Legal Medicine 37 (2019) 45–48

Contents lists available at ScienceDirect

Legal Medicine

journal homepage: www.elsevier.com/locate/legalmed

# Genotyping of 49-plex autosomal SNP panel in Iranian Turkmens ethnic group

Omid Yousef[a], Sayed Mostafa Hosseini[b,*]

[a] Genetic & Biotechnology Department, Faculty of Biology Science, Varamin-Pishva, Islamic Azad University, Tehran, Iran
[b] Human Genetic Research Center, Baqiyatallah University of Medical Sciences, Tehran, Iran

ARTICLE INFO

*Keywords:*
Autosomal SNPs
SNP*for*ID
Human identification
SNaPshot
Iran
Turkmens

ABSTRACT

A total of 94 unrelated individuals from Turkmens ethnic group in Iran were typed for forty-nine of the autosomal single nucleotide polymorphisms (SNPs) in the SNP*for*ID 52plex using the SNaPshot assay. Allele frequencies are presents for the 49 SNPs. No deviation from Hardy–Weinberg equilibrium (HWE) was observed in all but one of the 49 SNP systems and no significant linkage disequilibrium was detected for any SNP pairs. $F_{IS}$ and $F_{ST}$ were estimated. A statistically significant global $F_{ST}$ value was obtained when Turkmens ethnic group were compared with other 20 populations in Turkey, Israel, Pakistan, India, China, Taiwan, Japan, Thailand, Siberia, Algeria, Somali, Uganda, Mozambique, Angola, Nigeria, Russia, Slovenia, Sweden, France and Spain. All but 11 pairwise $F_{ST}$ values were statistically significant. Multidimensional scaling plot drawn based on the pairwise $F_{ST}$ values showed that the Turkmens ethnic grouped with populations geographically close to Iran and other Middle-Eastern populations. The cumulative values for the match probability using the 49 SNPs was $5.65 \times 10^{-19}$ consistent to a combined power of discrimination of > 99.99999% and the mean exclusion probability was 99.95%.

## 1. Introduction

### 1.1. Population

The Turkmens are a nation and Turkic ethnic group native of over 6,000,000 inhabitants to Central Asia, mainly the Turkmen nation state of Turkmenistan. Smaller communities exist are also found in Iran [1]. The Iranian Turkmans are about 1,328,585 and comprise two percent of Iranian population [2]. An ethnic minority who speak the Turkish language with the Eastern Oghoz accent. The same dialect has spoken in the Republic of Turkmenistan. They live in the Turkoman Sahra and in the Gorgan plains. The area is a fertile plain near the Iranian border with the Republic of Turkmenistan. It extends from the Atrak River in the north, to the Caspian Sea in the west, Quchan Mountains to the east and the Gorgan River to the south. Iranian Turkmans have been living in Iran since 550 CE. They are the descendants of Central Asian Turks, who retained their ethnic identity during the Mongol invasion. They now live in Golestan and are different from the other Iranian ethnic groups in appearance, language and culture [3].

So far, various ethnic groups of Iran have been investigated using lineage markers (mitochondrial DNA) [4,5], Y-chromosomal short tandem repeats (STRs) [6], high resolution Y-chromosomal genotyping [7,8], autosomal markers (STRs) [9] and single nucleotide polymorphisms (SNPs) [10,11]. However, the genetic structure of the Iranian Turkmens ethnic group from a SNP standpoint is still uncertain. The main contribution of the current study is analyzing of 94 Iranian Turkmens individuals for the 49 autosomal SNPs, and compare our data to other Individuals genotyped by 49 autosomal SNPs data in order to update the data in the Iranian population. To our best knowledge, this is the first study on the analysis of the allele frequency and genotyping of the 49 autosomal SNPs loci in Turkmens ethnic group of Iranian population.

## 2. Material and methods

### 2.1. Samples

Whole blood was collected from a total of 94 available unrelated healthy donors from Iranian Turkmen ethnic group with ancestry traced back at least two generations on 3 mm BD Vacutainer™ plastic blood collection tubes with $K_2$EDTA. The samples were collected at Medical Laboratory of Gorgan University of Medical Sciences, Golestan, Iran. The individuals were selected using town of origin, language and the ethnicity of the individuals and their parent's criteria. The work

* Corresponding author.
*E-mail address:* smhosseini@bmsu.ac.ir (S.M. Hosseini).

https://doi.org/10.1016/j.legalmed.2019.01.004
Received 20 June 2018; Received in revised form 22 September 2018; Accepted 1 January 2019
Available online 02 January 2019
1344-6223/ 

approved by the Baqiyatallah University Ethical Committee (IR.BMSU.RBC.1395.769).

