Y-STR Mutation Patterns in North Indian Male Relatives at 16 Loci: A Preliminary Study

Vikash Chandra Mishra*, Dinesh Chandra, Anoushka Raina, Ritu Sharma, Diksha Dhuliya, Vimarsh Raina

Department of Molecular Genetics and Transplant Immunology, Chimera Transplant Research Foundation, New Delhi, India

Corresponding author:

Vikash Chandra Mishra (PhD)
Department of Molecular Genetics and Transplant Immunology,
Chimera Transplant Research Foundation, New Delhi, India
Email id: vikashbiotech01@gmail.com

ABSTRACT

Y-chromosomal short tandem repeats (Y-STRs) are passed from father to son through the male line only. They change very little over generations and are widely used in forensic science and for tracing paternal ancestry. However, mutations at Y-STR loci can complicate relationship analyses, and may result in inaccurate exclusions during father-son or extended paternal lineage testing. In this preliminary study, we looked for possible Y-STR mutations in 292 cases from the North Indian population. These cases included sibling and grandparentage, uncles and nephews, first cousins, and second cousins. We found a total of 29 mutations across 11 Y-STR markers out of 16 markers. All mutations were single-step changes consistent with the stepwise mutation model. The highest mutation rates were at DYS458 (0.006) and DYS385 (0.005), while loci such as DYS19, DYS390, and DYS389II had lower rates (0.001). These results align with previous findings on mutation rates at different loci. They also support the observation that Y-STRs with rapid mutations provide better information for distinguishing closely related paternal lineages. Our findings give baseline estimates of mutation rates for the North Indian populace, highlighting the need for population-specific data in forensic work and genealogical research.

Keywords: Y-chromosomal short tandem repeats; Y-STR Mutation; Y-STR loci; Y-STR markers; forensic genetics

1. Introduction

Y-chromosomal short tandem repeats (Y-STRs) are commonly used in forensic genetics, paternity testing, and anthropological research because they are inherited from father to son and recombine little across most of the Y chromosome [1,2]. About 95% of the Y chromosome is made up of the non-recombining region (NRY). Mutational changes at Y-STR loci are the primary source of highly informative, individualizing variation for forensic analysis and the establishment of recent paternal kinship [2,3]. This makes Y-STRs very useful for distinguishing male relatives, identifying male contributions in mixed DNA samples, and tracing paternal family lineages. The primary mechanism of mutation at STRs is polymerase slippage during DNA replication. This usually results in either gaining or losing a single repeat unit [4,5]. Even though these mutations do not happen often, it is important to understand them for correct forensic genetics. Y-STRs can help distinguish closely related paternal lineages; however, if these mutations are not considered, they can lead to incorrect exclusions in father-son paternity tests [1,6]. Therefore, knowledge of mutation rates for different Y-STR markers is necessary for precise forensic conclusions and the development of population-specific databases. Recent large-scale pedigree studies and forensic reviews have highlighted the significance of population-specific mutation rate data for accurate kinship interpretation and likelihood ratio estimation [1,7]. The present study examines Y-STR mutations in 292 paternal cases from North India. By comparing Y-STR profiles of close and distant paternal relatives, the study provides insights into Y-STR mutation patterns, locus-specific differences, and their relevance in forensic genetic investigations.

2. Materials and methods

2.1 Study population

This study represents a retrospective analysis of anonymized data derived from routine casework conducted between August 2016 and December 2024 at an ISO 15189-accredited laboratory in North India. A total of 292 paternal lineage cases were included, all originating from the North Indian population. Institutional Ethics Committee approval was obtained prior to data analysis and manuscript preparation. Written informed consent had been obtained from all participants at the time of sample collection during routine casework.

2.2 Sample Collection and DNA Extraction

Three millilitres (3 mL) of peripheral blood were collected in EDTA vacutainers from BD Vacutainer®, USA, as part of routine casework for Y-STR testing. No additional sample collection or experimental intervention was performed specifically for this study. This was followed by DNA extraction using the Macherey-Nagel (MN) solid-phase extraction kit from Germany, according to the manufacturer's instructions. The amount of DNA was measured using a Qubit® 2.0 Fluorometer (Invitrogen, USA).

