ORİJİNAL ARAŞTIRMA ORIGINAL RESEARCH DOI: 10.5336/forensic.2022-89168

Determination of Gene Frequency of 10 New STR Loci in Turkey: Experimental Research

Yeni 10 STR Lokusunun Türkiye'deki Gen Sıklığının Belirlenmesi: Deneysel Çalışma

Tuğba SOHTORİK ÖZTÜRK[a], Özlem BÜLBÜL[a], Gönül FİLOĞLU TÜFEK[a]

[a]Department of Biology, İstanbul University-Cerrahpaşa Institute of Forensic Medicine and Forensic Sciences, İstanbul, Türkiye

Initial results of this study was presented as a poster in IAFS 2017 21st Triennial Meeting of the International Association of Forensic Sciences, August 21-25, 2017, Toronto, Canada.

ABSTRACT Objective: Short tandem repeats (STRs) have been widely used in human identification in forensics. Evaluation and design of new markers are useful tools to obtain additional information and complete conventional analysis. Also, it will be an alternative way to confirm the results in the problematic cases (complex kinship, degraded samples, etc.). Increasing the number of loci will result in reducing the risk of adventitious matches in countries that have a national DNA database. In this study, we investigated allele frequencies of the 10 non-Combined DNA Index System STR loci (D7S1517, D3S1744, D12S391, D2S1360, D6S474, D4S2366, D8S1132, D5S2500, D21S2055, D10S2325) by using Investigator® HDplex kit. **Material and Methods:** DNAs were extracted from 100 blood samples by using the QIAmp DNA Mini Kit (Qiagen). STR loci were amplified according to Investigator® HDplex kit. Polymerase chain reaction products were separated with ABI 3130 Genetic Analyzer and analyzed with GeneMapper IDX software. Forensic and population parameters were estimated with the Promega PowerStats Excel sheet. Allele frequencies, p-values for Hardy-Weinberg equilibrium, and population differentiation based on the loci (Fst) were calculated with Arlequin ver.3.5. **Results:** We observed that all STR loci showed high power of discrimination. For population comparison, we found statistically significant differences between Türkiye and African, East Asian, and American populations at a couple of STR loci. Besides, there was no statistically significant variation between Turkish and European populations as expected. Conclusion: This study provides Turkish population data for forensic laboratories and this kit can be used to support the existing STR loci.

Keywords: Non-CODIS; DNA analysis; new short tandem repeat loci

ÖZET Amaç: Otozomal kısa ardışık tekrar [short tandem repeat (STR)] lokusları, adli vakalarda kişi identifikasyonunda yaygın olarak kullanılmaktadır. Yeni belirteçlerin araştırılması ve dizayn edilmesi geleneksel analizleri tamamlayıcı olduğu gibi kardeşlik ve akrabalık ilişkilerinin söz konusu olduğu kompleks vakalarda ya da zorlukların yaşanabildiği bazı durumlarda (degrade örnekler vs.) sonuçların alternatif yollarla teyit edilmesini mümkün kılmaktadır. Araştırılan lokus sayısının artırılması veritabanı uygulaması olan ülkelerde yanlış eşleşme riskini de azaltmaktadır. Bu çalışmada, "Combined DNA Index System"de yer almayan, yaygın olarak kullanılan STR lokuslarından farklı, 10 yeni STR lokusunun (D7S1517, D3S1744, D12S391, D2S1360, D6S474, D4S2366, D8S1132, D5S2500, D21S2055, D10S2325) Türkiye'deki gen sıklığı Investigator® HDplex kiti kullanılarak incelendi. Gereç ve Yöntemler: DNA analizi için 100 g.nüllüden kan örneği alındı ve DNA izolasyonu QIAmp DNA Mini Kit (Qiagen) protokolüne uygun olarak gerçekleştirildi. STR lokusları Investigator® HDplex Kit protokolüne uygun olarak çoğaltıldı. Polimeraz zincir reaksiyonu ürünleri ABI 3130 Genetik Analizörde yürütüldü ve GeneMapper IDX ile analiz edildi. Popülasyon ve adli istatistik parametreleri Promega Power-Stats Excel kullanılarak hesaplandı. Alel frekansları, Hardy-Weinberg dengesi ve popülasyonlar arası lokus bazındaki farklılıklar (Fst) Arlequin v.3.5 programı kullanılarak hesaplandı. **Bulgular:** Tüm STR lokuslarının yüksek ayrım gücüne sahip olduğu gözlendi. Türkiye ile Afrika, Doğu Asya ve Amerika popülasyonları arasında bir kaç lokusta istatistiksel olarak anlamlı farklılıklar bulunurken; Avrupa popülasyonu ile istatistiksel olarak anlamlı bir farklılık gözlenmedi. **Sonuç:** Bu çalışma adli laboratuvarlar için Türk popülasyonu verileri sağlamaktadır ve çalışma bulgularına göre söz konusu kitin adli laboratuvarlarda mevcut STR lokuslarını desteklemek amacıyla güvenilirlikle kullanılabileceği gösterilmiştir.

