Allele frequency of DYS393 and DYS19 in Iranian Kurdish men

Gholamreza Homayounpour[1,2], Mojgan Hosseini[3], Fatemeh Keshavarzi[1,4*]

1. Iranian Legal Medicine Research Center, Legal Medicine Organization, Tehran, Iran
2. Kurdistan Province, Central office of Legal Medicine, Sanandaj, Iran
3. Department of Genetics, Payame Noor University, P,O,Box 19395-3697, Tehran, Iran
4. Department of biology, Sanandaj Branch, Islamic Azad University, Sanandaj, Iran

***Corresponding author:** Tel: +98 9183704918 Fax: +98 8733288677
Address: Department of biology, Sanandaj Branch, Islamic Azad University, Sanandaj, Iran
E-mail: f.keshavarzi@iausdj.ac.ir
Received: 25/02/2020 Revised: 12/05/2020 Accepted: 30/08/2020

Abstract

Introduction: Repetitive Y sequences in evolutionary studies and tracing of human migration are useful indicators. This study aimed to investigate the allelic frequency and differentiation of DYS393 and DYS19 markers in 200 Iranian Kurdish men.
Materials and Methods: The blood samples were collected from 200 unrelated Kurdish men that three generations of them were Kurd and lived in four Kurdish provinces of Iran (Kurdistan, Kermanshah, Ilam, and West Azerbaijan). After DNA extraction, two markers were evaluated using the Monoplex PCR technique, and then alleles were separated for each marker using the Real-time PCR technique. The obtained data were analyzed by HRM diagram and the final statistical calculations were analyzed by Genalex software (version 4.6).
Results: The genetic diversity of DYS393 and DYS19 markers were 0.69, 0.607 and the number of alleles were 5 and 4 for each of the markers, respectively. Also, the most allele frequency belonged to the alleles 14 and 12 in DYS393 and DYS19 markers, respectively. 12, 9, 8 and 7 haplotypes were belong to Kermanshah, Kurdistan, Azerbaijan and Ilam, respectively, and 2, 1 and 1 specific haplotypes were also observed in Kermanshah, Kurdistan and Azerbaijan provinces, respectively. In addition, Kermanshah province has the highest amount of N_e and H_e, therefore, it has the most influential haplotypes in the population.
Conclusion: The loci studied in the Kurdish population of the west of Iran has a high haplotype diversity and has close similarity with the Caucasian Kurdish immigrants, the Iranian population and the Turkish Kurds. The desired loci can also be used in forensic programs.

Keywords: DYS393, DYS19, Iranian Kurdish, HRM

Introduction

Singular properties of the Y chromosome such as male-specific and haploid alongside its particular characteristic of avoidance to recombination made it a proper candidate for studying the population genetic. Reconstruction of human paternal lineages' history can be done using the DNA polymorphisms of the Y chromosome (1). STRs (short tandem repeats) of the Y chromosome are hypervariable sequences. These are extensively used in kinship analysis and forensic DNA typing. Due to the amplification and visualization of PCR products, allelic genotyping of STRs could be accomplished without using complex molecular techniques (2-3). Identification of the population's ancestry is an important issue in realizing its genetic structure. Most importantly, understanding and discovery of the allelic variance or genetic diversity of a population could serve as a guideline for future researches (1- 6). The DYS393 and DYS19 are markers that have low diversity in the population (4-12). The genetic

diversity or GD value of the DYS393, DYS19 loci were shown over 0.6 in the population of South Tunisian, Iraq, Chinese Han, and Belize (4-10). This study was aimed to find genetic characteristic markers mentioned above in 200 Iranian Kurdish men of Kermanshah, Kurdistan, Ilam and West Azerbaijan provinces.

Materials and Methods

Sampling

This research is a case-control study. The study was ethically approved Iranian Legal Medicine Research Center, Legal Medicine Organization, Tehran, Iran. During this study, the peripheral blood samples were collected in tubes containing anticoagulant EDTA (K2). Kurdish men Iranian were selected randomly among individuals who referred to Blood Transfusion Organization in every 4 provinces. After obtaining informed consent, blood samples were obtained from 200 unrelated Iranian Kurdish males from Kurdistan, Kermanshah, Ilam, and West Azerbaijan provinces in Iran.

