Forensic Science International: Genetics 3 (2009) 179–184

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Mutation rates at Y chromosome short tandem repeats in Texas populations

Jianye Ge [a,d], Bruce Budowle [b], Xavier G. Aranda [c], John V. Planz [c], Arthur J. Eisenberg [c], Ranajit Chakraborty [d,*]

[a] Department of Biomedical Engineering, University of Cincinnati, Cincinnati, OH 45221, USA
[b] Federal Bureau of Investigation, Quantico, VA 22135, USA
[c] UNT Center for Human Identification, University of North Texas Health Science Center at Ft. Worth, Ft. Worth, TX 76107, USA
[d] Center for Genome Information, Department of Environmental Health, University of Cincinnati College of Medicine, Cincinnati, OH 45267, USA

ARTICLE INFO

Article history:
Received 30 July 2008
Received in revised form 23 December 2008
Accepted 6 January 2009

Keywords:
DNA forensics
DNA typing
Short tandem repeat (STR)
Y chromosome
Mutation rate

ABSTRACT

Father–son pairs from three populations (African American, Caucasian, and Hispanic) of Texas were typed for the 17 Y STR markers DYS19, DYS385, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS456, DYS458, DYS635, DYS448, and Y GATA H4 using the AmpFlSTR® YfilerTM kit. With 49,578 allele transfers, 102 mutations were detected. One three-step and four two-step mutations were found, and all others (95.1%) were one-step mutations. The number of gains (48) and losses (54) of repeats were nearly similar. The average mutation rate in the total population is 2.1×10^{-3} per locus (95% CI $(1.7–2.5) \times 10^{-3}$). African Americans showed a higher mutation rate (3.0×10^{-3}; 95% CI $(2.4–4.0) \times 10^{-3}$) than the Caucasians (1.7×10^{-3}; 95% CI $(1.1–2.5) \times 10^{-3}$) and Hispanics (1.5×10^{-3}; 95% CI $(1.0–2.2) \times 10^{-3}$), but grouped by repeat-lengths, such differences were not significant. Mutation is correlated with relative length of alleles, i.e., longer alleles are more likely to mutate compared with the shorter ones at the same locus. Mutation rates are also correlated with the absolute number of repeats, namely, alleles with higher number of repeats are more likely to mutate than the shorter ones (p-value = 0.030). Finally, occurrences of none, one, and two mutations over the father–son transmission of alleles were consistent with the assumption of independence of mutation rates across loci.

1. Introduction

The Y chromosome short tandem repeat (STR) markers are inherited along the paternal lineage (i.e., from father to son), and have been a useful supplement to the autosomal STRs in forensic and kinship analyses [1–3]. Mutations can impact both forensic kinship calculations and evolutionary studies [3,4] and thus, mutation rates for a number of markers have been estimated from multigeneration pedigrees [5,6] or simply independent father–son pairs [7–17]. Both locus-specific mutation rates and allele-specific mutation trends have been investigated [8,13]. Sample sizes have been limited in most studies and hence more data are needed. Additionally, there has been little data comparing mutation rates across populations.

In this study, father–son pairs from the three largest subpopulations from Texas, i.e., African Americans, Caucasians, and Hispanics, were typed by the commercial kit AmpFlSTR® Yfiler™ kit (Applied Biosystems, Foster City, CA), which includes 17 Y-STR loci: DYS19, DYS385 (comprised of two loci), DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS456, DYS458, DYS635, DYS448, and Y GATA H4. The observed mutations for each locus in each population and average mutation rates were summarized. The relationships between mutation rates and relative or absolute length of alleles, as well as motif sizes, were investigated. The occurrences of mutations in 2913 father–son pairs across the loci were also checked for consistency with independence of mutation rates across loci. Lastly, the mutation rates between the sample populations were compared.