Anahtar Kelimeler: Non-CODIS; DNA analizi; yeni kısa ardışık tekrar lokusları

Forensic analysis has been a very powerful technique for discrimination among individuals since the 1980s. It serves social incidents such as uncovering the crime and the criminal, investigating the crime tools, inheritance-paternity cases, and kinship relations. It also determines whether the suspect and DNA samples from the scene come from the same sources. For these purposes, it is possible to analyze a variety of the samples (such as blood, semen, saliva, hair, and stain) taken from the crime scene or individuals.[1,2]

Correspondence: Tuğba SOHTORİK ÖZTÜRK
Department of Biology, İstanbul University-Cerrahpaşa Institute of Forensic Medicine and Forensic Sciences, İstanbul, Türkiye
E-mail: tugbasohtorik@yahoo.com

Peer review under responsibility of Turkiye Klinikleri Journal of Forensic Medicine and Forensic Sciences.

Received: 25 Feb 2022 ***Received in revised form:*** 20 Apr 2022 ***Accepted:*** 20 Apr 2022 ***Available online:*** 28 Apr 2022

After Jeffreys discovered the polymorphism called "minisatellite" in the DNA molecule; more reliable, high discriminative, and sensitive DNA systems have been started to be used in forensic science.[3] Following the discovery of the polymorphic feature of DNA, Variable Number of Tandem Repeats (VNTRs), which showed much greater variability among people started to be used primarily in forensic studies. Over time, VNTRs have been replaced by short tandem repeats (STRs) and after the production of the first STR kit in 1994, 13 STR loci were defined in 1997.[4] Although new genetic markers [e.g. single-nucleotide polymorphisms (SNPs), insertion and deletion variants (InDels)] have been discovered with current technological progress, human identity testing is routinely made by using STR systems.[5]

Autosomal STR loci are applied successfully in many areas such as; identification, missing people, forensic cases, mass disasters, DNA database applications, and a wide variety of routine laboratory studies. STRs are accepted as ideal genetic markers as they can be easily amplified by multiplex polymerase chain reaction (PCR), high polymorphism features, short analysis time, presence of the standardized commercial kits, automated and error-minimized systems, allowing multiple analysis at the same time and DNA profiles can be obtained from a small amount and degraded samples.[1] STR loci are various in the genome and there are many high numbers of possible loci for DNA analysis. Evaluations and design of new STR markers are useful tools to obtain additional information and to complete conventional analysis.[6] Besides this, increasing the number of investigated STR markers also enhances the reliability of the study especially in the cases of fraternity and kinship relations.[7] Although there are lots of commercially available kits commonly used in forensic fields, increasing the number of data loci will result in reducing the risk of adventitious matches in the countries that have national DNA databases.[8] For these reasons, researchers continue their scientific studies in the new STR loci, and database applications are being studied. Investigator® HDplex Kit (Qiagen, Germany) was released by Qiagen in 2010 to aim to increase the power of discrimination (PD) for paternity tests and also specific forensic cases. The kit includes 10 new non-Combined DNA Index System (CODIS) STR loci (D7S1517, D3S1744, D12S391, D2S1360, D6S474, D4S2366, D8S1132, D5S2500, D21S2055, D10S2325), sex marker amelogenin and other common STR loci (D18S51 and SE33). Population studies on this set of markers have already been performed in different populations. However, no data has been reported for the Turkish population so far. This experimental study aims to investigate ten new non-CODIS STR loci and determine forensic data for the population of Türkiye and population comparison with data available for other populations.

MATERIAL AND METHODS

The DNA samples were collected from 100 unrelated healthy individuals from İstanbul, Türkiye. Written informed consent was obtained from all volunteers for use of the sample in research. Statement of compliance with standards of research involving humans as subjects. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. The study was also approved by the ethics committee dated 24.11.211 and numbered 40222 at the İstanbul University, Cerrahpaşa faculty of Medicine.

Genomic DNAs were isolated from whole blood by using Mini QIAamp® (Qiagen, Germany) and quantitative analysis of the samples was performed fluorometrically using the Quant-iT™ dsDNA HS Assay kit (Invitrogen, Thermo Fisher Scientific, USA) according to the manufacturer's protocol. DNA extracts were diluted to approximately 0.5 ng/µL for the required optimal amount of DNA. The amplifications of the DNA samples were performed in a 25 µL volume of the multiplex PCR reaction using the Investigator HD-plex kit (Qiagen) according to the manufacturer's instructions. Ten STR loci (D7S1517, D3S1744, D12S391, D2S1360, D6S474, D4S2366, D8S1132, D5S2500, D21S2055, D10S2325) and also amelogenin were amplified simultaneously in one PCR reaction. PCR products were detected by capillary electrophoresis and fluorescent detection in ABI PRISM® 3130 Genetic Analyzer (Thermo Fisher Sci-

