



Original Article

Genetic Analysis of 27 Y-STR Haplotypes in 11 Iranian Ethnic Groups

Somayeh Alinaghi¹, Marzieh Mohseni¹, Zohreh Fattahi¹, Maryam Beheshtian¹, Fatemeh Ghodratpour¹, Farzane Zare Ashrafi¹, Sanaz Arzhang¹, Khadijeh Jalalvand¹, Reza Najafipour¹, Hamid Reza Khorram Khorshid¹, Kimia Kahrizi¹, Hossein Najmabadi¹

¹Genetics Research Center, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran

Abstract

Background: The study of Y-chromosomal variations provides valuable insights into male susceptibility in certain diseases like cardiovascular disease (CVD). In this study, we analyzed paternal lineage in different Iranian ethnic groups, not only to identify developing medical etiology, but also to pave the way for gender-specific targeted strategies and personalized medicine in medical genetic research studies.

Methods: The diversity of eleven Iranian ethnic groups was studied using 27 Y-chromosomal short tandem repeat (Y-STR) haplotypes from Y-filer® Plus kit. Analysis of molecular variance (AMOVA) based on pair-wise R_{ST} along with multidimensional scaling (MDS) calculation and Network phylogenetic analysis was employed to quantify the differences between 503 unrelated individuals from each ethnicity.

Results: Results from AMOVA calculation confirmed that Gilaks and Azeris showed the largest genetic distance ($R_{ST}=0.35434$); however, Sistans and Lurs had the smallest considerable genetic distance ($R_{ST}=0.00483$) compared to other ethnicities. Although Azeris had a considerable distance from other ethnicities, they were still close to Turkmens. MDS analysis of ethnic groups gave the indication of lack of similarity between different ethnicities. Besides, network phylogenetic analysis demonstrated insignificant clustering between samples.

Conclusion: The AMOVA analysis results explain that the close distance of Azeris and Turkmens may be the effect of male-dominant expansions across Central Asia that contributed to historical and demographics of populations in the region. Insignificant differences in network analysis could be the consequence of high mutation events that happened in the Y-STR regions over the years. Considering the ethnic group affiliations in medical research, our results provided an understanding and characterization of Iranian male population for future medical and population genetics studies.

Keywords: Ethnic group, Haplotype, Iran, Y-STR, Yfiler® Plus

Cite this article as: Alinaghi S, Mohseni M, Fattahi Z, Beheshtian M, Ghodratpour F, Zare Ashrafi F, et al. Genetic analysis of 27 Y-STR haplotypes in 11 Iranian ethnic groups. Arch Iran Med. 2024;27(2):79-88. doi: 10.34172/aim.2024.13

Received: November 21, 2023, **Accepted:** January 23, 2024, **ePublished:** February 1, 2024

Introduction

The pathophysiology of several diseases, including cancer, neuro-developmental, neuro-degenerative, and cardiovascular diseases (CVDs), is influenced by sex, affecting male development and metabolic homeostasis.¹⁻³ Numerous studies have reported how Y-chromosome haplotypes and haplogroups may help in understanding the etiology of complex diseases or manifestation of a medical outcome.⁴⁻⁶ Recently published studies have shown that the Y-chromosome can play a role in CVDs and hypertension manifestation.⁷⁻⁹ Males with haplogroup I of the Y-chromosome would have 50% higher age-adjusted risk of CVDs than other Y chromosome lineages. Besides, transcriptome molecular pathway in males with haplogroup I revealed significant differential expression between genes related to inflammation and immunity. Furthermore, some of these genes are relevant to atherosclerosis.¹⁰ Consequently, the paternal lineage may inform about the increased risk of developing heart disease and the immune system's response, affecting coronary artery disease (CAD) occurrence.^{10,11} For

example, the *UTY* and *KDM5D* genes modulate the inflammatory and adaptive immunity processes that contribute to atherosclerosis and the resulting outcomes of CVD.^{2,12,13} Mutations in the *SLY* and *RBMV* genes impact experimental allergic encephalomyelitis (EAE) and experimental myocarditis.¹⁴ Autism and intellectual disabilities are associated with *NLGN4Y*.¹⁵ Additionally, male breast and prostate cancer manifest through deletion in the putative *TMSB4Y* and *TSPY2* tumor suppressor genes on the Y-chromosome.^{16,17} Exploring paternal lineage with Y-chromosomal short tandem repeat (Y-STR) markers provides a unique lens into the inheritance of sex-related medical outcomes, and offering valuable insights into disease risk assessment, personalized treatment plans, and advancing precision medicine initiatives.

Iran's population is a combination of various ethnic and cultural groups. A genetically diverse population was established in the area by a mixture of several linguistic communities with differing claims to common ancestry.¹⁸⁻²¹ Regarding gene flow, the Iranian plateau has

been limited by topological barriers like the Zagros and Alborz mountains, the Caspian Sea, the Dasht-e Kavir and the Dash-e Lut deserts, and the Persian Gulf.^{22,23} Furthermore, as a transcontinental route, Iran was influenced by past migration and expansion events.²⁴ With the emergence of modern humans out of Africa, Iran's territory played a pivotal role in pre-historic and historic migratory events between Africa, Asia, and Europe.²⁵ Over the centuries, the Iranian plateau was the destination for male-dominant directional expansion of Indo-European (IE) speakers from Europe, Arab Muslims from West Asia and North Africa (in the 7th century), and Turkic speakers from central Asia (in the 13th century).²⁶⁻²⁸ Investigation of sex-specific population patterns in paternal and maternal lineages would be ideal for determining potentially different degrees of genetic diversity and migration patterns.²⁹ Male population stratification assessment among single worldwide populations and between continental groups can be determined based on Y-STR markers using single markers and haplotype frequency distributions.³⁰ Y-chromosomal markers provide exciting insights into the past demographic events of a population, such as male-specific migration, especially since the Industrial Revolution, which blurs the picture of historical population structure and admixture events.³¹ Different studies have shown that Y-STR haplotype spectra are different in Western Asia compared with other geographical regions in the world, which is proposed to be a corridor for ancient human migrations.^{32,33}

In the past few years, a larger number of studies have investigated Y-STR data from countries and regions in vicinity of Iran and beyond. This includes studies on Armenia/Turkey/Caucasus/Georgia,³⁴⁻³⁶ Afghanistan,³⁷ and Arab countries,^{32,38-40} as well as Russia,⁴¹ whose southern part is considered to have hosted the speakers of Proto-Indo-European (PIE), Pakistan,⁴² India,⁴³ Greece,⁴⁴ the Levant,⁴⁵ Central Asian Uyghurs⁴⁶ and East and West Africans.^{47,48} Nevertheless, Y-STR data from Iran are sparse, both with respect to the geographical and ethnical coverage. A study from 2009 compiled Y-filer data of 259 males from six ethnic groups in Iran and Azerbaijan,⁴⁹ complemented by another 259 Y-filer samples from the East and Southeast of Iran.⁵⁰ More Y-filer samples were added from Mazandaran, Gilan (n=209)⁵¹ and Golestan (n=106) provinces.⁵² However, these studies neither cover large parts of Iran, nor was their sampling always strictly based on ethnic groups, but rather geography, with limited number of studied markers set. Therefore, they are of limited use in investigating the male population structure in Iran and inferring past migration events. Higher mutation rates of STRs can lead to a high diversity of haplotypes that share a recent ancestor. As frequency distributions of Y-STR haplotypes are applicable for studying genetic population differences, their profiles give a more accurate insight into the paternal history of mixed populations.^{53,54}

A genome-wide genotyping for 1021 DNA samples from eleven Iranian ethnicities demonstrated that the CIC (Central Iranian Cluster) comprises seven ethnic groups that show a large genetic overlap, and the four remaining ethnic groups show different ancestral populations.⁵⁵ In this regard, investigation of uniparental markers such as Y-chromosomal STR haplotypes in the Iranian population could establish the role of Y-chromosome in identifying group's susceptibility for developing health problems.⁵⁶ This genetic information can advise assessing variations in medical approaches. Previous studies on Iranian male samples demonstrated that different migration events and critical geographic barriers led mainly to large and significant diversities between Iranian samples.^{21,49,52,57} To adjust the interpretable findings of genetic variability of Y-STRs, we performed an adequate design and analysis of Y-STR markers and haplotypes in the Iranian population as an interpretation for future population and medical genetics studies like risk assessment, treatment strategies and prevention measures.

