Molecular Biology Research Communications 2020;9(1):11-16 DOI: 10.22099/mbrc.2020.35547.1462 MBRC
Original Article Open Access

Genetic polymorphisms of Y-chromosome short tandem repeats (Y-STRs) in a male population from Golestan province, Iran

Minoo Sayyari[1], Ali Salehzadeh[1,*], Mohammad Amin Tabatabaiefar[2, 3], Ali Abbasi[4]

1) Department of Biology, Rasht Branch, Islamic Azad University, Rasht, Iran
2) Department of Genetics and Molecular Biology, School of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran
3) Pediatric Inherited Diseases Research Center, Research Institute for Primordial Prevention of Noncommunicable Disease, Isfahan University of Medical Sciences, Isfahan, Iran
4) Iranian Legal Medicine Research Center, Legal Medicine Organization, Tehran, Iran

ABSTRACT

Short Tandem Repeats (STRs), which are located out of pseudo-autosomal parts of the human Y chromosome and passed-down from fathers to the male offspring in a non-recombinant form, are regarded as appropriate markers for forensic purposes and evolutionary investigations. Few studies concerned the genotyping of Y chromosome short tandem repeats (Y-STRs) among the ethnic groups of the north of Iran, especially the province of Golestan which is a multiethnic region of Iran. Thus, in this work we investigated the frequency of Y-STR haplotypes among the male population from Golestan province, to elucidate their identity and kinship patterns. A total number of 106 unrelated male individuals participated in this study. Genomic DNA was extracted from blood samples and the multiplex polymerase chain reaction was employed to amplify DNA fragments. Genotyping was performed using capillary electrophoresis and, finally, allele polymorphisms, haplotype diversity (HD) and haplotype discrimination capacity (DC) were determined using GenAlEXv6.5 and Arlequin v5.3.2 software and compared to other regions of Iran. A total number of 87 unique haplotypes were determined. The highest and least allelic polymorphism was observed for the DYS385b and DYS391 loci, respectively. HD and DC were 0.9962 and 0.8207, respectively. In the case of locus with the least allelic variation, we didn't observe any difference between the Gilan and Golestan but there was a difference between the Golestan and Mazandaran provinces. Our results indicated the efficiency of Y-STRs to be used as genetic markers for forensic medicine, and also the evolutionary comparison of different ethnic groups of Golestan, Iran. Also, a low genetic distance between the population of Golestan with other northern provinces was noticed.

Keywords: Allele; Haplotype diversity; Haplotype discrimination capacity

INTRODUCTION

Short Tandem Repeats (STRs) are DNA regions with repeating units of 2-6 bp, which are frequently found throughout of human genome including autosomal and both sex chromosomes X and Y. In addition, STRs are commonly found in the non-coding regions of DNA [1]. Unlike autosomal STRs, the Y chromosome STRs are in haplotype state due to the lack of counterpart

*Corresponding Author: Department of Biology, Rasht Branch, Islamic Azad University, Rasht, Iran
Tel: +98 1333424094 ; Fax:+98 1333447060
E. mail: salehzadeh@iaurasht.ac.ir

homolog on the X chromosome. In Y chromosome STRs (Y-STRs) are located out of the pseudo-autosomal region of the chromosome and thus, are passed down from father to the male offspring without recombination. In other words, identical Y-STR markers are passed down to the sons. However, Similar to autosomal chromosomes, a relatively high mutation rate of Y-STR markers (with an average mutation rate of 3.17×10^{-3}) leads to Y-STR diversity [2, 3].

Thus, polymorphisms in the Y-STRs are regarded as valuable tools for evolutionary studies which could be used as genetic markers for regional and population studies. In addition, such polymorphisms could be employed for forensic purposes including paternity testing, identification in sexual abuse cases, and proving male kinship with the survivors in mass disasters [4-6]. Golestan province is located in the north-east of Iran and consists of several ethnic groups including Turkmens, Mazandaranis, Sistanis, Baluchis, Qizilbash and others. Due to the high diversity of Y-STR markers in human, finding informatics markers in a specific population is crucially important for forensic and evolutionary investigations.