entific, USA). Positive, negative controls, and allelic ladder were run to ensure reliable allelic assignment. Allele allocation was carried out with GeneMapper ID-X v.1 software (Thermo Fisher Scientific, USA) by reference to DNA Size Standard 550 (BTO) (Qiagen). Forensic statistical parameters including polymorphism information content (PIC), matching probability (PM), PD, power of exclusion (PE), and typical paternity index was performed by using PowerStats Excelsheet v.1.2 (Promega, USA). Allele frequencies, observed and expected heterozygosities (H_o and H_e), p-values for Hardy-Weinberg equilibrium (HWE), population differentiation based on the loci, and F_{st} were calculated with Arlequin ver. 3.5.

RESULTS

In this study, we analyzed ten new non-CODIS STR loci by using the Investigator® HDplex Kit (Qiagen) from 100 unrelated healthy individuals from İstanbul, Türkiye. Allele frequencies and HWE were calculated in the Arlequin v.3.5.1.2 (http://cmpg.unibe.ch/software/arlequin35/, Switzerland) by using the obtained DNA data. As indicated in the Table 1, the highest and the lowest allele frequencies were obtained in the loci D3S1744 (allele 17) and D12S391 (allele 15), respectively. The most common alleles were 25 for D7S1517 (0.210), 17 for D3S1744 (0.320), 18 for D12S391 (0.195), 22 for D2S1360 (0.295), 16 for D6S474 (0.280), 9 for D4S2366 (0.300), 18 for D8S1132 (0.215), 11 for D5S2500 (0.270), 19.1 for D21S2055 (0.25), 12 for D10S2325 (0.190).

Bonferroni correction (p=0.05/10) was applied to the obtained p-values after HWE test and the level of significance was taken as p>0.005. Statistical deviation from HWE was observed for 6 loci (D7S1517, D12S391, D4S2366, D8S1132, D10S2325, D21S2055, D12S391) after Bonferroni correction (Table 1).

Forensic statistical parameters were determined for the data from allele frequencies by using PowerStats v.1.2 (Promega, USA) (Table 1). The entire set of markers showed locus-specific PD values greater than 0.918 (D6S474) with the highest value for D21S2055 (0.968) while PE ranged from 0.328 (D6S474) to 0.581 (D2S1360). On the other hand, the most distinguishing loci in terms of the PE were D2S1360 (PE: 0.581) and D5S2500 (PE: 0.545). D6S474 (PE: 0.328), D8S1132 (PE: 0.369), and D4S2366 (PE: 0.383) were observed as less distinguishing loci in terms of the PE (Table 1).

Locus-based differences between populations were determined by using allele frequency data from European, African, East Asian and American population and also the data obtained as a result of this study (Table 2 and Table 3). Table 3 represents the locus-based differences (F_{st}) between populations. F_{st} p-values below 0.005 were evaluated as significant after Bonferroni correction. F_{st} values indicate that no variation was observed between Türkiye and the European population (Table 3). On the other hand, statistically significant differences were found between Turkish and African populations at 3 loci (D4S2366, D10S2325, and D21S391), East Asian population at 4 loci (D2S1360, D4S2366, D6S474, and D21S2055), and American population at 4 loci (D2S1360, D4S2366, D7S1517, and D21S2055) (Table 3).

DISCUSSION

Many different issues such as detecting the crime and the criminal, investigating crime tools and identifying the criminal, determining whether the DNA samples obtained from the suspect and the crime scene are of the same origin, inheritance-paternity cases, detection of missing persons and kinship relations are clarified by forensic sciences and developing DNA technologies.[1]

Although new genetic markers have been identified with current technological advances, STR systems that are cheaper, easier to apply, standardized and proven reliable are preferred for routine identification studies.[4] In recent years, researches have been carried out to identify new STR loci that are not included in CODIS all over the world.[9] In 2017, the FBI announced that it had increased the number of 13 loci in CODIS to 20. The newly added loci are mostly composed of mini-STR loci that allow easy typing of degraded or trace DNA samples from the crime scene.[10] STR loci on both the Y chromosome and the X chromosome are used to elucidate some complex cases.

TABLE 1: Allele frequencies and statistical parameters of 10 STR loci in the Turkish population.