Materials and Methods

Iranian Samples

To collect data on age, sex, ethnicity, demographics, and health status, written informed consent was obtained from all participants prior to the study, following the rules of the Research Ethics Committee, University of Social Welfare and Rehabilitation Sciences (USWR), Tehran, Iran. Approval to undertake the work was obtained from USWR Research Ethics Committee (Approval number IR.USWR.REC.1395.376 and IR.USWR.REC.1400.136).

In the current study, we included 530 unrelated healthy males from the Iranome project samples (<http://www.iranome.ir>)⁵⁸ (49.57% of total samples) representing eleven Iranian ethnic groups, namely Iranian Arabs, Iranian Azeris, Iranian Baluchis, Iranian Kurds, Iranian Lurs, Iranian Gilaks, Iranian Mazanderanis, Iranian Sistanis, Iranian Persians, Iranian Turkmens, and Iranian Persian Gulf Islanders (PGI). The concentration of all DNA samples to use in STR analysis was quantified before application. Highly degraded samples and samples containing insufficient DNA were excluded from STR typing as they cannot be retrieved with capillary electrophoresis. A total number of 503 male DNA samples were selected for the subsequent procedures (Table 1). Previously in the Iranome Project, individuals who met the inclusion criteria were collected from different provinces of Iran.⁵⁸ The individuals were assigned to each ethnic group if they originated from the same ethnic background for at least two generations, whereas close relatives and individuals from mixed ethnic groups were excluded. Besides, individuals in each ethnic group were above 40 years of age at the time of admission to reduce the possibility of genetic diseases developing later in life. In general, the average age of all ethnic group's participants was 45 years, and sampling was performed based on total population, equally from males and females. In the final

Table 1. Samples Filtration and Quality Control

Ethnic Groups	Samples Before QC	Samples After QC
Iranian Arabs	52	50
Iranian Azeris	51	50
Iranian Baluchis	52	50
Iranian Gilaks	50	48
Iranian Kurds	42	36
Iranian Lurs	42	41
Iranian Mazanderanis	50	46
Iranian Persian Gulf Islanders	51	48
Iranian Persians	40	36
Iranian Sistanis	50	49
Iranian Turkmens	50	49
Total	530	503

QC, quality control.

stage, a clinician evaluated all the participants of the project.

Quality Control

All experimental procedures were operated entirely according to Yfiler® Plus kit controls (Cat. No. 4484678 and 4482730) and internal laboratory standards. To evaluate the efficiency of the amplification steps, a panel of standards was used for PCR amplification, PCR product sizing, and genotyping which were: Yfiler® Plus DNA Control 007, GeneScan™-600 LIZ™ Size Standard v2.0 (Cat. No. 4408399) (Thermo Fisher Scientific's Yfiler® Plus PCR Amplification Kit, Life Technologies Ltd, Woolston, United Kingdom), Yfiler® Plus Allelic Ladder. Additional purifications were performed before proceeding to STR typing to obtain required conditions that produced optimum PCR allele typing results and appropriate cycle number and avoid off-scale peaks.

The kit developed optimal performance to improve the resolution of paternal lineages and discriminate between closely related males based on the updated and revised guidelines from the Scientific Working Group on DNA Analysis Methods (SWGDM, December 2012 and March 2022). The kit's sensitivity and reliability were provided based on comparable studies, ensuring a robust foundation for accurate and consistent outcomes.⁵⁹⁻⁶¹ This paper follows the guidelines for the publication of population data proposed by the journal.

Y-STR Typing

DNA extraction was performed on blood samples following the standard salting out procedure.⁶² Samples were typed for all 27 Y-STR markers of the Y-filer Plus kit using the manufacturer's protocol (Thermo Fisher Scientific's Y-filer® Plus PCR Amplification Kit, Life Technologies Ltd, Woolston, United Kingdom). Although Duplicated markers DYS385a/b and DYS387S1a/b were typed in all samples, the representative genotypes were discarded in all subsequent analyses due to allele

assignment disruption. This kit examined ten significant loci; three loci were highly polymorphic (DYS460, DYS481, and DYS533), and seven loci (DYS387S1a/b, DYS449, DYS518, DYS570, DYS576, DYS627) were rapidly mutating (RM). The remaining 17 markers were referred to as the 'large marker set' (DYS389I/II, DYS635, DYS458, DYS19, YGATAH4, DYS448, DYS390, DYS391, DYS456, DYS438, DYS392, DYS437, DYS393, DYS439, DYS385 a/b). The combination of RM Y-STRs with an increased number of targeted loci has empowered discrimination in high levels of population haplotype diversity in the Y-filer Plus kit.⁵⁹ Amplified PCR products were produced in approximately 30 cycles of PCR on a Veriti® 96-Well Thermal Cycle, along with negative and positive controls, to optimize thermal cycling conditions. PCR product separation and detection were performed on an ABI 3500xL Genetic Analyzer machine (Thermo Fisher Scientific™, Massachusetts, United States) filled with POP-6™ polymer. Allele values were assigned using the reference ladder provided with each kit and repeat numbers indicated by the GeneMapper® ID-X v1.6 software (Thermo Fisher Scientific, Massachusetts, United States). In this study, the number of repetitions in DYS389I was deducted from DYS389II, referring to the repeat numbers at individual loci rather than the repeat numbers revealed by the multiplex genotyping method. A total of 403 samples (80.12%) were typed for all 27 Y-STR markers of the Y-filer Plus kit using the manufacturer's protocol. The remaining 19.88% (100 samples) were, with insight, excluded from complete retyping, as the data from the 'large marker set', comprising key markers in the Y-chromosome haplotype, had been previously obtained and considered to be sufficient for robust statistical analysis. This strategic approach ensures both efficiency in resource utilization and methodological velocity in the desire for meaningful results.

Statistical Analysis

Analysis of Molecular Variance

Analysis of population variation between eleven Iranian subpopulations was quantified based upon pair-wise sum of squared size difference (R_{ST}). In the analysis of molecular variance (AMOVA) calculation, for significance testing P values of $R_{ST} > 0$ and using randomization with 1000 permutations, we employed Arlequin v 3.5.2.2.⁶³ The DYS385a/b and DYS387S1a/b markers were not included in the AMOVA because they do not allow for easy calculation.

Multidimensional Scaling Analysis

Multidimensional scaling (MDS) was utilized to visualize the average distances in Y-STR genetic variation between ethnic groups. It was based upon pair-wise R_{ST} between sampling sites as estimated with Arlequin (see above) and was implemented in the *cmdscale* function of R v4.1.3.⁶⁴ MDS was used to investigate genetic similarities between populations.⁶⁵ Plots of the first two MDS components,

C1 and C2, capturing the most and second most R_{ST} -defined variation were generated with R using in-house scripts. For R_{ST} calculation and MDS analysis, multi-locus markers, haplotypes containing duplications, along with null alleles and micro-variants, were disregarded from the analysis of the respective marker set.

Median-Joining Network Analysis

To identify phylogenetic connections between samples containing unique and uncommon variations based on Y-STR haplotypes, we used the stepwise star-contraction, Median-Joining (MJ), and Maximum-Parsimony (MP) algorithm calculations with NETWORK v10 and NETWORK Publisher v2.1.2.5 as described on the Fluxus Engineering website (Fluxus Technology Ltd, London, United Kingdom) (<https://www.fluxus-engineering.com>).⁶⁶ All Y-STR loci's weights were adjusted based on the marker-specific mutation rate (https://yhrd.org/pages/resources/mutation_rates). In network construction, each marker was weighted by the inverse of its marker-specific mutation rate and these rates were used for calibration. Then, the ages of nodes within the network can be estimated. In the network phylogenetic analysis, the deleted alleles were coded '99' in input files and thereby considered as missing data. Additionally, duplicated loci (DYS385a/b; DYS387S1a/b) were removed for network construction, as it is not possible to associate particular alleles with specific copies. We applied the star-contraction, MJ, and MP options of the NETWORK software in our analyses.