To our knowledge, very few studies were performed on the Y-STR markers of the populations residing in the north of Iran and no investigation was performed on the population of Golestan province. Thus, the current study was conducted to investigate the frequency of Y-STR haplotypes among the male populations of Golestan province, Iran and to compare them with other parts of Iran.

MATERIALS AND METHODS

A total number of 106 unrelated male volunteers were included in this study. The exclusion criteria were as following :1) to be from Sadat ancestry (a subgroup of the population which thought to be descendants of the prophet Mohammad, the great messenger of Islam), 2) residing in Golestan for less than three previous generations, 3) having blood transfusion during 24 h prior to sampling. All participants signed the consent form and the study was approved by review boards and the ethics committee of the Iranian legal medicine research center.

Blood samples were obtained from individuals on Whatman FTA Classic cards. DNA extraction was performed using a commercial DNA extraction kit according to the manufacturer's protocol (Kousar biotech, Iran). Then, 17 Y-STR markers were amplified using the AmpFLSTR Yfiler PCR amplification kit (Applied Biosystems, Foster City, CA, USA), according to the manufacturer's instruction. Characteristics of Y-STR markers studied in this work were displayed in Table 1.

Table 1: Characteristics of Y-STR markers used in this study

Y-STR marker	Repeat motif	Allele range*
DYS456	AGAT	10-20
DYS389 I	$(TCTG)_3$(TCTA)n	11-15
DYS390	(TCTA)2(TCTG)n(TCTA)n(TCTG)nTCA(TCTA)2	21-31
DYS389 II	(TCTG)n(TCTA)Nn28(TCTG)2(TCTA)n	20-32
DYS458	GAAA	13-20
DYS19/ DYS394	$(TAGA)_3$TAGG(TAGA)n	12-18
DYS385a	GAAA	9-17
DYS385b	GAAA	11-22
DYS393	AGAT	11-16
DYS391	(TCTG)3(TCTG)n	8-11
DYS439	AGAT	9-24
DYS635	TSTA	12-26
DYS392	(TAT)n	10-16
Y GATA H_4	TAGA	9-15
DYS437	TCTR	10-17
DYS438	TTTTC	7-20
DYS448	AGAGAT	11-23

* Allele range represents the range of the number of repeat motifs for each Y-STR marker.

Amplification conditions were as following: primary incubation at 80 °C for 11 min, then 30 cycles of denaturation at 94°C for 1 min, annealing at 61°C for 1 min and extension at 72°C for 1 min, followed by a final extension at 60°C for 8 min. Separation of PCR products was performed using capillary electrophoresis on an ABI 3500 genetic analyzer (Applied Biosystems, Foster City, CA, USA) and the alleles were assigned by GeneMapper ID-X v.1.3 software (Applied Biosystems, Foster City, CA, USA). Allele frequency for each locus and haplotype frequency was determined and then, Gene and haplotype diversity were calculated based on Nei's suggestion [7] by the following formula using the Arlequin Software package v3.5.2.2 and GenAlEx 6.5 software [8, 9]:

$$h = \frac{N}{N-1\ (1-\sum_{Pi}2)}$$

where N is the population size and P_i is the frequency of each haplotype. Haplotype match probability (HMP) was calculated according to HMP = 1–HD formula and haplotype discrimination capacity (DC) was also calculated by dividing the total number of a unique haplotype in the population by the total number of individuals. Finally, genetic distance from the studied population with populations from other regions of Iran was compared via the F_{ST} index.

