ORİJİNAL ARAŞTIRMA ORIGINAL RESEARCH

DOI: 10.5336/forensic.2022-89115

Assessing Population Structure Using 46 AIM-Indel (Ancestry Informative Marker-Insertion/Deletion) Panel in Türkiye: Experimental Research (Population Study)

46 AIM-InDel (Soy Bilgilendirici Belirteçler-İnsersiyon/Delesyon) Paneli Kullanılarak Türkiye Popülasyon Yapısının Değerlendirilmesi: Deneysel Araştırma (Popülasyon Çalışması)

Cemaleddin ARSLAN[a], Özlem BÜLBÜL[a], Gönül FİLOĞLU[a]

[a]Department of Science, Institute of Forensic Science and Legal Medicine, İstanbul University-Cerrahpaşa, Istanbul, Türkiye

ABSTRACT Objective: Insertion/deletion (InDels), also known as "new generation genetic markers" can be an alternative to short tandem repeats (STRs) and single nucleotide polymorphisms (SNPs). Identification and inference of biogeographical ancestry analysis can be more accurate when InDels are used with STRs and SNPs. The aim of this study is to determine the allele frequencies of 46 AIM-InDel loci in Turkish population which can be used to infer biogeographical ancestry. We also compare our population results with other populations. **Material and Methods:** 46 AIM-InDel loci were typed in 148 Turkish volunteers. Allele frequencies, Hardy Weinberg equilibrium and pairwise F_{ST} values were calculated using Arlequin ver.3.5. Biogeographic ancestry was analyzed using Structure 2.3.4. Snipper program was used for individual based inference of biogeographical ancestry. **Results:** The lowest genetic distance was observed between Turkish population and European, Middle Eastern and Central South Asian populations. When the genetic distance are longer, we observed clearer differences between Turkish and African and East Asian populations. **Conclusion:** The study provides new InDel data on understanding population structure of Türkiye. This could be useful in forensic investigations and as well as molecular anthropology.

Keywords: Ancestry informative markers (AIMs); InDel; Turkish population; biogeographical ancestry (BGA); forensic genetics

ÖZET Amaç: "Yeni nesil genetik belirteçler" olarak da bilinen insersiyon/delesyonlar (InDel), kısa tekrar dizilerilerine [short tandem repeats (STRs)] ve tek nükleotid polimorfizmlerine [single nucleotide polymorphisms (SNP)] alternatif bir belirteç olabilir. InDel belirteçleri, STR'ler ve SNP'ler ile birlikte kullanıldığında, daha doğru bir biyocoğrafik soy analizi yapılabilir. Bu çalışmanın amacı, Türkiye popülasyonundaki 46 AIM-InDel lokusunun biyocoğrafik soy tahmini yapmak için kullanılabilecek alel frekanslarını belirlemektir. Bunun yanı sıra Türkiye popülasyonu sonuçlarını diğer popülasyonlarla da karşılaştırılmaktır. **Gereç ve Yöntemler:** 148 Türkiyeli gönüllüde 46 AIM-InDel lokusu çalışıldı. Alel frekansları, Hardy Weinberg dengesi ve ikili F_{ST} değerleri Arlequin ver.3.5 kullanılarak hesaplandı. Biyocoğrafik soy, Structure 2.3.4 kullanılarak analiz edildi. Snipper Programı ile de bireysel bazda biyocoğrasif soy tahmini yapıldı. **Bulgular:** En düşük genetik uzaklık Türkiye popülasyonu ile Avrupa, Orta Doğu ve Orta Güney Asya popülasyonları arasında gözlendi. Genetik mesafe arttıkça Türkiye ile Afrika ve Doğu Asya popülasyonları arasında daha net genetik farklılıklar gözlemlendi. **Sonuç:** Çalışma, Türkiye'nin popülasyon yapısını anlama konusunda yeni InDel verileri sunmaktadır. Bu, adli araştırmalarda ve moleküler antropolojide faydalı olabilir.

