Taylor & Francis
Taylor & Francis Group

Canadian Society of Forensic Science Journal

ISSN: 0008-5030 (Print) 2332-1660 (Online) Journal homepage: http://www.tandfonline.com/loi/tcsf20

# Genetic analysis of the Saraiki population living in Pakistan

Jilani A, Nadeem A, Tahir M A & Rasool N

**To cite this article:** Jilani A, Nadeem A, Tahir M A & Rasool N (2016) Genetic analysis of the Saraiki population living in Pakistan, Canadian Society of Forensic Science Journal, 49:3, 152-160, DOI: 10.1080/00085030.2016.1184851

**To link to this article:** https://doi.org/10.1080/00085030.2016.1184851

Published online: 31 May 2016.

CANADIAN SOCIETY OF FORENSIC SCIENCE JOURNAL, 2016
VOL. 49, NO. 3, 152–160
http://dx.doi.org/10.1080/00085030.2016.1184851

Taylor & Francis
Taylor & Francis Group

# Genetic analysis of the Saraiki population living in Pakistan

Jilani A[a], Nadeem A[a], Tahir M A[b] and Rasool N[b,c]

[a]Institute of Biochemistry and Biotechnology, University of Veterinary and Animal Sciences, Lahore, Pakistan; [b]DNA and Serology Department, Punjab Forensic Science Agency, Lahore, Pakistan; [c]Department of Chemistry, University of Management and Technology, Lahore, Pakistan

**ABSTRACT**

Autosomal short tandem repeat (STR) markers are a powerful tool used in forensic sciences for profile matching and paternity testing, sibship and kinship analysis. This study represents the allele frequency distribution of 15 autosomal STR multiplex of the Saraiki population living in Pakistan. Allele frequencies of this population were compared with the other populations living in Pakistan. Buccal swabs were taken from 150 unrelated individuals of the Saraiki population living in different regions of Pakistan and profiles were generated using an AmpFlSTR Identifiler PCR Amplification kit. Population genetic calculations were performed on this population. Allele frequencies of the Saraiki population showed that this population is at Hardy Weinberg Equilibrium except at loci D3S1358, TH01, D13S317, TPOX and FGA, with distinct differences in $H_{Obs}$ and $H_{Exp}$ values. The average heterozygosity and polymorphism information content (PIC) at all loci were 0.77 and 0.766, respectively. Each STR marker showed a high degree of polymorphism and a high power of discrimination (PD). A phylogenetic tree shows that the Saraiki population living in Pakistan is genetically distinct from other geographically neighboring populations of the country. The population data presented in this study can be used as a reference database for the Saraiki population in forensic casework.

**RÉSUMÉ**

Les marqueurs génétiques autosomiques de répétitions courtes en tandem (RCT) sont un outil puissant utilisé en sciences judiciaires pour la correspondance de profils, les tests de paternité et les analyses de fratrie et de parentalité. Cette étude présente la distribution des fréquences alléliques d'une population Saraiki vivant au Pakistan, en utilisant un multiplex de 15 marqueurs RCT. Les fréquences alléliques de cette population furent comparées avec d'autres populations vivant au Pakistan. Des écouvillons buccaux de 150 individus non apparentés de la population Saraiki vivant dans différentes régions du Pakistan furent analysés avec le système AmpFlSTR Identifiler PCR pour générer des profils génétiques. Des calculs de génétique des populations furent faits sur cette population. Les fréquences alléliques de cette population Saraiki montrent que cette population est en équilibre Hardy Weinberg sauf au loci D3S1358, TH01, D13S317, TPOX et FGA avec des différences distinctes pour les valeurs $H_{Obs}$ et $H_{Exp.}$ La moyenne

**ARTICLE HISTORY**
Received 3 March 2015
Accepted 20 January 2016

**KEYWORDS**
Autosomal short tandem repeats; multiplex PCR; Saraiki population; allele frequency distribution

**MOTS-CLÉS**
Répétitions courtes en tandem autosomiques; RPC multiplex; population Saraiki; distribution des fréquences alléliques

**CONTACT** Jilani A aeman.biotech@yahoo.com

d'hétérozygosité et le contenu informatif du polymorphisme (PIC) à tous les loci étaient de 0.77 et 0.766, respectivement. Chaque marqueur RCT montre un haut degré de polymorphisme et un haut pouvoir de discrimination. Un arbre phylogénétique montre que la population Saraiki vivant au Pakistan est génétiquement distincte des autres populations géographiquement voisines de ce pays. Les données de population présentées dans cette étude peuvent être utilisées comme base de données de référence en sciences judiciaires pour la population Saraiki.

