Available online at www.medicinescience.org

ORIGINAL ARTICLE

Medicine Science International Medical Journal

Medicine Science 2024;13(4):763-8

The distribution of allele frequencies of the 15 Short Tandem Repeats loci in Elazığ population, Türkiye

Aziz Aksoy[1], Mehmet Tokdemir[2], Vahit Konar[3]

[1]Malatya Turgut Özal University, Faculty of Engineering and Natural Sciences, Department of Bioengineering, Malatya, Türkiye
[2]İzmir Katip Çelebi University, Faculty of Medicine, Department of Forensic Medicine, İzmir, Türkiye
[3]Amasya University, Science and Literature Faculty, Department of Biology, Amasya, Türkiye

Received 12 August 2024; Accepted 14 September 2024
Available online 27 September 2024 with doi: 10.5455/medscience.2024.08.092

Abstract

The DNA within the Human Genome involves loci with high polymorphism such as short tandem repeats (STR). The STR polymorphism analysis is extensively used for numerous purposes such as identification studies in forensic sciences, paternity tests, genome mapping, and population studies. In this study, blood and swab samples were obtained from 200 unrelated healthy individuals living in Elazig province. DNA was amplified via the AmpFlSTR Identifiler PCR Amplification Kit (Applied Biosystems, USA) to analyze the 15 STR loci. Using the ABI PRISM310 Genetic Analyzer capillary electrophoresis, the samples were analyzed and the allelic ladder was used to compare and identify the alleles of the samples, the distribution of the allele frequencies and peak values were examined. For statistical analysis Arlequin v.3.11 and PowerStats v.1.2 were used to assess the data. To determine if the allele frequencies were compatible with the Hardy-Weinberg (HW) equilibrium, the Markov chain and the Fisher's exact test p-value ($p<0.05$) were examined. It was determined that the D8S1179 (0.00002), D21S11 (0.00593) and vWA (0.04939) loci were not compatible with the HW equilibrium, but the rest of the loci were compatible. For the loci, the combined power of exclusion is 0.9999995, the combined power of discrimination is 0.9999999999998. The probability of matching for the 15 STR loci was calculated to be 1 in $1.364x10^{-17}$. The power of discrimination was quite high when all the loci were considered. The results of this study indicate that the populations residing in the Elazığ province may benefit from paternity testing, forensic identification, and genomic mapping using the data as an appropriate marker.

Keywords: Short tandem repeat, AmpflSTR, polymorphism, genetic analyzer, human identification

Introduction

In the human genome, there are sequences having a length of 2-7 base pairs, repeating continuously, and usually functionless. These sequences are named as microsatellites or Short Tandem Repeats (STRs) [1,2]. Since the STR loci are highly heterozygotic, the diversity of these loci can be effectively used for studies in population genetics, chromosome mapping and forensic purposes [3]. The STR loci were first used effectively in the early 1990s for identification purposes [4]. 13 genetic markers (D8S1179, D21S11, D7S820, CSF1PO, D3S1358, THO1, D13S317, D16S539, VWA, TPOX, D18S51, D5S818, FGA) were added to these loci by the FBI in 1997, and they have been used as DNA databases for forensic case solving for decades [5,6]. In addition to the loci such as the FGA, THO1 VWA, D3S1358, D8S1179, D16S539, D18S51, and D21S11, which started to be used in Europe, the D2S1338 and D19S433, consisting of tetra-nucleotide repeating sequences, were also used [6,7]. In early 2001, with the multiplex Polymerase Chain Reaction (PCR) system for 15 STR loci, (the 13 genetic markers aforementioned, D2S1338, D19S433 and Amelogenin) markers were amplified in a single tube and tested [8]. The effectiveness of this procedure was approved for its use in identification and paternity tests [9,10]. This study is conducted to determine the gene frequency data in the population of the Elazig province, which is located in the Eastern Anatolian Region (Figure 1), one of the seven geographical regions in Türkiye [11].

CITATION
Aksoy A, Tokdemir M, Konar V. The distribution of allele frequencies of the 15 Short Tandem Repeats loci in Elazığ population, Türkiye. Med Science. 2024;13(4):763-8.