DNA Extraction

Genomic DNA was extracted from the whole peripheral blood samples using the standard extraction protocols kit (DN8115c.cn) and then be electrophoresed on %1 agarose gel together with allelic ladders. The obtained results were analyzed using Gel Girgi Doc((51)2000C). Sharp bands represent a high rate of extraction. The isolated DNAs were then placed in separate microtubes and stored at -24 ° C until PCR was performed.

Polymerase Chain Reaction

STR primers (DYS393 Fwd: 5'-GTGGTCTTCTACTTGTGTCAATAC-3 'Rev: 5'- CCTCATTTTTTGGACTTGAGTT-3' and DYS19: Fwd: 5'-CTACTGAGTTTCTGTTATAGT-3' Rew: 5'-TGTCCTCACTACATGCCAT-3') were taken from the STRbase DNA database (https://www.cstl.nist.gov/strbase). Monoplex-PCR to determining annealing temperature was performed for each STR in 30 cycles, separately. PCR reaction carried out in total volume of containing 25µl(12/5µl Master Mix,7/5µlH2O,1µl Forward Primer, 1µl Reverse Primer,3µl DNA) and by a Rotor-Gene Q thermal by the gradient program of the thermal cycler. After amplification, 3 µl of the samples were transferred to 1% agarose gel and the best bond was selected as the analytical temperature. Analyzing temperature was selected for markers 393(60) and 19 (57) degrees.

Real Time-PCR and HRM

To amplification the target region of the genome, the Real time-PCR method was used. Each DNA sample was run in a total volume of 20 µl containing (4µl Master Mix containing the fluorescent dye Eva Green _14µl H2O_ 1µl Forward Primer_ 1µl Reverse Primer) for up to 40 cycles. The thermal profile consists of 95℃ for 12 and then 95 °C for 15s, 60 °C for the 20s, 72 °C for 20s for 40 cycles in the following 72℃ for 10 min for DYS393 and so 95℃ for 12 and then 95 °C for 15s, 60 °C for 20s, 72 °C for 20s for 40 cycles in the following 72 °C for 10 min about DYS19. Amplification was followed by HRM(high-resolution melting) method for identifying genetic polymorphism as a Post-PCR process (14-15). To determining the precise nucleotide length and the number of observed alleles from each of the curve types, a sample was selected, and the desired DNA was purified. The purification of DNA from samples was carried out with the use of the KIA GENE PCR template purification Kit (DENA Gene, IRAN). The purified DNA was forwarded to Kowsar laboratory for sequencing and it was sequenced with an ABI3700 capillary sequencer machine (Figure 1). The genetic characteristic of DYS393 and DYS19 markers are depicted in Table 1.