Allele	D7S1517	D3S1744	D12S391	D2S1360	D6S474	D4S2366	D8S1132	D5S2500	D21S2055	D10S2325
7										0.130
8						0.000				0.055
9						0.300				0.080
10						0.145		0.075		0.135
11						0.085		0.270		0.160
12					0.000	0.175		0.165		0.190
13		0.005			0.255	0.165		0.060		0.135
14		0.009			0.230	0.120		0.050		0.070
15		0.075	0.003		0.130	0.010		0.245		0.040
16	0.005	0.145	0.015		0.280		0.015	0.105		0.005
16.1									0.070	
17	0.005	0.320	0.115		0.090		0.080	0.025		
17.1									0.020	
18	0.045	0.185	0.195		0.015		0.215	0.005		
18.1									0.010	
19	0.130	0.115	0.140	0.010			0.155			
19.1									0.250	
20	0.105	0.045	0.130	0.110			0.13			
20.1									0.055	
21	0.105	0.020	0.125	0.075			0.125			
21.1									0.015	
22	0.105		0.105	0.295			0.13			
22.1									0.005	
23	0.120		0.095	0.155			0.095		0.005	
24	0.130		0.040	0.095			0.045			
25	0.210		0.010	0.085			0.005		0.110	
26	0.025			0.090			0.005		0.115	
27	0.010			0.040					0.020	
28	0.005			0.035					0.010	
29				0.010					0.035	
30									0.030	
31									0.030	
32									0.020	
33									0.065	
34									0.080	
35									0.050	
36									0.005	
PM	0.041	0.061	0.036	0.048	0.082	0.075	0.041	0.064	0.032	0.045
PD	0.959	0.939	0.964	0.952	0.918	0.925	0.959	0.936	0.968	0.955
PIC	0.860	0.790	0.860	0.830	0.740	0.780	0.850	0.790	0.880	0.850
PE	0.428	0.476	0.460	0.581	0.328	0.3833	0.69	0.545	0.428	0.493
TPI	1.670	1.850	1.790	2.380	1.350	1.520	1.470	2.170	1.670	1.920
No.of alleles	13	9	11	11	6	7	11	9	20	10
H_o	0.700	0.730	0.720	0.790	0.630	0.670	0.660	0.770	0.700	0.740
H_e	0.87618	0.81704	0.87814	0.84809	0.78236	0.81347	0.86693	0.82055	0.89085	0.87337
p value	**0.00000**	0.16202	**0.00255**	0.03432	0.07197	**0.00032**	**0.00016**	0.00617	**0.00000**	**0.0001**

STR: Short tandem repeat; PM: Matching probability; PD: Power of discrimination; PIC: Polymorphism information content; PE: Power of exclusion; TPI: Typical paternity index; Ho: Observed heterozygosity; He: Expected heterozygosity.

TABLE 2: Allele frequencies of 10 new STR loci in different populations.

Locus	Alleles	Türkiye (n=100)	Europe* (n=158)	Africa* (n=112)	East Asia* (n=232)	USA* (n=65)	Locus	Alleles	Türkiye (n=100)	Europe* (n=158)	Africa* (n=112)	East Asia* (n=232)	USA* (n=65)
	17	0	0	0.009	0	0		9	0	0	0.024	0.008	0
	19	0.01	0.006	0.004	0.013	0		10	0.075	0.072	0.155	0.024	0.117
	20	0.11	0.101	0.019	0.017	0		11	0.27	0.269	0.247	0.298	0.312
	21	0.075	0.072	0.063	0.161	0.273		12	0.165	0.167	0.111	0.174	0.156
	22	0.295	0.294	0.194	0.648	0.46	D5S2500	13	0.06	0.069	0.077	0.035	0.031
	23	0.155	0.167	0.155	0.097	0.14		14	0.05	0.053	0.082	0.077	0.023
D2S1360	24	0.095	0.082	0.165	0.031	0.07		15	0.245	0.253	0.179	0.287	0.343
	25	0.085	0.069	0.194	0.015	0.023		16	0.105	0.088	0.077	0.064	0.015
	26	0.09	0.075	0.087	0.006	0.031		17	0.025	0.025	0.024	0.024	0
	27	0.04	0.053	0.058	0.004	0		18	0.005	0	0.019	0.004	0
	28	0.035	0.044	0.019	0.004	0		15.1	0	0	0	0.002	0
	29	0.01	0.009	0.024	0	0		16.1	0.07	0.085	0.019	0.13	0.023
	30	0	0.022	0	0	0		17.1	0.02	0.012	0	0.006	0
	31	0	0	0.004	0	0		18.1	0.01	0.006	0	0	0
	15	0	0	0.009	0	0		19.1	0.25	0.246	0	0.024	0
	16	0.005	0.006	0	0	0		20.1	0.055	0.05	0	0	0.007
	17	0.005	0.006	0.009	0.008	0		21	0	0	0	0	0.007
	18	0.045	0.025	0.019	0.019	0		21.1	0.015	0.003	0	0	0
	19	0.13	0.139	0.092	0.064	0.023		22	0	0	0.014	0	0.015
	20	0.105	0.094	0.131	0.117	0.07		22.1	0.005	0.003	0	0	0
	21	0.105	0.107	0.131	0.123	0.273		23	0.005	0.003	0.004	0.002	0.007
	22	0.105	0.113	0.087	0.135	0.304		24	0	0.009	0.004	0.017	0.039
D7S1517	23	0.12	0.139	0.14	0.159	0.242	D21S2055	25	0.11	0.117	0.038	0.17	0.359
	24	0.13	0.107	0.106	0.112	0.031		26	0.115	0.12	0.082	0.194	0.359
	25	0.21	0.208	0.155	0.163	0.031		26.1	0	0.003	0	0	0
	26	0.025	0.025	0.034	0.057	0.023		27	0.02	0.022	0.014	0.022	0.07
	27	0.01	0.012	0.014	0.026	0		28	0.01	0.009	0.009	0.044	0.023
	28	0.005	0.012	0.019	0.008	0		29	0.035	0.038	0.092	0.028	0
	29	0	0	0.014	0.002	0		30	0.03	0.025	0.082	0.022	0.023
	31	0	0	0.029	0	0		31	0.03	0.034	0.111	0.024	0.007