2. Materials and methods

2.1. Samples and analysis

DNA was obtained from unrelated male donors and their confirmed sons from paternity testing cases submitted to the DNA Identification Lab at the University of North Texas Health Science Center, Ft. Worth, Texas. Population affinity was ascribed by self-declaration. The samples are from African Americans, $N = 950$

* Corresponding author at: Center for Genome Information, Department of Environmental Health, University of Cincinnati College of Medicine, 3223 Eden Avenue, Kettering Laboratories, Room K-110, Cincinnati, OH 45267-0056, USA. Tel.: +1 513 558 4925; fax: +1 513 558 4505.
E-mail addresses: chakrar@ucmail.uc.edu, ranajit.chakraborty@uc.edu (R. Chakraborty).

1872-4973/$ – see front matter © 2009 Published by Elsevier Ireland Ltd.
doi:10.1016/j.fsigen.2009.01.007

father–son pairs; Caucasians, $N = 957$ father–son pairs; and Hispanics, $N = 1006$ father–son pairs. Paternity was confirmed by using autosomal STR typing prior to and independent of Y-STR typing. Father–son relationship was asserted from no exclusion based on 13 CODIS autosomal markers and with a minimum PI of 1000, prior to and independent of Y-STR typing.

Sample collection, DNA extraction, and Y-STR typing were performed as described previously [19].

2.2. Statistical analysis

The Y-STR mutation rate at each locus was calculated as the number of mutations divided by the number of allele transfers from father to son. In mutation counting, there were several father–son pairs where one-step slippages seen for both DYS389I and DYS389II, for instance (14, 29) → (15, 30). These were treated as one mutation instead of two because these two loci are the result of a tandem duplication. We followed the common practice of reporting point estimates of the mutation rates, along with their confidence intervals (CIs) [2,4–7,9–14,17], the later estimated using the theory of exact binomial probability distribution (see http://statpages.org/confint.html).

To assess the relationship of relative allele sizes and mutation rates, the allele sizes were arbitrarily categorized into 25%, 50%, and 25% categories for short, moderate and long allele sizes, respectively. Two-way ANOVA was conducted to examine if the population differences of average mutation rate differences over all loci persist when the progenitor alleles (i.e., alleles in fathers) are grouped by repeat size of alleles. Finally, expected numbers of mutations over the 16 loci for 2913 father–son transmission of alleles were computed with the assumption that mutations are independent across loci using the following algorithm.

Let $m_1, m_2, \ldots m_{16}$ represent the estimated mutation rate over the 16 loci. Under the assumption of independence of mutations at each locus, the expected numbers of father–son pairs involving zero, one, and two or more mutations are given by

$$E(N_0) = 2913 \times \prod_{i=1}^{16}(1 - m_i),$$

$$E(N_1) = 2913 \times \sum_{i=1}^{16} m_i \prod_{j \neq i}(1 - m_j) = E(N_0) \times \sum_{i=1}^{16} \frac{m_i}{1 - m_i}, \quad \text{and}$$

$$E(N_{\geq 2}) = 2913 - E(N_0) - E(N_1)$$

respectively, which were compared with the observed counts of father–son pairs with no, one, and two or more mutations (N_0, N_1, and $N_{\geq 2}$) to check the assumption of independence. Comparison of this trinomial classification of occurrences of mutations with the expectation was done by computing the simultaneous 95% confidence interval estimates of cell frequencies by using the algorithm developed by Sison and Glaz [18], which takes into account the dependencies of the trinomial cell frequencies. Note that Dupuy et al. [8] used a similar approach to validate occurrences of two or more mutations for father–son pairs, while the above approach is applicable for the entire distribution of number mutations for each father–son pair.

3. Results and discussion

This study analyzed 2913 father–son pairs from three sample populations residing in Texas. While the numbers of allele transfers were nearly the same for the African American (16,165), Caucasian (16,275), and Hispanic (17,138) samples, 49, 28, and 25 mutations were observed in the respective populations (Table 1). In total, 102 mutations were observed in 49,578 allele

Table 1
Mutation count, rates and 95% confidence interval (CI) for the 17 Y-STR loci studied in Texas populations.