### 2.2. DNA extraction and quantification

DNA was extracted from 500 µl Donor's blood samples using RGDE [12] assay extraction. DNA concentrations were determined by measuring optical density and a semi-quantitative method using agarose gels.

### 2.3. SNP genotyping

Forty-nine of autosomal SNP*for*ID SNPs were analyzed followed the protocol as previously described [13] by one PCR and two single base extension (SBE) using the SNaPshot kit (AB/LT: Applied Biosystems by Life Technologies, Carlsbad, CA, USA). The SBE products were electrophoresed in an ABI 3130xl Genetic TM Analyzer (AB/LT) using POP-4 polymer (AB/LT) and 36 cm capillary arrays. The GenMapper v3.2 software (AB) were used to analyze the resulting electropherograms. Details on the protocol and criteria for homozygote and heterozygote allele calling were used as previously described [14]. All experiments were performed in the Noor Human Genetics Research Center, Tehran, Iran which is accredited according to the ISO 17,025 standard [15].

### 2.4. Quality control

Two analysts reviewed data independently after genotyping and automated analysis with GeneMapper v3.2, and then results were scrutinized. This study followed the recommendations of the International Society of Forensic Genetics (ISFG) on the analysis of DNA polymorphisms, which included the use of recommended nomenclature and guidelines regarding quality control and statistical calculations [16].

### 2.5. Statistical analysis

Population genetic parameters like allele frequencies, observed and expected heterozygosities, Hardy–Weinberg equilibrium and linkage disequilibrium were calculated by Arlequin (v.3.5.2.2) software [17]. In order to obtain the global amount of population genetic structure a two hierarchical AMOVA analysis was performed. Data regarding the 49 autosomal SNPs analyzed in 20 populations from Turkey, Israel, Pakistan, India, China, Taiwan, Japan, Thailand, Siberia, Algeria, Somali, Uganda, Mozambique, Angola, Nigeria, Russia, Slovenia, Sweden, France and Spain were collected from the SNP*for*ID browser [18]. SNP profiles where results from more than two missing SNP systems removed from the analysis. Moreover, fixation indices $F_{IS}$ and $F_{ST}$ and pairwise $F_{ST}$ values were calculated and their significance was tested using 10,000 permutations. A multi-dimensional scaling (MDS) was drawn from the pairwise $F_{ST}$ values using the SPSS statistical package 24.0 (SPSS Inc., Chicago, IL, USA). The Holm–Sˇidak correction was applied to adjustment for multiple statistical analyses [19]. Estimation of forensic informative metrics included: random match probability and discrimination power plus paternity exclusion power and typical paternity indices were calculated using the Power Stats v. 1.2 (http://www.promega.com/geneticidtools/powerstats/).

## 3. Results and discussion

The allele frequencies, observed and expected heterozygosities and the Hardy–Weinberg equilibrium test results for the 49 SNPs in the Turkmens ethnic are presented in (Supplementary Material SM1). The minor allele frequency were evaluated from 0.005 (rs2056277) to 0.495 (rs727811) and only one another loci (rs938283) of the 49 SNP loci showed minor allele frequencies under 10%. The mean heterozygosity reaching a value of 42.6%, and it was the lowest value among other Iranian ethnic groups [10]. The average gene diversity index observed for the 49 loci was 0.435. The SNP loci (rs2056277) and (rs737681) showed the lowest (0.01) and the highest (0.59) gene diversities, respectively (Supplementary Material SM1). A non-significant ($P < 0.05$), positive $F_{IS}$ value was observed for the 49 SNPs in the Turkmens ethnic ($F_{IS} = 0.023$). (not given)

After Holm–Sˇidak correction, p values $< 0.0104$ were considered statistically significant. All SNPs were in equilibrium except (rs354439). A significant low number of heterozygotes were not observed for the (rs354439) system either in any of the 20 populations collected from the SNP*for*ID browser [18] or previous Iranian ethnic groups studies[10,11]. Considering that Turkmens ethnic known as an isolation population due to their cultural and linguistic in Iran, a certain degree of inbreeding is expected. In addition, migrant rate in this population already estimated as "-1.8" (2017) [2]. Moreover, based on their Ethnogenesis, the population of Turkmen divided to multiple subpopulations including Uzbek, smaller population of Kazakh, Russian, Tatar, and etc. Due to Turkmen population structure, it leads to the Wahlund effect, which can be defined as an overall deficit of heterozygotes and increasing of homozygotes observed for the SNP marker (rs354439) even when the local Turkmens populations are randomly mating.