Anahtar Kelimeler: Soy bilgilendirici markırlar (AIMs); InDel; Türkiye popülasyonu; biyocoğrafik soy (BGA); adli genetik

When a complete gene or a specific nucleotide sequence is added (insertion) or removed (deletion) from DNA helices, these type of genetic variations are named as insertion or deletion, commonly abbreviated "InDel".[1-3] InDel polymorphisms, which are biallelic length polymorphims, occur as a result of sequential mutations and are inherited to the offsprings according to Mendel's laws.[4] Although a

Correspondence: Özlem BÜLBÜL
Department of Science, Institute of Forensic Science and Legal Medicine, İstanbul University-Cerrahpaşa, İstanbul, Turkiye
E-mail: ozlem.bulbulercan@iuc.edu.tr

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

Received: 21Feb 2022 ***Received in revised form:*** 29 Mar 2022 ***Accepted:*** 29 Mar 2022 ***Available online:*** 11 Apr 2022

2619-9459 / Copyright © 2022 by Türkiye Klinikleri. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

short InDel locus is 1-5 bp long, rarely it may be up to 25 bp. Sometimes, only 3 or 4 allele can be observed in an InDel locus.[5]

The InDel is an alternative genetic marker for the widely used forensic short tandem repeats (STRs) identification method.[6] In many forensic cases, degraded or low amount of biological samples are collected from the crime scene and these samples generally do not meet the criteria for the rutine laboratory analysis. Polymorphic markers that have short amplicons are preferred in such cases. InDels can be designed shorter than 150 bp amplicon size where the classic STR amplicon sizes are 150-450 bp in length. Therefore, InDel polymorphisms provide an alternative to STR polymorphisms in DNA profiling.[5,7]

Biogeographic ancestry of an unknown individual can be estimated using population specific ancestry informative markers (AIMs).[8] This information may assist forensic investigations of unknown contributors or identification of missing persons and disaster victims.[9,10] There are recent studies on multiplexes including different numbers of InDel loci developed for forensic identification and inference of biogeographical ancestry.[1,6,7,11-15] InDel loci that are used in inference of biogeographical ancestry are named as ancestry informative markers-InDel (AIM-InDel). Some AIM-InDel multiplexes are used alone for estimating ancestry, whereas some of them are combined with ancestry informative SNPs (AISNPs) for finer ancestry assignment.[1,11,15,16] The AIM-InDel panel which includes 46 InDel was designed to cluster four continental (Africa, Europe, East Asi and America) populations.[1] In addition, Oceanian population sample in the collection of "Human Genome Diversity Project-Centre d'Étude du Polymorphisme Humain" (HGDP-CEPH) diversity panel were tested and the results indicate that Oceanian population can also be discriminated with high accurucy using 46 AIM-InDel panel.[1] In a later study, Middle Eastern and Central South Asian populations in HGDP-CEPH diversity panel were analyzed with 46 AIM-InDel and 34-plex SNP panels. It was aimed to increase the reliability and accuracy of the results by increasing the number of analyzed loci.[16]

In this study, we aim to investigate patterns of genetic variation of 46 AIM-InDel loci for Turkish population and compare it with world populations.

MATERIAL AND METHODS

EXPERIMENTAL STUDIES

DNA Sample Collection

This study was approved by İstanbul University Cerrahpaşa Medicine Faculty Clinical Research Ethics Committee (date: November 5, 2014, no: 235829). We collected blood or buccal swab samples from 148 unrelated volunteers with written informed consent. These samples were selected to represents all regions of Türkiye. A total of 700 reference InDel profiles from the HGDP-CEPH diversity panel database were accessed using the *SPSmart forInDel* browser and were used as comparison reference samples (Table 1).[17]

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.

Amplification and genotyping

Isolation of DNA was carried out by QIAamp® DNA Mini Kit (Qiagen, Germany). Amounts of DNA was measured by Qubit dsDNA HS Assay (Invitrogen by Thermo Fisher Scientific, ABD) kit. Primer amounts and polymerase chain reaction (PCR) conditions were applied according to the procedure recommended by Pereira et al.[1] PCR products were run on ABI PRISM® 310 Genetic Analyzer (Thermo Fisher Scientific, ABD). Final data were analyzed via GeneMapper programme (Thermo Fisher Scientific, ABD).

STATISTICAL ANALYSIS

The allele frequencies for each locus and deviation from Hardy-Weinberg equilibrium (HWE) were tested using Arlequin ver. 3.5.[18] For the population comparison, the allele frequencies at each locus were compared with the reference populations from Africa, Europe, East Asia, Oceania, America, Middle

TABLE 1: The details of the HGDP-CEPH reference populations.

