MINI REVIEW

Estimation of the Y-Chromosomal Short Tandem Repeat (Y-STR) Mutation Rates in Uzbekistan

Kurganov Sardarkhodja, Muxamedov Rustam, Axmedova Dilobar, Filatova Viktoriya, Akhmedov Bakhodir

Sardarkhodja K. Estimation of the Y-chromosomal short tandem repeat (Y-STR) mutation rates in Uzbekistan. J Genet Mutat 2018;1(1):11-4.

ABSTRACT

In the practice of forensic genetic examination, there are cases when the biological trace on physical evidence is left by an unknown male person. In such cases, along with the definition of the genotype of an unknown person, studies on the Y-chromosome are performed on the STR loci of nuclear DNA. In most cases, these studies lead to the successful disclosure of the crime. But sometimes, despite the large number of samples examined, the suspect cannot be identified. In most such cases, DNA samples of individuals living in the same area are examined. To narrow the circle of suspects, calculate the lifetime of a common ancestor, through which it is possible to obtain information about related affinity and narrow the circle of suspects. Later, these results will be used to determine the relationship on the paternal line, where a reliable knowledge on mutation properties is necessary for correct data interpretation. When determining the degree of solution on the paternal line, if discrepancies between the child's father and other paternal relatives are not taken into account, population-specific mutation rates should be used to determine if this is a mutation or a true exception. Therefore in this study, we aim to determine the mutation rates of 17 Y-STR loci in Uzbekistan.

Keywords: *DNA analysis; Mutation; Y chromosome*

Abbreviations: Y-STR Y-Chromosomal Short Tandem Repeat; *NYR*- Non-Recombinant Region of the Y-chromosome

The human Y-chromosome represents about 2% of the total human genome and is approximately 60 Mb in length (1-14). Most of the Y-chromosome consists of a non-recombinant region of the Y-chromosome (NRY) (11). The NRY is inherited intact through paternal lineages unless mutation/s has occurred. Because of such inheritance pattern, STR markers located in the NRY region have become useful for applications including genetic structure studies, paternity testing, identification of disasters male victims, identification of male lineages for anthropology purposes, and the identification of male perpetrators in sexual assault criminal cases (4-6,10,13,15). The first use of STR on the Y chromosome occurred in 1992 and this STR is now known as *DYS19*. Since then the potential use of Y-STR analysis for forensic casework has been recognized and well documented. Y-STRs, which have an average mutation rate of about 10−3 per locus per generation (2), have proven to be useful for forensic applications and have been included in several commercial Y-STR kits.

Although the greatest value of Y-STRs is male specificity, this also turns into a major limitation due to the existence of an identical haplotype within a male lineage (9,16). This means that while currently used Y-STRs are able to reliably differentiate between different male lineages, they cannot resolve these lineages down to individual level in case of paternal relatives (3). When determining the degree of solution on the paternal line, if discrepancies between the child's father and other paternal relatives are not taken into account, population-specific mutation rates should be used to determine if this is a mutation or a true exception. Therefore in this study, we aim to determine the mutation rates of 17 Y-STR loci in Uzbekistan.

MATERIALS AND METHODS

Objects of the research

The subjects of the study were blood samples and dried saliva on sterile gauze tampons, selected from 1170 individuals.

Samples collection

Samples were collected from people living in the districts of Chirchik (200 unrelated males), Angren (150 unrelated males), Zangiota (100 paternal relatives) and Kibray (120 paternal relatives), also samples were collected from 300 (fathers and sons) pairs unrelated males throughout Uzbekistan.

DNA extraction

Genomic DNA was extracted from peripheral blood and dried saliva samples using the salting-out method (8). 600 objects from 1.170 samples were collected from paternity cases data tests involving males whose biological relationship were previously confirmed by autosomal STRs using AmpF/STR Identifiler kit (Thermo Fisher Scientific).

DNA quantification

After isolation, the quantity of genomic DNA of each sample was determined by quantitative real-time polymerase chain reaction (PCR) using the Quantifiler™ Human Male DNA Quantification kit (Thermo Fisher Scientific), which includes internal positive control to test for the presence of PCR inhibitors in the DNA extracts. Quantitative real-time PCR was performed on 7500 Real-Time PCR System (Applied Biosystems).

PCR amplification and detection

To ensure successful amplification, 0.5 ng to 1 ng of DNA was used for each multiplex amplification reaction. All thermal cycling was conducted on Applied Biosystems® GeneAmp® PCR System 9700 thermal cyclers. PCR amplification using Y-filer PCR Amplification Kit (Thermo Fisher Scientific) was performed as recommended by the manufacturer, although half of the recommended reaction volume (12.5 μl) was used.