No significant linkage disequilibrium ($P < 0.01$) was found after correction for multiple comparisons of the 1176 pairwise comparisons in the Turkmens ethnic. However, these results were contrary to the results of a single SNP pair (rs1015250–rs251934) previous studies in Kurds ethnic [10].

Analyzing a two-hierarchical AMOVA in 21 populations from Turkmens, Turkey, Israel, Pakistan, India, China, Taiwan, Japan, Thailand, Siberia, Algeria, Somali, Uganda, Mozambique, Angola, Nigeria, Russia, Slovenia, Sweden, France and Spain showed a significant overall $F_{ST}$ value ($F_{ST} = 0.0696$, $P < 0.00001$) (Supplementary Table SM2) which underlines the fact that none of the SNP markers has any likely functional relevance. Significant pairwise $F_{ST}$ values were obtained after step-down Holm–Sˇidak correction, and P values under 0.0099 were significant at the 0.05 level; for all population pairs. Eleven comparisons from the 210 pairwise $F_{ST}$ values were statistically non-significant. The lowest $F_{ST}$ values was observed between the Pakistan population and Turkmen, and afterwards the closest $F_{ST}$ value belongs to the Turkey population, respectively (Supplementary Material SM3).

A multi-dimensional scaling (MDS) was drawn from the pairwise $F_{ST}$ values between 21 populations included in the analyses (Fig. 1). The plot visualized a clearly differentiates between populations in East Asians and Sub-Saharan Africa from the other populations. Europeans, North Africans, and Western Asia and South Asia did not show a distinct cluster. The populations are distributed according to their geographic origin from the east to the west and, as expected, Turkmens grouped together with populations geographically close to Iran and West Asia populations included in the study and interestingly, in agreement to other Iranian ethnicities[10] or the Iraqi population that already have been investigated [20]. The relative position of the Turkmens ethnic in this study segregates in close proximity to the Central Asia and populations in Turkey and the Caucasus.

Several previous studies using mtDNA[4], Y-chromosome [6,7] specific markers, STRs[9] and SNPs [10,11], demonstrated that the genetic distances between Iranians population and populations in the Indus Valley and Central Asia were higher than those between Iranians population and populations in Turkey and the Caucasus. It was in contrast with our results and relative place for the Turkmens ethnic. We performed a pairwise $F_{ST}$ between Turkmens ethnic and eight Pakistanian ethnics groups contain Balochi, Brahui, Burusho, Hazara, Kalash, Makrani, Pathan and Sindhi. The result indicated that the lowest $F_{ST}$ values was observed between the Pakistan ethnic groups and Turkmen ethnic belongs to the Makrani, and afterwards Balochi, respectively (not given). It was closely similar to distance between Turkmens ethnic