Africa (n=105)	Europe (n=160)	East Asia (n=128)	America (n=61)	Oceania (n=28)	Middle East (n=116)	Central-South Asia (n=97)
South African Bantu	Orkney Islands-Orcadian	Siberia-Yakut	Brazil- Karitiana	Bougainville-Melanesian	Algeria (Mzab)-Mozabite	Pakistan-Makrani
Namibia-San	France-French	Cambodia-Cambodian	Brazil-Surui	New Guinea-Papuan	Israel (Carmel)-Druze	Pakistan-Sindhi
Kenya-Bantu N.E.	France-Basque	China	Colombia-Colombians		Israel (Central)-Palestinian	Pakistan-Burusho
D. R. of Congo-Mbuti Pygmy	Italy-Sardinian	Japan-Japanese	Mexico-Maya		Israel (Negev)-Bedouin	
C. African Republic-Biaka Pygmy	Italy-Tuscan		Mexico-Pima			
Nigeria - Yoruba	Italy-from Bergamo					
Senegal-Mandenka	Russia-Russian					
	Russia Caucasus-Adygei					

HGDP-CEPH: Human Genome Diversity Project-Centre d'Étude du Polymorphisme Humain.

East, and Central South Asia by using Arlequin ver.3.5.[18] Statistical analysis on biogeographical ancestry was carried out via Snipper 2.5 web online application (http://mathgene.usc.es/snipper/).[16] Genetic characteristics of the Turkish population among other populations were tested by using Structure 2.3.4.[19] In Structure analysis, clusters between populations are represented with letter "K".[19] The K values ranging from 2-6 were analyzed in order to determine the best K value. The other Structure parameters were applied as 10,000 burn-in period followed by 10,000 Markov Chain Monte Carlo (MCMC) repetitions; admixture model; correlated allele frequencies; 10 independent replicates. Graphical results were created by using CLUMPAK (http://clumpak.tau.ac.il).[20]

Ancestry assignments of each individual were established via USC Bayesian forensic SNP classifier Snipper 2.5 (http://mathgene.usc.es/snipper/). The ancestry inference was done according to the 5 major populations (Europe, Africa, America, East Asia and Oceania), which are embedded in the web application.[16] principal component analysis were also created using Snipper 2.5 (http://mathgene.usc.es/snipper/).

RESULTS

A total of 148 individuals have succesfully been typed for all 46 AIM-InDel markers. In this study, allele frequencies of 46 AIM-InDel loci in Turkish population were determined. Also the genetic distance analysis is performed between Turkish and other populations. A sample electropherogram of 46 AIM-InDel loci is illustrated in Figure 1.

Allele frequencies and HWE ($p<0.05$) values of 46 AIM-InDel loci in Turkish population were calculated using Arlequin 3.5. Bonferroni correction was conducted for p-values and statistical significance level was determined as $p<0.00108$. After Bonferroni correction, it was observed that all 46 AIM-InDel loci were in HWE (Table 2). Allele frequencies, expected and observed heterozygosities (H_o, H_e) and HWE p-values for Turkish population are shown in Table 2. Allele frequencies of 46 AIM-InDel loci were compared to seven reference (African, European, East Asian, Oceanian, American, Middle Eastern, Central South Asian) populations data. Analysis of pairwise genetic differences (F_{ST} values) between these populations are represented in Table 3.

Snipper programme establishes the closest population for a queried sample profile. The Snipper classification result of a sample profile is illustrated in Figure 2. On the left side, the triangle vertices represent the 3 most likely populations of origin for the classifed sample with the vertix representing an assignment probability of 1 and the opposite side a probability of 0 (Figure 2A). The classification was shown to be highly discriminative for the Europe, America and East Asia. The tested individual (in grey color) is located on the European side of the triangular and assigned as European. Three dimensional