continue →

TABLE 2: Allele frequencies of 10 new STR loci in different populations *(continued)*.

Locus	Alleles	Türkiye (n=100)	Europe* (n=158)	Africa* (n=112)	East Asia* (n=232)	USA* (n=65)
D3S1744	11	0	0	0	0.002	0
	13	0.005	0	0.004	0.013	0
	14	0.09	0.088	0.058	0.092	0.062
	15	0.075	0.072	0.131	0.081	0.195
	16	0.145	0.132	0.116	0.117	0.195
	17	0.32	0.341	0.344	0.345	0.296
	18	0.185	0.18	0.228	0.192	0.195
	19	0.115	0.132	0.082	0.099	0.046
	20	0.045	0.038	0.034	0.046	0.007
	21	0.02	0.012	0	0.008	0
D6S474	12	0	0	0.014	0.004	0.031
	13	0.255	0.262	0.262	0.36	0.359
	14	0.23	0.221	0.184	0.342	0.14
	15	0.13	0.136	0.16	0.121	0.179
	16	0.28	0.281	0.286	0.112	0.14
	17	0.09	0.085	0.077	0.057	0.14
	18	0.015	0.012	0.014	0	0.007
D8S1132	15	0	0.003	0.024	0	0
	16	0.015	0.028	0.101	0.008	0.062
	17	0.08	0.094	0.189	0.068	0.093
	18	0.215	0.212	0.15	0.203	0.14
	19	0.155	0.164	0.126	0.216	0.203
	20	0.13	0.148	0.126	0.157	0.179
	21	0.125	0.117	0.14	0.123	0.109
	22	0.13	0.098	0.038	0.13	0.14
	23	0.095	0.098	0.068	0.059	0.054
	24	0.045	0.025	0.029	0.028	0.007
	25	0.005	0.006	0.004	0.002	0.007
	26	0.005	0.003	0	0	0
D4S2366	8	0	0.003	0	0	0
	9	0.3	0.284	0.092	0.236	0.171
	10	0.145	0.142	0.335	0.057	0.007
	11	0.085	0.075	0.218	0.354	0.179
	11.2	0	0.003	0.16	0.002	0
	12	0.175	0.189	0.135	0.161	0.304
	12.2	0	0	0.019	0	0
	13	0.165	0.177	0.029	0.092	0.078
	14	0.12	0.113	0.004	0.09	0.242
	15	0.01	0.009	0.004	0.004	0.015
D10S2325	32	0.02	0.012	0.194	0.044	0.015
	33	0.065	0.066	0.135	0.077	0.007
	34	0.08	0.082	0.111	0.095	0
	35	0.05	0.034	0.043	0.064	0.015
	36	0.005	0.009	0.014	0.017	0
	37	0	0.003	0.009	0.011	0.015
	38	0	0	0.009	0	0
	39	0	0	0.004	0	0
	6	0	0.003	0.082	0.013	0.023
	7	0.13	0.161	0.116	0.194	0.07
	8	0.055	0.038	0.053	0.017	0.007
	9	0.08	0.066	0.194	0.101	0.101
	10	0.135	0.129	0.165	0.099	0.242
	11	0.16	0.183	0.208	0.15	0.125
	12	0.19	0.161	0.145	0.161	0.21
	13	0.135	0.145	0.019	0.139	0.109
	14	0.07	0.053	0.009	0.068	0.062
	15	0.04	0.047	0	0.028	0.031
	16	0.005	0.006	0.004	0.011	0.007
	17	0	0	0	0.002	0.007
	17.2	0	0	0	0.002	0
	18	0	0.003	0	0	0
	19	0	0	0	0.002	0
	21	0	0	0	0.002	0
D12S391	15	0.03	0.035	0.021	0.016	0.032
	16	0.015	0.019	0.012	0.006	0.029
	17	0.115	0.107	0.249	0.106	0.059
	17.3	0	0.019	0	0.001	0.012
	18	0.195	0.215	0.196	0.278	0.182
	18.3	0	0.007	0	0.001	0.007
	19	0.14	0.121	0.145	0.221	0.178
	19.3	0	0.016	0	0	0.01
	20	0.13	0.117	0.071	0.155	0.197
	21	0.125	0.093	0.055	0.109	0.104
	22	0.105	0.114	0.113	0.051	0.073
	23	0.095	0.072	0.078	0.028	0.061
	24	0.04	0.04	0.009	0.015	0.026
	25	0.01	0.021	0.014	0.007	0.016
	26	0	0.002	0.002	0.001	0.003