## Introduction

In the last two decades, autosomal short tandem repeat (STR) loci have been widely used in forensic identity. Variable number tandem repeats (VNTR) are generally about 400−1000bp long as compared with STRs, which are about 100−400bp [1]. The small fragment size of STRs gives them an edge and makes them much more suitable for forensic casework because smaller fragments have less chance of being lost by degradation.

Many countries have developed large databases containing allele frequencies of STR loci. These databases are used for establishing match statistics for forensic use. These population databases are further categorized into sub populations. For example, the US population database is subdivided into African American, US Caucasian, Hispanics, Far East Asian and Native American[2].

Those countries that are attempting to develop their forensic capacity may not have their own established population databases. With the lack of population databases, even though DNA retrieved from the evidence at a crime scene matches that of a suspect, the strength of the match cannot be reported statistically [3]. In such a scenario it is difficult to find an appropriate existing population database of any other population to apply statistical weight to the DNA evidence. Sometimes an analyst has to rely on the population databases of other countries to use their allele frequencies for forensic case work. It is pertinent to mention that each country has its own population and population substructures [4]. Therefore, there is a need to calculate allele frequencies for specific loci in each population sub-structure to develop a national database that can be used for forensic analysis.

Pakistan is geographically divided into five provinces, named Punjab, Sindh, Gilgit-Baltistan, Khybar Pakhtoonkha and Balochistan. The history of Pakistan shows that there were many migrations and invasions by foreigners to this region. These migrations in this region were from West Asia, The Levant, and Sub-Saharan Africa. The population of Pakistan can be divided by a North-South division into two distinct regions. Balti, Hazara, Kalash, Kashmiri, Pathan and Punjabi constitute the major population living in northern parts of Pakistan. The populations living in the southern region of Pakistan are comprised of Balochi, Brahuis, Makrani, and Sindhi [5]. Brahuis belong to Kalat, Balochistan and are believed to be descendants of Turkish Iranian origin [6]. Balochis, residents of Balochistan, migrated from Aleppo (Syria) via Iran [6, 7]. Makarni living in coastal areas of Balochistan and Sindh have mainly negroid features and it is believed that they came from Africa [6]. Sindhis belong to the native residents of the South Eastern province of Pakistan, which is historically known as Sindomana by the Greeks. This population was influenced by the immigrants and settlers from Arabia, Turkey and Persia [8, 9]. Balti

have Tibetan origins and settled in the North-East Himalayan region of Pakistan [10]. Kalash Kafir are a well-isolated population living in the mountains of Hindukash in North-West Pakistan. It is speculated that they came to this region accompanying Alexander the Great when he invaded India [11]. Pathans living in the Khyber Pakhtonkha province are offspring of the lost tribes of Israel and soldiers of Alexander the Great [12, 13]. Hazara people have Mongoloid features and are considered descendants of Tartars [13]. Punajbis live in the province of Punjab, which is the most populous region of Pakistan. This can be further divided into two regions: North Punjab and Southern Punjab. Major ethnic groups living in Punjab are Arain, Jat, Rajput, Gujar, Awana and Meos [14]. The Saraiki belt is in the southern part of Punjab, their language is Saraiki. This belt is further extended to a few parts of the north Sindh, South Khyber Pakhtunkha and North East Balochistan. In the census of 1998, it was reported that out of a national population of 132 million, a figure of 13.9 million Saraiki speakers are resident in Pakistan. According to an estimate, the majority of Saraiki-speaking individuals live in the southern Punjab (www.census.gov.pk).

It is believed that religion, language and geography are considerable hurdles in gene flow within a population. The aim of this study is to calculate allele frequencies of 15 STR loci in the Saraiki population and to establish whether this sub-population is genetically distinct from the other populations living in Pakistan. Allele frequencies of these 15 loci were also compared to other populations of the world. Since there is no population database for the Saraiki population of Punjab, Pakistan, there is a dire need to develop one in order to provide accurate statistical weights to DNA evidence for the criminal justice system.