RESULTS

This study was conducted for genotyping of 17 Y-STR markers among 106 unrelated males from Golestan province to characterize the diversity of the haplotypes. A total number of 87 unique haplotypes were determined in the studied population. The frequency of different Y-STR markers were presented in Table 2. Haplotypes were deposited to the Y-Chromosome Haplotype Reference Database (YHRD: www.yhrd.org) with the accession number of YA004259.

The highest allele diversity was observed for DYS385b with 9 alleles, while the least diversity was recorded for DYS391 (3 alleles). Also, the allele 10 of DYS391 locus was the most frequent allele with a frequency of 0.764. The studied population exhibited gene diversity ranging from 0.368 to 0.825 for different Y-STR markers. The highest gene diversity and thus, the highest polymorphism was associated with the DYS385b locus while the least gene diversity was noticed for the DYS391 locus (Table 3). The average gene diversity for 17 Y-STR markers was calculated 0.687. The HD and DC values were determined 0.9962 and 0.8207, respectively. In addition, comparison of the genetic distance of the studied population with data from Gilan and Mazandaran provinces revealed F_{ST} values of 0.0020 and 0.0021(*p-value*=0.00), respectively.

Table 2: Allele frequency of 17 Y-STR markers in the studied population

Y-STR markers (n=106)														
DYS19			DYS389I			DYS389II			DYS390			DYS391		
allele	n	frequency	allele	n	frequency	allele	n	frequency	allele	n	frequency	allele	n	frequency
12	0	0.000	11	4	0.038	20	0	0.000	21	1	0.009	8	0	0.000
13	10	0.094	12	22	0.208	22	0	0.000	22	27	0.255	9	1	0.009
14	46	0.434	13	60	0.566	27	1	0.009	23	35	0.330	10	81	0.764
15	26	0.245	14	18	0.170	28	16	0.151	24	21	0.198	11	24	0.226
16	20	0.189	15	2	0.019	29	31	0.292	25	22	0.208			
17	4	0.038				30	33	0.311	26	0	0.000			
18	0	0.000				31	22	0.208	31	0	0.000			
						32	3	0.028						

Y-STR markers (n=106)															
DYS392			DYS393			DYS385a			DYS385b			DYS438			
allele	n	frequency	allele	n	frequency	allele	n	frequency	allele	n	frequency	allele	n	frequency	
10	6	0.057	11	6	0.057	9	2	0.019	11	0	0.000	7	4	0.038	
11	66	0.623	12	32	0.302	10	0	0.000	12	3	0.028	8	1	0.009	
12	5	0.047	13	51	0.481	11	23	0.217	13	6	0.057	9	21	0.198	
13	10	0.094	14	12	0.113	12	11	0.104	14	14	0.132	10	38	0.358	
14	14	0.132	15	5	0.047	13	36	0.340	15	17	0.160	11	37	0.349	
15	5	0.047	16	0	0.000	14	20	0.189	16	35	0.330	12	4	0.038	
16	0	0.000				15	9	0.085	17	11	0.104	13	0	0.000	
						16	5	0.047	18	12	0.113	14	0	0.000	
						17	0	0.000	19	5	0.047	20	1	0.009	
									20	3	0.028				
									21	0	0.000				
									22	**0**	**0.000**				

Y-STR markers (n=106)															
DYS439			DYS437			DYS448			DYS456			DYS458			
allele	n	frequency	allele	n	frequency	allele	n	frequency	allele	n	frequency	allele	n	frequency	
9	0	0.000	10	1	0.009	11	0	0.000	10	0	0.000	13	0	0.000	
10	30	0.283	13	0	0.000	15	0	0.000	12	2	0.019	14	5	0.047	
11	34	0.321	14	53	0.500	17	1	0.009	13	3	0.028	15	31	0.292	
12	24	0.226	15	39	0.368	18	7	0.066	14	17	0.160	16	27	0.255	
13	17	0.160	16	13	0.123	19	38	0.358	15	61	0.575	17	22	0.208	
14	0	0.000	17	0	0.000	20	42	0.396	16	15	0.142	18	16	0.151	
15	0	0.000				21	16	0.151	17	6	0.057	19	5	0.047	
24	1	0.009				22	2	0.019	18	1	0.009	20	0	0.000	
						23	0	0.000	20	1	0.009				