2.3 PCR Amplification and Genotyping

Genotyping was carried out using the Yfiler™ PCR Amplification Kit (Thermo Fisher Scientific, USA) with ~1 ng of template DNA. PCR cycling was performed on a Veriti Thermal Cycler under standard conditions: initial denaturation at 95°C for 11 minutes, followed by 30 cycles of 94°C for 1 minute, 61°C for 1 minute, and 72°C for 1 minute, and final extension at 60°C for 80 minutes. PCR products were analyzed on a SeqStudio Genetic Analyzer after mixing with Hi-Di™

formamide and LIZ® size standard. Allele calling was performed using GeneMapper® v5 software.

2. 4 Statistical Analyses

Mutation rates were calculated using a direct counting method, whereby the number of observed mutational events was divided by the total number of meioses analyzed per locus. Per-locus and overall mutation rates were reported to three decimal places. Exact 95% confidence intervals were estimated using the Clopper–Pearson method under a binomial distribution framework. Data tabulation and preliminary calculations were performed using Microsoft Excel.

3. Results

A total of 292 pedigreed cases were analyzed. This included 66 sibling and grandparentage cases, 133 uncle-nephew cases, 70 first-cousin cases, and 23 second-cousin cases. Across these cases, 29 mutational events were identified at 11 Y-STR loci out of the 16 Y-STR loci examined, as shown in Table 1. No mutations were found at five loci: DYS389I, DYS392, DYS437, DYS438, and DYS448. This indicates low mutation rate/stability of these five loci (DYS389I, DYS392, DYS437, DYS438, DYS448) in the studied population. All observed mutations were single-step changes, which fit the stepwise mutation model, as summarized in Table 2 The proportion of cases showing at least one mutation increased with greater genetic distance. An increasing proportion of mutation-positive cases was observed with increasing meiotic distance (siblings < uncle–nephew < first cousins < second cousins), consistent with stepwise accumulation of mutations across generations. The highest proportion was observed in second-cousin pairs (13.04%), as shown in Table 1. This increasing trend across relationship categories is illustrated in Figure 1. Per-locus mutation rates ranged from 0.001 to 0.006, as presented in Table 3. The highest rates were found

at DYS458 (0.006) and DYS385 (0.005), followed by DYS391, GATA-H4, and DYS365.The distribution of per-locus mutation rates is illustrated in Figure 2.

4. **Discussion**

This study provides data on Y-STR mutational patterns in the North Indian populace. It contributes to the global dataset on paternal lineage markers. All the Y-STR mutations seen in this study were single-step changes, consistent with the stepwise mutation model of STR evolution [5,8]. In comparison, five Y-STR markers, DYS389I, DYS392, DYS437, DYS438, and DYS448, showed no mutations. This shows that these markers are stable and reliable for confirming paternal relationships in forensic work. The high number of single-step mutations is also similar to results reported in many other populations. This suggests that Y-chromosomal STRs undergo a similar mutation process across diverse human groups [1,2].

Mutation rates at the generally used Y-STR loci are largely comparable across global populations. However, regional validation remains important in kinship analysis [1,3]. Even small population-specific differences in mutation rates can affect likelihood ratio calculations and kinship interpretation in relationship cases [3,6]. Therefore, estimates of mutations from regional populations help improve the statistical standardization and reliability of paternal lineage testing in Indian forensic casework. The mutation rate observed in this study is between 0.001 and 0.006 per meiosis. This range is consistent with what has been reported in studies of European and East Asian lineages, including large datasets comparing fathers and sons [1,2,9]. Moreover, haplogroup-stratified analyses have demonstrated modest inter-haplogroup variability in mutation behavior, supporting the value of region-specific datasets for improved forensic standardization [3,7,8]. Rapidly mutating loci such as DYS458 and DYS385 showed higher mutation frequencies, supporting their greater discriminatory power for discriminating closely related paternal ancestries [10,11,12]. On the other hand, unchanging Y-STR loci are still important for confirming paternal ancestry. Y chromosomal STR loci such as DYS19 and DYS389II exhibited low mutation rates,

making them more consistent for paternal lineage confirmation but less useful for discriminating between closely related males. Interestingly, more mutations were seen in cases involving more distant paternal relatives, such as first and second degree cousins. An increase in mutation frequency with greater meiotic separation was observed across different relationship categories (siblings & grandparentage, uncle–nephew, first cousins and second cousins). The observed trend confirms the expected pattern of gradual mutation buildup over generations. While the current sample size restricts formal modeling of mutation rates, the pattern is consistent with results from extended lineages analyses, including large family studies that have demonstrated inconsistency in Y-STR mutation behavior across haplogroups [2].