## Materials and methods

### *Sample collection*

Buccal swab samples were collected using Fisherbrand™ Cotton-Tipped Applicators [15] from 150 unrelated Saraiki volunteers living in the Saraiki belt of Pakistan. For this study, four swabs were taken from each individual. The collection of samples was done in accordance with the ethical guidelines outlined by the institutes involved in this study.

### *DNA extraction and quantitation*

DNA was extracted using an organic extraction method. Isolated human DNA was quantified on RT-PCR ABI 7500 using Plexor® HY System [16].

### *Genotyping*

DNA was amplified using the AmpFlSTR Identifiler PCR Amplification kit (Applied Biosystems) on a GeneAmp PCR system 9700. PCR amplification was carried out following the instructions provided by the manufacturer. The amplified products were analyzed on the ABI Prism Genetic Analyzer-3130 by Applied Biosystems. Alleles were called by running an allelic ladder in each run and analyzing the data on GeneMapper® ID-X V 1.2 [17].

### *Data analysis*

Statistical tools used to calculate the allele frequencies, power of discrimination (PD), power of exclusion (PE), polymorphism information content (PIC), matching probability (MP), typical paternity index (TPI) and chi-square ($\chi^2$) test for Hardy-Weinberg equilibrium were the PowerStat V121 Excel spread sheet and GenePop web user interface [18].

A Neighbour-Joining phylogenetic tree was developed by using Poptree2 software and data were taken from the online STR database PopSTR [19]. The Pakistani populations used to generate the phylogenetic tree are Pathan, Brahui, Balochi, Kalash, Makrani, Hazara, Burusho and Sindhi.

## Results

Allele frequencies calculated on 15 autosomal STR loci of the Saraiki populations of Pakistan are listed in Table 1. The allele frequencies of different populations geographically distributed in distinct regions of Pakistan are used to generate a Neighbour-Joining phylogenetic tree (Figure 1). Genetic distance ($F_{ST}$) is calculated in this tree. Analysis of this phylogenetic tree shows a cluster of Pathan with Balochi as a group while Brahui clusters with Makrani as another group. Sindhi, Saraiki and Kalash, although living in distanced places, more likely form a close cluster. This phylogenetic tree shows that the Saraiki population is a distinct population compared with other populations of Pakistan.

By looking at Table 2, there is significant difference between observed heterozygosity ($H_{Obs}$) and expected heterozygosity ($H_{Exp}$) at loci D3S1358, TH01, D13S317, TPOX and FGA, which shows that these loci are not in Hardy Weinberg Equilibrium (HWE). By comparing the allele types from other populations, it was observed that the Saraiki population had alleles different from other populations, such as allele 34.2 at D21S11, allele 14 at D7S820, alleles 8 and 14 at CSF1PO, alleles 13 and 20 at D3S1358, allele 7 at D13S317, allele 14 at D16S539, alleles 10 and 21 at vWA, alleles 10 and 22 at D18S51, alleles 7 and 15 at D5S818 and alleles 17 and 18 at FGA. The inbreeding parameter within populations $F_{IS}$ is 0.0205 and the total inbreeding coefficient $F_{IT}$ is 0.0345 for the Saraiki population. The co-ancestry coefficient ($F_{ST}$) between the Saraiki and the Caucasian populations is 0.025.

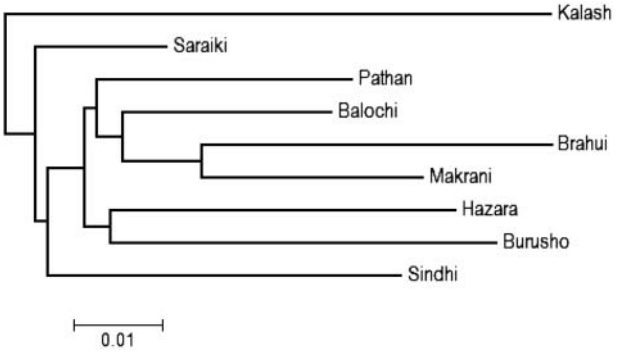


**Figure 1.** Neighbor-Joining phylogenetic tree representing $F_{ST}$ distances using allele frequencies of the STR loci of the different populations of Pakistan.