Corresponding Author: Aziz Aksoy, Malatya Turgut Özal University, Faculty of Engineering and Natural Sciences, Department of Bioengineering, Malatya, Türkiye
Email: aksoy_aziz@hotmail.com

DOI: 10.5455/medscience.2024.08.092

Figure 1. Elazığ province which is in the Eastern Anatolian Region in the Elazığ, one of the seven geographical regions in Türkiye

Material and Methods

Population: The study was comprised of 200 unrelated healthy individuals of the province of Elazığ. Blood and swab samples were taken from 200 volunteers in 2008-2009 and kept in the cold chain in the laboratory for study.

DNA extraction: Using the Invisorb Spin Blood and Tissue micro kit (Invitek GmbH, Berlin, Germany), DNA was extracted from each blood and swab samples [10].

Each of the 200 participants in the study had a 5 ml blood sample obtained. 200 swab samples were also collected from the same individuals. Up until they were examined, all samples were kept at -20 C.

PCR Amplification: DNA specimens were amplified using AmpFlSTR Identifiler kit (Applied Biosystems, Foster City, CA, USA [12]. For PCR, 10 µl AmpFlSTR PCR Reaction Mix, 5.5 µl AmpFlSTR Identifiler Primer Set, 0.5 µl AmpliTaq Gold DNA Polymerase, and 9 µl DNA specimens were taken into 0.5 µl PCR tubes. Using a 310 Genetic Analyzer (Applied Biosystems) and the GeneScan500-LIZ internal lane size standard (Applied Biosystems, USA), PCR fragments were separated by capillary electrophoresis in accordance with manufacturer's instructions. A Thermal Cycler 3200 (MyGenieTM 96/384 Gradient Thermal Block, Bioneer Corporation,USA) device was adjusted to 25 µl volume, initial denaturation for 11 minutes at 95°C; for 28 cycle, 1 minute denaturation at 94°C; bonding of the primers for 1 minute at 59°C; primer elongation for 1 minute at 72°C; and 60 minutes at 60°C for the termination step.

DNA Analysis

PCR products were processed in capillary in a Macintosh version capillary electrophoresis ABI PRISM 310 Genetic Analyzer (GeneScan software version 3.1.2, Applied Biosystems). The specimens were processed in Performance Optimized Polymer (POP-4), which formulates DNA repeat regions at a desired solubility with low viscosity, and they were analyzed in Data Collection software. Peak values of the specimens were determined and compared according to reference ladder included in the PCR amplification kit. Allele frequencies and the statistical data was shown in Table 1.

Statistical Analysis

Allele data of each individual was recorded in TextPad file format. The observed and expected heterozygosity, and P-Values were calculated in Arlequin Software v3.11 [13]. Heterozygosity (h) and homozygosity (H) percentage values, power of exclusion (PE), power of discrimination (PD), probability of matching (PM), typical paternity index (TPI), polymorphic information content (PIC) data for the population was calculated in PowerStats version 1.2 software [14,15] and were shown in an allelic table (Table 1). Data was analyzed separately for each locus, based on the Hardy-Weinberg equilibrium P>0.05 value [16].

Table 1. 15 STR loci allele frequencies in Elazığ and statistical analyses

Loci Allele	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	THO1	D13S317	D16S539	D2S1338	D19S433	vWA	TPOX	D18S51	D5S818	FGA
6						**0.255**									
7			0.040			0.195						0.010			
8	0.015		0.200			0.135	0.105	0.035				**0.507**			
9	0.005		0.085	0.028		0.240	0.065	0.148				0.080		0.073	
9,3						0.170									
10	0.078		**0.281**	0.245			0.085	0.093		0.020		0.098	0.005	0.098	
11	0.075		0.223	**0.339**		0.005	0.305	**0.312**		0.013		0.262	0.015	**0.357**	
12	0.105		0.148	0.323			**0.342**	0.279		0.085		0.043	0.135	0.354	
12,2										0.003					
13	**0.272**		0.015	0.060	0.008		0.078	0.107		0.223	0.008		0.134	0.113	
13,2										0.013					
14	0.240		0.005	0.005	0.073		0.020	0.023		**0.331**	0.100		**0.212**	0.005	
14,2										0.030					
15	0.150		0.003		0.243			0.003		0.112	0.100		0.152		
15,2										0.073					
16	0.050				0.255				0.025	0.063	0.239		0.125		
16,2										0.013					
17	0.010				*0.258*				0.173	0.018	**0.277**		0.090		
17,2										0.003					
18					0.153				0.068		0.175		0.053		0.003
19					0.010				0.119		0.085		0.043		0.040
20									0.157		0.013		0.025		0.085
21									0.055		0.003		0.003		0.184
22									0.053						0.183
23									**0.193**				0.005		**0.214**
24									0.104				0.003		0.168
25		0.003							0.040						0.080