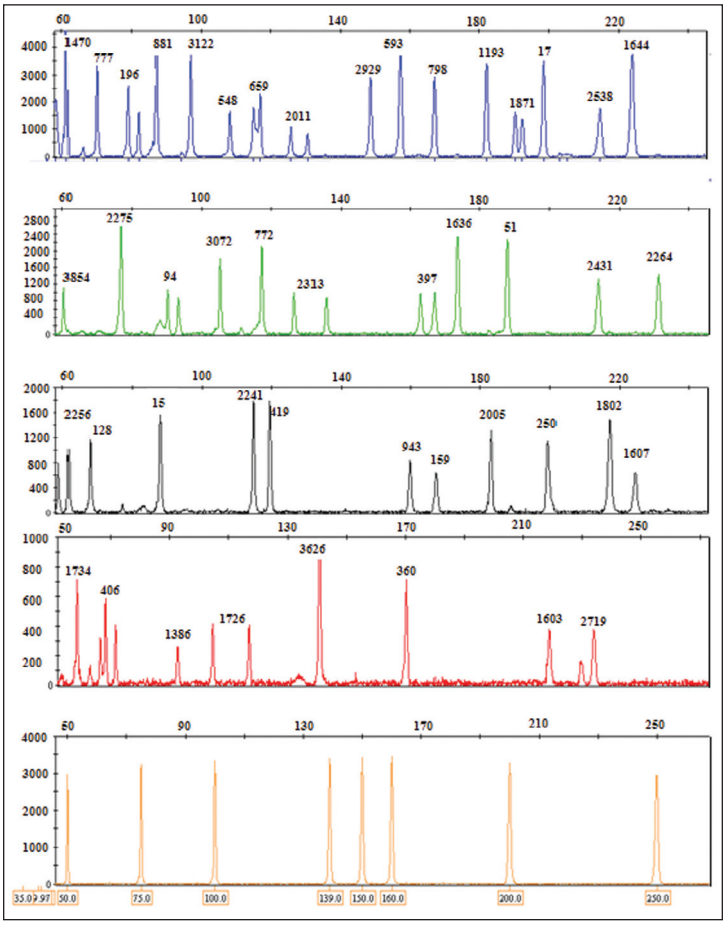


FIGURE 1: Electropherogram of 46-plex AIM-InDel loci. AIM-InDel: Ancestry informative marker-insertion/deletion.

analysis shows the position of queried sample profile within the three closest populations (Figure 2B). Europe displays in red, East Asia in green and America in blue. Arrow points at the queried individual in yellow. Again queried individual clustered with European populations. Therefore; according to Bayesian forensic SNP classifier, we can conclude that this individual is a billion (10^9) times more likely to have a European ancestor than an East Asian ancestor and more than a billion (10^9) times more likely European than American. In other words, it can be concluded that this individual is classified as European. Moreover, all Turkish samples (n=148) were classified as Europe with high likelihoods when we used the following function "Binary AIM classification of individuals as Europe-East Asia-Africa-America-Oceania (34 SNPs, 46 Indels, or both sets)" with fixed 5-group training sets on Snipper web tool (http://mathgene.usc.es/snipper/).

Figure 3 indicates Structure cluster plots and ancestral membership proportions of 8 populations (Africa, East Asia, Europe, America, Oceania, Middle East, Central South Asia, and Türkiye) for K=2-6. The K=2 plot distinguishes East Asian and American (orange) from all other population groups. Then, at K=3 African population (dark purple) is seperated from other populations resulting occurance of the three main continental population groups. STRUCTURE clearly separated Americans (green) into distinct subset at K=4. A new cluster (purple) is occurred at K=5 and showed admixture patterns among Europe, Middle East, Central South Asia and Türkiye. These populations show admixed patterns from co-ancestries of Europe (blue) and Middle East/Central South Asia/Türkiye (purple) at comparable levels. The Middle East, Central South Asia and Türkiye (blue and purple) remain clustered together with various European admixture ratios. In European

TABLE 2: Allele frequencies, expected and observed heterozygosities (Ho, He) and HWE p-values of 46 AIM-InDel loci in Turkish population.