Locus	African American				Caucasian				Hispanic				Total			
	Count	Allele	Mutation rate $\times 10^{-3}$	95% CI $\times 10^{-3}$	Count	Allele	Mutation rate $\times 10^{-3}$	95% CI $\times 10^{-3}$	Count	Allele	Mutation rate $\times 10^{-3}$	95% CI $\times 10^{-3}$	Count	Allele	Mutation rate $\times 10^{-3}$	95% CI $\times 10^{-3}$
DYS19	2	950	2.1	0.3–7.6	0	958	0	0–3.8	0	1,010	0	0–3.7	2	2,918	0.7	0.1–2.5
DYS385	7	1,900	3.7	1.5–7.6	3	1,914	1.6	0.3–4.6	3	2,012	1.5	0.3–4.4	13	5,826	2.2	1.2–3.8
DYS389I	3	950	3.2	0.7–9.2	1	957	1.0	0.0–5.8	1	1,016	1.0	0.0–5.5	5	2,923	1.7	0.6–4.0
DYS389II	5	951	5.3	1.7–12.2	1	958	1.0	0.0–5.8	4	1,009	4.0	1.1–10.1	10	2,918	3.4	1.6–6.3
DYS390	1	950	1.1	0.0–5.9	1	957	1.0	0.0–5.8	0	1,007	0	0–3.7	2	2,914	0.7	0.1–2.5
DYS391	2	950	2.1	0.3–7.6	2	957	2.1	0.3–7.5	0	1,006	0	0–3.7	4	2,913	1.4	0.4–3.5
DYS392	0	950	0	0–3.9	0	957	0	0–3.9	0	1,006	0	0–3.7	0	2,913	0	0–1.3
DYS393	2	950	2.1	0.7–7.6	2	957	2.1	0.3–7.5	0	1,006	0	0–3.7	4	2,913	1.4	0.4–3.5
DYS437	1	950	1.1	0.0–5.9	2	957	2.1	0.3–7.5	0	1,013	0	0–3.6	3	2,920	1.0	0.2–3.0
DYS438	0	950	0	0–3.9	0	957	0	0–3.9	0	1,006	0	0–3.7	0	2,913	0	0–1.3
DYS439	6	951	6.3	2.3–13.7	0	958	0	0–3.9	6	1,017	5.9	2.2–12.8	12	2,926	4.1	2.1–7.2
DYS448	3	962	3.1	0.6–9.1	2	959	2.1	0.3–7.5	3	1,006	3.0	0.6–8.7	8	2,927	2.7	1.2–5.4
DYS456	4	950	4.2	1.2–10.7	8	957	8.4	3.6–16.4	0	1,006	0	0–3.7	12	2,913	4.1	2.1–7.2
DYS458	9	951	9.5	4.3–17.9	1	957	1.0	0.0–5.8	6	1,006	6.0	2.2–12.9	16	2,914	5.5	3.1–8.9
DYS635	3	950	3.2	0.7–9.2	3	958	3.1	0.7–9.1	0	1,006	0	0–3.7	6	2,914	2.1	0.8–4.5
GATA_H4	1	950	1.1	0.0–5.9	2	957	2.1	0.3–7.5	2	1,006	2.0	0.2–7.2	5	2,913	1.7	0.6–4.0
Summary	49	16,165	3.0	2.2–4.0	28	16,275	1.7	1.1–2.5	25	17,138	1.5	0.9–2.2	102	49,578	2.1	1.7–2.5