Hom: homozygosity percentage, Het: heterozygosity percentage, Ho: observed heterozygosity, He: expected heterozygosity, P: Fisher's Exact test P-value for Hardy-Weinberg Equilibrium, MP: matching probability, PD: power of discrimination, PIC: polymorphic information content, PE: power of exclusion, TPI: typical paternity index

DOI: 10.5455/medscience.2024.08.092

Med Science 2024;13(4):763-8

Table 1. 15 STR loci allele frequencies in Elazığ and statistical analyses

Loci Allele	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	THO1	D13S317	D16S539	D2S1338	D19S433	vWA	TPOX	D18S51	D5S818	FGA
26		0.003							0.013						0.033
27		0.020													0.010
28		**0.196**													
29		0.193													
29,2															
30		0.186													
30,2		0.023													
31		0.040													
31,2		0.150													
32		0.018													
32,2		0.100													
33															
33,2		0.058													
34,2		0.010													
Total	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Statistical parameters of 15 STR loci Elazığ province															
Hom	11.5%	18.0%	19.5%	28.5%	13.5%	21.5%	23.5%	14.0%	17.5%	20.5%	**10.5%**	**34.5%**	12.0%	27.0%	13.5%
Het	88.5%	82.0%	80.5%	71.5%	86.5%	78.5%	76.5%	86.0%	82.5%	79.5%	**89.5%**	**65.5%**	88.0%	73.0%	86.5%
Ho	0.88500	0.82500	0.80500	0.71500	0.86500	0.78000	0.76500	0.86000	0.82500	0.79500	**0.89500**	**0.65500**	0.88000	0.73000	0.86500
He	0.82219	0.85198	0.80185	0.71778	0.78193	0.79410	0.76269	0.78229	**0.87363**	0.80971	0.80935	**0.65784**	0.86791	0.72053	0.84366
PM	0.073	0.044	0.075	0.133	0.097	0.078	0.096	0.091	**0.033**	0.058	0.078	**0.165**	0.037	0.136	0.052
PD	0.927	0.956	0.925	0.867	0.903	0.922	0.904	0.909	**0.967**	0.942	0.922	**0.835**	0.963	0.864	0.948
PIC	0.80	0.83	0.77	0.66	0.74	0.76	0.73	0.75	**0.86**	0.79	0.78	**0.61**	0.85	0.67	0.82
PE	0.765	0.637	0.608	0.452	0.725	0.572	0.536	0.715	**0.646**	0.590	0.785	**0.362**	0.755	0.476	0.725
TPI	4.35	2.78	2.56	1.75	3.70	2.33	2.13	3.57	**2.86**	2.44	4.76	**1.45**	4.17	1.85	3.70
P	0.99609	0.17009	0.56891	0.47312	1.00000	0.35777	0.57283	0.99902	**0.03226**	0.31574	1.00000	0.53372	0.73118	0.72141	0.80938

Hom: homozygosity percentage, Het: heterozygosity percentage, Ho: observed heterozygosity, He: expected heterozygosity, P: Fisher's Exact test P-value for Hardy-Weinberg Equilibrium, MP: matching probability, PD: power of discrimination, PIC: polymorphic information content, PE: power of exclusion, TPI: typical paternity index

DOI: 10.5455/medscience.2024.08.092

DOI: 10.5455/medscience.2024.08.092

Results

The present investigation compared all pertinent studies that examined genotypic data in Türkiye with respect to allele frequencies, rates of heterozygosity and homozygosity, power of exclusion, power of discrimination, probability of matching, polymorphic information content, and typical paternity index. The comparisons did not yield any significant differences ($p>0.05$). The loci's power of discrimination was determined to be 0.835-0.967, its power of exclusion to be 0.362-0.785, its combined power to be 0.9999995, and its combination power to be 0.9999999999998. The probability of matching for the 15 STR loci was calculated to be 1 in $1.364x10^{-17}$. The compatibility of the allele frequencies with the HW equilibrium was tested using the Markov chain and the Fisher's exact test p-value ($p<0.05$). The p value was found to be significant for the locus D2S1338 (P=0.03226*) (*Through the Bonferroni correction ($\alpha=0.05/15=0.00333$), it was observed that the p-value ($P>0.0033$) was compatible with the HW equilibrium). The p-values were insignificant ($p>0.05$) for the other loci. Our findings for the Hardy-Weinberg equilibrium are consistent, which suggests that it is unaffected by selection, mutation, and other evolutionary processes and that it still mates randomly. It also implies that the ratios can stay constant over generations and that future generations will be protected in Elazığ.