Y-STR markers (n=106)					
DYS635			GATA H4		
allele	n	frequency	allele	n	frequency
12	0	0.000	9	1	0.009
19	1	0.009	10	9	0.085
20	6	0.057	11	42	0.396
21	23	0.217	12	46	0.434
22	26	0.245	13	4	0.038
23	28	0.264	14	3	0.028
24	17	0.160	15	1	0.009
25	4	0.038			
26	1	0.009			

Table 3: Effective allele, Shannon's index, and genetic diversity in the studied population

Y-STR locus	Population size	Number of alleles	Effective allele	Shannon's index	Genetic diversity	Total genetic diversity
DYS19	106	5	3.397	1.368	0.706	0.712
DYS389I	106	5	2.537	1.148	0.606	0.612
DYS389II	106	6	4.013	1.479	0.751	0.758
DYS390	106	5	3.901	1.405	0.744	0.751
DYS391	106	3	1.574	0.586	0.365	0.368
DYS392	106	6	2.371	1.236	0.578	0.584
DYS393	106	5	2.934	1.267	0.659	0.665
DYS385a	106	7	4.575	1.676	0.781	0.789
DYS385b	106	9	5.470	1.917	0.817	0.825
DYS438	106	7	3.417	1.391	0.707	0.714
DYS439	106	5	3.845	1.396	0.740	0.747
DYS437	106	4	2.497	1.016	0.600	0.605
DYS448	106	6	3.194	1.318	0.687	0.693
DYS456	106	8	2.622	1.315	0.619	0.624
DYS458	106	6	4.531	1.608	0.779	0.787
DYS635	106	8	4.818	1.696	0.792	0.800
GATA H4	106	7	2.817	1.251	0.645	0.651

DISCUSSION

The provision of population information of Y-STR markers is highly valuable for forensic applications and identifying ancestral lineage. The present study provides information on the genetic polymorphisms of 17 Y-STR alleles in a population from Golestan province, a region previously poorly studied at the genetic level. In addition, the information provided for the 17 Y-STR markers was compared to those for other regions of Iran, including the northern provinces of Gilan and Mazandaran.

Characterization of Y-STR markers in different regions of Iran, including Gilan, Mazandaran, Isfahan, and Tehran has been performed [10-12]. Our study showed that the

highest allele diversity was associated with the DYS385b locus with a frequency of 0.825. This was similar to the results from Gilan and Mazandaran, which presented DYS385b locus as the most polymorphic marker with frequencies of 0.824 and 0.863 respectively [10]. However, these findings were not similar to those from Isfahan province in which the DYS390 was the most polymorphic locus (with frequency of 0.71), and DYS391 showed the least polymorphic content with frequency of 0.44) [11]. Also, the least allele diversity among the studied population was associated with the DYS391 locus (3 alleles) with a frequency of 0.368. This finding was similar to the result of a study on the Gilan population, which presented the DYS391 locus as the least diverse locus with 3 alleles and a frequency of 0.379. However, it was not similar to the population from Mazandaran in which the DYS437 locus indicated the least diversity (3 alleles) with a frequency of 0.612 [10].

Furthermore, the mean gene diversity value found in the studied population (GD= 0.687) was very similar to the values reported for the northern provinces of Iran, including Mazandaran (GD= 0.665) and Gilan (GD= 0.683) [10]. Comparing the F_{ST} of Y-STR markers for the studied population with those reported for Mazandaran and Gilan revealed that the genetic distance between Golestan with Gilan and Mazandaran was 0.0020 and 0.0021(*P-value*=0.00), respectively which were not significantly different.