transfers. Among these father–son pairs, 96 mutations (94.1%) occur singly in different father–son pairs and three father son pairs had two mutations. There were 98 mutations that resulted in a one repeat difference (Table 2), which is consistent with the general notion that the majority of mutations comprise single step repeat gain or loss due to strand slippage during replication [9,14]. Four mutations were two-step changes and only one three-step mutation was observed, which was at the DYS456 locus. Gains and losses of repeats (48 and 54, respectively) were approximately equal. Hence, the data herein support that mutations at these Y chromosome microsatellites do not have any contraction or expansion bias. In contrast, Gusmao et al. [13] reported an expansion bias from a smaller number of mutations at these loci. However, with classifying the progenitor alleles into short, moderate, and long allele sizes (discussed later), more mutations with allele size gain were observed for short and moderate alleles and more losses for long alleles (Fig. 1) in total population, which is consistent with the results of Dupuy et al. [8].

The average mutation rate of the total population sample across all markers is 2.1×10^{-3} (95% CI $(1.7–2.5) \times 10^{-3}$). This rate is higher than the 1.57×10^{-3} by Budowle et al. [14] also on the Texas subpopulations but on fewer loci, and similar to 2.1×10^{-3} by Heyer et al. [5], 2.3×10^{-3} by Dupuy et al. [8], 1.89×10^{-3} by Hohoff et al. [11] and 2.36×10^{-3} by Lee et al. [12], 2.2×10^{-3} by Sánchez-Diz et al. [17], but significantly lower than 2.8×10^{-3} by Kayser et al. [7] and 3.1×10^{-3} by Decker et al. [16]. The mutation rate, based on the confidence interval bounds, was significantly higher in African Americans compared with Hispanics and approaching significance between African Americans and Caucasians. Decker et al. [16] also observed such higher mutation rates in African Americans.

Point estimates of mutation rates vary across the markers, and locus specific mutation rates are not significantly different from previous investigations of this type [11,12]. However, when inter-locus variations are examined with their respective confidence-bounds, most of these differences of point estimates are not significant. The skewed distribution of motif sizes in these 17 loci (Table 3) prevents any meaningful analysis of relationship between mutation rate and motif size. The average mutation rates of 14 tetranucleotide loci and one hexanucleotide locus (e.g., DYS448) were 2.5×10^{-3} (95% CI $(2.0–3.0) \times 10^{-3}$) and 2.7×10^{-3} (95% CI $(1.2–5.4) \times 10^{-3}$), respectively. Considering the point estimates alone, the DYS458 locus had the highest mutation rate at 5.5×10^{-3} in the total population, 9.5×10^{-3} in African Americans and 6.0×10^{-3} in Hispanics. The highest mutation rate in Caucasians was 8.4×10^{-3} observed at the DYS456 locus (Table 1). Gusmao et al. [13] also made observations similar to this. Neither of these loci has allele sizes that are larger compared with the other loci in terms of range of repeats as well as average length (Table 3).

To assess the relationship between allele sizes and mutation rates, alleles were simply classified into short, moderate and long categories by approximately partitioning them into 25%, 50% and 25% quantiles, respectively (Table 4). In the total population, the mutation rate of long alleles (2.7×10^{-3}) is significantly greater than short (2.1×10^{-3}) and moderate (2.2×10^{-3}) alleles. The lowest mutation rates generally were associated with short alleles in the Hispanic sample, and long alleles were associated with the highest mutation rates in the African American sample. The differences were not significant between the Caucasian and Hispanic samples.

Table 2
Stepwise distribution of mutation repeats.

Step	Population			
	African	Caucasian	Hispanic	Total
−3		1		1
−2	2		1	3
−1	23	11	16	50
1	24	16	8	48
Sum	49	28	25	102

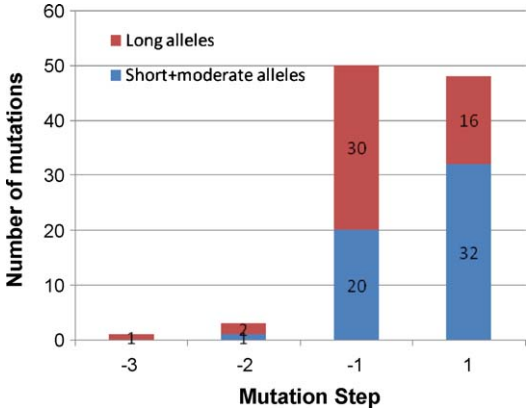

Fig. 1. Distribution of mutations with mutation steps for long and short + moderate alleles.