Results

Sample Characteristic

The 503 male samples originated from eleven Iranian ethnic groups, ranging from 36 to 50 per group (Table 1). A number of 435 unique haplotypes were derived from 503 individuals using the 27 Y-STRs Y-filer® Plus kit. There were 34 haplotypes shared among two individuals.

AMOVA and MDS Analysis

In population distance analysis, all pair-wise comparisons resulted in R_{ST} values which were significantly different from zero ($P < 0.001$) and genetic distance (R_{ST}) values for pair-wise comparisons of haplotypes between ethnic groups were further explored (Table 2). The results indicated that Gilak and Azeri ethnicities are located at the largest genetic distance ($R_{ST} = 0.35434$) in comparison with other ethnic groups. In contrast, the least considerable genetic distances are seen in Sistanis and Lurs ($R_{ST} = 0.00483$) and Arabs and Lurs ($R_{ST} = 0.00521$). The largest R_{ST} values were observed for Azeris in comparison with other ethnic groups, which shows the considerable difference of Azeris with other ethnicities. Meanwhile, in the value of R_{ST} , among the other ethnic groups, the Turkmen ethnicity demonstrated the least distance from Azeris ($R_{ST} = 0.13284$). The degree and significance of differentiation between groups were assessed by AMOVA. The AMOVA results demonstrated that the Y-STR haplotypes differ significantly across ethnic groups; 11.57% of the genetic variance reflects differences among population groups, whereas 88.43% reflects differences within ethnic groups, with a Fixation Index (F_{ST}) of 0.11574 (Table 3). The MDS analysis compared all eleven ethnic groups and was performed using R_{ST} values (Figure 1). Ultimately, MDS analysis of pair-wise R_{ST} from eleven Iranian ethnic groups revealed the absence of resemblance between distinct Iranian ethnicities.

Y-STR Network Analysis

To reveal the detailed structures of Iranian male populations, a median-joining network using Y-STRs was constructed, derived from the data on all markers with mutation rates under 0.0035 (Table 4). Within both MJ and MP, no portioning of populations was observed by ethnic groups, and all male individuals were distributed throughout the network. No phylogenetic clustering of ethnic groups within network analysis became apparent

Table 2. R_{ST} Values Between Pairs of Ethnic Groups

	Iranian Arabs	Iranian Azeris	Iranian Baluchis	Iranian Gilaks	Iranian Kurds	Iranian Lurs	Iranian Mazanderanis	Iranian Persian Gulf Islanders	Iranian Persians	Iranian Sistanis	Iranian Turkmens
Iranian Arabs	0	0.23142	0.03069	0.06957	0.11522	0.00521	0.03895	0.08141	0.06524	0.04776	0.01662
Iranian Azeris	0.23142	0	0.27394	0.35434	0.15842	0.24810	0.18979	0.18743	0.28124	0.22494	0.13284
Iranian Baluchis	0.03069	0.27394	0	0.09219	0.10617	0.03040	0.06570	0.08968	0.09231	0.06049	0.07288
Iranian Gilaks	0.06957	0.35434	0.09219	0	0.18996	0.09102	0.16729	0.19571	0.13030	0.10057	0.12947
Iranian Kurds	0.11522	0.15842	0.10617	0.18996	0	0.08427	0.06935	0.06093	0.14353	0.06422	0.08559
Iranian Lurs	0.00521	0.24810	0.03040	0.09102	0.08427	0	0.01618	0.07674	0.03898	0.00483	0.02712
Iranian Mazanderanis	0.03895	0.18979	0.06570	0.16729	0.06935	0.01618	0	0.08485	0.10991	0.02458	0.02600
Iranian Persian Gulf Islanders	0.08141	0.18743	0.08968	0.19571	0.06093	0.07674	0.08485	0	0.07213	0.08535	0.08374
Iranian Persians	0.06524	0.28124	0.09231	0.13030	0.14353	0.03898	0.10991	0.07213	0	0.06267	0.08639
Iranian Sistanis	0.04776	0.22494	0.06049	0.10057	0.06422	0.00483	0.02458	0.08535	0.06267	0	0.04099
Iranian Turkmens	0.01662	0.13284	0.07288	0.12947	0.08559	0.02712	0.02600	0.08374	0.08639	0.04099	0

Table 3. Analysis of Molecular Variance (AMOVA) Results

Source of Variation	df	Sum of Squares	Variance Components	Percentage of Variation (%)	P Value
Among populations (Va)	10	2327.729	4.36654	11.57	<0.001
Within populations (Vb)	492	16414.096	33.36198	88.43	<0.001
Total	502	18741.825	37.72852	100	
Fixation Index	F_{ST} : 0.11574				

Va, Variance for among populations; Vb, Variance for within population; df, degrees of freedom. Significance of tests was based on 1023 permutations.

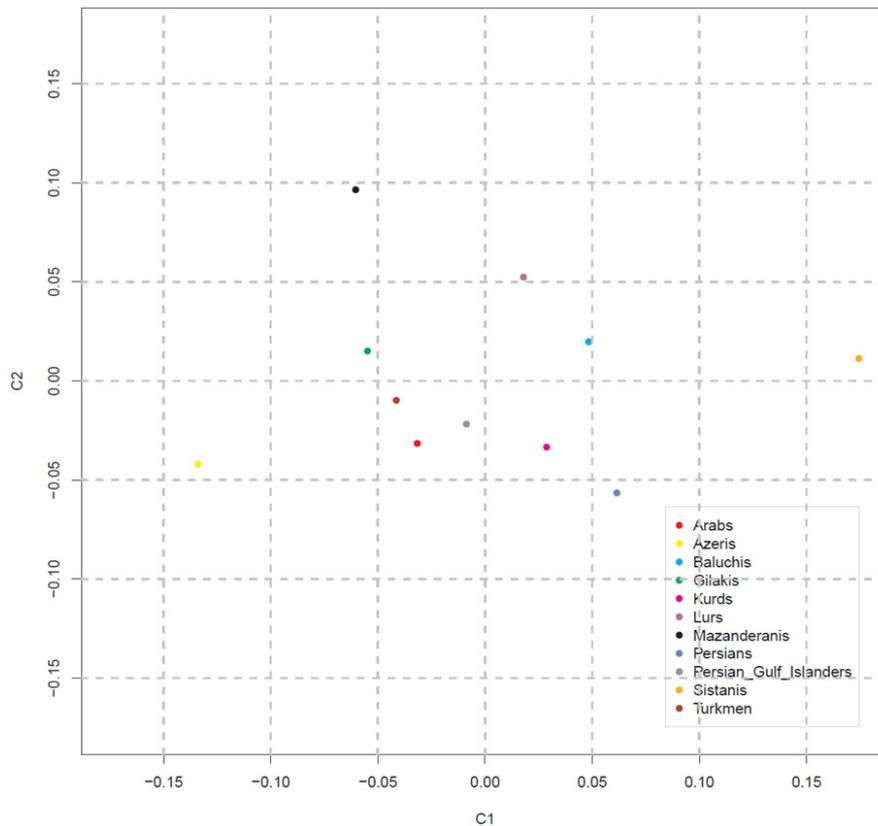


Figure 1. Multidimensional Scaling (MDS) Analysis of 11 Iranian Ethnic Groups Based on Pairwise Genetic Distances (RST). MDS plot shows the relationships between all Iranian male individuals

(Figure 2). It is evident that in the network results overall, many mutation events occurred between the samples, which indicates that Iranian male samples do not have a specific clustering structure. Absence of substructure between samples could be the result of high mutation events that happened in the Y-STR regions during past years.