**Table 1.** Allele frequencies of the Saraiki Population (* = Frequency of allele is less than 0.016) [1].

| | D8S1179 | D21S11 | D7S820 | CSF1P0 | D3S1358 | TH01 | D13S317 | D16S539 | D2S1338 | D19S433 | vWA | TPOX | D18S51 | D5S818 | FGA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 6 | – | – | – | – | – | 0.26 | – | – | – | – | – | – | – | – | – |
| 7 | – | – | 0.0233 | – | – | 0.19 | 0.01* | – | – | – | – | – | – | 0.0033* | – |
| 8 | 0.0167 | – | 0.1733 | 0.0033* | – | 0.13 | 0.1567 | 0.08 | 0.0033* | – | – | 0.39 | – | – | – |
| 9 | 0.0067* | – | 0.0667 | 0.0133* | – | 0.2233 | 0.0633 | 0.11 | 0.0033* | – | – | 0.1 | – | 0.0333 | – |
| 9.3 | – | – | – | – | – | 0.1867 | – | – | – | – | – | – | – | – | – |
| 10 | 0.1533 | – | 0.1933 | 0.1967 | – | 0.01* | 0.08 | 0.0967 | – | – | 0.0033* | 0.0933 | 0.01* | 0.1133 | – |
| 11 | 0.0633 | – | 0.2967 | 0.3033 | – | – | 0.2533 | 0.4133 | – | – | – | 0.4067 | 0.03 | 0.3167 | – |
| 12 | 0.12 | – | 0.2033 | 0.37 | – | – | 0.27 | 0.1933 | – | 0.06 | – | 0.01* | 0.0467 | 0.3333 | – |
| 13 | 0.1767 | – | 0.0333 | 0.867 | 0.0033* | – | 0.1267 | 0.0967 | – | 0.2967 | – | – | 0.0967 | 0.1833 | – |
| 13.2 | – | – | – | – | – | – | – | – | – | 0.02 | – | – | – | – | – |
| 14 | 0.23 | – | 0.01* | 0.0267 | 0.0333 | – | 0.0367 | 0.01* | – | 0.2367 | 0.1333 | – | 0.2333 | 0.0133* | – |
| 14.2 | – | – | – | – | – | – | – | – | – | 0.0533 | – | – | – | – | – |
| 15 | 0.1667 | – | – | – | 0.3067 | – | 0.0033* | – | – | 0.1533 | 0.14 | – | 0.2367 | 0.0033* | – |
| 15.2 | – | – | – | – | – | – | – | – | – | 0.08 | – | – | – | – | – |
| 16 | 0.04 | – | – | – | 0.3033 | – | – | – | 0.01* | 0.0467 | 0.23 | – | 0.1367 | – | – |
| 16.2 | – | – | – | – | – | – | – | – | – | 0.0533 | – | – | – | – | – |
| 17 | 0.0233 | – | – | – | 0.2533 | – | – | – | 0.1067 | – | 0.2467 | – | 0.1067 | – | 0.0033* |
| 18 | 0.0033* | – | – | – | 0.09 | – | – | – | 0.1433 | – | 0.15 | – | 0.07 | – | 0.0133* |
| 19 | – | – | – | – | 0.0067* | – | – | – | 0.1467 | – | 0.07 | – | 0.0133* | – | 0.04 |
| 20 | – | – | – | – | 0.0033* | – | – | – | 0.0967 | – | 0.0233 | – | 0.0133* | – | 0.0967 |
| 21 | – | – | – | – | – | – | – | – | 0.0467 | – | 0.0033* | – | 0.0033* | – | 0.1667 |
| 21.2 | – | – | – | – | – | – | – | – | – | – | – | – | – | – | 0.0133* |
| 22 | – | – | – | – | – | – | – | – | 0.05 | – | – | – | 0.0033* | – | 0.1533 |
| 22.2 | – | – | – | – | – | – | – | – | – | – | – | – | – | – | 0.02 |
| 23 | – | – | – | – | – | – | – | – | 0.17 | – | – | – | – | – | 0.12 |
| 23.2 | – | – | – | – | – | – | – | – | – | – | – | – | – | – | 0.0033* |
| 24 | – | – | – | – | – | – | – | – | 0.1333 | – | – | – | – | – | 0.2 |
| 25 | – | – | – | – | – | – | – | – | 0.0633 | – | – | – | – | – | 0.13 |
| 26 | – | 0.0033* | – | – | – | – | – | – | 0.03 | – | – | – | – | – | 0.03 |
| 27 | – | 0.0133* | – | – | – | – | – | – | – | – | – | – | – | – | 0.0067* |
| 28 | – | 0.1533 | – | – | – | – | – | – | – | – | – | – | – | – | – |
| 29 | – | 0.2467 | – | – | – | – | – | – | – | – | – | – | – | – | – |
| 29.2 | – | 0.0067* | – | – | – | – | – | – | – | – | – | – | – | – | – |
| 30 | – | 0.1533 | – | – | – | – | – | – | – | – | – | – | – | – | – |
| 30.2 | – | 0.01* | – | – | – | – | – | – | – | – | – | – | – | – | – |
| 31 | – | 0.0333 | – | – | – | – | – | – | – | – | – | – | – | – | – |
| 31.2 | – | 0.1167 | – | – | – | – | – | – | – | – | – | – | – | – | – |
| 32 | – | 0.0133* | – | – | – | – | – | – | – | – | – | – | – | – | – |
| 32.2 | – | 0.1767 | – | – | – | – | – | – | – | – | – | – | – | – | – |
| 33.2 | – | 0.0633 | – | – | – | – | – | – | – | – | – | – | – | – | – |
| 34.2 | – | 0.01* | – | – | – | – | – | – | – | – | – | – | – | – | – |