Furthermore, the STR alleles with 7-32 repeats were observed in the studied population which was similar to those from Gilan and Mazandaran. This work showed that the allele 10 of DYS391 locus was the most frequent allele in the populations from Golestan with frequency of 0.764, which was similar to the results from Gilan and Mazandaran provinces, in which the mentioned allele was the most frequent one with frequency of 0.767 and 0.672, respectively. However, among the population from Mazandaran, the allele 11 of DYS392 was also the most frequent allele with the frequency of 0.672 [10].

Thus, comparing the studied population with data reported for Gilan and Mazandaran, revealed the low genetic differences for the populations from the north of Iran.

Locus diversity by itself is not the most meaningful value for estimating the power of discrimination and thus, haplotype diversity is a more valuable measure for this purpose. In this study, 87 unique haplotypes were identified. HD value for the population from Golestan was calculated 0.9962 which was comparable to the values from other regions of Iran including Gilan (0.9998), Mazandaran (0.9993), eastern provinces of Iran (0.9999), Tehran (0.997) and Isfahan (0.997) [10-12]. Moreover, the DC value of the population from Golestan province was calculated (0.8207) which was lower than previously DC values reported for other regions of Iran, including Gilan (0.9988), Mazandaran (0.9633), Tehran (0.90), Isfahan (DC=0.938 and the population from east of Iran (DC=0.9884) [12]. These findings show lower haplotype diversity of the population of Golestan in comparison with other regions of Iran.

Comparing the genetic diversity of the population from Golestan with those from other northern regions of Iran, showed that there was not significant genetic distance between people residing in the northern parts of Iran. In other words, it seems that the people from the north of Iran have just been separated from each other only by geographical boundaries, while they are genetically similar.

In this study, Y-STR markers were characterized. Our results showed that genotyping of Y-STR markers could be used as a valuable tool for evolutionary studies and forensic applications. However, studying greater population and additional Y-STR markers would provide more detailed information on the characteristics of STR markers in this region to be used for different purposes. Also, to investigate the role of Y-STR in father-son kinship, it is necessary to compare the population of the father-son pair and the rate of mutation.

Acknowledgements: The authors would like to thanks Dr. Arash alipour, for collaborating in some analysis.

Conflict of Interest: The authors declare that they have no conflict of interest exist.

REFERENCES

1. Butler JM. Advanced Topics in Forensic DNA Typing Methodology. Elsevier Inc 2012; 704p.
2. Butler JM. Fundamentals of Forensic DNA Typing: Lineage Markers: Y Chromosome. Elsevier Inc 2010;520p.
3. Butler JM. Recent Developments in Y Short tandem repeat and Y-single nucleotide polymorphism analysis. Forensic Sci Rev 2003;15:91-111.
4. Jobling MA. In the name of the father: surnames and genetics. Trends Genet 2001;17:353-357.
5. Roewer L, Arnemann J, Spurr NK, Grzeschik KH, Epplen JT. Simple repeat sequences on the human Y-chromosome are equally polymorphic as their autosomal counterparts. Hum Genet 1992;89:389-394.
6. Jeffreys AJ. Highly variable minisatellites and DNA fingerprints. Biochem Soc Trans 1987; 15:309-317.
7. Nei M. Molecular evolutionary genetics. New York: Columbia university press 1987;512p.
8. Peakall R, Smouse PE. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 2012;28:2537-2539.
9. 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;10:564-567.
10. Sayyari M, Salehzadeh A, Tabatabai MA, Abbasi A. Profiling of 17 Y-STR loci in Mazandaran and Gilan provinces of Iran. Turk J Med Sci 2019;49:1277-1286.
11. Vatandoost N, Salehi AR, Kazemi M, Khosravi S, Eslami G, Kamali S. Genetic polymorphism of 8 Y-STR loci in native population of Isfahan province in central part of Iran. Ann Hum Biol 2017;44:175-179.
12. Tabrizi AA, 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 2014;14:6-7.