Separation and detection of the 17 Y-STR loci were performed using the 3130xl Genetic Analyser (Applied Biosystems) 16-capillary array system and filter set G5. Each sample was prepared by adding 1 mL PCR product to 14 mL of Hi-DiTM formamide and 0.4 mL GeneScanTM-500 LIZTM internal size standard (Thermo Fisher Scientific).

The sample run data were analyzed, together with an allelic ladder and positive and negative controls, using GeneMapper ID-X v3.2 (Applied Biosystems) software.

Statistical analysis

Comparison information of the sample data was generated using an in-house software program involving DNA-expert macros designed to check for allele sharing across all loci. Obtained mutations were compared with those available at the YHRD (Y-Chromosome Haplotype Reference Database) (17).

Republican Centre of Forensic Expertise, Chilonzor, Tashkent, Republic of Uzbekistan

Correspondence: Sardarkhodja K, Republican Centre of Forensic Expertise, Chilonzor, Tashkent, Republic of Uzbekistan, Telephone (371) 2773541, e-mail sardorbioinformatik@mail.ru

Received: May 29, 2018, Accepted: June 22, 2018, Published: January 29, 2018

RESULTS AND DISCUSSION

We assessed the mutation rates of 17 Y-STR loci in Uzbek population. Samples of 300 father-son pairs from all geographic regions of Uzbekistan were typed, and each Y-STR locus pair was compared.

We observed a total of seven mutations (*DYS389 II, DYS458, DYS385 a/b* (three in *DYS385 a/b locus*), *DYS439 and DYS438*). The highly polymorphic Y-STR locus *DYS385* was observed to have a higher mutation rate compared to all other Y-STRs loci analyzed. In this study, the observed higher specific locus mutation rate for Y-STR locus *DYS385a/b* (if treated as single locus) was 1.0×10-2. Among the three mutations, two repeat losses and one repeat gain were observed.

The mutation rates of this study were compared with those available at the *YHRD* (Table 1). It was found that the *DYS438* (3.75×10−3), *DYS458* (6.36×10−3) and *DYS389 II* (4.12×10−3) loci had the highest mutation rate in both YHRD and in our study (Table 1). A low mutation rate was obtained in *DYS385 a/b* (2.45×10−3) and *DYS439* (0.38×10−3) loci.

In our study, since no mutations were observed in *DYS456, DYS389 I, DYS390, DYS19, DYS393, DYS391, DYS635, DYS392, Y GATA H4, DYS437 and DYS448* loci, the mutation rate of those loci was determined as zero (0).

All observed mutation events were characterized by single step mutations, these results are in accordance with the generally accepted mutation model for microsatellites, in which the alleles are known to mutate primarily through the gain and loss of single repeat units (18,19).

Within the framework of the research, the speed of mutations in 300 pairs of fathers and sons throughout Uzbekistan was studied. Among 5100 translations of alleles, mutations were detected in seven cases, giving an average mutation rate of 7/5100 = 1.373×10-3 at all loci by generation is of limited reliability. For this reason, the next study will use the value of the mutation frequency from the YHRD database. According to YHRD, the average mutation rate for the 17 marker loci is 0.031 mutations per haplotype (Table 1).

Also, the rate of mutations in two large related families from districts Zangiota and Kibray was studied. According to the data obtained the first family consisted of 15 generations, and the second family consisted of 14 generations. As a result of our analysis, one mutation was detected in each family of the locus *DYS385a*.

As a result of our study of 150 people living in the city of Angren, it was revealed that they had a genetic similarity of the Y chromosome. Among this random sample, 80 people who had close haplotypes, turned out to be relatives. Among them, in the study of loci with a high mutation frequency, 19 individual mutations were found (Table 2).

TABLE 1

Mutations and allele transmissions per Y-STR

Locus	Y-STR mutations of 300 father-son pairs			Total (YHRD and this study mutation rates)		
	No. Mutations	Allele Transmission	Mutation rate (×10−3) (No. Mutations/Meiosis)	No. Mutations	Allele Transmission	Mutation rate (×10−3) (No. Mutations/Meiosis)
DYS456	-	300	-	31	7229	4,29
DYS389 I	-	300	-	42	14339	2,93
DYS390	-	300	-	33	15612	2,11
DYS389 II	1	300	3,3	59	14310	4,12
DYS458	1	300	3,3	46	7228	6,36
DYS19	-	300	-	36	16090	2,24
DYS385 a/b	3	300	10	64	26171	2,45
DYS393	-	300	-	15	14264	1,05
DYS391	-	300	-	38	15486	2,45
DYS439	1	300	3,3	58	10647	5,45
DYS635	-	300	-	35	8076	4,33
DYS392	-	300	-	8	15418	5,19
Y GATA H4	-	300	-	25	8260	3,03
DYS437	-	300	-	13	10652	1,22
DYS438	1	300	3,3	4	10673	3,75
DYS448	-	300	-	11	7229	1,52
Mean value of mutation frequency			1,4			3,08