Discussion

In the largest male genetic study of the Iranian population to date, we investigated the pattern of Y-chromosome STR markers in 503 samples from eleven different Iranian ethnic groups. The sampling sites were chosen to represent different ethnic groups. Our study revealed that regarding the Y-STRs, Iranian ethnic groups show largely overlapping genetic variations with no or only subtle substructure ($P > 0.05$).

The overall results from AMOVA calculations indicate that the Azeris ethnic group showed high R_{ST} values among

all the studied sub-populations; meanwhile, among Azeris ethnic group allocated R_{ST} values, the Turkmen ethnicity has the lowest value comparing other ethnic groups. In a previous study on 1021 Iranian samples,⁵⁵ the authors demonstrated that the Iranian population comprises a group of seven ethnicities (CIC) among other groups with largely overlapping autosomal variations, including Iranian Azeris. Meanwhile, the remaining four ethnic groups that showed large degrees of distance were Iranian Baluchis, Iranian Sistanis, Iranian Turkmens, and Iranian inhabitants of the islands in Persian Gulf. Notably, in terms of paternal Y-STR markers, the Azeri ethnicity, which with respect to autosomal data belongs to CIC, has a considerable distance from other ethnicities but is still close to Turkmens. Based on our results of Y-chromosome haplotypes, the Iranian population has a different structure in Y-STR markers compared to autosomal variants. In this area, if we could assume Y-STR markers as representatives of disease-causing genetic variants on

Table 4. Y-STR Mutation Rates

Mutation	Rate
DYS19	2.1e-03
DYS389I	2.39e-03
DYS389II	4.60e-03
DYS390	2.0e-03
DYS391	2.4e-03
DYS392	0.52e-03
DYS393	1.2e-03
DYS437	1.3e-03
DYS438	0.35e-03
DYS439	5.1e-03
DYS448	1.4e-03
DYS456	4.3e-03
DYS458	6.6e-03
DYS635	4.2e-03
YGATAH4	2.5e-03
DYS576	12.1e-03
DYS481	4.6e-03
DYS533	3.0e-03
DYS570	9.2e-03
DYS627	14.7e-03
DYS460	4.4e-03
DYS518	11.9e-03
DYS449	1.02e-03

Source: https://yhrd.org/pages/resources/mutation_rates

Since it is not possible to associate particular alleles with particular copies of duplicated loci (DYS385a/b; DYS387S1a/b), the mutation rates of these loci were not included.

Y-chromosome, we would expect the carrier frequencies of Y-chromosome variants to be approximately the same for all ethnic groups, excluding the Azeris.

During the 6th century, the Turkic peoples of Central Asia gradually spread westward; they were dominating male warriors that often took local women.⁶⁷ These expansions eventually played a significant role in shaping the demographic features of the inhabitants.⁶⁸ Our observations highlight this possibility that the Azeri and Turkmen ethnicities have been influenced by male-dominant expansions across Central Asia over the past thousand years, leading to the spread of their genes through the male lineage. To our knowledge, based on their earliest origins in Central Asia to their present-day genetic and linguistic diversity, these two ethnicities have the same linguistic history. These migrations of Turkic speakers and interactions with other genetic ancestries had a profound influence on the development of the Turkic languages, leading to the emergence of numerous dialects and subgroups within the language family. Therefore, the genetic admixture of people in the North West and East of Iran, i.e. the Azeri and Turkmen ethnicities, might have originated from Turkic-speaking genetic ancestries.^{69,70} On the other hand, the gradual influx of westward male and female migrations in the East Eurasia could be the reason that the Turkmen ethnic group is genetically close to Azeris but does not have significantly large genetic differences from other ethnicities,^{71,72} which could be considered in medical genetics applications.^{73,74}

The results of network analysis demonstrated a large variability with subtle substructure of Iranian male

Geography

- IRANIAN ARABS
- IRANIAN AZERIS
- IRANIAN BALUCHIS
- IRANIAN GILAKS
- IRANIAN KURDS
- IRANIAN LURDS
- IRANIAN MAZANDERANIS
- IRANIAN PERSIANS
- IRANIAN PGI
- IRANIAN SISTANIS
- IRANIAN TURKMEN

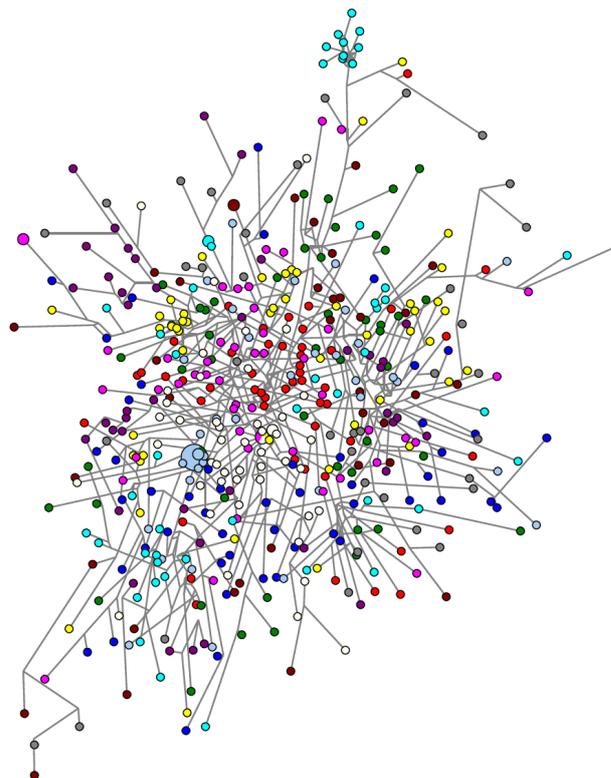


Figure 2. Median-Joining Haplotype Network for 27 Y-STR for 11 Iranian Ethnic Groups

Y-STR marker variation. STR loci on Y-chromosome can exhibit high complex repeat motif variability and a unique inheritance pattern due to lack of recombination and the evolutionary mechanisms driving the increased mutation rates at these markers.^{75,76} The high level of mutation rates between samples in the Network analysis results confirmed the same genetic structure result for the different ethnic groups of the Iranian male population. The results indicate high variability with subtle substructure of Y-STR marker variation, which could be the cause of unique inheritance pattern and high mutation rates at these markers.⁷⁷ Y-STR markers are passed down from father to son in a lineage-specific manner, which means that each lineage may have different patterns of Y-STR marker alleles. Additionally, the Y chromosome is not subject to recombination events like the other chromosomes, which means that mutations in Y-STR markers can accumulate over generations without being shuffled or diluted by recombination events.⁷⁸ On the other hand, Y-STR mutations can occur at a higher rate compared to mutations in other genetic markers, which can result in rapid changes in Y-STR haplotypes over time. These high mutation rate of Y-STR markers may also contribute to their variability.⁷⁹

It must be noted that high mutation rate in Y-STR loci can also induce problems in interpreting the results of genetic profiles and create a major limitation, particularly when comparing profiles across different groups of a population or time period.⁸⁰ Because the high mutation rate results in different haplotypes arising from a common ancestor, they can complicate the analysis and interpretation of genetic data and limit the investigation of possible substructures and comparison of genetic profiles.⁸¹ Additionally, Y-STR loci are prone to high mutation rates, and reduce the accuracy of genetic profiling.⁸² According to previous research, the high mutation rate complicates establishing the time of the most recent common ancestor (TMRCA) in genealogical applications. It may also have an effect on the variability of the human Y-chromosome and, as a result, on its application in forensic sciences, genetic genealogy, human population genetics, and molecular anthropology.^{83,84}

Populations from regions that included Iranian samples were previously analyzed using the Yfiler® marker panel.^{49,51,52} Their results showed that there was little genetic distance between socio-geographical populations in Iran. However, there was a relatively subtle substructure among groups when populations were grouped by geography. Parallel to the present study, the authors concluded that insignificant distances exist between the bulk of Iranian samples. By analyzing the patterns of Y-STR variation, they demonstrated that Iranian samples have negligible genetic differences, which may suggest that they have diverged from a common ancestor a long time ago and experienced a variety of migratory events. This variation among different ethnic groups is defined by different factors, including both geographic and linguistic ones.²¹ It is important to note that for further research, it will

be beneficial to confirm our observations with a larger sample size along with assessment of mutation rates and TMRCA ensuring a comprehensive understanding of the genetic landscape in the specified area.

