# Allele Frequencies of 15 Short Tandem Repeats (STR) Loci in Libyan Human Population

**Ashrf Aboaleid ***
**Othman El-Ansari ****
**Mohamed A. Al-Griw****

## Abstract:

The present work is to describe STR markers polymorphism with some parameters of forensic interest in the general Libyan population in order to contribute to the analysis of its genetic diversity for forensic purposes. Genotype and allele frequencies distribution for 15 polymerase chain reaction (PCR)-based loci (D3S1358, THO1, D21S11, D18S51, Penta E, D5S818, D13S317, D7S820, D16S539, CSF1PO, Penta D, VWA, D8S1179, TPOX, and FGA) were determined for a Libyan population. The 15 loci amplified by the PowerPlex® 16 HS System (Promega). The results showed that the allele frequencies ranged from 0.0033 for the rarest alleles found at the most studied loci to 0.463 for the most common allele (allele 8 at the TPOX locus). The study also showed 151 different alleles in Libyan population. Conventional statistics between-population comparison has been performed.

**Keywords:** STRs; Population data; PowerPlex® 16 HS System; Forensic parameters.

*Technical department for identification of missing persons Lab, Ministry of Martyrs and Missing Persons, Tripoli, Libya

** Develomental Biology Division, Zoology Department, Faculty of Science, University of Tripoli, Tripoli, Libya

## Introduction:

Short tandem repeats (STRs) are currently the most commonly analyzed genetic polymorphism in forensic genetics (Butler, 2012).

They were introduced into casework in the mid-1990s and are now the main tool for just about every forensic laboratory in the world the vast majority of forensic genetic casework involves the analysis of STR polymorphisms (Gill *et al*., 2004). There are thousands of STRs that can potentially be used for forensic analysis. STR loci are spread throughout the genome including the 22 autosomal chromosomes and the X and Y sex chromosomes. They have a core unit of between 1 and 6 bp and the repeats typically range from 50 to 300 base pare. The majority of the loci that are used in forensic genetics are tetranucleotide repeats, which have a four base pair repeat motif.

The use of polymorphic short tandem repeats (STRs) has become important in genetics applications, such as gene mapping, identification and paternity. In forensics, STRs are mainly used for paternity testing and personal identification.

Genotype and allele frequencies distribution for 15 autosomal genetic loci were determined for a Libyan population sample of 150 unrelated individuals (300 alleles for each locus). Conventional statistics between population comparison have been performed. The highest frequent allele was allele 8 at the TPOX locus (46.3%), moreover the rarest alleles found at most loci studied were frequent 0.33% in studied population.

## Materials and methods:

### Sample preparation:

Blood samples from 90 unrelated male individuals were deposited on FTA paper and directly amplified without quantification using the PowerPlex® 16 HS System (Promega) kit. A classical phenol/chloroform-based DNA extraction protocol

is used for 60 bone samples were collected from mass graves in Libya. In the International Commission on Missing persons lab (ICMP) the Extracted DNA from 60 bone samples quantified by real-time PCR (Rt-PCR) using the Plexor® HY system (Promega). Separation and detection of the alleles for blood and bone samples were carried out using capillary electrophoresis on an ABI PRISM 310 Genetic analyzer instrument (Applied Biosystems).

**Statistical analysis:**

The calculating of allele frequencies based on the number of the detected alleles for each specific locus, the genotyping information was then converted into allele frequencies (Microsoft Excel) by counting the number of times each allele was observed in all samples. To assess departures from Hardy-Weinberg Equilibrium (HWE) observed heterozygosity (Ho) and expected heterozygosity (He), forensic and population genetic parameters including power of discrimination (PD), polymorphic Information Content (PIC), matching probability (MP), typical paternity index (TPI), and probability of exclusion (PE).