TABLE 2

STR Locus

S. No	DYS456	DYS389-I	DYS390	DYS389-II	DYS458	DYS19	DYS385a/b	DYS393	DYS391	DYS439	DYS635	DYS392	H4_Y_GATA_	DYS437	DYS3438	DYS448	Quantity
1	15	14	25	30	19	16	12;13	13	10	10	21	11	11	14	10	22	8
2	15	13	25	29	18	16	12;13	13	10	10	21	11	11	14	10	21	9
3	15	13	25	29	19	16	12;13	13	10	10	21	11	11	14	10	22	7
4	15	13	25	29	17	16	12;13	13	10	10	21	11	11	14	10	22	7
5	15	14	25	29	17	16	12;13	13	10	10	21	11	11	14	10	22	8
6	15	13	25	29	17	16	12;14	13	10	10	21	11	11	14	10	22	9
7	15	13	24	29	18	16	12;13	14	10	10	21	11	11	14	10	22	8
8	15	13	24	29	18	16	12;13	13	10	10	21	11	11	14	10	22	6
9	15	13	24	29	18	16	12;12	14	10	10	21	11	11	14	10	22	8
10	15	13	25	30	18	16	12;13	13	10	10	21	11	11	14	10	22	6
11	14	13	25	29	19	16	12;13	13	10	10	21	11	11	14	10	21	4
Mut.	1	2	3	2	5	0	2	2	0	0	0	0	0	0	0	1	19

TABLE 3

STR Locus

S. No	DYS456	DYS389-I	DYS390	DYS389-II	DYS458	DYS19	DYS385a/b	DYS393	DYS391	DYS439	DYS635	DYS392	H4_Y_GATA_	DYS437	DYS3438	DYS448	Quantity
1	15	15	19	16	17	14	12;13	13	11	13	24	13	9	14	10	18	6
2	15	14	19	16	17	14	13;13	13	11	13	24	13	9	14	10	20	7
3	16	15	19	15	17	14	13;13	13	11	13	24	13	9	14	10	19	9
4	15	14	19	16	17	14	13;13	13	11	13	24	13	9	14	10	19	7
5	15	13	19	15	17	14	12;13	13	11	13	24	13	9	14	10	20	8
6	15	15	19	17	17	14	12;13	13	11	13	24	13	9	14	10	19	9
7	14	14	19	16	17	14	13;13	13	11	13	24	13	9	14	10	19	3
8	15	14	19	16	17	14	13;13	13	11	13	24	13	9	14	9	19	8
9	15	14	19	16	17	14	13;13	13	11	13	24	13	9	14	10	19	9
10	15	16	19	17	17	14	13;13	13	11	13	24	13	10	14	10	19	10
11	15	14	19	16	17	14	13;13	13	11	13	24	13	9	14	10	19	8
12	15	15	19	16	17	14	13;14	13	11	13	24	13	9	14	10	19	9
13	15	14	19	16	17	14	13;13	13	11	13	24	13	9	14	10	19	7
14	15	14	19	15	17	14	12;13	13	11	13	24	13	9	14	10	19	9
Mut.	2	5	0	5	0	0	5	0	0	0	0	0	1	0	1	5	24

The total number of mutations per sample (19 mutations) is divided by the number of haplotypes (11 haplotypes) and the number of markers in the haplotype (17 markers).

$$\frac{19}{11*17} = 0.101604 \text{ (Mutations per haplotype)}$$

In this sample study, the average observed number of mutations per marker is 0.101604.

Now we can approximately estimate the age of the common ancestor. The rate of mutations for our 17-marker haplotype: 0.031 mutations per haplotype, or 0.0018 mutations per marker. The duration of one generation is assumed to be 25 years (20).

The age of a common ancestor is obtained by dividing the average observed number of mutations per marker by the rate of mutations (also on the haplotype) (20): $\frac{0.101604}{0.0018} = 55.71$ Generations

Multiplying 55.71 generations by the duration of one generation of 25 years, we obtain 1392 years. Since the calculation is approximate, we recommend rounding the result down to tens of years. In our case rounding gives 1390 years. 1390 years is an approximate, rough estimate of the age of a common ancestor.

The same method was used to investigate the genetic set of 200 people living in the town of Chirchik. Among this random sample, 109 people were related on the paternal line. Among them, when studying a set of loci with a high frequency of mutations, 24 individual mutations were found (Table 3).