*The data for other countries except Türkiye taken from Phillips C, Fernandez-Formoso L, Gelabert-Besada M, García-Magariños M, Amigo J, Carracedo A, et al. Global population variability in Qiagen Investigator HDplex STRs. Forensic Sci Int Genet. 2014;8(1):36-43; STR: Short tandem repeat.

TABLE 3: Locus-based differences between populations.

	Türkiye-Europe		Türkiye-Africa		Türkiye-East Asia		Türkiye-USA	
Locus	F_{st}	p value	F_{st}	p value	F_{st}	p value	F_{st}	p value
D2S1360	-0.00732	0.99099+-0.0030	0.01070	0.05405+-0.0201	0.11082	**0.00000+-0.0000**	0.04373	**0.00901+-0.0091**
D3S1744	-0.00717	0.99099+-0.0030	-0.00313	0.73874+-0.0446	-0.00585	0.98198+-0.0096	0.00459	0.22523+-0.0546
D4S2366	-0.00786	0.99099+-0.0030	0.08542	**0.00000+-0.0000**	0.04687	**0.00000+-0.0000**	0.03910	**0.00000+-0.0000**
D5S2500	-0.00754	0.99099+-0.0030	0.00022	0.40541+-0.0563	-0.00114	0.48649+-0.0667	0.00364	0.18018+-0.0332
D6S474	-0.00811	0.99099+-0.0030	-0.00745	0.97297+-0.0125	0.02776	**0.00000+-0.0000**	0.01501	0.08108+-0.0212
D7S1517	-0.00695	0.99099+-0.0030	-0.00423	0.78378+-0.0490	-0.00098	0.51351+-0.0731	0.06650	**0.00000+-0.0000**
D8S1132	-0.00654	0.98198+-0.0096	0.00963	0.02703+-0.0194	-0.00279	0.77477+-0.0474	-0.00252	0.62162+-0.0345
D10S2325	-0.00596	0.94595+-0.0205	0.01686	**0.00901+-0.0091**	-0.00169	0.63063+-0.0407	0.00043	0.40541+-0.0493
D12S391	-0.00604	0.96396+-0.0142	0.00774	0.09009+-0.0192	0.00782	0.06306+-0.0194	-0.00465	0.64865+-0.0618
D21S2055	-0.00703	0.99099+-0.0030	0.05557	**0.00000+-0.0000**	0.03018	**0.00000+-0.0000**	0.11010	**0.00000+-0.0000**

Besides the current CODIS loci, the determination of new STR markers will provide supplementary extra information for conventional STR analyses and the reliability of the study will also increase especially in difficult cases where kinship relationship is concerned.[11] In some special cases such as paternity cases involving 2 siblings, additional autosomal STR loci need to be studied besides the conventional STR loci to increase the possibilities of inclusion/exclusion. The Investigator Hdplex kit developed for this purpose includes at least 9 non-CODIS STR loci unlike other commercial STR kits.[12-20] This means that these additional markers can be worked to increase the probability in cases of close kinship or father-child mismatch.

Before using a polymorphic marker in case resolution routinely, gene frequencies in the relevant population should be determined and an appropriate database should be used. For this purpose, 10 non-CODIS STR loci were typed for 100 people from İstanbul by using the Investigator Hdplex kit in this study. Obtained genotype data, expected and observed heterozygosity values and p-values were calculated by using Arlequin v.3.5.1.2. Level of significance was taken as p>0.005 according to Bonferroni correction (Table 1).[21] Deviation from HWE was observed for 6 loci. The reason for this can be explained by the low number of individuals included, the high number of alleles of loci, high migration or mutation rates, or the inability of the sampling to be representative of the entire population.[22]

Compare with the global population (Europe, Africa, Asia and America) data, the gene frequencies of Türkiye were generally similar to the European population, but showed small and moderate differences with other populations. Similar results were obtained in the Turkish population in which STR, SNP and InDel markers were studied by various researchers before.[23-25]

The Investigator HDplex Kit (Qiagen, Germany) has been studied in many populations to expand STR data.[8,26-29] However, no study has been conducted with the kit in question for the Turkish population so far.

