

RESEARCH

Open Access



Short tandem repeat (STR) variation from 6 cities in Iraq based on 15 loci

Majeed A. Sabbah¹, Mohammed M. Al-Zubaidi^{1*} , Thooalnoon Y. Al-janabi¹, Dhuha S. Namaa¹, Haider K. Al-rubai¹ and Hala K. Ibrahim¹

Abstract

Background One thousand sixty-one individuals were sampled from the cities of Anbar, Baghdad, Basra, Diyala, Najaf, and Wasit in Iraq and typed for 15 forensic STRs to explore the genetic structure of Iraq and develop a forensic DNA database. The total number of alleles that were identified was 203.

Result Analyses of molecular variance (AMOVA) were then conducted Baghdad provides a good representation of the rest of the country, while Anbar is the most genetically distinct. The average heterozygosities of these loci was 0.779, homozygosities was 0.221, polymorphism information content was 0.77, power of discrimination was 0.927, and power of exclusion was 0.563. At these loci, a matching genotype will occur, on average, in 1 in 8.152 × 1017 individuals. For paternity tests, the average paternity probability for a matching profile is 99.9997%.

Conclusions These loci are appropriate for use in forensic and paternity testing for this population. Iraq is similar to other countries in the Middle East, particularly Iran and Turkey, and is more similar to Europe than either Asia or Africa.

Keywords STR, Iraq population, Forensic DNA, Baghdad

Background

Short tandem repeats, also known as microsatellites, are repeated sequences of DNA, usually consisting of 2 to 6 bases. They are highly polymorphic and distributed throughout the genome. These sequences vary in the number of repeats, resulting in them being multi-allelic [1]. Consequently, STRs prove very useful for forensic and paternity testing when multiple STR loci are employed [2]. The distribution of allele frequencies within the tested population enables the calculation of match probabilities. The objective of this study was to report allele frequencies for these STRs in the Iraqi population, facilitating their application in forensic

and paternity investigations. The population of Iraq was estimated at 39 million in July 2017 [3]. Arabs constitute 75–80% of the population, while Kurds make up 15–20% [3, 4]. Other minorities, including Turkmen, Assyrian, Shabak, Yazidi, and more, are also present [3]. Samples were collected from various Iraqi cities: Diyala, Anbar, Wasit, Najaf, Baghdad, and Basra. Among these, Baghdad, Diyala, and Najaf are situated in the center of Iraq; Wasit is in the east; Basra is in the southeast; and Al-Anbar is in the west [4] and subjected to STR profiling.

Methods

Ethics committee approval and patient approval

The collection of samples used in the research (blood, buccal swab, saliva, and fingernails) was approved by the Center's Scientific Research Ethics Committee before starting work. The volunteers signed a written consent to participate in this study.

*Correspondence:

Mohammed M. Al-Zubaidi
molecular_fdna@yahoo.com

¹ Forensic DNA Research and Training Center, Al-Nahrain University, Jadriya, Baghdad, Iraq

Sample collection

A total of 1061 unrelated individuals were sampled from six Iraqi cities: Diyala ($n = 139$), Anbar ($n = 132$), Wasit ($n = 120$), Najaf ($n = 119$), Baghdad ($n = 354$), and Basra ($n = 198$). The samples were collected as Buccal swabs from patients and laboratory workers at private laboratories. These samples were used to study the population genetic diversity in Iraq. The population genetics of Iraqis is important due to their ethnic diversity. Several studies have been conducted to analyze the genetic diversity of Iraqi populations, including studies on the distribution of Y chromosome haplotypes, 23-YSTR markers, and 15 STRs. These studies aimed to analyze the genetic structure of different cities in Iraq and to create a forensic DNA database for the country.

DNA extraction

Samples were extracted using a PrepFiler Forensic DNA Extraction Kit (Applied Biosystems, Foster City, CA), and their DNA content was quantified with NanoDrop (Thomson, Wilmington, DE).

PCR amplification

Fifteen autosomal STR markers (the 13 CODIS core loci and D19S433 and D2S1338) were genotyped along with the amelogenin locus on the X and Y chromosomes using the Applied Biosystems AmpFiSTR® Identifiler™ kit (3). Approximately 1 ng of template DNA was amplified for each sample following the protocols described in the user's manual (Applied Biosystems). The samples were amplified with an Applied Biosystems Veriti® PCR System (Applied Biosystems).