Loci	Deletion (-)	Insertion (+)	He	Ho	p value
MID-1470	0.520	0.480	0.541	0.501	0.41136
MID-777	0.466	0.534	0.541	0.499	0.33057
MID-196	0.453	0.547	0.541	0.497	0.32521
MID-881	0.821	0.179	0.277	0.295	0.56848
MID-3122	0.986	0.014	0.014	0.027	0.02047
MID-548	0.233	0.767	0.385	0.359	0.49208
MID-659	0.213	0.787	0.318	0.336	0.46915
MID-2011	0.669	0.331	0.459	0.444	0.71515
MID-2929	0.696	0.304	0.473	0.425	0.17182
MID-593	0.142	0.858	0.189	0.244	0.01113
MID-798	0.537	0.463	0.480	0.499	0.74182
MID-1193	0.209	0.791	0.311	0.332	0.45779
MID-1871	0.365	0.635	0.446	0.465	0.72241
MID-17	0.348	0.652	0.439	0.455	0.71979
MID-2538	0.392	0.608	0.500	0.478	0.61077
MID-1644	0.861	0.139	0.182	0.239	0.00874
MID-3854	0.034	0.966	0.068	0.066	1.00000
MID-2275	0.172	0.828	0.304	0.286	0.56898
MID-94	0.274	0.726	0.493	0.399	0.00319
MID-3072	0.973	0.027	0.054	0.053	1.00000
MID-772	0.882	0.118	0.169	0.209	0.03078
MID-2313	0.368	0.632	0.453	0.467	0.72967
MID-397	0.716	0.284	0.365	0.408	0.22390
MID-1636	0.784	0.216	0.338	0.340	1.00000
MID-51	0.642	0.358	0.500	0.461	0.37393
MID-2431	0.122	0.878	0.203	0.214	0.45190
MID-2264	0.338	0.662	0.554	0.449	0.00501
MID-2256	0.166	0.834	0.304	0.277	0.36942
MID-128	0.493	0.507	0.568	0.502	0.13776
MID-15	0.483	0.517	0.466	0.501	0.41223
MID-2241	0.341	0.659	0.507	0.451	0.15060
MID-419	0.780	0.220	0.358	0.344	0.80876
MID-943	0.682	0.318	0.405	0.435	0.43707
MID-159	0.628	0.372	0.486	0.469	0.72130
MID-2005	0.601	0.399	0.486	0.481	1.00000
MID-250	0.645	0.355	0.453	0.459	0.85776
MID-1802	0.030	0.970	0.034	0.059	0.00483
MID-1607	0.169	0.831	0.284	0.282	1.00000
MID-1734	0.711	0.289	0.442	0.412	0.41879
MID-406	0.676	0.324	0.419	0.440	0.57007
MID-1386	0.250	0.750	0.378	0.376	1.00000
MID-1726	0.666	0.334	0.493	0.447	0.26517
MID-3626	0.713	0.287	0.372	0.411	0.30971
MID-360	0.818	0.182	0.243	0.299	0.02670
MID-1603	0.236	0.764	0.419	0.362	0.06641
MID-2719	0.247	0.753	0.399	0.373	0.49964

HWE: Hardy-Weinberg equilibrium; AIM-InDel: Ancestry informative marker-insertion/deletion.

TABLE 3: Population differentiation tests between Turkish with seven population (F_{ST} values).

	Africa	Europe	East Asia	America	Oceania	Middle East	Central-South Asia	Türkiye
Africa	*							
Europe	0.3652	*						
East Asia	0.3923	0.2839	*					
America	0.4423	0.2974	0.2199	*				
Oceania	0.3742	0.2249	0.2248	0.3098	*			
Middle East	0.2889	0.0155	0.2576	0.2679	0.1750	*		
Central-South Asia	0.2982	0.0336	0.2093	0.2112	0.1758	0.0229	*	
Türkiye	0.3107	**0.0169**	0.2399	0.2438	0.1704	**0.0123**	**0.0158**	*

Bold values are indicated lowest F_{ST} values. * indicates no calculations.

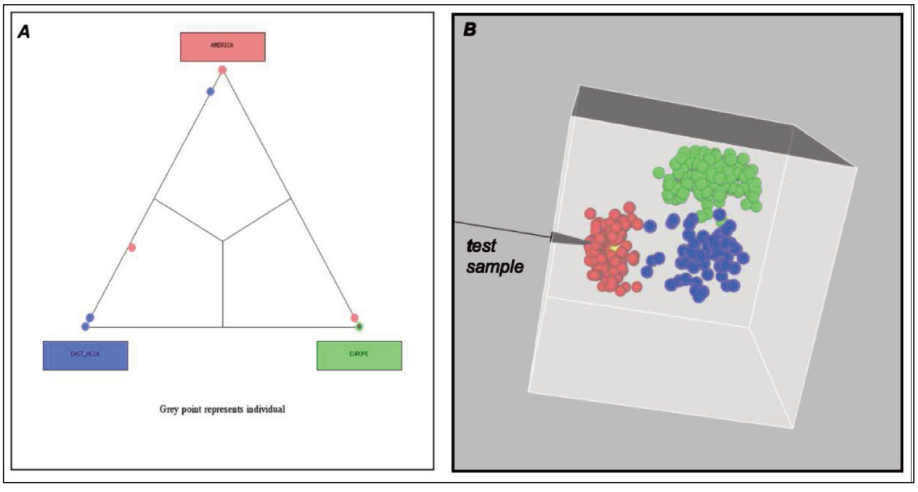


FIGURE 2: (A) Triangle and **(B)** three dimensional plots visualising the classifcation of a Turkish individual with the Snipper software.