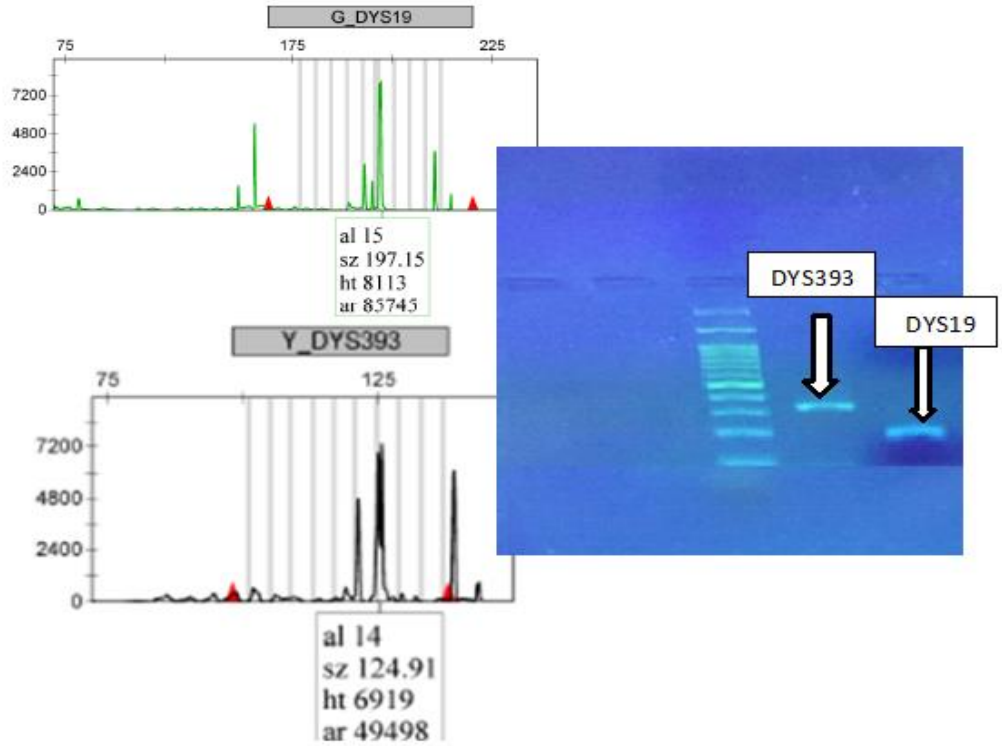

Figure 1. To determine the precise nucleotide length and the number of observed alleles from each of the curve types, a sample forwarded to sequencing and it was sequenced with an ABI3700 capillary sequencer machine and compared to bands belong to two markers on gel agarose.

Table 1. The genetic characteristic of DYS393 and DYS19 markers.

YSTR	Physical position	Allele Repeat	Repeat Motif	Allele Range	Observed Alleles	GenBank Accession
DYS393	ChrYp 14.103 Mb	12	AGAT	7-18	17	AC006152
DYS19	ChrYp 10.13 Mb	12	TAGA	10 – 19	15	AC017019

Data Analysis

Allele frequencies, Gene diversity, and haplotype diversity for each locus were computed using the Arlequin software (version3.5). Gene Alex software(version4.6) was used to calculate the F-st value between the population.

Results

The Figure 2 was presented as an example of an HRM diagram of DYS393 and DYS19 loci. The results in comparison with control samples showed that the genetic diversity of DYS393 and DYS19 markers were 0.69, 0.607 and the number of alleles for each of the two markers were 5 and 4, respectively. Also, the most allelic frequency belonged to the alleles 14 and 12 in DYS393 and DYS19 markers, respectively. 12, 9, 8 and 7 haplotypes were belong to Kermanshah, Kurdistan, Azerbaijan and Ilam, respectively. and 2, 1 and 1 specific haplotypes were also observed in Kermanshah, Kurdistan and Azerbaijan provinces, respectively. In addition Kermanshah province has the highest amount of N_e and H_e, therefore, it has the most influential haplotypes in the population(table 2,3). In Figure 3 the number of haplotypes was compared among of all provinces, in which haplotype 3 had the highest frequency in 3 provinces.