Table 3
Motif and motif sizes for the 17 Y STR studied.

Marker	Motif size	Motif	Mutation count	Avg. allele size (range)[a]
DYS19	4	TAGA	2	14.50 (<10–18)
DYS385	4	GAAA	13	14.27 (8–22)
DYS389I	4	(TCTG) (TCTA)	5	13.00 (9–15)
DYS389II	4	(TCTG) (TCTA) (TCTG) (TCTA)	10	29.58 (25–34)
DYS390	4	(TCTA) (TCTG)	2	23.11 (20–26)
DYS391	4	TCTA	4	10.37 (7–13)
DYS392	3	TAT	0	12.25 (8–18)
DYS393	4	AGAT	4	13.19 (9–17)
DYS437	4	TCTA	3	14.59 (13–18.2)
DYS438	5	TTTTC	0	11.10 (8–>13)
DYS439	4	AGAT	12	11.79 (8–15)
DYS448	6	AGAGAT	8	19.60 (15-24)
DYS456	4	AGAT	12	15.36 (11–>18)
DYS458	4	GAAA	16	16.62 (13–22)
DYS635	4	T**S**TA compound	6	22.12 (19–>26)
GATA_H4	4	TAGA	5	11.52 (8–15)

[a] DYS19, DYS437, DYS438, DYS448 and DYS456 contain null alleles; null alleles were not counted in calculation of average allele sizes.

Table 4
Mutation counts and rates by relative allele sizes (short, moderate, and long) for each locus in the Texas populations. The allele sizes are generally categorized into 25%, 50%, and 25% quantiles for short, moderate and long allele sizes, respectively.

Locus	African			Caucasian			Hispanic			Total		
	Short	Moderate	Long	Short	Moderate	Long	Short	Moderate	Long	Short	Moderate	Long
DYS19			2									2
DYS385		2	5	1	1	1	1	1	1	2	4	7
DYS389I		2	1		1				1		3	2
DYS389II		1	4			1		1	3		2	8
DYS390		1		1						1	1	
DYS391		2				2					2	2
DYS392												
DYS393			2		2						2	2
DYS437		1			2						3	
DYS438												
DYS439		3	3				1	2	3	1	5	6
DYS448	1		2	1	1			2	1	2	3	3
DYS456		1	3	1	1	6				1	2	9
DYS458	1	2	6	1			1	4	1	3	6	7
DYS635		2	1	3						3	2	1
GATA_H4	1			1	1			2		2	3	0
Total	3	17	29	9	9	10	3	12	10	15	38	49
No. of allele	3209	7951	5005	4308	8378	3589	4560	7882	4696	12,077	24,211	13,290
Mutation rate $\times 10^{-3}$	0.9	2.1	5.8	2.1	1.1	2.8	0.7	1.5	2.1	1.2	1.6	3.7
95% CI $\times 10^{-3}$	0.2–2.7	1.3–3.4	3.9–8.3	1.0–4.0	0.5–2.0	1.3–5.1	0.1–1.9	0.8–2.7	1.0–3.9	0.7–2.1	1.1–2.2	2.7–4.9

Categorization based on absolute length of allele repeat size (i.e. the number of motif repeats) with the same general quantile proportions as above also show that the long and moderate alleles are more likely to mutate than short alleles in the total population. No notable difference between moderate and long alleles was observed (Table 5). In spite of a somewhat (relatively) lower proportions of small alleles (<13 repeats) in African Americans (29.6%, as opposed to 37.2% in Caucasians), they exhibited the largest difference in mutation rates between short and long alleles. The short and long categories separated by allele 11 have significantly different mutation rates, which is comparable to results as reported previously by Hohoff et al. [11].