DNA typing

Amplification products were diluted 1:15 in Hi-Di™ formamide and GS500-LIZ internal size standard (Applied Biosystems) and analyzed on the 16-capillary ABI Prism® 3100 Genetic Analyzer. On a 36-cm array, POP™-4 (Applied Biosystems) was used for higher-resolution separations.

Data collection

Data collection was performed with Data Collection v. 2.0 software (Applied Biosystems), and samples were analyzed with GeneMapper v. 3.2 software (Applied Biosystems) at the Forensic DNA Center for Research and Training of Al-Nahrain University.

Statistical analyses

Allele frequencies for each locus, heterozygosities, homozygosities, polymorphic information anthropolgy, and content (PIC) measure the formativeness of a

genetic marker, indicating how well it can distinguish between different alleles in a population. $PIC = 1 - \sum (p_i^2)$ PIC ranges from 0 to 1, with higher values indicating greater allelic diversity and formativeness of the marker.

Matching probability (MP), powers of discrimination (PD) is a measure of how well a genetic marker can discriminate between individuals within a population.

P_i and p_j are the frequencies of the i -th and j -th alleles at the locus, summed over all possible pairwise allele combinations.

PD also ranges from 0 to 1, with higher values indicating better discriminatory power. $PD = \sum (p_i * p_j)$.

Powers of exclusion (PE) measures the probability that two individuals randomly chosen from a population will have different genotypes at a specific locus.

$PE = 1 - P$ (same genotype)

P (same genotype) is the probability that two individuals chosen at random will have the same genotype at the locus.

PE can also range from 0 to 1, with higher values indicating a higher probability of distinguishing between individuals.

And typical paternity index (TPI) was calculated at each locus using Power Stats v1.2 [5]. The significance value for the HWE test was set at $\alpha = 0.05$ and p values were calculated using Monte Carlo methods with 100,000 permutations. The Holm-Bonferroni method [6] was used to account for multiple testing. Multidimensional scaling (MDS) was done using the "stats" package (R Core Team, 2016) to compare the Iraqi population to other Middle Eastern countries as well as countries from Europe, Asia, and Africa [7].

Results

A total of 203 alleles were identified in this study. The distribution of these alleles across the 15 short tandem repeat (STR) loci is detailed in Table 1. Notably, notable occurrences include allele 8 of TPOX, which had the highest frequency at 51.0%, allele 12 of CSF1PO at 32.9%, allele 12 of D5S818 at 31.9%, allele 12 of D5S818 at 30.8%, allele 11 of CSF1PO at 30.7%, allele 11 of D16S539 at 30.6%, and allele 12 of D13S317 at 30.4%. It is noteworthy that the most diverse loci in terms of the number of distinct alleles were FGA with 23 alleles, D18S51 with 20 alleles, D2S1338 with 19 alleles, D19S433 with 17 alleles, and D21S11 with 17 alleles.

The city of Baghdad boasted the largest sample size, comprising 354 individuals. Interestingly, five alleles were exclusive to Baghdad and were not found in the other cities or neighboring countries. Three of these five alleles were absent from the other five Iraqi cities and all countries used for comparison, including Turkey, Iran