From a forensic casework perspective, locus-specific mutation rate estimates are critical for accurate statistical interpretation of paternal lineage testing. Rapidly mutating Y-STR loci enhance discrimination among closely related males, particularly in complex kinship investigations involving extended paternal relatives. However, single-step mutations may result in apparent mismatches in father–son comparisons. Incorporating empirically derived mutation rates into likelihood calculations reduces the risk of false exclusions and strengthens the reliability of forensic conclusions [3,6,13] Future studies with larger numbers of samples and detailed haplogroup information will help improve population-specific Y-STR mutation rate estimates for India. Additionally, the use of high-resolution Y-STR sequencing may also help explain specific mutation patterns [5]. These efforts will improve the practical use of Y-STRs in forensic work and also support anthropological and family history studies in North Indian populations.

5. Conclusion

To conclude, this study gives Y-STR mutation rate estimates for specific Y-STR loci in the North Indian population. All events observed were single-step mutations. The highest rates were observed at DYS458 and DYS385, which align with reports from other populations. These findings highlight the importance of accounting for rapidly mutating Y-STR loci when differentiating closely related paternal lineages, while

recognising the potential for mutational mismatches in kinship testing. Establishing population-specific Y-STR mutation data will enhance the reliability of forensic casework and genealogical applications in India.

6. Limitation

This study has certain limitations. First, haplogroup testing was not performed as part of routine casework, and therefore, haplogroup-specific differences in mutation rates could not be assessed. Given that Y-STR mutation behaviour can differ among haplogroups, this may influence the generalizability of our findings. Second, the use of the 16-locus Y-filer kit restricts the evaluation of rapidly mutating Y-STRs that are included in the Y-filer Plus system. Future studies incorporating haplogroup classification and extended marker sets will help strengthen population-specific mutation estimates.

Funding: This research received no external funding.

Conflicts of interest There are no conflicts of interest to declare.

List of Abbreviations

Y-STRs: Y-chromosomal short tandem repeats
NRY: Non-Recombining Region

Author Contributions

Conceptualization: V.C.M and V.R; Methodology: D. C and A. R; Formal Analysis: V.C.M., D.C. and R.S; Data curation: D.C and D. D; Writing-original draft, Writing-review & editing: V. C. M, D. D, A. R and R.S; Visualization: V.R; Supervision: V.R and V.C.M; Project administration: V.C.M. and V.R. All authors read and approved the submitted version.

Availability of Data and Materials

The data supporting the findings of this study are included within the article.

Ethics Committee Approval and Consent to Participate

This retrospective study was approved by the institutional review board (CH/002/2025, approval date: 22 March 2025). Written informed consent had been obtained from all participants at the time of routine casework, including consent for participation and for publication purposes.

Human Rights Statement: The study was conducted in accordance with the Declaration of Helsinki.

Consent for Publication

All data presented in this manuscript are anonymized and do not contain any identifiable personal information or human images.

Acknowledgements: We are grateful to our laboratory colleagues for their kind support**.**

AI-Declaration

While preparing this manuscript, the authors used Grammarly to improve readability and language. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References:

[1] Lee DG, Kim SJ, Cho WC, Cho Y, Park JH, Lee J, Jung JY. Analysis of mutation rates and haplotypes of 23 Y-chromosomal STRs in Korean father–son pairs. Forensic Science International: Genetics. 2023 Jul 1;65:102875.

[2] Claerhout S, Vandenbosch M, Nivelle K, Gruyters L, Peeters A, Larmuseau MH, Decorte R. Determining Y-STR mutation rates in deep-routing genealogies: Identification of haplogroup differences. Forensic Science International: Genetics. 2018 May 1;34:1-0.

[3] Kayser M. Forensic use of Y-chromosome DNA: a general overview. Human genetics. 2017 May;136(5):621-35.

[4] Schlötterer C, Tautz D. Slippage synthesis of simple sequence DNA. Nucleic acids research. 1992 Jan 25;20(2):211-5.

[5] Antão-Sousa S, Amorim A, Gusmão L, Pinto N. Mutation in Y STRs: Repeat motif gains vs. losses. Forensic Science International: Genetics Supplement Series. 2019 Dec 1;7(1):240-2.

[6] Pinto N, Gusmao L, Amorim A. Mutation and mutation rates at Y chromosome specific short tandem repeat polymorphisms (STRs): a reappraisal. Forensic Science International: Genetics. 2014 Mar 1;9:20-4.