As the mutation rate across the 16 loci differed by locus, and there was an apparent inter-population rate difference as well, all observed mutations were organized in a two-way table by their occurrences in populations and repeat length of alleles. Two-way ANOVA showed that when progenitor allele size of the 102 mutations were measured in terms of the repeat size of alleles, inter-population difference was insignificant (F-ratio = 0.02, $p = 0.995$), with repeat size effect being predominant factor of inter-locus variation of mutation rate (F-ratio = 59350.5, $p < 0.001$). Therefore, merging the three populations together, Fig. 2 shows the correlation between mutation rates and repeat number of alleles, in which 18 null alleles and 51 alleles with ">" or "<" were not counted since the lengths of these allele are zeros or not precisely determined. 173 alleles whose counts are less than 100 (except "33" because one mutation occurs at "33") were also not counted because of limited number of alleles, which include 66 fractional alleles as well, none of which is involved with mutations. The p-value for the hypothesis of no correlation between the mutation rate and repeat number is 0.030. The highest mutation rate of alleles was found at the DYS389II marker with one mutation in 22 alleles of "33", which is 20-fold greater than the average mutation rate. These results are consistent with the concept that longer alleles have generally higher mutation rates. Because the major mechanism of STR mutations is replication slippage [20] and 99.5% alleles were included in this comparison, removing 107 minor integer alleles and 66 fractional alleles would not significantly change the correlation between mutation rates and repeat number. In fact, when these 173 alleles were included, though the correlation is somewhat weaker (0.394, as opposed to 0.453) the p-value was even smaller (i.e., 0.01).

Finally, of the 2913 father–son pairs, 2814 showed no mutation at any of the 16 loci, 96 showed one mutation each, and three pairs showed mutations at two loci each (not necessarily the same). Comparison of this distribution with the expected under the assumption of independence of mutation rates across loci by the method described earlier is shown in Table 6, along with the simultaneous 95% confidence limits of the three classes. No significant deviation for any of the three classes was noted suggesting that mutations across the 16 loci are independent. Absence of significant deviation of occurrences of number of mutations for each father–son pair also reasserts the accuracy of father–son status determination based on observations of no exclusion at 13 CODIS autosomal STR loci prior to and independent of Y-STR typing.

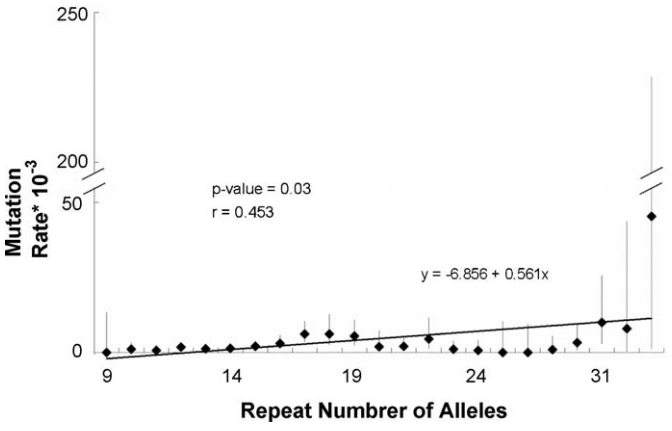

Fig. 2. Correlation between mutation rates and repeat number of alleles (see text for details).

Table 5
Mutation counts and rates by absolute allele sizes across all markers. There are two categorizations. The first partitioning is the approximate 25%, 50%, 25% quantiles by count; the second partitioning is divided into those less than or equal to allele 11 and those greater than allele 11.