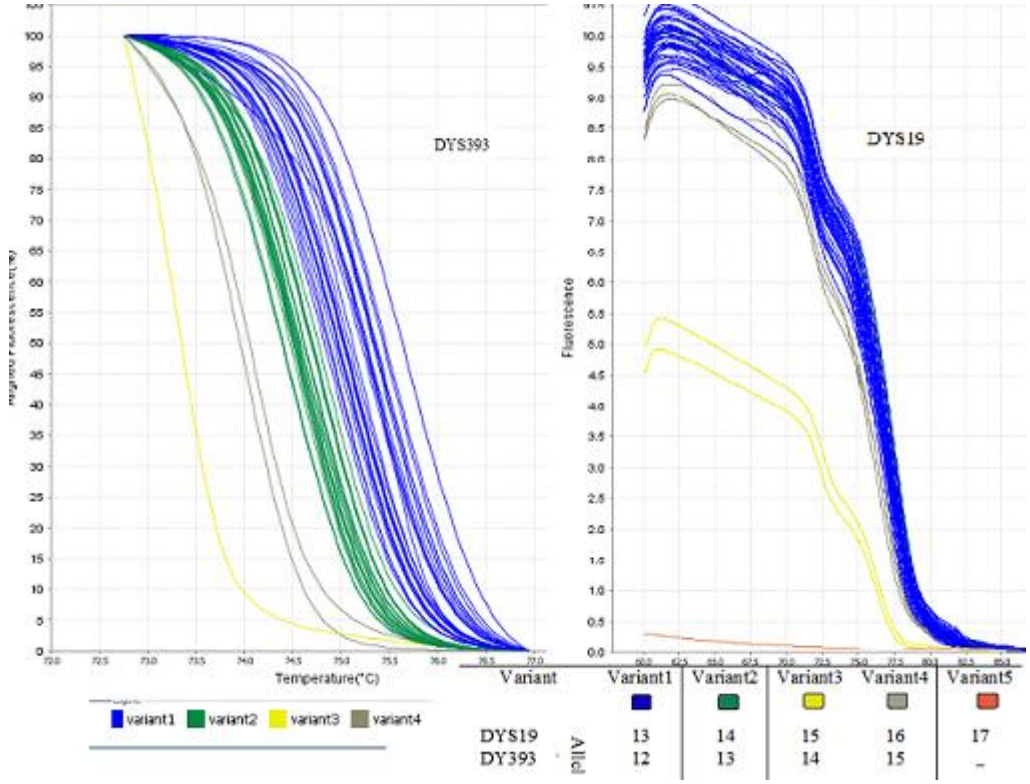

Figure 2. An example of HRM diagram of DYS393 and DYS19 loci. As shown, the HRM technique classified samples into different variants based on differences in Tm. In Figures, different variants are separated by different colors in the graph, and each color represents a variant. People in a variant are displayed in the same color and are thus separated from each other. Since Known samples with a specific profile were among the run samples, based on the number and code of Known samples, the allele number of each variant and its color were determined in the relevant diagrams, and the allele size of other samples was determined based on placement in each variant. In this method, samples were classified according to the chart, and the alleles frequency was obtained by matching with control samples.

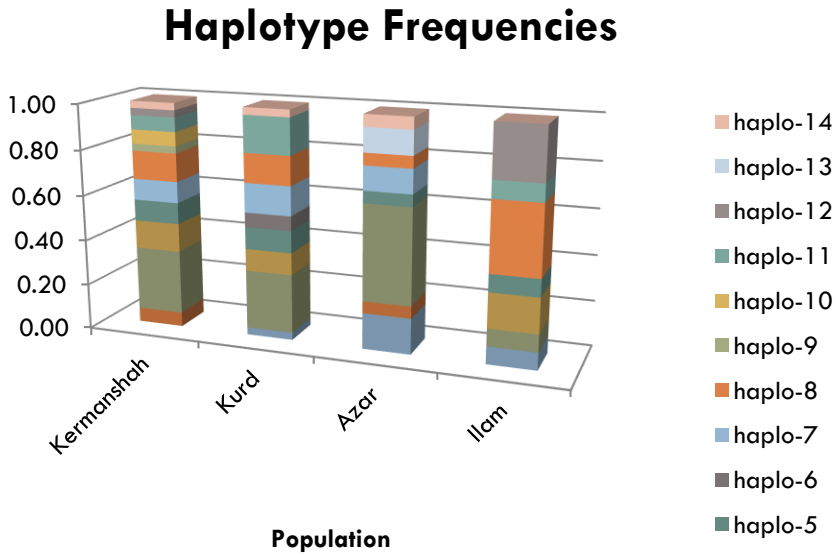

Figure 3. The number of haplotypes is compared among of all provinces, in which the haplotype No. 3 has the highest frequency in the provinces.

Table 2. The specific haplotypes and frequency of them in each province.

Haplotype Code	Kermanshah	Kurdistan	Azerbaijan	Ilam
haplo-1	0	1	3	1
haplo-2	2	0	1	0
haplo-3	9	8	8	1
haplo-4	4	3	0	2
haplo-5	3	3	1	1
haplo-6	0	2	0	0
haplo-7	3	4	2	0
haplo-8	4	4	1	4
haplo-9	1	0	0	0
haplo-10	2	0	0	0
haplo-11	2	5	0	1
haplo-12	1	0	0	3
haplo-13	0	0	2	0
haplo-14	1	1	1	**0**

Table 3. Comparison of the genetic impact on the studied populations.

Population	N	A	P	N_e	R_h	H_e
Kermanshah	64	12	2	7.014	6.575	0.885
Kurdistan	62	9	1	6.628	5.896	0.877
Azarbaijan	38	8	1	4.247	5.541	0.807
Ilam	26	7	0	5.121	6.000	0.872
Mean	50	8.750	1.000	5.752	6.003	0.860

N:sample size; A: Number of haplotypes detected in each population; P: Number of private haplotypes; N_e: Effective number of haplotypes; R_h: Haplotypic richness; H_e: Genetic diversity.

Table 2 shows the specific haplotypes and frequency of them in each province. As can be seen, 12, 9, 8 and 7 haplotypes belong to Kermanshah, Kurdistan, Azerbaijan and Ilam, respectively, and 2, 1 and 1 specific haplotypes were observed in Kermanshah, Kurdistan and Azerbaijan provinces, respectively. Also, Kermanshah province has the highest amount of N_e and H_e, therefore has the most influential haplotypes on the population (Table 3).