The total number of mutations per sample (24 mutations) is divided by the number of haplotypes (14 haplotypes) and the number of markers in the haplotype (17 markers).

$$\frac{0.10084}{0.0018} = 0.10084 \text{ (Mutations per haplotype).}$$

In this sample study, the average observed number of mutations per marker is 0,10084.

The age of a common ancestor is obtained by dividing the average observed number of mutations per marker by the rate of mutations (also on the haplotype): $\frac{0.10084}{0.0018} = 50.42$ Generations

Multiplying 55.30 generations by the duration of one generation of 25 years, we obtain 1382 years. Since the calculation is approximate, we recommend rounding the result down to tens of years. In our case rounding gives 1380 years. 1380 years is an approximate, rough estimate of the age of a common ancestor.

CONCLUSION

The present study shows the locus specific mutation rate estimate of 17 Y-STRs loci. Considering the results obtained in the present study, it can be concluded that, the reliable knowledge about mutation rates of 17 Y-STRs loci used in forensics and paternity testing involving males is very important for a correct interpretation of results.

ETHICAL APPROVAL

This study was reviewed and approved by the ethics of the collective council of the Republican Center for Forensic Expertise (Republic of Uzbekistan).

CONFLICT OF INTEREST

There are no conflicts of interest to disclose from all authors.

GRANT SUPPORT

This work was supported by Ministry of Innovation Development of The Republic of Uzbekistan, Grant No. A-2-089+(A-2-056).

REFERENCES

1. Buehler EM (1980) A synopsis of the human Y-chromosome. Hum Genet 55:145–175.
2. Goedbloed M, Vermeulen M, Fang RN et al. (2009) Comprehensive mutation analysis of 17 Y-chromosomal short tandem repeat polymorphisms included in the AmpF/STR® Yfiler® PCR amplification kit. Int J Legal Med 123:471–482.
3. Gusmão L, Sánchez-Diz P, Alves C, et al. (2003) Grouping of Y-STR haplotypes discloses European geographic clines. Forensic Sci Int 134: 172–179.

4. Hammer MF, Karafet TM, Redd AJ, et al. (2001) Hierarchical patterns of global human Y-chromosome diversity. Mol Biol Evol 18:1189–203.
5. Jobling MA, Tyler-Smith C (2003) The human Y-chromosome: An evolutionary marker comes of age. Nat Rev Genet 4:598–612.
6. Kayser M, Caglià A, Corach D, et al. (1997) Evaluation of Y-chromosomal STRs: A multicenter study. Int J Legal Med 110:125–33.
7. Gusmão 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 Legal Med 120:191–200.
8. Miller SA, Dykes DD, Polesky HF (1988) A single salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 16:1215
9. Mulero JJ, Chang CW, Calandro LM, et al. (2006) Development and validation of the AmpFLSTR® Yfi ler™ PCR amplification kit: A male specifi c, single amplification 17 Y-STR multiplex system. J Forensic Sci 51:64–75.
10. Oota H, Settheetham-Ishida W, Tiwawech D, et al. (2001) Human mtDNA and Y-chromosome variation is correlated with matrilocal versus patrilocal residence. Nat Genet 29:20–1.
11. Quintana-Murci L, Krausz C, McElreavey K (2001) The human Y-chromosome: Function, evolution and disease. Forensic Sci Int 118:169–81.
12. Roewer L, Croucher PJ, Willuweit S, et al. (2005) Signature of recent historical events in the European Y-chromosomal STR haplotype distribution. Hum Genet 116:279–91.
13. Shi W, Ayub Q, Vermeulen M, et al. (2010) A worldwide survey of human male demographic history based on Y-SNP and Y-STR data from the HGDP-CEPH populations. Mol Biol Evol 27:385–93.
14. Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, et al. (2003) The male-specific region of the human Y-chromosome is a mosaic of discrete sequence classes. Nature 423:825–37.
15. Underhill PA, Shen P, Lin AA, et al. (2000) Y-chromosome sequence variation and the history of human populations. Nat Genet 26:358–61.
16. Thompson R, Zoppis S, McCord B (2012) An overview of DNA typing methods for human identification: past, present, and future. Methods Mol Biol 830:3–16.
17. YHRD: Y-Chromosome STR haplotype reference database. Available at: https://yhrd.org/pages/resources/mutation_rates. Accessed: March 2018.
18. Weber JL, Wong C (1993) Mutation of human short tandem repeats. Hum Mol Genet 2:1123-8.
19. Zhivotovsky LA, Feldman MW (1995) Microsatellite variability and genetic distances. Proc Natl Acad Sci USA 92:11549–52.
20. Zhivotovsky LA, Underhill PA, Cinnoglu C, et al. (2004) The effective mutation rate at Y-chromosome short tandem repeats, with application to human population-divergence time. Am J Hum Genet 74:50-61.