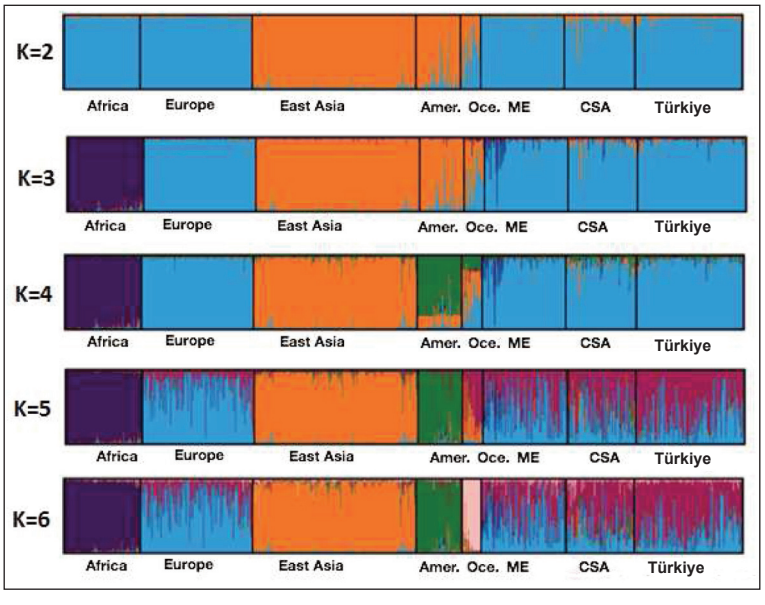


FIGURE 3: Structure cluster plots of populations for values K=2-6 using 46 AIM-InDel. Each vertical bar represents one individual and the colours represent the individual admixture proportions based on K assumed clusters. AIM-InDel: Ancestry informative marker-insertion/deletion.

individuals, small proportions of Middle East/Central South Asia/Türkiye (purple) co-ancestries were also observed. The highest level of distinction resulted in Africa, Europe, East Asia, America and Ocenia clusters for K=6.

DISCUSSION

InDels have demonstrated their considerable potential as a forensic marker. InDels have very simple direct genotyping approach and have small amplicon size which provides succesful typing highly degraded DNA. They can also be analyzed using current technologies in the forensic genetic laboratories. InDels also show high allele frequency differentiation amongst population groups, hence this make them a good marker for the forensic analysis of biogeographic ancestry.[5] In this study, we determined allele frequencies of 46 AIM-InDel loci in Turkish population. We used different population genetic analysis (F_{ST} Structure and Snipper) to establish degree of genetic differences between Turkish and other populations (African, European, East Asian, Oceanian, American, Middle Eastern, and Central South Asian).

The allele frequencies of 46 AIM-InDel loci in Turkish population are shown in Table 2. Deletion and insertion frequencies of MID-128 locus were observed as 0.493 and 0.507, respectively and MID-128 showed the highest heterozygosity. Similarly, heterozygosity ratios of MID-1470 and MID-15 loci were also high. F_{ST} analysis, which converts differentiation of populations to numerical values, is a frequently used analysis method in population genetics. F_{ST} values between 0-0.05 correspond to lower degrees of genetic differentiation; F_{ST} values between 0.05-0.15 correspond to moderate degrees of differentiation; F_{ST} values between 0.15-0.25 correspond to higher degrees of differentiation and F_{ST} values greater than 0.25 correspond to very high degrees of genetic differentiation.[21] When F_{ST} values are examined, slight genetic differentiation is observed between Turkish population and Middle Eastern, Central South Asian, and European populations (Table 3). On the other hand, large genetic differentiation with American, Oceanian and East Asian populations are obtained and a very large genetic differentiation with African population is also observed. Our results corroborate the findings of the previous studies.[1,16 22]

Snipper analysis is an online forensic application that can be used in inference of ancestry.[23,24] This application, which is based on Bayesian approach, determines the closest population for an individual by comparing the sample profile with reference data.[24,25] Each of the population sample profiles were estimated for individual ancestry assignment using reference HGDP-CEPH diversity panel genetic data from the 5 population groups (Africa, Europe, Asia, America, Oceania). Three-dimensional depiction of genetic likelihood of a sample profile for different populations exhibit the position of this profile in the European reference samples (Figure 2). Thus, we can deduce that sample profile is classified as European. Moreover, the Bayesian ancestry assignments for the all Turkish individuals were assigned as European (data not shown). Since this data set does not contain Middle Eastern and Central South Asian populations, samples of Turkish population were observed only close to the European population.