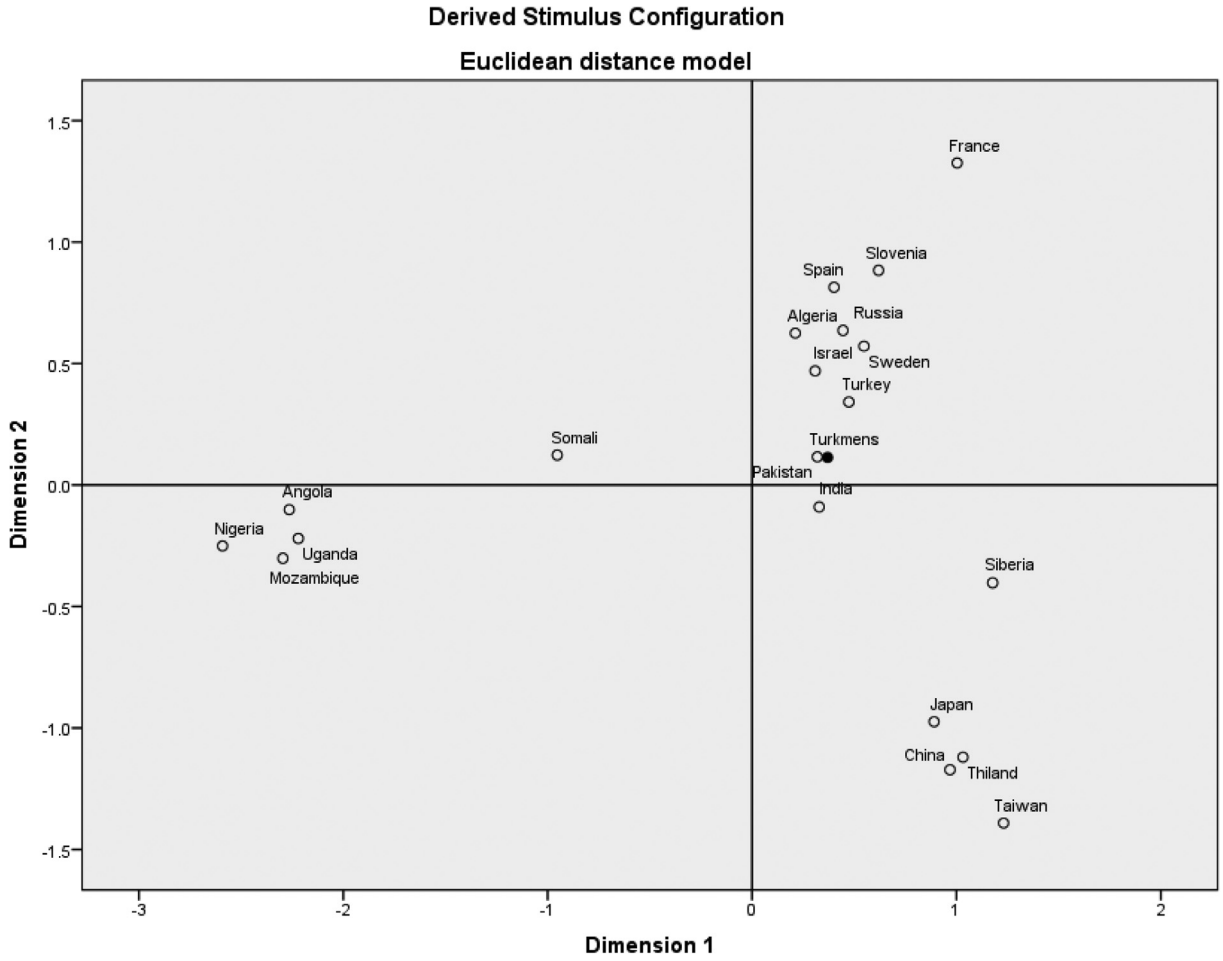


**Fig. 1.** Multidimensional scaling plot drawn from $F_{ST}$ values calculated between 21 worldwide populations analyzed for 49 autosomal SNPs from the SNP*for*ID 49plex. The Turkmen Iranian ethnic groups evaluated in the current study are shown by filled circles.

and Turkey population (Supplementary Material SM3). Interestingly, this point was conflicted in the literature by studies using mtDNA lineage markers [5] and Y-Chromosome specific markers [8] on people who living in Golestan area.

Random match probability, discrimination power, exclusion power and typical paternity index values for the Turkmens ethnic group were calculated for each SNP and the results are presented in (Supplementary Material SM4). The cumulative values for the match probability using the 49 SNPs was $5.65 \times 10^{-19}$ corresponding to a combined power of discrimination of > 99.99999% and the mean exclusion probability was 99.95%. These values were compared with previously calculated values for Asians, Africans and Europeans [10,13,20]. In relation to forensic productivity, the level of informativeness of the 49plex SNP obtained in the Turkmens ethnic group were high overall and reached comparable levels to those previously found for the other population groups. The genotyping results of the 49plex SNP indicates that these assay will provide valuable information both for population studies and for forensic applications.

## 4. Conclusion

The implementation of population data, as presented in this work revealed that the 49plex SNP assay is robust and sufficiently sensitive to be used for forensic purposes for Turkmens ethnic groups. Therefore, the obtained data could be used for forensic calculations for the Iranian Turkmens ethnic group. Overall, this study suggests that populations from geographically intermediate regions like Iran which high ethnic heterogeneity require careful characterization in order to choose the appropriate dataset to base forensic calculations.

## Acknowledgements

The authors would like to thank the contribution of Medical Laboratory of Gorgan, staff for help with sample collection and very grateful to the indigenous people of Turkmens who volunteered samples for this study. We appreciate Baqiyatallah Hospitals for providing the collaboration. We thank all the staff at the Human Genetics Research Center and constant feedback of members greatly acknowledged.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.legalmed.2019.01.004.

