Published online: 9 November 2016 © Springer-Verlag Berlin Heidelberg 2016

Abstract In this study, a novel multiplex polymerase chain reaction (PCR) assay was developed for amplifying the newly introduced 13 rapidly mutating Y-STR markers (RM Y-STRs) including DYF387S1, DYF399S1, DYF403S1a/b, DYF404S1, DYS449, DYS518, DYS526b, DYS547, DYS570, DYS576, DYS612, DYS626, and DYS627. In addition, a survey for mutation rates of the 13 RM Y-STRs in Chinese Han population was performed to make sure of the mutation characteristic and application in Chinese group. With 14,476 allele transfers in 1034 father-son pairs, 221 mutation events occurred, of which 215 were one-step mutations and 6 were two-step mutations. Nineteen father-son pairs were found to have mutations at two loci and one pair at three loci. Based upon our research data, 18.96 % of all 1034 father-son pairs were successfully differentiated, and the estimated locus-specific mutation rates varied from 4.84×10^{-3} to 6.29×10^{-2} , with an average estimated mutation rate $1.53 \times$ 10^{-2} (95 % CI 1.33×10^{-2} to 1.74×10^{-2}). Among the 13 Y-STR markers, eight loci (DYF399S1, DYF403S1a, DYF404S1, DYS449, DYS518, DYS547, DYS576, and DYS612) had mutation rates higher than 1.00×10^{-2} , and the rest loci lower than 1.00×10^{-2} in Chinese samples.

Electronic supplementary material The online version of this article (doi:10.1007/s00414-016-1489-y) contains supplementary material, which is available to authorized users.

Daixin Huang huangdaixin@mails.tjmu.edu.cn

- ¹ Department of Forensic Medicine, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, People's Republic of China
- ² Qingdao Licang District People's Procuratorate, Qingdao 266100, People's Republic of China
- ³ Wuhan Blood Center, Wuhan 430033, People's Republic of China

Keywords Rapidly mutating Y-STR \cdot Multiplex assay \cdot Mutation \cdot Lineage differentiation

Introduction

Y-chromosome short tandem repeats (Y-STRs) are widely used in tracing human evolution through male lineages [1] as well as forensic applications including investigating sexual assault cases [2, 3] and paternity testing of male offspring [4, 5] because of their paternal inheritance characteristics. Currently used single multiplex Y-STR system showed considerable haplotype diversities in many populations as reported [6, 7]. However, their use is limited to differentiating male lineages in some isolated and inbred populations [8-10] and they fail to differentiate between individuals from the same paternal lineage in some cases. Adding additional highly polymorphic Y-STR markers would improve the discrimination capacity in populations with low genetic diversity [7, 11], but the newly discovered Y-STR set with high mutation rates has been verified to be able to solve the two limitations at the same time [11, 12].

The new set of 13 novel "rapidly mutating" Y-STR markers (RM Y-STRs) was selected from 186 Y-STRs with exceptionally high mutation rates (> 10^{-2} per locus per generation) raised by Ballantyne et al. [12]. This set has been proved not only to be able to greatly increase the discrimination capacity of male lineage differentiation but also to be extremely useful in close male relative resolution [11, 13–15].

To make it more convenient in the use of forensic daily casework and investigate the mutation characteristic of the 13 RM Y-STRs in Chinese Han population, we developed a novel multiplex amplification assay gathering all the 13 markers and tested their mutation rates among 1034 Chinese father–son pairs in our study.

Materials and methods

DNA samples

A total of 1034 father-son pair samples were collected from Han population of Hubei province of China for the mutation rate investigation of the 13 RM Y-STRs after informed consent. All father-son pairs had been confirmed biological paternity with autosomal STR kit (AGCU EX22 STR kit, AGCU ScienTech, Wuxi, China). Two male control DNA 007 (Thermo Fisher Scientific, MA, USA) and 9948 (AGCU ScienTech, Wuxi, China), one female control DNA 9947A (Thermo Fisher Scientific, MA, USA), and two female samples (F1 and F2) were used for the multiplex assay development. In addition, five cigarette butts and epithelial cells found on the comb were collected from five male volunteers, respectively. Genomic DNA of male samples was isolated from whole blood using the Chelex-100 method [16]. Female DNA samples (F1, F2) were isolated using phenolchloroform extraction method and subsequently quantified with the Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific, MA, USA).