Conclusion

Genetic ancestry, partially mirrored by ethnic group affiliation, is an important factor to consider in medical research, as differences in genetic background can impact disease prevalence and response to treatments. Based on our results, we predict that among the eleven Iranian ethnic groups studied, there is no significant difference in the prevalence average in phenotypes that are influenced by genetic variants on the Y-chromosome. Moreover, regarding allele frequency, these conditions can be influenced by the frequency of specific alleles in the population. By characterizing the genetic makeup of the Iranian male population, our study provides a valuable resource for future medical research and population genetics studies. By doing so, we can improve our ability to predict and prevent diseases and may develop more personalized treatment options for patients based on their genetic background and ethnic identity.

Acknowledgements

We would like to express our sincere gratitude to all those who contributed to the successful completion of this research article and extend our deepest appreciation to the participants from the eleven Iranian ethnic groups who generously provided their DNA samples for this study. Their invaluable contributions made this research possible and provided crucial insights into the genetic diversity of Iranian population. Also, we are grateful to the research staff and technicians of 'Iranome project' who diligently carried out the laboratory work.

Authors' Contribution

Conceptualization: Hossein Najmabadi, Kimia Kahrizi, Marzieh Mohseni, Maryam Beheshtian, Zohreh Fattahi, Hamid Reza Khorram Khorshid, Reza Najafipour.

Data curation: Somayeh Alinaghi, Maryam Beheshtian, Marzieh Mohseni, Zohreh Fattahi, Hossein Najmabadi.

Formal analysis: Somayeh Alinaghi, Zohreh Fattahi, Maryam Beheshtian.

Funding acquisition: Hossein Najmabadi.

Investigation: Somayeh Alinaghi, Maryam Beheshtian, Marzieh Mohseni, Farzane Zare Ashrafi.

Methodology: Somayeh Alinaghi, Marzieh Mohseni, Farzane Zare Ashrafi, Maryam Beheshtian, Zohreh Fattahi, Khadijeh Jalalvand, Fatemeh Ghodrathpour, Sanaz Arzhanghi.

Project administration: Hossein Najmabadi.

Resources: Hossein Najmabadi, Kimia Kahrizi, Hamid Reza Khorram Khorshid, Reza Najafipour.

Supervision: Hossein Najmabadi.

Validation: Zohreh Fattahi.

Visualization: Hossein Najmabadi.

Writing—original draft: Somayeh Alinaghi, Maryam Beheshtian.

Writing—review & editing: Somayeh Alinaghi, Maryam Beheshtian.

Competing Interests

The authors declare that they have no conflicts of interest with the contents of this article.

Consent for Publication

We wish to confirm that informed consent for the publication

of genetic data has been obtained from all individuals directly involved in the study, or from their legal guardians when applicable. The identities of these individuals have been protected to ensure confidentiality, and any potentially identifying information has been de-identified in accordance with ethical guidelines.

Data Availability Statement

The datasets analyzed during the current study available from the corresponding author on reasonable request. Data access is restricted exclusively to population genetic analyses. Data access requests should be sent to: Dr. Hossein Najmabadi, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran (hnajm12@yahoo.com). Requests have to include a description of the planned project and the planned use of the provided data set.

Ethical Approval

We confirm that all experimental protocols conducted in this study received prior approval from the Research Ethics Committee, University of Social Welfare and Rehabilitation Sciences (USWR), Tehran, Iran. (Approval number IR.USWR.REC.1395.376 and IR.USWR.REC.1400.136) in accordance with established ethical and regulatory guidelines according to the Declaration of Helsinki. We further affirm that informed consent was obtained from all subjects, and when applicable, from their respective legal guardian(s), participating in this study according to Research Ethics Committee, University of Social Welfare and Rehabilitation Sciences (USWR), Tehran, Iran guidelines ensuring their full understanding and voluntary participation in this research study.

Funding

No specific funding was received for this work.

Web Sources

R software: <https://www.r-project.org>
 Arlequin software: <http://cmpg.unibe.ch/software/arlequin35/>
 NETWORK software: <https://www.fluxus-engineering.com/sharenet.htm>
 YHRD: <https://yhrd.org/>