**Table 2.** Statistical parameters of forensic value for 15 autosomal loci from the Saraiki population.

| | D8S1179 | D21S11 | D7S820 | CSF1P0 | D3S1358 | TH01 | D13S317 | D16S539 | D2S1338 | D19S433 | vWA | TPOX | D18S51 | D5S818 | FGA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| P | 0.429 | 0.6837 | 0.7523 | 0.1849 | 0.3564 | 0.0021 | 0.326 | 0.6885 | 0.5317 | 0.5617 | 0.4712 | 0.7116 | 0.3918 | 0.5492 | 0.1472 |
| $H_{OBS}$ | 0.833 | 0.833 | 0.767 | 0.7 | 0.8 | 0.68 | 0.853 | 0.74 | 0.893 | 0.845 | 0.813 | 0.613 | 0.8 | 0.707 | 0.787 |
| $H_{EXP}$ | 0.8411 | 0.8369 | 0.7939 | 0.7228 | 0.7379 | 0.7921 | 0.8065 | 0.7507 | 0.8774 | 0.8112 | 0.819 | 0.6622 | 0.8378 | 0.7383 | 0.8627 |
| PD | 0.953 | 0.952 | 0.93 | 0.875 | 0.877 | 0.924 | 0.926 | 0.91 | 0.968 | 0.933 | 0.938 | 0.833 | 0.956 | 0.893 | 0.964 |
| PE | 0.662 | 0.662 | 0.539 | 0.428 | 0.599 | 0.398 | 0.701 | 0.493 | 0.782 | 0.684 | 0.624 | 0.307 | 0.599 | 0.439 | 0.575 |
| PIC | 0.82 | 0.82 | 0.77 | 0.68 | 0.7 | 0.76 | 0.78 | 0.73 | 0.87 | 0.79 | 0.8 | 0.6 | 0.83 | 0.7 | 0.85 |
| MAF | 0.0166 | 0.0166 | 0.0166 | 0.0166 | 0.0166 | 0.0166 | 0.0166 | 0.0166 | 0.0166 | 0.0166 | 0.0166 | 0.0166 | 0.0166 | 0.0166 | 0.0166 |

P = P-value for HWE, $H_{obs}$ = Observed heterozygosity, $H_{exp}$ = Expected heterozygosity, PD = Power of Discrimination, PE = Probability of Exclusion, PIC = Polymorphism Information Content, MAF = Minimum Allowable Frequency.

P-values are greater than 0.05 at loci D21S11, D7S820, D16S539, D5S818 and TPOX, while the P-value at locus TH01 is less than 0.05 (Table 2). Forensic efficiency parameters indicate that locus TPOX gives the highest match probability (0.167); locus D2S1338 has the highest power of exclusion (0.782), highest typical paternity index value (2.60), highest polymorphism information content value (0.87) and highest power of discrimination (0.968). Locus D2S1338 showed the highest heterozygosity percentage of 89.3% and locus TPOX showed the highest homozygosity percentage of 38.7%, as shown in Table 2.