Primer design and multiplex development

The multiplex assay was designed to include the Y-STR loci DYF387S1, DYF399S1, DYF403S1a/b, DYF404S1, DYS449, DYS518, DYS526b, DYS547, DYS570, DYS576, DYS612, DYS626, and DYS627. Allele ranges of the 13 RM Y-STRs referred from our previous frequency study [17] to help leaving sufficient space between loci when arranging multiplex combinations of each dye. New primers were designed for the construction of the multiplex assay except for DYF403S1, DYS627, and DYS526b. Primers for loci DYF403S1 and DYS627 followed the report by Ballantyne et al. [13], and primers for the locus DYS526b followed the report by Rogalla et al. [15]. Primer 3 software (http://bioinfo.ut. ee/primer3/) was used for designing new primers, and AutoDimer software was used to test possible primer-dimers after primer designing. All primer sequences were retested by BLAST to ensure the specificity of amplification products in genome. In addition, we added a single G on the 5' end of the unlabeled primer within a locus-specific primer pair to promote full adenylation of PCR products amplified from that locus, and if the 5' end of the unlabeled primer was G itself, then G was not added (for detailed data for each marker, see Electronic Supplementary Materials Table S1).

This multiplex assay was designed as a five-dye assay; one forward or reverse primer for each locus was labeled at the 5'-end with 6-FAMTM, VICTM, NEDTM, or PETTM fluorescent dye for the detection by ABI 3130 Genetic Analyzer (Thermo Fisher Scientific, MA, USA). Then the 13 RM Y-STRs were organized by allele size ranges and assigned to each of the four dyes to achieve a single multiplex assay (see Electronic Supplementary Materials Fig. S1).

Multiplex amplification

PCR amplification was performed in a total reaction volume of 20 μ L containing 10 μ L of Platinum[®] Multiplex Master Mix (Thermo Fisher Scientific, MA, USA), 5.5 μ L of the 13 RM Y-STRs primer mixture (see Electronic Supplementary Materials Table S1), and 2 μ L of DNA template. Thermal cycling was performed on GeneAmp 2720 (Thermo Fisher Scientific, MA, USA) under the following conditions: 96 °C for 5 min; 10 cycles of 95 °C for 30 s, 61 °C for 45 s and 72 °C for 1 min, 20 cycles of 95 °C for 30 s, 59 °C for 45 s, and 72 °C for 1 min, and a final extension hold at 72 °C for 10 min.

Detection and genotyping

PCR products were electrophoresed on ABI 3130 Genetic Analyzer (Thermo Fisher Scientific, MA, USA) following manufacturer's protocols. Samples were prepared as a mixture of 0.3 μ L GeneScanTM 600 LIZ[®] size standard (Thermo Fisher Scientific, MA, USA) with 8.7 μ L Hi-DiTM Formamide (Thermo Fisher Scientific, MA, USA) and 1 μ L PCR products. Samples were analyzed using GeneMapper ID v3.2 software (Thermo Fisher Scientific, MA, USA) after data collection. Allele nomenclature followed by the sequencing result of our previous study [17, 18] according to the recommendations of the International Society of Forensic Genetics (ISFG) Commission [19].

Multiplex performance studies

Sensitivity and specificity were conducted on the RM Y-STR multiplex to test its performance characteristic. Sensitivity testing was performed using a series of diluted male control DNA 9948 (31.25 pg to 1 ng). To test if the multiplex assay is specific for male DNA only, we amplified different concentration of female control DNA 9947A and two female samples (F1, F2). We also tested the male–female mixture resolution ability with 1 ng male control DNA 9948 mixed with 250 to 1000 ng female DNA (F1). Assessing species specificity encompassed testing performance of the assay in amplifying 2 ng of template DNA from animals most likely to appear on the crime scene, including chicken, duck, sheep, rabbit, cattle, pig, dog, hamster, and cat.