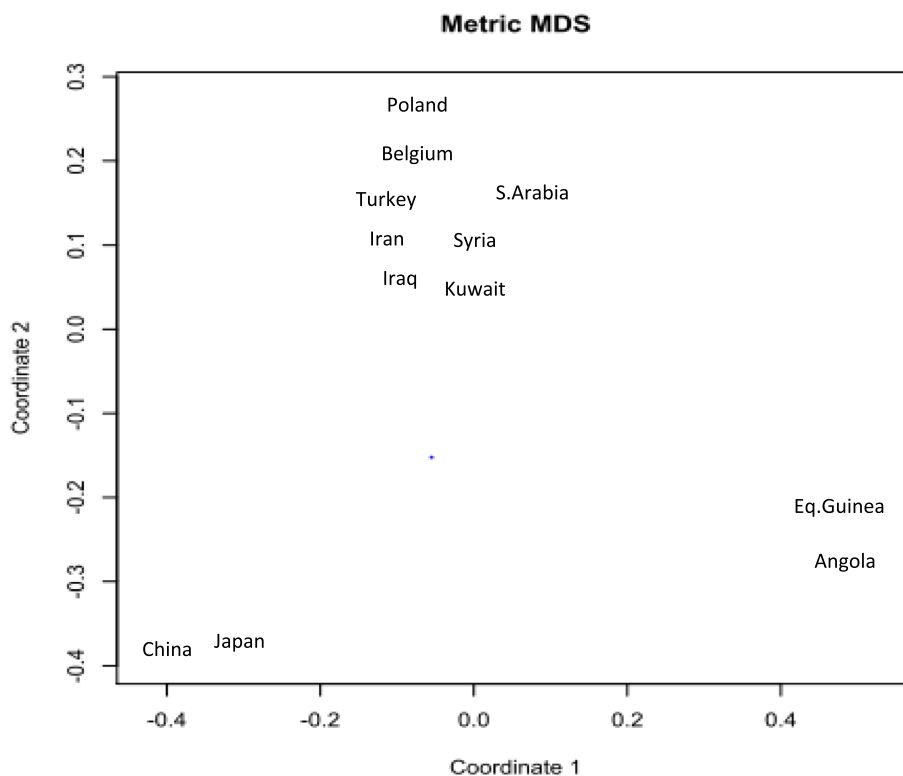


Fig. 1 Multidimensional scaling plot of the genetic distances between Iraq, Turkey, Iran, Syria, Kuwait, Saudi Arabia, Poland, Belgium, China, Japan, Equatorial Guinea, and Angola

[8], Syria [7], Kuwait [9], Saudi Arabia [10], Poland, Belgium [11], China [12], Japan [13], Equatorial Guinea, and Angola [14] (Fig. 1). Specifically, these alleles were allele 8 at vWA (found in Baghdad at 0.3%), allele 25 at vWA (found in Baghdad at 0.1%), and allele 16 at FGA (found in Baghdad at 0.3%). Baghdad also exhibited two other rare alleles: allele 11 at vWA (0.3% occurrence) and allele 15 at FGA (0.1% prevalence). Importantly, these alleles were absent in other Iraqi cities, Middle Eastern nations, and the European and Asian countries used for comparison [15]. However, they were present in the African countries used for comparison, such as Equatorial Guinea and Angola [16, 17]. Specifically, allele 11 at vWA was observed at 3.5% frequency in Equatorial Guinea and 0.4% in Angola, while allele 15 at FGA was found in Equatorial Guinea at 0.4%.

An additional rare allele, allele 22 at D13S317, was identified in the city of Wasit, with a frequency of 0.4%, but was not observed in any other Iraqi cities or countries in the comparison set [7].

To assess whether the observed genotype frequencies adhere to the expected frequencies, the population's adherence to the Hardy-Weinberg equilibrium (HWE) was examined [18]. Table 2 displays the log-likelihood

ratio p values for the HWE test [19, 20], which, after Holm-Bonferroni adjustment for multiple testing, were not deemed significant [21].

Details concerning each locus are provided in Table 3, encompassing heterozygosities (He), homozygosities (Ho), polymorphism information content (PIC), matching probabilities (MP), powers of discrimination (PD), powers of exclusion (PE), and the typical paternity index (TPI). All values fall within a range of 0.0 to 1.0, where 0.0 denotes the absence (or presence) of heterozygotes (or homozygotes) and 1.0 indicates full heterozygosity (or homozygosity) across the sampled individuals. The significance of these metrics varies with the genotype; high He, low Ho, substantial PIC, PD, and PE values characterize loci pertinent to forensic and paternity analyses. The average values for the 15 loci are He = 0.791, Ho = 0.209, PIC = 0.72, PD = 0.923, and PE = 0.587. Composite metrics, namely composite matching probability (CMP) and composite paternity index (CPI), are obtained by multiplying each locus's MP and TPI, respectively.