References

- García-González P, de Rojas I, Moreno-Grau S, Montreal L, Puerta R, Alarcón-Martín E, et al. Mendelian randomisation confirms the role of Y-chromosome loss in Alzheimer's disease aetiopathogenesis in men. *Int J Mol Sci.* 2023;24(2):898. doi: [10.3390/ijms24020898](https://doi.org/10.3390/ijms24020898).
- Maan AA, Eales J, Akbarov A, Rowland J, Xu X, Jobling MA, et al. The Y-chromosome: a blueprint for men's health? *Eur J Hum Genet.* 2017;25(11):1181-8. doi: [10.1038/ejhg.2017.128](https://doi.org/10.1038/ejhg.2017.128).
- Zagni E, Simoni L, Colombo D. Sex and gender differences in central nervous system-related disorders. *Neurosci J.* 2016;2016:2827090. doi: [10.1155/2016/2827090](https://doi.org/10.1155/2016/2827090).
- Greyn FP, Makariou MB, Bandres-Ciga S, Iwaki H, Singleton AB, Nalls MA, et al. Analysis of Y-chromosome haplogroups in Parkinson's disease. *Brain Commun.* 2022;4(6):fcac277. doi: [10.1093/braincomms/fcac277](https://doi.org/10.1093/braincomms/fcac277).
- Lassen O, Tabares S, Bertolotto P, Ojeda S, Sembaj A. Preliminary study between Y-chromosome haplogroups and chagasic cardiomyopathy manifestations in patients with Chagas disease. *Rev Soc Bras Med Trop.* 2020;53:e20190566. doi: [10.1590/0037-8682-0566-2019](https://doi.org/10.1590/0037-8682-0566-2019).
- Westra WM, Rygiel AM, Mostafavi N, de Wit GM, Roes AL, Moons LM, et al. The Y-chromosome F haplogroup contributes to the development of Barrett's esophagus-associated esophageal adenocarcinoma in a white male population. *Dis Esophagus.* 2020;33(9):doaa011. doi: [10.1093/dote/doaa011](https://doi.org/10.1093/dote/doaa011).
- Khan SI, Andrews KL, Jennings GL, Sampson AK, Chin-Dusting JP. Y-chromosome, hypertension and cardiovascular disease: is inflammation the answer? *Int J Mol Sci.* 2019;20(12):2892. doi: [10.3390/ijms20122892](https://doi.org/10.3390/ijms20122892).
- Mas-Peiro S, Abplanalp WT, Rasper T, Berkowitsch A, Leistner DM, Dimmeler S, et al. Mosaic loss of Y-chromosome in monocytes is associated with lower survival after transcatheter aortic valve replacement. *Eur Heart J.* 2023;44(21):1943-52. doi: [10.1093/eurheartj/ehad093](https://doi.org/10.1093/eurheartj/ehad093).
- Umar S, Cunningham CM, Itoh Y, Moazeni S, Vaillancourt M, Sarji S, et al. The Y-chromosome plays a protective role in experimental hypoxic pulmonary hypertension. *Am J Respir Crit Care Med.* 2018;197(7):952-5. doi: [10.1164/rccm.201707-1345LE](https://doi.org/10.1164/rccm.201707-1345LE).
- Charchar FJ, Bloomer LD, Barnes TA, Cowley MJ, Nelson CP, Wang Y, et al. Inheritance of coronary artery disease in men: an analysis of the role of the Y-chromosome. *Lancet.* 2012;379(9819):915-22. doi: [10.1016/s0140-6736\(11\)61453-0](https://doi.org/10.1016/s0140-6736(11)61453-0).
- Charchar FJ, Tomaszewski M, Padmanabhan S, Lacka B, Upton MN, Inglis GC, et al. The Y-chromosome effect on blood pressure in two European populations. *Hypertension.* 2002;39(2 Pt 2):353-6. doi: [10.1161/hy0202.103413](https://doi.org/10.1161/hy0202.103413).
- Bloomer LD, Nelson CP, Eales J, Denniff M, Christofidou P, Debiec R, et al. Male-specific region of the Y-chromosome and cardiovascular risk: phylogenetic analysis and gene expression studies. *Arterioscler Thromb Vasc Biol.* 2013;33(7):1722-7. doi: [10.1161/atvbaha.113.301608](https://doi.org/10.1161/atvbaha.113.301608).
- Walport LJ, Hopkinson RJ, Vollmar M, Madden SK, Gileadi C, Oppermann U, et al. Human UTY(KDM6C) is a male-specific Nc-methyl lysyl demethylase. *J Biol Chem.* 2014;289(26):18302-13. doi: [10.1074/jbc.M114.555052](https://doi.org/10.1074/jbc.M114.555052).
- Case LK, Wall EH, Dragon JA, Saligrama N, Kremontsov DN, Moussawi M, et al. The Y-chromosome as a regulatory element shaping immune cell transcriptomes and susceptibility to autoimmune disease. *Genome Res.* 2013;23(9):1474-85. doi: [10.1101/gr.156703.113](https://doi.org/10.1101/gr.156703.113).
- Ross JL, Bloy L, Roberts TPL, Miller J, Xing C, Silverman LA, et al. Y-chromosome gene copy number and lack of autism phenotype in a male with an isodicentric Y-chromosome and absent NLGN4Y expression. *Am J Med Genet B Neuropsychiatr Genet.* 2019;180(7):471-82. doi: [10.1002/ajmg.b.32745](https://doi.org/10.1002/ajmg.b.32745).
- Li Y, Zhang DJ, Qiu Y, Kido T, Lau YC. The Y-located proto-oncogene TSPY exacerbates and its X-homologue TSPX inhibits transactivation functions of androgen receptor and its constitutively active variants. *Hum Mol Genet.* 2017;26(5):901-12. doi: [10.1093/hmg/ddx005](https://doi.org/10.1093/hmg/ddx005).
- Wong HY, Wang GM, Croessmann S, Zabransky DJ, Chu D, Garay JP, et al. TMSB4Y is a candidate tumor suppressor on the Y-chromosome and is deleted in male breast cancer. *Oncotarget.* 2015;6(42):44927-40. doi: [10.18632/oncotarget.6743](https://doi.org/10.18632/oncotarget.6743).
- Amanolahi S. A note on ethnicity and ethnic groups in Iran. *Iran Cauc.* 2005;9(1):37-42. doi: [10.1163/1573384054068105](https://doi.org/10.1163/1573384054068105).
- Bahmanimehr A, Eskandari G, Nikmanesh F. The congruence between matrilineal genetic (mtDNA) and geographic diversity of Iranians and the territorial populations. *Iran J Basic Med Sci.* 2015;18(1):30-7.
- Farjadian S, Sazzini M, Tofanelli S, Castrì L, Taglioli L, Pettener D, et al. Discordant patterns of mtDNA and ethno-linguistic variation in 14 Iranian Ethnic groups. *Hum Hered.* 2011;72(2):73-84. doi: [10.1159/000330166](https://doi.org/10.1159/000330166).
- Grugni V, Battaglia V, Hooshyar Kashani B, Parolo S, Al-Zahery N, Achilli A, et al. Ancient migratory events in the Middle East: new clues from the Y-chromosome variation of modern Iranians. *PLoS One.* 2012;7(7):e41252. doi: [10.1371/journal.pone.0041252](https://doi.org/10.1371/journal.pone.0041252).
- Quintana-Murci L, Chaix R, Wells RS, Behar DM, Sayar H, Scozzari R, et al. Where west meets east: the complex mtDNA landscape of the southwest and Central Asian corridor. *Am J Hum Genet.* 2004;74(5):827-45. doi: [10.1086/383236](https://doi.org/10.1086/383236).
- Wells RS, Yuldasheva N, Ruzibakiev R, Underhill PA, Evseeva