Data analysis

Mutations were counted directly and the locus-specific mutation rate was calculated as the number of observed mutations divided by the number of father–son pairs. The confidence interval (CI) was estimated using exact binomial probability distribution (http://statpages.org/confint.html). Comparison of mutation rates between different population was performed by the χ^2 -test of a R × C contingency table analysis, and comparison of the mean age of fathers displaying and not displaying mutations was performed by Student's *t* test.

Results and discussion

Construction of the RM Y-STR multiplex assay

Although Alghafri et al. [20] had already developed a multiplex assay for simultaneously analyzing 13 RM Y-STRs, the primers for DYS526a/b in their multiplex assay would lead to some imbalance of the amplification peaks due to a primer mismatch at the second primer site. In fact, the DYS526a is only a part of the DYS526b, and its mutation rate is also relatively low. Therefore, in order to simplify the multiplex system, in this study, we used the same primers with Rogalla et al. [15] which made it possible to amplify a single allele of the DYS526b. After redesigning primers, labeling fluorescence dye and optimizing experiment conditions, a novel 13-plex fluorescent multiplex PCR system was successfully developed, and all of the 13 RM Y-STRs were amplified with satisfactory results (Fig. 1).

Overall assay's performance

A few pre-validation studies were conducted for demonstrating the effectiveness of the multiplex assay. A series of amounts of the male control DNA 9948 (31.25 pg to 1 ng) was used as template to test the sensitivity of this multiplex assay. The quantities of 9948 used were 1 ng, 500 pg, 250 pg, 125 pg, 62.5 pg, and 31.25 pg. Full profiles could be observed from 62.5 pg to 1 ng of 9948 while peak height of a few markers was below the detecting threshold (50 rfu) at the amount of 31.25 pg (see Electronic Supplementary Materials Fig. S2). In forensic practice, about 0.5–2 ng DNA is routinely recommended for typing, although 62.5 pg DNA is enough.

Different amount of female control DNA 9947A and female DNA (F1, F2) were used for the male specific test. When female DNA (1 to 5 ng) was added to the multiplex assay while amplification, no female product was detected, indicating good male specific with normal amount of female DNA.

Since this RM Y-STR multiplex assay would be mostly used in sexual assault cases, male-female DNA mixtures are



Fig. 1 The electrophoretogram of the 13 RM Y-STRs typing system from the control DNA 9948

very likely to be faced when detecting male genotype. To test the assay's ability to amplify male component mixed with female materials in different ratio, we tried a series of mixtures with male to female ratio ranging from 1:250 to 1:1000. Male DNA 9948 was kept 1 ng while female DNA was 250 ng, 500 ng, 750 ng and 1 μ g in each reaction. Each of the mixtures was amplified in triplicate for consistency. Our result showed that full profiles for the 13 markers of male DNA were detected in every ratio tested (see Electronic Supplementary Materials Fig. S3).

We tested the assay's specificity for human DNA template by making an attempt to amplify non-human samples collected from chicken, duck, sheep, rabbit, cattle, pig, dog, hamster, and cat—species most probable of being present at the crime scene in our region. As expected, amplification of DNA from all the above-mentioned species did not yield detectable products in any locus.

In order to evaluate the application effect of the multiplex assay on the routine biological samples encountered in daily forensic cases, five cigarette butts and epithelial cells found on the comb were also analyzed using the multiplex assay, and satisfactory genotyping results were obtained from all of them (see Electronic Supplementary Materials Fig. S4).