Exploring STR analysis within Iraqi populations has been somewhat limited, and the studies that have been conducted employed distinct commercial STR kits, posing challenges for direct comparisons with the current

Table 2 Hardy-Weinberg Equilibrium p -values (for each five city) at each locus

| Loci | Diyala | Anbar | Wasit | Najaf | Baghdad | South of Iraq |
|-------------|--------|-------|-------|-------|---------|---------------|
| D8S.1179 | 0.97 | 0.94 | 0.69 | 0.07 | 0.15 | 0.56 |
| D2.1S11 | 0.19 | 0.06 | 0.05 | 0.49 | 0.02 | 0.38 |
| D7.S820 | 0.72 | 0.70 | 0.55 | 0.91 | 0.19 | 0.87 |
| CS.F1PO | 0.09 | 0.66 | 0.77 | 0.79 | 0.16 | 0.35 |
| D.3S1358 | 0.04 | 0.95 | 0.10 | 0.37 | 0.08 | 0.07 |
| TH.01 | 0.84 | 0.14 | 0.09 | 0.005 | 0.18 | 0.61 |
| D13.S317 | 0.47 | 0.40 | 0.17 | 0.25 | 0.08 | 0.84 |
| D16.S539 | 0.01 | 0.19 | 0.21 | 0.83 | 0.03 | 0.17 |
| D2S.1338 | 0.03 | 0.64 | 0.98 | 0.37 | 0.001 | 0.07 |
| D19.S433 | 0.45 | 0.40 | 0.77 | 0.54 | 0.68 | 0.64 |
| vWA | 0.22 | 0.53 | 0.93 | 0.65 | 0.11 | 0.39 |
| TPOX | 0.31.0 | 0.12 | 0.07 | 0.78 | 0.01 | 0.52 |
| D1.8S51 | 0.10 | 0.82 | 0.39 | 0.56 | 0.02 | 0.87 |
| D.5S818 | 0.99 | 0.08 | 0.50 | 0.53 | 0.80 | 0.12 |
| FG.A | 0.26 | 0.82 | 0.48 | 0.04 | 0.01 | 0.15 |
| Sample size | 139 | 132 | 120 | 119 | 354 | 198 |

Table 3 Forensic efficiency parameters for 15 STR loci (1061 samples) included matching probabilities, powers of discrimination, polymorphism information content, powers of exclusion, a typical paternity index, homozygosities, and heterozygosities

| | MP* | PD* | PIC* | PE* | TPI* | Ho* | He* |
|-------------|-------|-------|------|-------|------|-------|-------|
| D8S1179 | 0.052 | 0.948 | 0.80 | 0.619 | 2.64 | 0.189 | 0.811 |
| D21S11 | 0.048 | 0.952 | 0.82 | 0.647 | 2.87 | 0.174 | 0.826 |
| D7S820 | 0.072 | 0.928 | 0.76 | 0.561 | 2.27 | 0.221 | 0.779 |
| CSF1PO | 0.130 | 0.870 | 0.67 | 0.439 | 1.71 | 0.293 | 0.707 |
| D3S1358 | 0.091 | 0.909 | 0.73 | 0.513 | 2.02 | 0.248 | 0.752 |
| TH01 | 0.076 | 0.924 | 0.76 | 0.531 | 2.11 | 0.238 | 0.762 |
| D13S317 | 0.790 | 0.921 | 0.75 | 0.570 | 2.32 | 0.216 | 0.784 |
| D16S539 | 0.075 | 0.925 | 0.76 | 0.543 | 2.17 | 0.231 | 0.769 |
| D2S1338 | 0.026 | 0.974 | 0.87 | 0.669 | 3.07 | 0.163 | 0.837 |
| D19S433 | 0.045 | 0.955 | 0.82 | 0.629 | 2.72 | 0.184 | 0.816 |
| vWA | 0.068 | 0.932 | 0.77 | 0.520 | 2.05 | 0.244 | 0.456 |
| TPOX | 0.165 | 0.835 | 0.61 | 0.384 | 1.52 | 0.330 | 0.670 |
| D18S51 | 0.029 | 0.971 | 0.86 | 0.646 | 2.85 | 0.175 | 0.825 |
| D5S818 | 0.102 | 0.898 | 0.71 | 0.497 | 1.94 | 0.257 | 0.743 |
| FGA | 0.033 | 0.967 | 0.85 | 0.682 | 3.19 | 0.157 | 0.843 |
| Sample size | 1061 | 1061 | 1061 | 1061 | 1061 | 1061 | 1061 |

*MP matching probabilities, PD powers of discrimination, PIC polymorphism information content, PE powers of exclusion, TPI typical paternity index, Ho homozygosities, and He heterozygosities

study's outcomes. Nevertheless, there is an overlap of the 15 STR markers between the aforementioned studies and neighboring countries, as outlined in Table 4.