[7] Parson W, Roewer L. Publication of population data of linearly inherited DNA markers in the International Journal of Legal Medicine. International Journal of Legal Medicine. 2010 ;124(5):505-9.[8] Valdes AM, Slatkin M, Freimer NB. Allele frequencies at microsatellite loci: the stepwise mutation model revisited. Genetics. 1993 Mar 1;133(3):737-49.

[9] Otagiri T, Sato N, Shiozaki T, Harayama Y, Hayashi T, Kobayashi K, Asamura H. Mutation analysis for 25 Y-STR markers in Japanese population. Legal Medicine. 2021 May 1;50:101860.

[10] Ding Q, Hu Y, Koren A, Clark AG. Mutation rate variability across human Y-chromosome haplogroups. Molecular biology and evolution. 2021 Mar 1;38(3):1000-5.

[11] Javed F, Shafique M, McNevin D, Javed MU, Shehzadi A, Shahid AA. Empirical Evidence on Enhanced Mutation Rates of 19 RM-YSTRs for Differentiating Paternal Lineages. Genes. 2022 May 26;13(6):946.

[12] Ballantyne KN, Goedbloed M, Fang R, Schaap O, Lao O, Wollstein A, et al. Mutability of Y-chromosomal microsatellites: rates, characteristics, molecular bases, and forensic implications. *Am J Hum Genet.* 2010;87(3):341–353.

[13] Xue Y, Wang Q, Long Q, Ng BL, Swerdlow H, Burton J, Skuce C, Taylor R, Abdellah Z, Zhao Y, MacArthur DG. Human Y chromosome base-substitution mutation rate measured by direct sequencing in a deep-rooting pedigree. Current Biology. 2009 Sep 15;19(17):1453-7.

Table 1: Distribution of observed Y-STR mutations among studied cases

Relationship type	**No. of cases**	**Total meioses**	**Mutations observed**	**% of cases with mutation**
Siblings & grandparentage	66	2,112	5	7.58%
Uncle-Nephew	133	6,384	13	9.77%
First cousins	70	4,480	8	11.43%
Second cousins	23	2,208	3	13.04%
Total	292	15,184	29	—

Percentages represent the proportion of cases with at least one mutation within each relationship group. Total meioses were calculated as the product of the number of meiosis steps and the number of loci examined per case.

Table 2 Summary of mutation events by locus among studied cases

Y-STR locus	No. of mutations	Mutation type	Example allele shift
DYS385	5	Single-step	16→15, 17→16,15→16,13→14,19→18
DYS391	3	Single-step	10→11, 11→10
DYS458	6	Single-step	16→17, 18→19,17→16,18→17,18→19
DYS19	1	Single-step	18→19
DYS393	2	Single-step	14→13,13→14
GATA-H4	3	Single-step	13→12,12→13
DYS439	2	Single-step	11→10, 12→13
DYS365	3	Single-step	20→21, 22→23
DYS456	2	Single-step	13→14, 16→17
DYS390	1	Single-step	24→25
DYS389II	1	Single-step	32→31
Total	29	—	—

All mutations were single-step changes, consistent with the stepwise mutation model. Example allele shifts are provided for representative cases.

Table 3 Per-locus mutation rates with exact 95% confidence intervals

Y-STR Locus	Mutation Count (x)	Mutation Rate	95% CI (Clopper–Pearson)
DYS458	6	0.006	0.0023–0.0137
DYS385	5	0.005	0.0017–0.0123
DYS391	3	0.003	0.0007–0.0092
GATA-H4	3	0.003	0.0007–0.0092
DYS365	3	0.003	0.0007–0.0092
DYS393	2	0.002	0.0003–0.0077
DYS439	2	0.002	0.0003–0.0077
DYS456	2	0.002	0.0003–0.0077
DYS19	1	0.001	0.0000–0.0059
DYS390	1	0.001	0.0000–0.0059
DYS389II	1	0.001	0.0000–0.0059

[*] *Mutation rates are reported to three decimal places; confidence intervals are presented to four decimal places.*