- I, Blue-Smith J, et al. The Eurasian heartland: a continental perspective on Y-chromosome diversity. *Proc Natl Acad Sci U S A*. 2001;98(18):10244-9. doi: [10.1073/pnas.171305098](https://doi.org/10.1073/pnas.171305098).
24. Lamberg-Karlovsky C. Archaeology and language: the Indo-Iranians. *Curr Anthropol*. 2002;43(1):63-88. doi: [10.1086/324130](https://doi.org/10.1086/324130).
 25. Richards M, Macaulay V, Hickey E, Vega E, Sykes B, Guida V, et al. Tracing European founder lineages in the Near Eastern mtDNA pool. *Am J Hum Genet*. 2000;67(5):1251-76. doi: [10.1016/s0002-9297\(07\)62954-1](https://doi.org/10.1016/s0002-9297(07)62954-1).
 26. Haak W, Lazaridis I, Patterson N, Rohland N, Mallick S, Llamas B, et al. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature*. 2015;522(7555):207-11. doi: [10.1038/nature14317](https://doi.org/10.1038/nature14317).
 27. Luis JR, Rowold DJ, Regueiro M, Caeiro B, Cinnioglu C, Roseman C, et al. The Levant versus the Horn of Africa: evidence for bidirectional corridors of human migrations. *Am J Hum Genet*. 2004;74(3):532-44. doi: [10.1086/382286](https://doi.org/10.1086/382286).
 28. Regueiro M, Cadenas AM, Gayden T, Underhill PA, Herrera RJ. Iran: tricontinental nexus for Y-chromosome driven migration. *Hum Hered*. 2006;61(3):132-43. doi: [10.1159/000093774](https://doi.org/10.1159/000093774).
 29. Underhill PA, Kivisild T. Use of Y-chromosome and mitochondrial DNA population structure in tracing human migrations. *Annu Rev Genet*. 2007;41:539-64. doi: [10.1146/annurev.genet.41.110306.130407](https://doi.org/10.1146/annurev.genet.41.110306.130407).
 30. Purps J, Siegert S, Willuweit S, Nagy M, Alves C, Salazar R, et al. A global analysis of Y-chromosomal haplotype diversity for 23 STR loci. *Forensic Sci Int Genet*. 2014;12(100):12-23. doi: [10.1016/j.fsigen.2014.04.008](https://doi.org/10.1016/j.fsigen.2014.04.008).
 31. Ellegren H, Galtier N. Determinants of genetic diversity. *Nat Rev Genet*. 2016;17(7):422-33. doi: [10.1038/nrg.2016.58](https://doi.org/10.1038/nrg.2016.58).
 32. Khubrani YM, Wetton JH, Jobling MA. Extensive geographical and social structure in the paternal lineages of Saudi Arabia revealed by analysis of 27 Y-STRs. *Forensic Sci Int Genet*. 2018;33:98-105. doi: [10.1016/j.fsigen.2017.11.015](https://doi.org/10.1016/j.fsigen.2017.11.015).
 33. Singh M, Sarkar A, Nandineni MR. A comprehensive portrait of Y-STR diversity of Indian populations and comparison with 129 worldwide populations. *Sci Rep*. 2018;8(1):15421. doi: [10.1038/s41598-018-33714-2](https://doi.org/10.1038/s41598-018-33714-2).
 34. Balanovsky O, Chukhryaeva M, Zaporozhchenko V, Urasin V, Zhabagin M, Hovhannisyan A, et al. Genetic differentiation between upland and lowland populations shapes the Y-chromosomal landscape of West Asia. *Hum Genet*. 2017;136(4):437-50. doi: [10.1007/s00439-017-1770-2](https://doi.org/10.1007/s00439-017-1770-2).
 35. Lowery RK, Herrera K, Uribe G, Reguero M, Herrera RJ. Sub-population structure evident in forensic Y-STR profiles from Armenian geographical groups. *Leg Med (Tokyo)*. 2013;15(2):85-90. doi: [10.1016/j.legalmed.2012.10.003](https://doi.org/10.1016/j.legalmed.2012.10.003).
 36. Nasidze I, Schädlich H, Stoneking M. Haplotypes from the Caucasus, Turkey and Iran for nine Y-STR loci. *Forensic Sci Int*. 2003;137(1):85-93. doi: [10.1016/s0379-0738\(03\)00272-x](https://doi.org/10.1016/s0379-0738(03)00272-x).
 37. Älgenäs C, Tillmar AO. Population genetics of 29 autosomal STRs and 17 Y-chromosomal STRs in a population sample from Afghanistan. *Int J Legal Med*. 2014;128(2):279-80. doi: [10.1007/s00414-013-0902-z](https://doi.org/10.1007/s00414-013-0902-z).
 38. Nazir M, Alhaddad H, Alenizi M, Alenizi H, Taqi Z, Sanqoor S, et al. A genetic overview of 23Y-STR markers in UAE population. *Forensic Sci Int Genet*. 2016;23:150-2. doi: [10.1016/j.fsigen.2016.04.009](https://doi.org/10.1016/j.fsigen.2016.04.009).
 39. Regueiro M, Garcia-Bertrand R, Fadhlouzi-Zid K, Álvarez J, Herrera RJ. From Arabia to Iberia: a Y-chromosome perspective. *Gene*. 2015;564(2):141-52. doi: [10.1016/j.gene.2015.02.042](https://doi.org/10.1016/j.gene.2015.02.042).
 40. Taqi Z, Alenizi M, Alenizi H, Ismael S, Bin Dukhyil AA, Nazir M, et al. Population genetics of 23 Y-STR markers in Kuwaiti population. *Forensic Sci Int Genet*. 2015;16:203-4. doi: [10.1016/j.fsigen.2015.01.007](https://doi.org/10.1016/j.fsigen.2015.01.007).
 41. Roewer L, Willuweit S, Krüger C, Nagy M, Rychkov S, Morozowa I, et al. Analysis of Y-chromosome STR haplotypes in the European part of Russia reveals high diversities but non-significant genetic distances between populations. *Int J Legal Med*. 2008;122(3):219-23. doi: [10.1007/s00414-007-0222-2](https://doi.org/10.1007/s00414-007-0222-2).
 42. Lee EY, Shin KJ, Rakha A, Sim JE, Park MJ, Kim NY, et al. Analysis of 22 Y chromosomal STR haplotypes and Y haplogroup distribution in Pathans of Pakistan. *Forensic Sci Int Genet*. 2014;11:111-6. doi: [10.1016/j.fsigen.2014.03.004](https://doi.org/10.1016/j.fsigen.2014.03.004).
 43. Ghosh T, Kalpana D, Mukerjee S, Mukherjee M, Sharma AK, Nath S, et al. Genetic diversity of 17 Y-short tandem repeats in Indian population. *Forensic Sci Int Genet*. 2011;5(4):363-7. doi: [10.1016/j.fsigen.2010.12.007](https://doi.org/10.1016/j.fsigen.2010.12.007).
 44. Tomas C, Skitsa I, Steinmeier E, Poulsen L, Ampati A, Børsting C, et al. Results for five sets of forensic genetic markers studied in a Greek population sample. *Forensic Sci Int Genet*. 2015;16:132-7. doi: [10.1016/j.fsigen.2015.01.001](https://doi.org/10.1016/j.fsigen.2015.01.001).
 45. El-Sibai M, Platt DE, Haber M, Xue Y, Youhanna SC, Wells RS, et al. Geographical structure of the Y-chromosomal genetic landscape of the Levant: a coastal-inland contrast. *Ann Hum Genet*. 2009;73(Pt 6):568-81. doi: [10.1111/j.1469-1809.2009.00538.x](https://doi.org/10.1111/j.1469-1809.2009.00538.x).
 46. Bian Y, Zhang S, Zhou W, Zhao Q, Siqintuya, Zhu R, et al. Analysis of genetic admixture in Uyghur using the 26 Y-STR loci system. *Sci Rep*. 2016;6:19998. doi: [10.1038/srep19998](https://doi.org/10.1038/srep19998).
 47. Abdel-Hafez AF, El-Alfy SH, Swelim HH, Hassan NH. Genetic variation at 15 autosomal STR loci among seven Egyptian populations. *Biochem Genet*. 2019;57(1):170-91. doi: [10.1007/s10528-018-9879-0](https://doi.org/10.1007/s10528-018-9879-0).
 48. Kofi AE, Agyemang DA, Ghansah A, Awandare GA, Hakim HM, Khan HO, et al. A genetic study of the Ghanaian population using 15 autosomal STR loci. *Biochem Genet*. 2023;61(5):1850-66. doi: [10.1007/s10528-023-10347-3](https://doi.org/10.1007/s10528-023-10347-3).
 49. Roewer L, Willuweit S, Stoneking M, Nasidze I. A Y-STR database of Iranian and Azerbaijani minority populations. *Forensic Sci Int Genet*. 2009;4(1):e53-5. doi: [10.1016/j.fsigen.2009.05.002](https://doi.org/10.1016/j.fsigen.2009.05.002).
 50. Alipour Tabrizi A, Hedjazi A, Kerachian MA, Honarvar Z, Dadgarmoghaddam M, Raoofian R. Genetic profile of 17 Y-chromosome STR haplotypes in East of Iran. *Forensic Sci Int Genet*. 2015;14:e6-7. doi: [10.1016/j.fsigen.2014.10.010](https://doi.org/10.1016/j.fsigen.2014.10.010).
 51. Sayyari M, Salehzadeh A, Tabatabaiefar MA, Abbasi A. Profiling of 17 Y-STR loci in Mazandaran and Gilan provinces of Iran. *Turk J Med Sci*. 2019;49(5):1277-86. doi: [10.3906/sag-1808-179](https://doi.org/10.3906/sag-1808-179).
 52. Sayyari M, Salehzadeh A, Tabatabaiefar MA, Abbasi A. Genetic polymorphisms of Y-chromosome short tandem repeats (Y-STRs) in a male population from Golestan province, Iran. *Mol Biol Res Commun*. 2020;9(1):11-6. doi: [10.22099/mbrc.2020.35547.1462](https://doi.org/10.22099/mbrc.2020.35547.1462).
 53. Willems T, Gymrek M, Highnam G, Mittelman D, Erlich Y. The landscape of human STR variation. *Genome Res*. 2014;24(11):1894-904. doi: [10.1101/gr.177774.114](https://doi.org/10.1101/gr.177774.114).
 54. Xu H, Wang CC, Shrestha R, Wang LX, Zhang M, He Y, et al. Inferring population structure and demographic history using Y-STR data from worldwide populations. *Mol Genet Genomics*. 2015;290(1):141-50. doi: [10.1007/s00438-014-0903-8](https://doi.org/10.1007/s00438-014-0903-8).
 55. Mehrjoo Z, Fattahi Z, Beheshtian M, Mohseni M, Poustchi H, Ardalani F, et al. Distinct genetic variation and heterogeneity of the Iranian population. *PLoS Genet*. 2019;15(9):e1008385. doi: [10.1371/journal.pgen.1008385](https://doi.org/10.1371/journal.pgen.1008385).
 56. Bloomer LD, Nelson CP, Denniff M, Christofidou P, Debiec R, Thompson J, et al. Coronary artery disease predisposing haplogroup I of the Y-chromosome, aggression and sex steroids--genetic association analysis. *Atherosclerosis*. 2014;233(1):160-4. doi: [10.1016/j.atherosclerosis.2013.12.012](https://doi.org/10.1016/j.atherosclerosis.2013.12.012).
 57. Malyarchuk B, Derenko M, Wozniak M, Grzybowski T. Y-chromosome variation in Tajiks and Iranians. *Ann Hum Biol*. 2013;40(1):48-54. doi: [10.3109/03014460.2012.747628](https://doi.org/10.3109/03014460.2012.747628).