The highest heterozygosity of 89.3% was shown by locus D2S1338 and the lowest heterozygosity of 61.3% was shown at the TPOX locus. A low heterozygosity value reflects less genetic variability while high heterozygosity reflects high genetic variation among the individuals at that locus. Loci D3S1358, D13S317, D2S1338 and D19S433 showed observed heterozygosity values greater than expected heterozygosity values under HWE at these loci (Table 2), while other loci D8S1179, D21S11, D7S820, CSF1PO, TH01, D16S539, vWA, TPOX, D18S51, D5S818 and FGA showed observed heterozygosity lower than expected heterozygosity values (Table 2). The average heterozygosity at all loci was 0.77.

The frequencies of some alleles at 14 different loci were very low in this population (Table 1). Among these loci, six had more than one rare allele (the frequency of the allele is less than 0.016). Polymorphism information content (PIC) refers to the number of polymorphic forms of an allele in a population. As the PIC value increases, the more informative the respective panel becomes. Locus D2S1338 has the highest polymorphism information content value (0.87) and the highest power of discrimination (0.968). The PIC values at loci CSF1PO, D3S1358, D16S539, vWA, TPOX and D5S818 were low in this population, reflecting low discriminatory power compared with the other loci and the average PIC value observed at all loci was 0.766.

## Conclusion

The frequency database presented in this study can be used for reporting match probabilities in the Saraiki population of Pakistan, which provides forensic significance to DNA evidence. In comparison with Balochi, Brahui, Makrani, Hazara, Burusho, Sindhi and Pathan populations, the Saraiki population has different allele ranges [19]. The $F_{ST}$, which was calculated by generating a phylogenetic tree, also shows that this population is genetically distinct from other populations present in Pakistan. Owing to this discrete property of the allele frequencies of the Saraiki population, this population database can assist in identifying unknown persons as possibly belonging to the Saraiki population.

Except for locus TH01, the *P*-values at all loci were higher than the alpha value of 0.05 and were uniformly distributed between 0 and 1, while in the US Caucasian database [20] TH01, TPOX and D13S317 showed *P*-values lower than 0.05. The average observed heterozygosity ($H_{Obs}$) of 15 STR loci for the Saraiki population was 0.7776, which was close to the Caucasian population and Nepalese population, which are 0.7923 and 0.786 respectively [20, 21]. The average observed heterozygosity value ($H_{Obs}$) of 15 STR loci for the Sariaki population was also higher than that of the Brahui population ($H_{Obs} = 0.715$) and Balochi ($H_{Obs} = 0.756$) as mentioned in the STR database [19]. The average $H_{Obs}$ for the Saraiki population was less than the observed $H_{Obs}$ of the Makrani and Pathan populations, which are 0.782 and 0.789, respectively, while the average $H_{Obs}$ of the Sindhi and the Kalash populations were similar to that of the Saraiki population. Another study,

which conducted DNA typing and mapping of Asian populations, had an average heterozygosity (0.69) that was less than the Saraiki population results [22]. This population is in Hardy Weinberg equilibrium and can be used for forensic case work [1].

The efficiency of a microsatellite marker panel can be evaluated by the calculation of the Combined Power of Exclusion (CPE). The CPE of the present study was found to be 0.999998, which was higher than the Japanese (0.9991), African-American (0.9996), the US Caucasian (0.9994), Spanish (0.9998) and French-Canadian Caucasian populations (0.9874) [2, 23, 24].

The knowledge of the presence of genetic substructures in a population is necessary before using any population database for forensic case work. The data show that the Saraiki population is in Hardy-Weinberg equilibrium with few exceptions, and is genetically distinct from other Pakistani populations. The allelic frequency data from the Saraiki population can be used for forensic identity and paternity casework.