Allele size	African American				Caucasian				Hispanic				Total			
	Count	Allele	Mutation rate $\times 10^{-3}$	95% CI $\times 10^{-3}$	Count	Allele	Mutation rate $\times 10^{-3}$	95% CI $\times 10^{-3}$	Count	Allele	Mutation rate $\times 10^{-3}$	95% CI $\times 10^{-3}$	Count	Allele	Mutation rate $\times 10^{-3}$	95% CI $\times 10^{-3}$
Short (7–<13)	6	4,779	1.3	0.5–2.7	5	6,052	0.8	0.3–1.9	6	4,935	1.2	0.5–2.6	17	14,633	1.2	0.7–1.9
Moderate (13–19.2)	30	7,825	3.8	2.6–5.5	18	6,993	2.6	1.5–4.1	14	8,697	1.6	0.9–2.7	62	24,626	2.5	1.9–3.2
Long (20–34)	13	3,561	3.7	2.0–6.2	5	3,230	1.5	0.5–3.6	5	3,506	1.4	0.5–3.3	23	10,319	2.2	1.4–3.3
Short (7–11)	3	3,447	0.9	0.2–2.5	4	3,229	1.2	0.3–3.2	1	3,041	0.3	0.0–1.8	8	9,465	0.8	0.4–1.7
Long (11.3–34)	36	12,718	2.8	2.0–3.9	24	13,046	1.8	1.2–2.7	24	14,097	1.7	1.1–2.5	94	40,113	2.3	1.9–2.9

Table 6
Observed and expected distributions of father–son pairs grouped by occurrences of mutations over 16 Y-STR loci.

Number of mutation events	Number of father–son pairs		
	Observed	Expected	95% CI
0	2813	2818.0	2793.8–2835.7
1	96	92.6	76.9–118.5
≥2	3[a]	1.4	0.7–25.6
Total	2912	2912.0	

[a]No father–son pair exhibited more than two mutations over the 16 loci.

4. Conclusion

This study adds to the growing data on Y-STR mutation rates. The average Y-STR mutation rate for the combined Texas populations is approximately 2.1×10^{-3} per locus (95% CI $(1.7–2.5) \times 10^{-3}$), which is comparable to other studies. Taken all observed mutations together, Y-STR mutation rates, however, differed between African Americans and those of Caucasians and Hispanics with the former being notably higher. Most mutations resulted in one-step repeat differences consistent with the strand slippage model for generating mutations with STRs. No bias for gains or losses of repeats was observed which differs from previous studies. Mutation rate is correlated with length of alleles with longer alleles generally more likely to mutate compared with shorter alleles. When the mutations (over all loci together) were grouped by absolute repeat sizes of the progenitor alleles (i.e., fathers' allele sizes), inter-population differences of mutation rates became insignificant, while the positive correlation between mutation rate and allele size ($r = 0.453$) remained significant ($p = 0.030$). Finally, mutations across the different Y-STR loci appear to be independent of each other.

Acknowledgement

Data analysis of this study was partially funded by US Public Health Service Research Grant GM41399 from the National Institutes of Health.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.fsigen.2009.01.007.

References

[1] B. Budowle, S.K. Sinha, H.S. Lee, R. Chakraborty, Utility of Y-chromosome STR haplotypes in forensic applications, Forensic Sci. Rev. 15 (2003) 153–164.
[2] M.A. Jobling, A. Pandya, C. Tyler-Smith, The Y chromosome in forensic analysis and paternity testing, Int. J. Legal Med. 110 (1997) 118–124.
[3] J. Buckleton, C. Triggs, S. Walsh, Forensic DNA Evidence Interpretation, CRC Press, Boca Raton, FL, 2005.
[4] M. Kayser, A. Sajantila, Mutations at Y-STR loci: implications for paternity testing and forensic analysis, Forensic Sci. Int. 118 (2001) 116–121.
[5] E. Heyer, J. Puymirat, P. Dieltjes, E. Bakker, P. de Knijff, Estimating Y chromosome specific microsatellite mutation frequencies using deep rooting pedigrees, Hum. Mol. Genet. 6 (1997) 799–803.
[6] M.A. Jobling, E. Heyer, P. Dieltjes, P. de Knijff, Y chromosome-specific microsatellite mutation rates re-examined using a minisatellite, MSY1, Hum. Mol. Genet. 8 (1999) 2117–2120.
[7] M. Kayser, L. Roewer, M. Hedman, L. Henke, J. Henke, S. Brauer, C. Krüger, M. Krawczak, M. Nagy, T. Dobosz, R. Szibor, P. de Knijff, M. Stoneking, A. Sajantila, 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 (2000) 1580–1588.
[8] B.M. Dupuy, M. Stenersen, T. Egeland, B. Olaisen, Y-chromosomal microsatellite mutation rates: differences in mutation rate between and within loci, Hum. Mutat. 23 (2004) 117–124.