Structure is developed based on the model-based clustering approach that utilizes MCMC simulations to simultaneously compute the most probable allele frequency distribution for a given set of populations K using the genotype information of the samples. This analysis has great advantages for many population genetic applications since clustering is performed based solely on the available genotype data with no prior information about population structure necessary.[19,26] We use different K values to determine the optimum population clustering in this study. Figure 3 illustrates population clusters among K values (K=2 to K=6). The K=2 analyses separate East Asia and America (orange color) as a single cluster from the other populations. At K=3, Africa (dark purple color) is differentiated, then at K=4 America (dark green color) is distinguishable as a unique group. Oceania (pink color) further separate as a distinct population cluster at K=6. Turkish population was clustered with Middle Eastern and Central South Asian populations and it also included a common cluster (max 20%) with European population at K=5 and K=6. This result is in concordance with F_{ST} and Snipper results. It also corroborates with another

study in which Turkish population is represented with a small data set.[22]

CONCLUSION

This study shows that Turkish population is genetically closer to European, Middle Eastern and Central South Asian populations when 46 AIM-InDel markers are analyzed. These markers initially selected for differentiating 5 major groups namely, Africa, Europe, East Asia, Native America, and Oceania.[1] In order to increase genetic differentiation between these populations, larger number of samples and new or additional marker (e.g. AISNPs) are needed.

InDel loci can be used in establishment of biogeographical ancestry studies, degraded and low amount of biological crime scene samples, anthropological and archaeological studies and cases in which STR typing cannot be used. InDel loci also enables more successful PCRs with shorter period of time and low cost compared to SNPs.[5] All of these advantages will increase the use of InDel loci in the future. More InDel loci located on the other somatic and gamete chromosomes should be investigated to make InDels alternative for STRs and SNPs.

Source of Finance

This work was supported by the [Istanbul University Research Fund] under Grant [Project no: 50135].

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: *Cemaleddin Arslan, Özlem Bülbül, Gönül Filoğlu;* ***Design:*** *Cemaleddin Arslan, Özlem Bülbül, Gönül Filoğlu;* ***Control/Supervision:*** *Cemaleddin Arslan, Özlem Bülbül, Gönül Filoğlu;* ***Data Collection and/or Processing:*** *Cemaleddin Arslan, Özlem Bülbül, Gönül Filoğlu;* ***Analysis and/or Interpretation:*** *Cemaleddin Arslan, Özlem Bülbül, Gönül Filoğlu;* ***Literature Review:*** *Cemaleddin Arslan, Özlem Bülbül, Gönül Filoğlu;* ***Writing the Article:*** *Cemaleddin Arslan, Özlem Bülbül, Gönül Filoğlu;* ***Critical Review:*** *Özlem Bülbül, Gönül Filoğlu;* ***References and Fundings:*** *Cemaleddin Arslan, Özlem Bülbül, Gönül Filoğlu;* ***Materials:*** *Özlem Bülbül, Gönül Filoğlu;*

REFERENCES

1. Pereira R, Phillips C, Pinto N, Santos C, dos Santos SE, Amorim A, et al. Straightforward inference of ancestry and admixture proportions through ancestry-informative insertion deletion multiplexing. PLoS One. 2012;7(1):e29684. [Crossref] [PubMed] [PMC]
2. Mills RE, Luttig CT, Larkins CE, Beauchamp A, Tsui C, Pittard WS, et al. An initial map of insertion and deletion (INDEL) variation in the human genome. Genome Res. 2006;16(9):1182-90. [Crossref] [PubMed] [PMC]
3. Mullaney JM, Mills RE, Pittard WS, Devine SE. Small insertions and deletions (INDELs) in human genomes. Hum Mol Genet. 2010;19(R2):R131-6. [Crossref] [PubMed] [PMC]
4. Weber JL, David D, Heil J, Fan Y, Zhao C, Marth G. Human diallelic insertion/deletion polymorphisms. Am J Hum Genet. 2002;71(4):854-62. [Crossref] [PubMed] [PMC]
5. Phillips C. Application of autosomal SNPs and indels in forensic analysis. Forensic Sci Rev. 2012;24(1):43-62. [PubMed]
6. Pereira R, Phillips C, Alves C, Amorim A, Carracedo A, Gusmão L. A new multiplex for human identification using insertion/deletion polymorphisms. Electrophoresis. 2009;30(21):3682-90. [Crossref] [PubMed]
7. Friis SL, Børsting C, Rockenbauer E, Poulsen L, Fredslund SF, Tomas C, et al. Typing of 30 insertion/deletions in Danes using the first commercial indel kit--Mentype® DIPplex. Forensic Sci Int Genet. 2012;6(2):e72-4. [Crossref] [PubMed]
8. Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, et al. Intraspecific phylogeography: the mitochondrial dna bridge between population genetics and systematics. Annu Rev Ecol Syst. 1987;18(1):489-522. [Crossref]
9. Tishkoff SA, Kidd KK. Implications of biogeography of human populations for 'race' and medicine. Nat Genet. 2004;36(11 Suppl):S21-7. [Crossref] [PubMed]
10. Kidd JM, Gravel S, Byrnes J, Moreno-Estrada A, Musharoff S, Bryc K, et al. Population genetic inference from personal genome data: impact of ancestry and admixture on human genomic variation. Am J Hum Genet. 2012;91(4):660-71. [Crossref] [PubMed] [PMC]
11. Gross TE, Zaumsegel D, Rothschild MA, Schneider PM. Combined analysis of two different ancestry informative assays using SNPs and Indels in Eurasian populations. Forensic Science International: Genetics Supplement Series. 2013;4(1):e25-6. [Crossref]
12. Pereira R, Gusmão L. Capillary electrophoresis of 38 noncoding biallelic mini-Indels for degraded samples and as complementary tool in paternity testing. Methods Mol Biol. 2012;830:141-57. [Crossref] [PubMed]
13. Pereira R, Pereira V, Gomes I, Tomas C, Morling N, Amorim A, et al. A method for the analysis of 32 X chromosome insertion deletion polymorphisms in a single PCR. Int J Legal Med. 2012;126(1):97-105. [Crossref] [PubMed]