In order to understand which locus is more useful in case resolution, the forensic statistical parameters of each locus should be known. High PIC, het and PD values indicate the superiority of investigated locus in terms of forensic genetics. Allele frequencies and statistical parameters of the genotypes obtained for the investigated 10 loci were calculated using PowerStats v.1.2 (Promega, USA) in this study. PowerStats v.1.2. has been preferred because of its advantages such as ease of data entry, summarization of the data and the ability to obtain the results graphically.[30] When all PIC, PD and heterozygosity values were compared with the global population data; we observed that PIC and PD values were similar for all loci but heterozygosity was slightly lower for all loci.[8,31] Observed heterozygosity of the 10 STR loci varies from 0.630 to 0.79. When forensic statistical parameters were evaluated for each locus, there was

equilibrium between them and because the PD values for all investigated loci were between 0.918 and 0.968, those loci have required exclusion power and can be safely used to establish a DNA-based database for the Turkish population.

This study shows that there is a high level of polymorphism of these ten non-CODIS STR loci, allele frequency distributions, and statistical parameters are similar with the European populations as expected.[7,8,32] Therefore, the investigated kit can be reliably used to support the existing STR loci in forensic laboratories.

CONCLUSION

The findings of this study indicate that Investigator HDplex shows a high level of discrimination power for forensic genetics applications such as personal identification, paternity, and kinship testing. The Investigator HDplex kit has additional loci different from the other commercial kits, it represents excellent support for databases i difficult cases. Applying new these non-CODIS STRs will provide better discrimination power addition to the core CODIS STR markers. Although current STRs provide enough power for databases, new additional STR markers are very informative for complex kinship analysis. Hence, this study is the first database compiled for this non-CODIS STR loci in the Turkish population and the Investigator HD-plex kit can be used in Turkish forensic laboratories.

Source of Finance

This research was supported by Istanbul University-Cerrahpaşa Scientific Research Projects Unit (Project no:18427)

Conflict of Interest

No conflicts of interest between the authors and / or family members of the scientific and medical committee members or members of the potential conflicts of interest, counseling, expertise, working conditions, share holding and similar situations in any firm.

Authorship Contributions

Idea/Concept: *Gönül Filoğlu Tüfek;* ***Design:*** *Gönül Filoğlu Tüfek;* ***Control/Supervision:*** *Gönül Filoğlu Tüfek;* ***Data Collection and/or Processing:*** *Tuğba Sohtorik Öztürk;* ***Analysis and/or Interpretation:*** *Tuğba Sohtorik Öztürk;* ***Literature Review:*** *Tuğba Sohtorik Öztürk;* ***Writing the Article:*** *Tuğba Sohtorik Öztürk;* ***Critical Review:*** *Gönül Filoğlu Tüfek, Özlem Bülbül;* ***References and Fundings:*** *İstanbul Üniversitesi Cerrahpaşa Adli Tıp Enstitüsü BAP projesi;* ***Materials:*** *Gönül Filoğlu Tüfek, Özlem Bülbül, Tuğba Sohtorik Öztürk.*

REFERENCES

1. Butler JM. Forensic DNA Typing: Biology, Technology and Genetics of STR Markers. 2nd ed. USA: Elsevier Academic Press; 2005.
2. Robertson J, Ross AM, Burgayne LA. DNA in Forensic Science: Theory, Techniques and Applications. 3rd ed. Longooddon: Ellis Howard Ltd; 1990. [Crossref]
3. Jeffreys AJ, Wilson V, Thein SL. Hypervariable 'minisatellite' regions in human DNA. Nature. 1985;314(6006):67-73. [Crossref] [PubMed]
4. Reno J, Marcus D, Leary ML, Samuels JE. The future of Forensic DNA Testing: Predictions of the Research and Development Working Group. National Institutes of Justice, National Commision on the Future of DNA Evidence; 2000. Report no.: NCJ 18397. [Link]
5. Filoglu G, Altuncul H, Bulbul O. Adli Genetik ve Genetik Kimliklendirme. 1. Baskı. Ankara: Seçkin Yayınevi; 2021.
6. Butler JM, Hill CR, Decker AE, Kline MC, Vallone PM. New autosomal STR loci. Forensic Science International: Genetics Supplement Series. 2008;1:95-6. [Crossref]
7. Qiagen [Internet]. © QIAGEN 2013-22. All rights reserved [Cited: April 5, 2022]. Population data for Investigator HDplex - (EN). Available from: [Link]
8. Phillips C, Fernandez-Formoso L, Gelabert-Besada M, García-Magariños M, Amigo J, Carracedo A, et al. Global population variability in Qiagen Investigator HDplex STRs. Forensic Sci Int Genet. 2014;8(1):36-43. [Crossref] [PubMed]
9. Gettings KB, Borsuk LA, Steffen CR, Kiesler KM, Vallone PM. Sequence-based U.S. population data for 27 autosomal STR loci. Forensic Sci Int Genet. 2018;37:106-15. [Crossref] [PubMed] [PMC]
10. Hares DR. Selection and implementation of expanded CODIS core loci in the United States. Forensic Sci Int Genet. 2015;17:33-4. [Crossref] [PubMed]
11. Gettings KB, Kiesler KM, Faith SA, Montano E, Baker CH, Young BA, et al. Sequence variation of 22 autosomal STR loci detected by next generation sequencing. Forensic Sci Int Genet. 2016;21:15-21. [Crossref] [PubMed] [PMC]
12. AmpFISTR® Identifiler® Plus PCR Amplification Kit User Guide, Applied BiosystemsTM. 2015. Cited: April 5, 2022. Available from: [Link]
13. ForenSeq™ DNA Signature Prep Reference Guide, Verogen. 2018. Cited: April 5, 2022. Available from: [Link]
14. GlobalFilerTM PCR Amplification Kit User Guide, Applied BiosystemsTM, 4476135(4477604). 2016. p.1-166. Cited: April 5, 2022. Available from: [Link]
15. Qiagen [Internet]. © QIAGEN 2013-22. All rights reserved [Cited: April 5, 2022]. Investigator® ESSplex SE QS Kit Handbook. Available from: [Link]