## Acknowledgements

I would like to express my deepest gratitude to all those individuals who volunteered for sample collection and to my instructors Ms Maryam Javed, Ms Tehseen Fatima and Ms Huma Mujahid. Without their cooperation, this population data collection, processing and interpretation would not have been possible.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## References

1. Butler JM. Forensic DNA typing. Biology, technology and genetic of STR markers. 2nd ed. NY, USA: Elsevier Academic Press; 2005.
2. Arce B, Heinrichs B, Armenteros MF, Carrasco F, Lorente JA, Budowle B. Spanish population data on nine STR loci. J Forensic Sci. 2001;46(4):1003−4.
3. Jeffreys AJ, Wilson V, Thein SL. Individual specific "Fingerprints" of human DNA. Nature. 1985;316(6023):76−9.
4. Buckleton JS, Triggs CM, Walsh SJ. Forensic DNA evidence interpretation. 1st ed. New York: CRC Press; 2004.
5. Mehdi SQ, Qamar R, Ayub Q, Khaliq S, Mansoor A, Ismail M, Hammer MF, Underhill PA, Cavalli-Sforza LL. Genomic diversity: applications in human population genetics. New York: Kluwer Academic/Plenum publishers; 1999. Chapter 7, The origins of Pakistani populations; 83−90.
6. Quddus SA. The tribal Baluchistan. 1st ed. Lahore, Pakistan: Ferozsons (Pvt.) Ltd; 1990.
7. Hughes-Buller R. Imperial gazetteer of India: provincial series: Calochistan. Lahore, Pakistan: Sang-e-Meel Publication; 1991.
8. Ansari SS A. The Musalman races found in Sindh, Balochistan and Afghanistan. 1st ed. Karachi, Pakistan: Indus Publications; 1996.
9. Burton RF. Sindh and the races that inhabit the valley of the Indus. London, UK: WH Aleen & Co. Ltd; 1851.
10. Grimes BF. Ethnologue: languages of the World. 1st ed. Dallas, Texas, USA: Summer Institute of Linguistics Inc; 1992.
11. Robertson GS. The Kafirs of the Hindu-Kush. 1st ed. Karachi and New York: Oxford University Press; 1974.

12. Caroe O. The Pathans. 1st ed. Karachi, Pakistan: Oxford University Press; 1992.
13. Bellew HW. The races of Afghanistan. 1st ed. Lahore, Pakistan: Sang-e-Meel Publications; 1979.
14. Ibbetson D. Punjab castes. 1st ed. Lahore, Pakistan: Sang-e-Meel Publications; 1994.
15. Fisher RA. Statistical methods for research workers. London: Oliver and Boyd; 1932.
16. Fox JC, Sangha J, Douglas EK, Schumm JW. New device and method for buccal cell collection and processing. Proceedings of the Thirteenth International Symposium on Human Identification; 2002.
17. Applied Biosystems. Quantifiler Human DNA Quantification Kit User's Manual. Foster City (CA): Applied Biosystems; 2003.
18. Raymond M, Rousset F. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Hered. 1995;86(3):248−9.
19. Amigo J, Phillips C, Salas T, Fernández Formoso L, Carracedo A, Lareu M. pop.STR - An online population frequency browser for established and new forensic STRs. Forensic Sci Int. 2009;2(1):361−2.
20. Hill CR, Duewer DL, Kline MC, Coble MD, Butler JM. U.S. population data for 29 autosomal STR loci. Forensic Sci Int Genet. 2013;7(3):82−3.
21. Jha DK, Gonzalez LJM., Rijal JP, Tuladhar BS, Chhetri NT. Allele frequencies of 15 Ampflstr Identifiler loci in the Nepalese population. Scientific World. 2012;10(10):20−3.
22. Edwards A, Civitello A, Hammond HA, Caskey CT. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. Am J Hum Genet. 1991;49(4):746−56.
23. Yamamoto T, Uchihi R, Nozawa H, Huang XL, Leong YK, Tanaka M, Mizutani M, Tamaki K, Katsumata Y. Allele distribution at nine STR loci D3S1358, vWA, FGA, TH01, TPOX, CSF1PO, D5S818, D13S317 and D7S820-in the Japanese population by multi-plex PCR and capillary electrophoresis. J Forensic Sci. 1999;44(1):167−70.
24. Busque L, Desmarais D, Provost S, Schumm JW, Zhong Y, Chakraborty R. Analysis of allele distribution for six short tandem repeat loci in the French Canadian population of Quebec. J Forensic Sci. 1997;42(6):1147−53.