Among these 15 common STR markers, D2S1338 demonstrated the highest Polymorphic Information Content (PIC) value within this study, while FGA exhibited the highest PIC value among Jordanians,

succeeded by D19S433 in Turkey, D18S51 and FGA in Palestinians [25], and D19S433 and FGA in Saudi Arabia [27]. Conversely, the TPOX locus presented the lowest PIC value in this study, and within Jordanians [22], D13S317 displayed the lowest PIC value, followed by TPOX in Turkey [23], Palestinians, Saudi Arabia [27], and Iran [26].

Table 4 Comparison *p* values of HWE for STRs data for Arab-related populations

| Locus | Iraq in this study | Jordan [22] | Turkey [23, 24] | Palestinian in Gazan [25] | Saudi Arabiae [10] | Iran [26] |
|---------|--------------------|-------------|-----------------|---------------------------|--------------------|-----------|
| | 1061 | 95 | 350 | 125 | 190 | 274 |
| D8S1179 | 0.80 | 0.774 | 0.814 | 0.83 | 0.81 | 0.81 |
| D21S11 | 0.82 | 0.793 | 0.796 | 0.82 | 0.81 | 0.83 |
| D7S820 | 0.76 | 0.694 | 0.775 | 0.77 | 0.75 | 0.78 |
| CSF1PO | 0.67 | 0.745 | 0.701 | 0.66 | 0.66 | 0.65 |
| D3S1358 | 0.73 | 0.722 | 0.759 | 0.71 | 0.73 | 0.73 |
| TH01 | 0.76 | 0.734 | 0.780 | 0.74 | 0.76 | 0.76 |
| D13S317 | 0.75 | 0.654 | 0.750 | 0.73 | 0.72 | 0.75 |
| D16S539 | 0.76 | 0.787 | 0.754 | 0.76 | 0.72 | 0.78 |
| D2S1338 | 0.87 | NS | 0.877 | NS | 0.84 | 0.87 |
| D19S433 | 0.82 | NS | 0.881 | NS | 0.85 | 0.82 |
| vWA | 0.77 | 0.802 | 0.794 | 0.78 | 0.78 | 0.78 |
| TPOX | 0.61 | 0.706 | 0.656 | 0.65 | 0.57 | 0.63 |
| D18S51 | 0.86 | 0.878 | 0.864 | 0.86 | 0.82 | 0.87 |
| D5S818 | 0.71 | 0.672 | 0.703 | 0.72 | 0.71 | 0.72 |
| FGA | 0.85 | 0.893 | 0.849 | 0.86 | 0.85 | 0.84 |

NS not significant

(-): not typed loci

Discussion

In forensic investigations, CMP is frequently expressed as the likelihood that one individual in a certain population subset possesses a genotype matching the composite value. With the global population approximately at 7.5 billion, an average MP of 0.173 or lower across all 15 loci is necessary for the CMP to be interpreted as 1 in 7.5 billion or greater. The dataset's mean MP across 15 loci is 0.073. In paternity reports, the probability of paternity is determined by $CPI/(CPI + 1)$, where a CPI of 100 or higher indicates a probability of 99.0% or higher. For the average TPI to reach 1.43 for 15 loci, the average TPI over 1.5 locations was 2.36 [20].

The findings of this investigation highlight the utility of these 15 autosomal STR loci as valuable markers for forensic and paternity testing within the Iraqi population. Within certain loci, the presence of a wide range of repeated sequences suggests possible population admixture dynamics. To elucidate the origins of the less common alleles discovered in this study, it may be beneficial to conduct further allele frequency analyses within specific ethnic groups in Iraq.

Lastly, the comparative evaluation of forensic genetic efficiency parameters across diverse populations, including the Middle Eastern region, is vital for comprehending genetic diversity and potential intermingling among these populations. To achieve a more accurate comparison, it is strongly recommended to conduct comparative studies employing the same STR

kit across these populations, considering that different STR kits have been used previously to define forensic efficiency parameters.

Consistency and variation

Looking at the values within each population, you can observe that some alleles have consistent frequencies across populations, while others vary more. This variation could be due to a variety of factors, including historical migrations, genetic drift, and natural selection.