Efficacy for differentiating individuals

In order to test the efficacy of the assay in differentiating between individuals from the same male lineages, we performed a mutation investigation among 1034 father–son pairs from Chinese Han population in Hubei province. Genotype mismatch was observed in 196 father–son pairs out of 1034 corresponding to a differentiating rate of 18.96 % (95 % CI 16.61 to 21.48 %) which is close to the previous study [12, 14]. This value was more than 4.11 and 1.76 times higher than that estimated for the Yfiler set with 4.61 % (95 % CI 3.32 to 6.23 %) and the Yfiler Plus set with 10.75 % (95 % CI 8.92 to 12.80 %) in Chinese population, respectively [21, 22]. Of the 196 resolved pairs, 89.80 % showed allele discrepancies at one locus, 9.69 % at two loci, and 0.51 % showed allele mismatch at three loci.

Estimated mutation rates

A total of 14,476 allele transfers in 1034 father–son pairs were studied for the 13 RM Y-STRs. The mutation rates for these loci are shown in Table 1. A total of 221 mutation events were observed for 13 RM Y-STRs, with an estimated mutation rate ranged between 4.84×10^{-3} (95 % CI 1.60×10^{-3} to 1.12×10^{-2}) for DYS570 locus and 6.29×10^{-2} (95 % CI 4.88×10^{-2} to 7.94×10^{-2}) for DYF399S1 locus. The average mutation rate was estimated to be 1.53×10^{-2} (95 % CI 1.33×10^{-2} to 1.74×10^{-2}),

slightly lower than the estimate by Ballantyne et al. [12] of 1.97×10^{-2} (95 % CI 1.80×10^{-2} to 2.20×10^{-2}). Among these Y-STR markers, DYF399S1, DYF403S1a, DYF404S1, DYS449, DYS518, DYS547, DYS576, and DYS612 had mutation rates higher than 1.00×10^{-2} , and DYF387S1, DYF403S1b, DYS526b, DYS570, DYS626, and DYS627 showed relatively low mutation rates (< 1.00×10^{-2}) compared with the study by Ballantyne et al. [12, 13]. However, it should be noticed that although the mutation rates of these loci are lower than former reports, they are already much higher than that of traditional used Y-STRs. In addition, the mutation rates of loci DYF387S1 and DYS449 were slightly different with that of Guangdong Han population [22], but the difference was not significant (0.1).

Characteristic of mutations

Among all of the 221 mutations observed, 215 were one-step mutations, while 6 were two-step mutations, supporting the generally accepted stepwise mutation model raised by Ohta and Kimura [23]. Repeat gains and losses were almost equal with 112 gains and 109 losses, resulting in a repeat gain to loss mutation ratio of 1.03:1.

Molecular factors that influence RM Y-STR mutations concluded by our study were partially in accordance with former reports by Brinkmann et al. [24] and Ballantyne et al. [12]. First of all, the structure complexity of repetitive regions has a strong correlation with mutation events. Locus with complex repetitive structures had relatively higher mutation rates such as DYF399S1, DYS449, DYS547, and DYS612. Besides, multi-copy markers DYF387S1, DYF399S1, DYF403S1a, and DYF404S1 generally had more mutations than single copy ones. Secondly, father's childbearing age generally plays a role in the mutation occurrence. In this study, however, we did not observe significant differences in age. The mean age of fathers who were involved in mutation events was 34.47 (±9.00) years old while the mean age of those had no mutations was $32.42 (\pm 7.85)$ years old, and the difference is not significant (p > 0.05).

Conclusion

In this study, a multiplex assay for 13 RM Y-STRs was constructed and performed well after a series of multiplex assay tests. This assay would strongly promote the practical application of RM Y-STRs in daily forensic cases. A mutation survey was conducted upon the assay, and 18.96 % of all 1034 father–son pairs were successfully differentiated by at least one mutation with the RM Y-STR set. The average mutation rate of all the 13 Y-STR markers is as high as

Table 1 Estimated mutation rates for 13 RM Y-STRs in a Han population of Hubei Province, Central China