Discussion

The P value was tested for observed (Ho) and expected (He) values, and for the compatibility with the Hardy-Weinberg equilibrium. The outcomes were contrasted with the earlier research data from Turkish studies, and each locus's similarities and differences were assessed independently. The review of the literature indicates that the studies have determined the P-value for following loci to be significant ($p<0.05$). D2S1338 (P=0.00867) by Tokdemir et al. (2016) in Eastern Türkiye [17], FGA (P=0.027) by Ülküer et al. (2004) in Turkish population [18], D5S818 (P=0.001) by Çakır et al. (2003) in Van-Ağrı districts of Türkiye [19] and D16S539 (P=0.042) by Akbaşak et al. (2001) in Malatya, Türkiye [20]. Gürkan et al. (2015) conducted a study with multiplex CODIS among Turkish Cypriots and no significant deviations were found in the loci (significance level: $P<0.05$) except D7S820 (P=0.09259) and vWA loci (P=0.00852) [21]. In a study by Özkorkmaz et al [22], in Turkish individuals living in Türkiye conformity to HW equilibrium was observed in all analyzed loci, and allele 22 was found in vWA locus and allele 9 was found in D19S433 locus. However, 29.2 (0.004) allele, which we identified in our region for the first time in D21S11 locus, was also reported by Ülküer et al. [18] and Aşıcıoğlu et al [23]. Apart from these findings, no significant difference was observed when compared to the previous studies in Türkiye [24-28]. Therefore, it can be suggested that these loci can be used as a suitable marker in genome mapping, forensic identification and paternity testing for populations living in the Elazığ province of the Eastern Anatolia region of Türkiye.

Conclusion

The power of discrimination was quite high when all the loci were considered. The heterozygosity values were very high for all the loci; thus, these loci can precisely discriminate two unrelated individuals. This result confirms the applicability of the 15 STR loci in identification. The power of discrimination was quite high when all the loci were considered. To conclude, this study will be beneficial in numerous aspects including forensic analyses, paternity tests, and especially population genetics and genetic mapping. The findings of this study suggest that paternity testing, forensic identification, and genomic mapping utilizing the data as a suitable marker may be beneficial for the communities living in the Elazığ province.

Conflict of Interests

The authors declare that there is no conflict of interest in the study.

Financial Disclosure

We would like to thank the Fırat University Scientific Research Projects Unit (FÜBAP) for supporting this project with the 1828 project number.

Ethical Approval

This study received ethical approval from the Local Ethical Committee of the Fırat University Medical Faculty (Approval number: 2009/06-VII, date: 21.04.2009).

References

1. Butler JM. Forensic DNA typing: biology, technology, and genetics of STR markers. 2nd edition. New York: Elsevier. 2005;1-668.
2. Saiz M, Haarkötter C, Martinez-Gonzalez LJ, et al. Genetic population flows of Southeast Spain revealed by STR analysis. Genealogy. 2023;7:29.
3. Dönbak L. Kısa ardarda tekrar eden DNA Dizilerinin Adli Amaçlı DNA çalışmalarındaki yeri. Turkiye Klinikleri Journal of Medical Sciences. 2002;22:233-8.
4. Wyner N, Barash M, & McNevin D. Forensic autosomal short tandem repeats and their potential association with phenotype. Front Genet. 2020;11:884.
5. Sabbah MA, Al-Zubaidi MM, Al-Janabi TY, et al. Short tandem repeat (STR) variation from 6 cities in Iraq based on 15 loci. J Genet Eng Biotechnol. 2023;21:160.
6. Butler JM. Genetics and genomics of core short tandem repeat loci used in human identity testing. J Forensic Sci. 2006;51:253-65.
7. Silva NM, Pereira L, Poloni ES, et al. Human neutral genetic variation and forensic STR data. PLoS One. 2012;7:e49666.
8. Haddadan SG, Babaei E, Setareh M. Population-genetic and comparative interpopulation studies of the 15 autosomal STR markers in the population living in the Northwest of Iran. Ann Hum GeneT. Published online June 19, 2024. doi:10.1111/ahg.12564
9. Messina F, Finocchio A, Akar N, et al. Enlarging the gene geography of Europe and the Mediterranean area to STR loci of commonforensic use: longitudinal and latitudinal frequency gradients. Ann Hum Biol. 2018;45:77-85.
10. Invisorb Forensic Kit I, Technical Bulletin. Invitek. https://www.invitek.com/en/infectious-diseases/pdp-invisorb-spin-universal-kit access date 07.08.2024