16. Qiagen [Internet]. © QIAGEN 2013-22. All rights reserved [Cited: April 5, 2022]. Investigator® 24plex QS Handbook (EN). Available from: [Link]
17. PowerPlex® 21 System for Use on the Applied Biosystems® Genetic Analyzers. Promega. 2021. Cited: April 5, 2022. Available from: [Link]
18. PowerPlex® 16 HS System for Use on the Applied Biosystems® Genetic Analyzers. Promega Corporation. 2016. Cited: April 5, 2022. Available from: [Link]
19. PowerPlex® Fusion 6C System for Use on the Applied Biosystems® Genetic Analyzers. Promega Corporation. 2016. Cited: April 5, 2022. Available from: [Link]
20. Precision ID GlobalFiler NGS STR Panels. Thermo Fisher Scientific. 2018. Cited: April 5, 2022. Available from: [Link]
21. Hochberg Y. A sharper Bonferroni procedure for multiple tests of significance. Biometrika. 1988;75(4):800-902. [Crossref]
22. Zintzaras E. Impact of Hardy-Weinberg equilibrium deviation on allele-based risk effect of genetic association studies and meta-analysis. Eur J Epidemiol. 2010;25(8):553-60. [Crossref] [PubMed]
23. Bulbul O, Fernandez-Formoso L, Phillips C, Altuncul H, Filoglu G, Lareu MV, et al. Allele frequencies of the five new European Standard Set (ESS) STRs and 15 established STRs in a Turkish population. Forensic Sci Int Genet. 2014;9:e26. [Crossref] [PubMed]
24. Tavacı İ, Bülbül O, Filoglu G, Altuncul H. X kromozomunda bulunan 15 SNP lokusunun Türkiye populasyonundaki polimorfizmi [Polymorphism of 15 X-chromosome SNPs in the Turkish population]. J Foren Sci Leg Med. 2020;17(3):238-44. [Crossref]
25. Duvenci A, Bulbul O, Filoglu G. Evaluation of population data and forensic parameters of Turkish population on 30 autosomal insertion and deletion polymorphisms. Russian Journal of Genetics. 2019;55(2):246-52. [Crossref]
26. Tillmar A, Nilsson H, Kling D, Montelius K. Analysis of investigator HDplex markers in Swedish and Somali populations. Forensic Sci. Int. Genet. 2013;7(1):e21-e2. [Crossref] [PubMed]
27. Zieger M, Gouy A, Utz S. Expanding the Swiss autosomal marker set to 32 STRs. Int J Legal Med. 2021;135(6):2309-10. [Crossref] [PubMed] [PMC]
28. Jung JY, Kim EH, Oh YL, Park HC, Hwang JH, Lim SK. Evaluation of forensic genetic parameters of 12 STR loci in the Korean population using the Investigator® HDplex kit. Int J Legal Med. 2017;131(5):1247-9. [Crossref] [PubMed]
29. Turrina S, Ferrian M, Caratti S, De Leo D. Investigator HDplex markers: allele frequencies and mutational events in a North Italian population. Int J Legal Med. 2015;129(4):731-3. [Crossref] [PubMed]
30. Tereba A. Tools for analysis of population statistics. Promega Corporation. Profiles DNA. 1999;2:14-26. [Link]
31. Liu Q, Nan H, He X, Wu W, Lu D. Population genetic data of Investigator HDplex markers in Han population from Southern China. Int J Legal Med. 2019;133(1):77-9. [Crossref] [PubMed]
32. Schmid D, Anslinger K, Rolf B. Allele frequencies of the ACTBP2 (=SE33), D18S51, D8S1132, D12S391, D2S1360, D3S1744, D5S2500, D7S1517, D10S2325 and D21S2055 loci in a German population sample. Forensic Sci Int. 2005;151(2-3):303-5. [Crossref] [PubMed]