Genetic diversity

The range of allele frequencies across the different populations indicates the genetic diversity present in these regions. Higher diversity might be indicative of a more mixed or heterogeneous population.

The findings from this study highlight the effectiveness of the 15 autosomal STR loci as markers for forensics and paternity testing within the Iraqi population. It is worth noting that certain loci exhibit a wide range of repeats, which may be indicative of population admixture. To gain insight into the origins of the less common alleles identified in this study, further investigations should focus on analyzing allele frequencies within specific ethnic groups in Iraq.

A comprehensive assessment of forensic genetic efficiency parameters is crucial for understanding genetic diversity and admixture among different populations, including those in the Middle East. To ensure a robust

comparison, it is strongly recommended that comparative studies utilize the same STR kit across these populations. Previous studies have employed disparate STR kits, which has affected the description of forensic efficiency parameters.

Conclusion

The study discussed in the given text highlights the effectiveness of 15 autosomal STR loci as markers for forensics and paternity testing within the Iraqi population. The CMP represents the likelihood that an individual in a specific population subset possesses a genotype matching a composite value. With an average Match Probability (MP) of 0.073 across all 15 loci in a global population of 7.5 billion, the CMP can be interpreted as 1 in 7.5 billion or greater. The presence of a wide range of repeated sequences within certain loci suggests possible population admixture dynamics. To understand the origins of less common alleles, further allele frequency analyses within specific ethnic groups in Iraq are recommended. Conducting comparative evaluations of forensic genetic efficiency parameters across diverse populations, including those in the Middle East, is essential for comprehending genetic diversity and potential intermingling among these populations. Using the same STR kit for these studies is crucial for accurate comparisons. The study also highlights that some alleles have consistent frequencies across populations, while others exhibit more variation, which could be attributed to factors such as historical migrations, genetic drift, and natural selection.

Abbreviations

| | |
|-------|---------------------------------|
| STR | Short tandem repeat |
| AMOVA | Analysis of molecular variance |
| CODIS | Combined DNA index system |
| PIC | Polymorphic Information Content |
| MP | Matching probability |
| PD | Power of discrimination |
| PE | Probability of exclusion |
| PI | Paternity index |
| HWE | Hardy-Weinberg Equilibrium |

Acknowledgements

We extend our thanks to everyone who agreed to give a sample and contribute to the completion of this research. We would also like to thank Al-Nahrain University, the Forensic DNA Center, represented by the director of the center, who expressed his approval to complete this research. Thanks are extended to the researchers who participated in the completion of this research.

Authors' contributions

Mohammed M. Al-Zubaidi was responsible for the design, supervision, and preparation of the manuscript. Majeed A. Sabbah contributed to the planning, sampling, and statistical analysis. Thooalnoon Y. Al-Janabi reviewed and edited the manuscript. Dhuha S. Namaa, Haider K. Al-rubai, and Hala K. Ibrahim are responsible for sample collection, demographic data, and lab work. All authors read and approved the final manuscript.

Funding

Not applicable.

Availability of data and materials

All data analyzed during this study are included in this article.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The author declares that they have no competing interests.