Y-STR loci	Number of allele transfers	Number of mutations	Number of gains	Number of losses	Number of one-step mutations	Number of two-step mutations	Mutation rates	Binomial 95 % CIs
DYF387S1	1034	9	7	2	9	0	8.70×10^{-3}	4.00×10^{-3} to 1.65×10^{-2}
DYF399S1	1034	65	28	37	63	2	6.29×10^{-2}	$\begin{array}{c} 4.88 \times 10^{-2} \text{ to} \\ 7.94 \times 10^{-2} \end{array}$
DYF403S1a	1034	28	16	12	27	1	2.71×10^{-2}	$\begin{array}{c} 1.81 \times 10^{-2} \text{ to} \\ 3.89 \times 10^{-2} \end{array}$
DYF403S1b	1034	8	4	4	8	0	7.74×10^{-3}	$\begin{array}{c} 3.30 \times 10^{-3} \text{ to} \\ 1.52 \times 10^{-2} \end{array}$
DYF404S1	1034	11	4	7	9	2	1.06×10^{-2}	5.30×10^{-3} to 1.90×10^{-2}
DYS449	1034	13	8	5	13	0	1.26×10^{-2}	$\begin{array}{c} 6.70 \times 10^{-3} \text{ to} \\ 2.14 \times 10^{-2} \end{array}$
DYS518	1034	13	6	7	12	1	1.26×10^{-2}	$\begin{array}{c} 6.70 \times 10^{-3} \text{ to} \\ 2.14 \times 10^{-2} \end{array}$
DYS526b	1034	10	4	6	10	0	9.67×10^{-3}	$\begin{array}{c} 4.70 \times 10^{-3} \text{ to} \\ 1.77 \times 10^{-2} \end{array}$
DYS547	1034	14	7	7	14	0	1.35×10^{-2}	7.40×10^{-3} to 2.26×10^{-2}
DYS570	1034	5	3	2	5	0	4.84×10^{-3}	1.60×10^{-3} to 1.12×10^{-2}
DYS576	1034	14	8	6	14	0	1.35×10^{-2}	7.40×10^{-3} to 2.26×10^{-2}
DYS612	1034	14	6	8	14	0	1.35×10^{-2}	7.40×10^{-3} to 2.26×10^{-2}
DYS626	1034	8	4	4	8	0	7.74×10^{-3}	3.30×10^{-3} to 1.52×10^{-2}
DYS627	1034	9	7	2	9	0	8.70×10^{-3}	4.00×10^{-3} to 1.65×10^{-2}
Total	14476	221	112	109	215	6	1.53×10^{-2}	$\begin{array}{c} 1.33 \times 10^{-2} \text{ to} \\ 1.74 \times 10^{-2} \end{array}$

 1.53×10^{-2} (95 % CI 1.33×10^{-2} to 1.74×10^{-2}). The mutation rates at DYF399S1, DYF403S1a, DYF404S1, DYS449, DYS518, DYS547, DYS576, and DYS612 are higher than 1.00×10^{-2} in Hubei Han population, while the mutation rates at DYF387S1, DYF403S1b, DYS526b, DYS570, DYS626, and DYS627 are lower than expected. In addition to the extremely high haplotype diversity according to our investigation [17], RM Y-STRs would show great use in forensic practice in Chinese population.

Acknowledgments This research was supported by the National Natural Science Foundation of China (No. 81401560) and the Fundamental Research Funds for the Central Universities, HUST: No. 2016YXZD025.

References

1. Jobling MA, Tyler-Smith C (1995) Fathers and sons: the Y chromosome and human evolution. Trends Genet 11(11):449–56

- Parson W, Niederstätter H, Brandstätter A, Berger B (2003) Improved specificity of Y-STR typing in DNA mixture samples. Int J Legal Med 117(2):109–14
- Roewer L (2009) Y chromosome STR typing in crime casework. Forensic Sci Med Pathol 5(2):77–84
- 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 Legal Med 115(1):12–5
- Ayadi I, Mahfoudh-Lahiani N, Makni H, Ammar-Keskes L, Rebaï A (2007) Combining autosomal and Y-chromosomal short tandem repeat data in paternity testing with male child: methods and application. J Forensic Sci 52(5):1068–72
- de Knijff P, Kayser M, Caglià A, Corach D, Fretwell N, Gehrig C et al (1997) Chromosome Y microsatellites: population genetic and evolutionary aspects. Int J Legal Med 110(3):134–40
- Purps J, Siegert S, Willuweit S, Nagy M, Alves C, Salazar R et al (2014) A global analysis of Y-chromosomal haplotype diversity for 23 STR loci. Forensic Sci Int Genet 12:12–23
- Hedman M, Pimenoff V, Lukka M, Sistonen P, Sajantila A (2004) Analysis of 16 Y STR loci in the Finnish population reveals a local reduction in the diversity of male lineages. Forensic Sci Int 142(1):37–43
- Woźniak M, Derenko M, Malyarchuk B, Dambueva I, Grzybowski T, Miścicka-Sliwka D (2006) Allelic and haplotypic frequencies at