58. Fattahi Z, Beheshtian M, Mohseni M, Poustchi H, Sellars E, Nezhadi SH, et al. Iranome: a catalog of genomic variations in the Iranian population. *Hum Mutat.* 2019;40(11):1968-84. doi: [10.1002/humu.23880](https://doi.org/10.1002/humu.23880).
59. Gopinath S, Zhong C, Nguyen V, Ge J, Lagacé RE, Short ML, et al. Developmental validation of the Yfiler® Plus PCR amplification kit: an enhanced Y-STR multiplex for casework and database applications. *Forensic Sci Int Genet.* 2016;24:164-75. doi: [10.1016/j.fsigen.2016.07.006](https://doi.org/10.1016/j.fsigen.2016.07.006).
60. Henry J, Dao H, Scandrett L, Taylor D. Population genetic analysis of Yfiler® Plus haplotype data for three South Australian populations. *Forensic Sci Int Genet.* 2019;41:e23-5. doi: [10.1016/j.fsigen.2019.03.021](https://doi.org/10.1016/j.fsigen.2019.03.021).
61. Rapone C, D'Atanasio E, Agostino A, Mariano M, Papaluca MT, Cruciani F, et al. Forensic genetic value of a 27 Y-STR loci multiplex (Yfiler® Plus kit) in an Italian population sample. *Forensic Sci Int Genet.* 2016;21:e1-5. doi: [10.1016/j.fsigen.2015.11.006](https://doi.org/10.1016/j.fsigen.2015.11.006).
62. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* 1988;16(3):1215. doi: [10.1093/nar/16.3.1215](https://doi.org/10.1093/nar/16.3.1215).
63. Excoffier L, Lischer HE. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour.* 2010;10(3):564-7. doi: [10.1111/j.1755-0998.2010.02847.x](https://doi.org/10.1111/j.1755-0998.2010.02847.x).
64. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing; 2014.
65. Kruskal JB. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika.* 1964;29(1):1-27. doi: [10.1007/bf02289565](https://doi.org/10.1007/bf02289565).
66. Bandelt HJ, Forster P, Röhl A. Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol.* 1999;16(1):37-48. doi: [10.1093/oxfordjournals.molbev.a026036](https://doi.org/10.1093/oxfordjournals.molbev.a026036).
67. Zhabagin M, Wei LH, Sabitov Z, Ma PC, Sun J, Dyussenova Z, et al. Ancient components and recent expansion in the Eurasian heartland: insights into the revised phylogeny of Y-chromosomes from Central Asia. *Genes (Basel).* 2022;13(10):1776. doi: [10.3390/genes13101776](https://doi.org/10.3390/genes13101776).
68. Narasimhan VM, Patterson N, Moorjani P, Rohland N, Bernardos R, Mallick S, et al. The formation of human populations in South and Central Asia. *Science.* 2019;365(6457):eaat7487. doi: [10.1126/science.aat7487](https://doi.org/10.1126/science.aat7487).
69. Johanson L. On the roles of Turkic in the Caucasus area. In: *Linguistic Areas: Convergence in Historical and Typological Perspective.* London: Palgrave Macmillan; 2006. p. 160-81. doi: [10.1057/9780230287617_7](https://doi.org/10.1057/9780230287617_7).
70. Ragagnin E. Major and minor Turkic language islands in Iran with a special focus on Khalaj. *Iran Stud.* 2020;53(3-4):573-88. doi: [10.1080/00210862.2020.1740881](https://doi.org/10.1080/00210862.2020.1740881).
71. Järve M, Saag L, Scheib CL, Pathak AK, Montinaro F, Pagani L, et al. Shifts in the genetic landscape of the Western Eurasian steppe associated with the beginning and end of the Scythian dominance. *Curr Biol.* 2019;29(14):2430-41.e10. doi: [10.1016/j.cub.2019.06.019](https://doi.org/10.1016/j.cub.2019.06.019).
72. Yunusbayev B, Metspalu M, Metspalu E, Valeev A, Litvinov S, Valiev R, et al. The genetic legacy of the expansion of Turkic-speaking nomads across Eurasia. *PLoS Genet.* 2015;11(4):e1005068. doi: [10.1371/journal.pgen.1005068](https://doi.org/10.1371/journal.pgen.1005068).
73. Haghi M, Khorshidi S, Hosseinpour Feizi MA, Pouladi N, Hosseinpour Feizi AA. β -Thalassemia mutations in the Iranian Kurdish population of Kurdistan and West Azerbaijan provinces. *Hemoglobin.* 2009;33(2):109-14. doi: [10.1080/03630260902862020](https://doi.org/10.1080/03630260902862020).
74. Najmabadi H, Karimi-Nejad R, Sahebjam S, Pourfarzad F, Teimourian S, Sahebjam F, et al. The β -thalassemia mutation spectrum in the Iranian population. *Hemoglobin.* 2001;25(3):285-96. doi: [10.1081/hem-100105221](https://doi.org/10.1081/hem-100105221).
75. Claerhout S, Vandenbosch M, Nivelte K, Gruyters L, Peeters A, Larmuseau MH, et al. Determining Y-STR mutation rates in deep-routing genealogies: identification of haplogroup differences. *Forensic Sci Int Genet.* 2018;34:1-10. doi: [10.1016/j.fsigen.2018.01.005](https://doi.org/10.1016/j.fsigen.2018.01.005).
76. Willems T, Gymrek M, Poznik GD, Tyler-Smith C, Erlich Y. Population-scale sequencing data enable precise estimates of Y-STR mutation rates. *Am J Hum Genet.* 2016;98(5):919-33. doi: [10.1016/j.ajhg.2016.04.001](https://doi.org/10.1016/j.ajhg.2016.04.001).
77. de Knijff P. Messages through bottlenecks: on the combined use of slow and fast evolving polymorphic markers on the human Y-chromosome. *Am J Hum Genet.* 2000;67(5):1055-61. doi: [10.1016/s0002-9297\(07\)62935-8](https://doi.org/10.1016/s0002-9297(07)62935-8).
78. Ralf A, Lubach D, Kousouri N, Winkler C, Schulz I, Roewer L, et al. Identification and characterization of novel rapidly mutating Y-chromosomal short tandem repeat markers. *Hum Mutat.* 2020;41(9):1680-96. doi: [10.1002/humu.24068](https://doi.org/10.1002/humu.24068).
79. Boattini A, Sarno S, Bini C, Pesci V, Barbieri C, De Fanti S, et al. Mutation rates and discriminating power for 13 rapidly-mutating Y-STRs between related and unrelated individuals. *PLoS One.* 2016;11(11):e0165678. doi: [10.1371/journal.pone.0165678](https://doi.org/10.1371/journal.pone.0165678).
80. Kayser M. Forensic use of Y-chromosome DNA: a general overview. *Hum Genet.* 2017;136(5):621-35. doi: [10.1007/s00439-017-1776-9](https://doi.org/10.1007/s00439-017-1776-9).
81. Decker AE, Kline MC, Redman JW, Reid TM, Butler JM. Analysis of mutations in father-son pairs with 17 Y-STR loci. *Forensic Sci Int Genet.* 2008;2(3):e31-5. doi: [10.1016/j.fsigen.2007.08.016](https://doi.org/10.1016/j.fsigen.2007.08.016).
82. Jobling MA, Tyler-Smith C. The human Y-chromosome: an evolutionary marker comes of age. *Nat Rev Genet.* 2003;4(8):598-612. doi: [10.1038/nrg1124](https://doi.org/10.1038/nrg1124).
83. Balanovsky O. Toward a consensus on SNP and STR mutation rates on the human Y-chromosome. *Hum Genet.* 2017;136(5):575-90. doi: [10.1007/s00439-017-1805-8](https://doi.org/10.1007/s00439-017-1805-8).
84. Zhivotovsky LA, Underhill PA, Cinnioglu C, Kayser M, Morar B, Kivisild T, et al. The effective mutation rate at Y-chromosome short tandem repeats, with application to human population-divergence time. *Am J Hum Genet.* 2004;74(1):50-61. doi: [10.1086/380911](https://doi.org/10.1086/380911).