Discussion

In Kurdish populations, there have been very few studies. Previous genetic studies showed that there is generally the genetic similarity of the Kurds to other Middle Eastern populations (16). As can be seen in figure 3 and Table 2, in 3 provinces of Kurdistan, Kermanshah, West Azarbaijan, the highest frequency of DYS19 allele was related to allele 12 and in Ilam province was related to allele 13. Also, the highest frequency of DYS393 allele marker in all 4 provinces was related to the 14 alleles. Compared to other studies (13, 16), it is observed that in the Kurdish population of Iraq and the Caucasus, the highest frequency of DYS19 allele markers was related to 12 and 13 alleles, so it can be concluded that these two alleles in these populations can be important be a demographer. In the marker DYS393, the highest repetition of alleles is related to the 14 allele, and this abundance is maintained in the Kurdish populations of Iraq and the Caucasus in the same way, which indicates the intense regeneration within the population in these communities. In the Tehran population, the highest allelic replication in DYS19 is related to allele 16, in DYS393 is related to allele 10 (17). The results based on the allelic frequency obtained indicate that the Kurds of Iran are most similar to each other. This similarity between repetitive alleles in all 4 Kurdish

provinces in the west of the country indicates a high percentage of inbreeding among the Kurds. Of course, despite the similarity of the Kurds with each other compared to other regions, there are differences within the Kurdish population itself, and specific haplotypes indicate this. After that, the Iranian Kurds in the study population are most similar to the Iraqi Kurds and then to the immigrant Kurds living in the Caucasus. Also, the study population is similar to the population of Isfahan(18) and Tehran(17), but their similarity with each other is less than the similarity of the population of Kurdish areas; Which shows the ancient history of the Kurds and Persians. According to the history of Kermanshah, which was the capital of the Medes, at the time of the defeat of the Medes from the Achaemenids, the capital was transferred from the Kurds to the Persians. According to this history, the Kurds and the Persians lived together and had much less reproduction than the Kurds themselves during these long years. We can also talk about the migration of Kurds to these areas in the many years since the exchange of the capital. On a larger scale, the study population can be compared with other countries such as Turkey, China, and the Caucasus in the north. When compared with published data, other Kurdish groups were compared with Europe, the Caucasus, and the West and Central Asian groups, showing that Kurdish groups were genetically similar to other West Asian groups and more distant than Central Asian groups (18) and (19). The similarity of the Iranian Kurdish population with the Iraqi Kurds can be explained by the closer ethnic and racial proximity of this population to the Iraqi Kurdish tribes (preservation of common ancestral markers) (18). Also, the reason for the similarity with non-Kurdish populations of Iran such as Tehran and Isfahan is the geographical proximity of these populations and to some extent the mixing of these two populations (17) and (16); however, this population is significantly different from other populations in the world.

Conclusion

The loci studied in the Kurdish population of the west of the country have a high haplotype diversity and are very similar to the Caucasian Kurdish immigrants, the Iranian population and the Turkish Kurds. The desired loci can also be used in forensic programs.

Acknowledgments

The study was ethically approved by the Legal Medicine Research Center, Legal Medicine Organization, Tehran, Iran(IR. LMO. REC. 1396. 5). Also, the authors would like to thank the volunteers who contributed samples for this study.

References

1. Iida R, Kishi K. Identification, Characterization and forensic application of novel Y-STRs. Leg Med. 2005; 7 (4) :255-8. doi: 10.1016/j.legalmed.2004.12.004.
2. Larmuseau MH, Vanderheyden N, Van Geystelen A, Decorte R. A substantially lower frequency of uninformative matches between 23 versus 17 Y-STR haplotypes in north Western Europe. Forensic Sci Int Genet. 2014; 11:214-9. doi: 10.1016/j.fsigen.2014.04.002.
3. Hanson EK, Ballantyne J. A Highly Discriminating 21 Locus Y-STR 'Megaplex' System Designed to Augment the Minimal Haplotype Loci

for Forensic Casework. J Forensic Sci. 2004;49 (1): 40-51. doi: 10.1371/journal.pone.0000688.