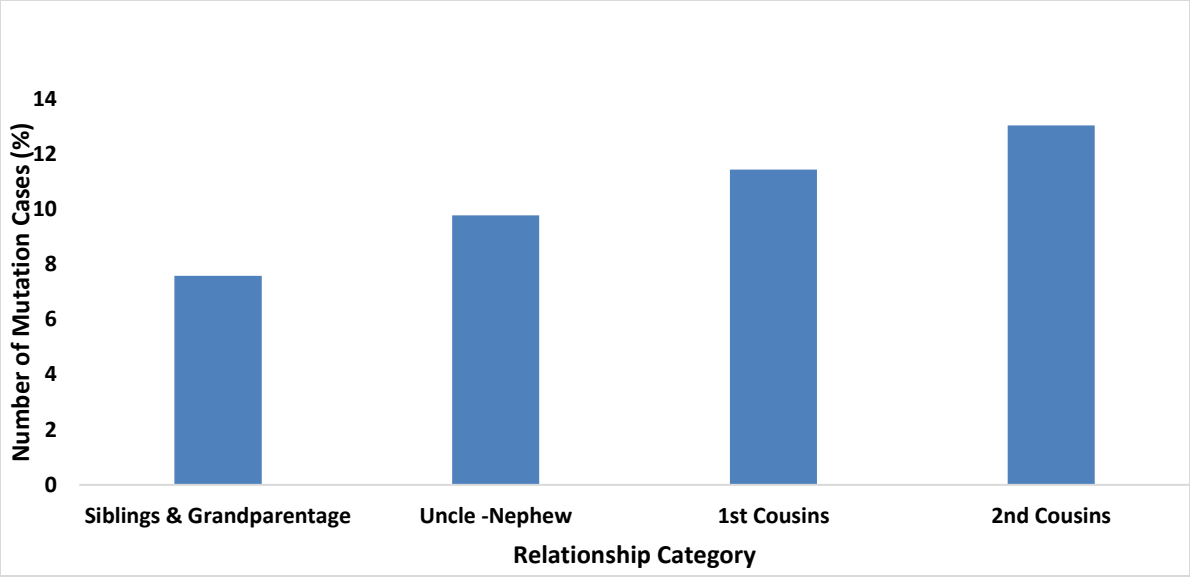

Figure 1: Mutation Frequency Across Relationship Categories

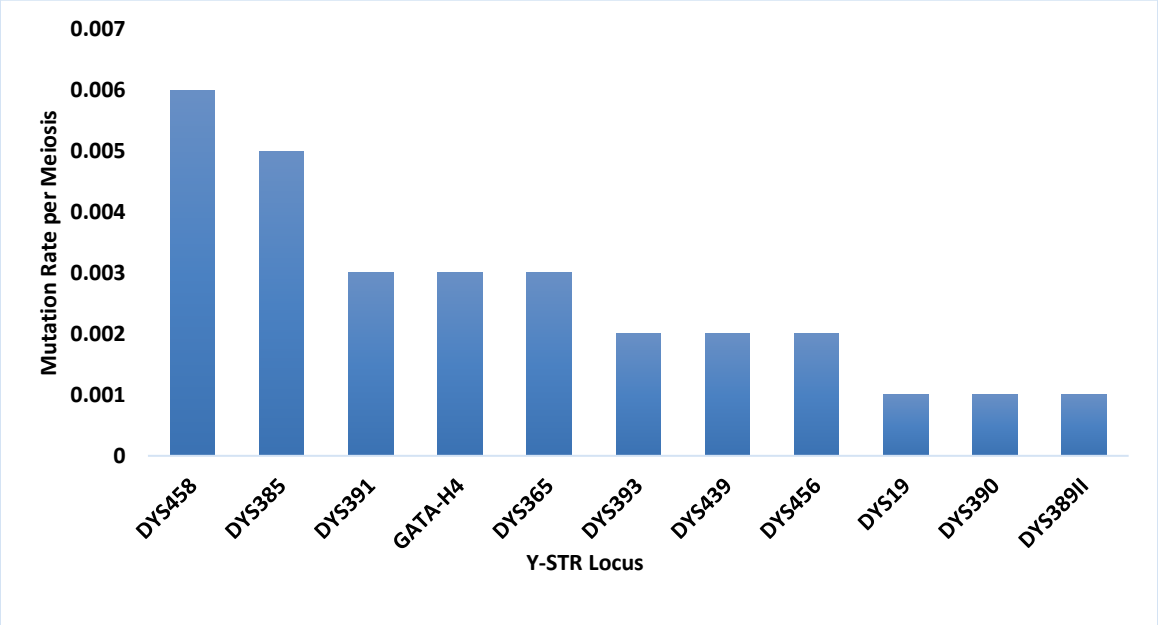

Figure 2: Per-locus Y-STR Mutation Rates Observed among the studied population