Received: 22 November 2022 Accepted: 26 October 2023

Published online: 05 December 2023

References

- Ellegren H (2004) Microsatellites: simple sequences with complex evolution. *Nat Rev Genet* 5:435–445
- Awad EO, Alsafar H, Tay GK, Theyab JBJM, Mubasher M, Sheikh NE-E, AlHarthi H, Crawford MH, El Ghazali G (2015) Autosomal short tandem repeat (STR) variation based on 15 loci in a population from the central region (Riyadh Province) of Saudi Arabia. *Forensic Res* 6:1
- World Factbook (2017) Central intelligence agency, Washington, DC, p 2017. <https://www.cia.gov/the-world-factbook/about/cover-gallery/2017-cover/>
- Kirmanj S (2013) Identity and Nation in Iraq. Lynne Rienner Publishers, Inc: Boulder, CO
- Tereba A (1999) Tools for analysis of population statistics. Profiles in DNA 2(3):14–16 PowerStats version 1.2, Promega Corporation
- Holm S (1979) A simple sequentially rejective multiple test procedure. *Scand J Stat* 6(2):65–70 <https://www.cia.gov/library/publications/the-world-factbook/geos/iz.html>
- Abdin L, Shimada I, Brinkmann B, Hohoff C (2003) Analysis of 15 short tandem repeats reveals significant differences between the Arabian populations from Morocco and Syria. *Leg Med* 5:S150–S155
- Avati MK, Akbari MT (2018) Allele frequency of 15 autosomal short tandem repeat loci in Iranian population with comparison to some other population. *J Human Gen Genome* 2(2):e87127
- Alenzi M, Goodwin W, Ismael S, Hadi S (2008) STR data for the AmpFSTR® Identifiler® loci in Kuwaiti population. *Leg Med* 10(6):321–325
- Al Idrissi E, Crawford MH, El Ghazali G (2015) Autosomal short tandem repeat (STR) variation based on 15 loci in a population from the central region (Riyadh Province) of Saudi Arabia. *J Forensic Sci* 6:267
- Decorte R, Gilissen A, Cassiman JJ (2003) Allele frequency data for 15 STR loci (AMPFISTR®SGM plus™ and AmpFISTR® Profiler™) in the Belgian population. *Int Congr Ser* 1239:219–222
- Wang ZY, Yu RJ, Wang F, Li XS, Jin TB (2005) Genetic polymorphisms of 15 STR loci in Han population from Shaanxi (NW China). *Forensic Sci Int* 147(1):89–91
- Hashiyada M (2000) Short tandem repeat analysis in Japanese population. *Electrophoresis* 21(2):347–350
- Babiker HMA, Schlebusch CM, Hassan HY, Jakobsson M (2011) Genetic variation and population structure of Sudanese populations as indicated by 15 Identifiler sequence-tagged repeat (STR) loci. *Investig Genet* 2(12)
- Szczerkowska Z, Kapińska E, Wycocka J, Cybulska L (2004) Northern Polish population data and forensic usefulness of 15 autosomal STR loci. *Forensic Sci Int* 144(1):69–71
- Alves C, Gusmão L, López-Parra AM, Mesa MS, Amorim A, Arruro-Pardo E (2005) STR allelic frequencies for an African population sample (Equatorial Guinea) using AmpFISTR Identifiler and Powerplex 16 kits. *Forensic Sci Int* 148(2-3):239–242
- Beleza S, Alves C, Reis F, Amorim A, Carracedo A, Gusmão L (2005) 17 STR data (AmpF/STR Identifiler and Powerplex 16 System) from Cabinda (Angola). *Forensic Sci Int* 141(2-3):193–196

18. Hardy HG (1908) Mendelian proportions in a mixed population. *Science* 28:49–50
19. Castle WE (1903) The laws of Galton and Mendel and some laws governing race improvement by selection. *Proc Am Acad Arts Sci* 35:233–242
20. Edwards AW, Hardy GH (2008) Anecdotal, Historical and Critical Commentaries on Genetics. *Genetics Society of America* 179:1143–1150
21. Emel HY, Yonar FC, Karatas O, Rayimoglu G, Engels WR (2009) Exact tests for Hardy-Weinberg proportions. *Genetics* 183(4):1431–1441 <https://CRAN.R-project.org/package=HWxtest>
22. Yasin SR, Hamad MM, Elkarmi AZ, Jaran J AS (2005) African Jordanian population genetic database on fifteen short tandem repeat genetic loci. *Croat Med J* 46(4):587–92
23. Ascioglu F, Canpolat E, Ozturk O, Erkan I (2023) Population Data and Internal Validation of the 21 Short Tandem Repeat Loci in Turkish Population. *Pak J Zool* 55(2):501–512
24. Yavuz I, Sarikaya AT (2005) Turkish Population Data for 15 STR Loci by Multiplex PCR. *J Forensic Sci* 50(3):737–738
25. Halimah MSA (2009) Genetic Variation of 15 Autosomal Short Tandem Repeat (STR) Loci in the Palestinian Population of Gaza Strip. *Leg Med* 11(4):203–204
26. Shepard EM, Herrera RJ (2006) Iranian STR variation at the fringes of biogeographical demarcation. *Forensic Sci Int* 158(2-3):140–148
27. Osman AE, Alsafar H, Tay GK, JBJM T, Mubasher M, Eltayeb-El Sheikh N, Al Harthi H, R Core Team (2016) R: A language and environment for statistical computer. R Foundation for Statistical Computing, Vienna Austria <https://www.R-project.org/>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