DOI: 10.5455/medscience.2024.08.092

11. NN Vikipedi, Göktürk. Türkiye Doğu Anadolu Bölgesinin iller haritası, Elazığ İli. http://www.nenedirvikipedi.com/cografya/dogu-anadolu-bolgesinin-illeri-7859.html. access date 07.08.2024.

12. Frank WE, Llewellyn BE, Fish PA, et al. Validation of the AmpFℓSTR™ profiler plus PCR amplification kit for use in forensic casework. J Forensic Sci. 2001;46:642-6.

13. Excoffier L, Laval G, Schneider S. Arlequin (version 3.0): An integrated software package for population genetics data analysis. Evol Bioinform Online. 2005;1:47-50.

14. Tereba A. Tools for analysis of population statistics. Profiles in DNA, 1999;2:14-6. https://www.bioz.com/result/powerstats%20v1%202%20software/product/Promega access date 07.08.2024.

15. Power-Stats for Windows: http://download.cnet.com/Power-Stats/3000-2648_4-74105.html access date 07.08.2024

16. Wigginton JE, Cutler DJ, Abecasis GR. A note on exact tests of Hardy-Weinberg equilibrium. Am J Hum Genet. 2005;76:887-93.

17. Tokdemir M, Tunçez FT, Vicdanli NH. Population genetic data for 15 autosomal STR markers in Eastern Turkey for 15 autosomal STR markers in Eastern Turkey. Gene. 2016;586:36-40.

18. Ülküer U, Kurtuluş-Ülküer M, Elma C, et al. Short tandem repeat polymorphism in Turkish population. J Genet. 2004;83:197-9.

19. Çakır AH, Çelebioğlu A, Altunbaş S, et al. Allele frequencies for 15 STR loci in Van-Agri districts of the Eastern Anatolia region of Turkey. Forensic Sci Int. 2003;135:60-3.

20. Akbaşak BS, Budowle B, Reeder DJ, et al. Turkish population data with the CODIS multiplex short tandem repeat loci. Forensic Sci Int. 2001;123:227-9.

21. Gürkan C, Demirdov DK, Yamacı RF, et al. Population genetic data for 15 autosomal STR markers in Turkish Cypriots from Cyprus. Forensic Sci Int Genet. 2015;14:e1-3.

22. Özkorkmaz A, Baransel Isır A, Pehlivan S, et al. Forensic value of ten short tandem repeat loci in Turkey compared to other etnic groups. BJMG. 2009;12;69-72.

23. Asıcıoğlu F, Akyüz F, Çetinkaya U, et al. Allele distribution data of nine short tandem repeat loci for Turkish population: D3S1358, vWa, FGA, D8S1179, D21S11, D18S51, D13S317, D7S820. Forensic Sci Int. 2002;129:75-7.

24. Erkol Z, Tuğ A, Alakoç DY, et al. STR data for the AmpFISTR Identifiler loci from an old settlement in northwestern Turkey. Forensic Sci Int. 2007;173:238-40.

25. Çakir AH, Çelebioğlu A, Şimşek F. STR data for the AmpFISTR SGM Plus from Aegean region of Turkey. Forensic Sci Int. 2002;129:137-9.

26. Çakir AH, Şimşek F, Altunbaş S, et al. Distributions of THO1, TPOX, CSF1PO, vWA, D13S317, D7S820, and D16S539 alleles in a Turkish population sample. Forensic Sci Int. 2001;124:224-5.

27. Çakir AH, Şimşek F. Katırcı N, et al. STR data for the AmpFISTR SGM Plus from the eastern and western sections of Mediterranean region of Turkey. Forensic Sci Int. 2004;142:55-7.

28. Baransel Isir A, Özkorkmaz A, Pehlivan S. Allele frequencies for 13 STRs loci in a Western Anatolia population and their forensic evaluation. Ann Hum Biol. 2015;42:494-7.