[9] R. Kurihara, T. Yamamoto, R. Uchihi, et al., Mutations in 14 Y-STR loci among Japanese father–son haplotypes, Int. J. Leg. Med. 118 (2004) 125–131.
[10] D.J. Ballard, C. Phillips, G. Wright, et al., A study of mutation rates and the characterisation of intermediate, null and duplicated alleles for 13 Y chromosome STRs, Forensic Sci. Int. 155 (2005) 65–70.
[11] C. Hohoff, K. Dewa, U. Sibbing, K. Hoppe, P. Forster, B. Brinkmann, Y-chromosomal microsatellite mutation rates in a population sample from northwestern Germany, Int. J. Legal Med. 121 (2007) 359–363.
[12] H.Y. Lee, M.J. Park, U. Chung, H.Y. Lee, W.I. Yang, S.H. Cho, K.J. Shin, Haplotypes and mutation analysis of 22 Y-chromosomal STRs in Korean father–son pairs, Int. J. Legal Med. 121 (2007) 128–135.
[13] L. Gusmao, P. Sanchez-Diz, F. Calafell, et al., Mutation rates at Y chromosome specific microsatellites, Hum. Mutat. 26 (2005) 520–528.
[14] B. Budowle, M. Adamowicz, X.G. Aranda, C. Barna, R. Chakraborty, D. Cheswick, B. Dafoe, A. Eisenberg, R. Frappier, A.M. Gross, C. Ladd, H.S. Lee, S.C. Milne, C. Meyers, M. Prinz, M.L. Richard, G. Saldanha, A.A. Tierney, L. Viculis, B.E. Krenke, Twelve short tandem repeat loci Y chromosome haplotypes: genetic analysis on populations residing in North America, Forensic Sci. Int. 150 (2005) 1–15.
[15] M.S. Shi, J.P. Tang, R.F. Bai, X.J. Yu, J.Y. Lv, B. Hu, Haplotypes of 20 Y-chromosomal STRs in a population sample from southeast China (Chaoshan area), Int. J. Legal Med. 21 (2007) 9827–9837.
[16] A. Decker, M. Kline, J. Redman, T. Reid, J. Butler, Analysis of mutations in father–son pairs with 17 Y-STR loci, Forensic Sci. Int. Genet. 2 (2008) e31–e35.
[17] P. Sánchez-Diz, Population and segregation data on 17 Y-STRs: results of a GEP-ISFG collaborative study, Int. J. Legal Med. 122 (2008) 529–533.
[18] C.P. Sison, J. Glaz, Simultaneous confidence intervals and sample size determination for multinomial proportions, J. Am. Statist. Ass. 90 (1995) 366–369.
[19] B. Budowle, J. Ge, X. Aranda, J. Planz, A. Eisenberg, R. Chakraborty, Texas population substructure and its impact on estimating the rarity of Y STR haplotypes from DNA evidence, J. Forensic Sci., in press.
[20] D. Pumpernik, B. Oblak, B. Borštnik, Replication slippage versus point mutation rates in short tandem repeats of the human genome, Mol. Genet. Genomics 279 (1) (2008) 53–61.