11 Y-STR loci in Buryats from South-East Siberia. For ensic Sci Int 164(2--3):271--5

- Hedman M, Neuvonen AM, Sajantila A, Palo JU (2011) Dissecting the Finnish male uniformity: the value of additional Y-STR loci. Forensic Sci Int Genet 5(3):199–201
- Ballantyne KN, Ralf A, Aboukhalid R, Achakzai NM, Anjos MJ, Ayub Q et al (2014) Toward male individualization with rapidly mutating Y-chromosomal short tandem repeats. Hum Mutat 35(8): 1021–32
- Ballantyne KN, Goedbloed M, Fang R, Schaap O, Lao O, Wollstein A et al (2010) Mutability of Y-chromosomal microsatellites: rates, characteristics, molecular bases, and forensic implications. Am J Hum Genet 87(3):341–53
- Ballantyne KN, Keerl V, Wollstein A, Choi Y, Zuniga SB, Ralf A, Vermeulen M, de Knijff P, Kayser M (2012) A new future of forensic Y-chromosome analysis: rapidly mutating Y-STRs for differentiating male relatives and paternal lineages. Forensic Sci Int Genet 6(2):208–18
- Robino C, Ralf A, Pasino S, De Marchi MR, Ballantyne KN, Barbaro A et al (2015) Development of an Italian RM Y-STR haplotype database: results of the 2013 GEFI collaborative exercise. Forensic Sci Int Genet 15:56–63
- Rogalla U, Woźniak M, Swobodziński J, Derenko M, Malyarchuk BA, Dambueva I, Koziński M, Kubica J, Grzybowski T (2015) A novel multiplex assay amplifying 13 Y-STRs characterized by rapid and moderate mutation rate. Forensic Sci Int Genet 15:49–55
- 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 10(4):506–13

- Zhang W, Xiao C, Wei T, Pan C, Yi S, Huang D (2016) Haplotype diversity of 13 RM Y-STRs in Chinese Han population and an update on the allele designation of DYF403S1. Forensic Sci Int Genet 23:e1–9
- Zhang W, Xiao C, Yu J, Wei T, Pan C, Huang D (2016) Allele repeat structure and designation update for the Y-STR loci DYF399S1, DYS526b and DYS626. Forensic Sci Int Genet 20: 147–8
- Gusmão L, Butler JM, Carracedo A, Gill P, Kayser M, Mayr WR 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. Forensic Sci Int 157(2–3): 187–97
- Alghafri R, Goodwin W, Ralf A, Kayser M, Hadi S (2015) A novel multiplex assay for simultaneously analysing 13 rapidly mutating Y-STRs. Forensic Sci Int Genet 17:91–8
- Wu WW, Hao HL, Ren WY, Su YJ, Lv DJ (2012) Mutation analysis of 17 Y-STR loci in Chinese Han population. Chin J Forensic Med 27(6):455–7
- 22. Wang Y, Zhang YJ, Zhang CC, Li R, Yang Y, Ou XL, Tong DY, Sun HY (2016) Genetic polymorphisms and mutation rates of 27 Ychromosomal STRs in a Han population from Guangdong Province, Southern China. Forensic Sci Int Genet 21:5–9
- 23. Ohta T, Kimura M (1973) A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. Genet Res 22(2):201–4
- 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(6):1408–15