## References

[1] https://www.cia.gov/library/publications/the-world-factbook/geos/ir.html, The World Factbook — Central Intelligence Agency. 2018 (accessed SEP 22.2018).
[2] https://www.cia.gov/library/publications/the-world-factbook/geos/tx.html#People, The World Factbook — Central Intelligence Agency. 2018 (accessed SEP 22. 2018).
[3] V. Rashidvash, Turkmen status within Iranian ethnic identity (cultural, geographical, political), Res. Humanities Social Sci. 3 (22) (2013) 88–93.
[4] L. Quintana-Murci, R. Chaix, R.S. Wells, D.M. Behar, H. Sayar, R. Scozzari, C. Rengo, N. Al-Zahery, O. Semino, A.S. Santachiara-Benerecetti, Where west meets east: the complex mtDNA landscape of the southwest and Central Asian corridor, Am. J. Human Genetics 74 (5) (2004) 827–845.

[5] S. Farjadian, M. Sazzini, S. Tofanelli, L. Castrì, L. Taglioli, D. Pettener, A. Ghaderi, G. Romeo, D. Luiselli, Discordant patterns of mtDNA and ethno-linguistic variation in 14 Iranian Ethnic groups, Hum. Hered. 72 (2) (2011) 73–84.
[6] L. Roewer, S. Willuweit, M. Stoneking, I. Nasidze, A Y-STR database of Iranian and Azerbaijanian minority populations, Forensic Sci. Int. Genet. 4 (1) (2009) e53–e55.
[7] M. Regueiro, A. Cadenas, T. Gayden, P. Underhill, R. Herrera, Iran: tricontinental nexus for Y-chromosome driven migration, Hum. Hered. 61 (3) (2006) 132–143.
[8] V. Grugni, V. Battaglia, B.H. Kashani, S. Parolo, N. Al-Zahery, A. Achilli, A. Olivieri, F. Gandini, M. Houshmand, M.H. Sanati, Ancient migratory events in the Middle East: new clues from the Y-chromosome variation of modern Iranians, PloS one 7 (7) (2012) e41252.
[9] E. Shepard, R. Herrera, Iranian STR variation at the fringes of biogeographical demarcation, Forensic Sci. Int. 158 (2–3) (2006) 140–148.
[10] M.S. Farzad, C. Tomas, C. Børsting, Z. Zeinali, M. Malekdoost, S. Zeinali, N. Morling, Analysis of 49 autosomal SNPs in three ethnic groups from Iran: Persians, Lurs and Kurds, Forensic Sci. Int.: Genetics 7 (4) (2013) 471–473.
[11] L. Poulsen, M.S. Farzad, C. Børsting, C. Tomas, V. Pereira, N. Morling, Population and forensic data for three sets of forensic genetic markers in four ethnic groups from Iran: Persians, Lurs, Kurds and Azeris, Forensic Sci. Int.: Genetics 17 (2015) 43–46.
[12] S.M. Ali, S. Mahnaz, T. Mahmood, Rapid genomic DNA extraction (RGDE), Forensic Sci. Int.: Genetics Suppl. Ser. 1 (1) (2008) 63–65.
[13] J.J. Sanchez, C. Phillips, C. Børsting, K. Balogh, M. Bogus, M. Fondevila, C.D. Harrison, E. Musgrave-Brown, A. Salas, D. Syndercombe-Court, A multiplex assay with 52 single nucleotide polymorphisms for human identification, Electrophoresis 27 (9) (2006) 1713–1724.
[14] C. Børsting, E. Rockenbauer, N. Morling, Validation of a single nucleotide polymorphism (SNP) typing assay with 49 SNPs for forensic genetic testing in a laboratory accredited according to the ISO 17025 standard, Forensic Sci. Int. Genet. 4 (1) (2009) 34–42.
[15] I.S.O.I.E. Commission, General requirements for the competence of testing and calibration laboratories, ISO/IEC 17025 (1999).
[16] P.M. Schneider, Scientific standards for studies in forensic genetics, Forensic Sci. Int. 165 (2–3) (2007) 238–243.
[17] L. Excoffier, H.E. Lischer, Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows, Mol. Ecol. Resour. 10 (3) (2010) 564–567.
[18] http://spsmart.cesga.es/snpforid.php?dataSet=snpforid52.
[19] S. Holm, A simple sequentially rejective multiple test procedure, Scand. J. Stat. (1979) 65–70.
[20] C. Tomas, I.E. Diez, E. Moncada, C. Børsting, N. Morling, Analysis of 49 autosomal SNPs in an Iraqi population, Forensic Sci. Int. Genet. 7 (1) (2013) 198–199.