4. Hall A, Ballantyne J. The development of an 18-locus Y-STR system for forensic casework. Anal Bioanal Chem. 2003;376(8):1234-46. doi:10.1007/s00216-003-2039-2.
5. Bosch E, Lee AC, Calafell F, Arroyo E, Henneman P, de Knijff P, Jobling MA. High-resolution Y chromosome typing: 19 STRs amplified in three multiplex reactions. Forensic Sci Int. 2002;125 : 42-51 doi: 10.1016/s0379-0738(01)00627-2.
6. Nuzzo F, Caviezel F, De Carli L. Lancet. 1996;2 (7457): 260-2. doi: 10.1016/s0140-6736(66)92544-x.
7. Roewer L, Epplen JT. Rapid and sensitive typing of forensic stains by PCR amplification of polymorphic simple repeat sequences in casework. Forensic Sci Int. 1992;53 (2):163-71. doi.org/10.1016/0379-0738(92)90193-Z.
8. Roewer L, Krawczak M, Willuweit S, Nagy M, et al. A new method for the evaluation of matches in non-recombining genomes: application to Y-chromosomal short tandem repeat (STR) haplotypes in European males. Forensic Sci Int. 2001;118 (23) : 106-13. doi: 10.1016/s0379-0738(00)00287-5.
9. Kayser M, Brauer S, Willuweit S, Schädlich H, Batzer MA, Zawacki J, Prinz M, Roewer L, Stoneking M. Online Y-chromosomal short tandem repeat haplotype reference database (YHRD) for U.S. populations. J Forensic Sci. 2002; 47 (3):513-9. doi.org/10.1520/JFS15290J.
10. Lessig R, Willuweit S, Krawczak M, Wu FC. Asian online Y-STR Haplotype Reference Database. Leg Med. 2003;160-163. doi: 10.1016/s1344-6223(02)00100-1.
11. Gao T, Yun L, Gao S, Gu Y. Population genetics of 23 Y-STR loci in the Mongolian minority population in Inner Mongolia of China. Int J Legal Med. 2016;130 (6):1509-11. doi: 10.1007/s00414-016-1433-1.
12. Mohd Yussup SS, Marzukhi M, Md-Zain BM, Mamat K, Mohd Yusof FZ. Polymorphism of 11 Y Chromosome Short Tandem Repeat Markers among Malaysian Aborigines. Evol Bioinform. 2017;17;13:1176934317735318. doi: 10.1177/1176934317735318.
13. Ozbas-Gerceker F, Bozman N, Arslan A, Serin A. Population Data for 17 Y-STRs in Samples from Southeastern Anatolia Region of Turkey. Int J Hum Genet. 2013;13 (2) : 105-11. doi: 10.1080/09723757.2013.11886204.
14. Deng JQ, Liu BQ, Wang Y, Liu W, Cai JF, Long R, Li WH. Y-STR genetic screening by high-resolution melting analysis. Genet Mol Res. 2016;15(1). doi: 10.4238/gmr.15017266.
15. Nicklas JA, Noreault-Conti T, Buel E. Development of a Fast, Simple Profiling Method for Sample Screening Using High Resolution Melting (HRM) of STRs. J Forensic Sci. 2012;57 (2): 478-88. doi: 10.1111/j.1556-4029.2011.01981.x.
16. Kareem M, Jabor M, Hamed I. Allele frequency present within the DYS635, DYS437, DYS448, DYS456, DYS458, YGATAH4, DYS389I, DYS389II, DYS19, DYS391, DYS438, DYS390, DYS439, DYS392, DYS393, DYS385 of unrelated individuals in Iraq. African J Biotechnol. 2015: 14(10): 851-58.
17. Farazmand A, Sokhansanj A, Rafee M, Mehrabani YH. Comparative analysis of Y-chromosomal short tandem repeats (YSTRs) polymorphism in an Iranian

Sadat subpopulation. JUST. 2009;35 (1):7-12. doi: 10.4103/2277-9175.188493.

18. Vatandoost N, Salehi AR, Kazemi M, Khosravi S, Eslami G, Kamali S, et al. Genetic polymorphism of 8 Y-STR loci in the native population of Isfahan province in the central part of Iran. Ann Hum Biol. 2017;44 (2) :175-9. doi: 10.1080/03014460.2016.1200671.
19. Reed GH, Kent JO, Wittwer CT. High-resolution DNA melting analysis for simple and efficient molecular diagnostics. Pharmacogenomics. 2007; 8 (6): 597-608. doi: 10.2217/14622416.8.6.597.