14. Ünsal T, Filoğlu G, Aşıcıoğlu F, Bülbül Ö. Population data of new 21 mini-InDels from Turkey. Forensic Sci Int Genet Suppl Ser. 2017;6:e189-91. [Crossref]
15. Santos NP, Ribeiro-Rodrigues EM, Ribeiro-Dos-Santos AK, Pereira R, Gusmão L, Amorim A, et al. Assessing individual interethnic admixture and population substructure using a 48-insertion-deletion (INSEL) ancestry-informative marker (AIM) panel. Hum Mutat. 2010;31(2):184-90. [Crossref] [PubMed]
16. Santos C, Phillips C, Oldoni F, Amigo J, Fondevila M, Pereira R, et al. Completion of a worldwide reference panel of samples for an ancestry informative Indel assay. Forensic Sci Int Genet. 2015;17:75-80. [Crossref] [PubMed]
17. Amigo J, Salas A, Phillips C, Carracedo A. SPSmart: adapting population based SNP genotype databases for fast and comprehensive web access. BMC Bioinformatics. 2008;9:428. [Crossref] [PubMed] [PMC]
18. Excoffier L, Lischer HE. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour. 2010;10(3):564-7. [Crossref] [PubMed]
19. Rosenberg NA, Pritchard JK, Weber JL, Cann HM, Kidd KK, Zhivotovsky LA, et al. Genetic structure of human populations. Science. 2002;298(5602):2381-5. [Crossref] [PubMed]
20. Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. Mol Ecol Resour. 2015;15(5):1179-91. [Crossref] [PubMed] [PMC]
21. Wright S. The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution. 1965;19(3):395-420. [Crossref]
22. Bulbul O, Duvenci A, Zorlu T, Gurkan C, Santos C, Phillips C, et al. Studies of East European populations with a 46-plex ancestry-informative indel set. Forensic Sci Int Genet Suppl Ser. 2015;e16-8. [Crossref]
23. Phillips C, Freire Aradas A, Kriegel AK, Fondevila M, Bulbul O, Santos C, et al. Eurasiaplex: a forensic SNP assay for differentiating European and South Asian ancestries. Forensic Sci Int Genet. 2013;7(3):359-66. [Crossref] [PubMed]
24. Yun LB, Gao TZ, Sun K, Gu Y, Hou YP. Comparison of two online algorithm methods for forensic ancestry inference. Forensic Sci Int Genet Suppl Ser. 2015;5:e559-60. [Crossref]
25. Phillips C, Salas A, Sánchez JJ, Fondevila M, Gómez-Tato A, Alvarez-Dios J, et al; SNPforID Consortium. Inferring ancestral origin using a single multiplex assay of ancestry-informative marker SNPs. Forensic Sci Int Genet. 2007;1(3-4):273-80. [Crossref] [PubMed]
26. Rosenberg NA, Li LM, Ward R, Pritchard JK. Informativeness of genetic markers for inference of ancestry. Am J Hum Genet. 2003;73(6):1402-22. [Crossref] [PubMed] [PMC]