**Results:**

In this study, the genotype and allele frequencies distribution for 15 PCR-based loci (D3S1358, THO1, D21S11, D18S51, Penta E, D5S818, D13S317, D7S820, D16S539, CSF1PO, Penta D, VWA, D8S1179, TPOX and FGA) were determined for Libyan population samples of about 150 unrelated individuals.

The allele frequencies ranged from 0.0033 (33%) for the rarest alleles found at most loci studied to 0.463 (46.3%) for the most common allele (allele 8 at the TPOX locus) (Table 1). And conventional statistics have been compared and performed between population. Statistical forensic parameters are

importance. The PD, observed and expected heterozygosity values, homozygosity (h), PIC, PM, PE and PI were calculated for the loci (Table 2).

These parameters indicated the usefulness of the loci in forensic personal identification and paternity testing among Libyan population in this study.

**General Discussion**:

In this study, we found that the allele frequencies ranged from 0.003 (0.33%) for the rarest alleles found at most loci studied to 0.463 (46.3%) for the most common allele (allele 8 at the TPOX locus). The level of heterozygosity was found to be high with average 0.79 (79%), the lowest observed heterozygotes frequency being 0.68 at the TPOX locus and the highest being 0.90 for the Penta E locus pointing to a random mating population. The results also showed that the frequency of observed homozygotes was low across all studied loci. Specifically, Penta E locus showing the lowest frequency (0.096) and locus CSF1PO having the highest frequency of homozygotes (0.279) with average 0.20 (20%), indicating that inbreeding is insignificant in the study population. Inbreeding reduces the frequency of heterozygous genotypes while increasing the homozygous genotypes as opposed to random mating (Jeffrey's *et al*., 1985).

This is because inbreeding results in an excess of homozygotes while random mating leads to an excess of heterozygotes. Table 1 where the heterozygotes frequencies are above 89%, indicating that mating is random.

It is well known that the possibility to find two persons with the same DNA profile if chosen at random in a population is defined as the matching possibility (MP) (Goodwin *et al*., 2007).

The results of this study revealed that the MP was low, ranging from 0.0216 in Penta E locus to 0.136 in CSF1PO locus.

The lower of MP indicating that the lower of chances of picking someone whose alleles are the same with someone else (Linacre, 2001). The higher of power of discrimination (PD) that mean the higher of chances that someone will be picked randomly whose alleles are different from the first person (Goodwin *et al*., 2007). As shown earlier (Table 1), the PD values imply that chances of picking out two people randomly that are genetically different are high.

There is less chance of a random sample matching because a higher heterozygosity means that more allele diversity would occur. For a given population, the quality of a polymorphism as a genetic marker was measured by the heterozygosity, the frequency

This indicates the probability that a given offspring of a parent carrying a rare allele at a locus will allow deduction of the parental genotype at the locus. This is determined by adding the mating frequencies multiplied by the probability that an offspring will be formed.

The TPOX locus is the least polymorphic marker while Penta E is the most polymorphic marker based on the degree of polymorphism of every marker, expressed in heterozygosity and PIC terms (Table 1). The usefulness of the findings for genetic polymorphism studies and linkage mapping programs in humans is confirmed by the high PIC values of the selected markers.

The PE can be computed to demonstrate how rare it is to find a random man who could not be excluded as the biological father of the child. The PE for every locus in this study was calculated and tabled as planned. For all the microsatellites analyzed the PE ranged from 0.438 (TPOX) to 0.775 (D18S51), with an average of 0.583.

The values PD for all tested loci were 91.4% for the THO and D3S1358 loci and 96.5%, 97.8%, 90.7%, and 91.5%, for D18S51, Penta E, D5S818, and D13S317 loci, respectively and

91.8%, 86.3%, 96.3, and 93.6% for the D7S820, CSF1PO, Penta D, and VWA loci respectively, and 95.3%, 85.7%, 96.6%, and 92% for D8S1179, TPOX, FGA, and D21S11 loci respectively, and ranged from 85.7% to 97.8%, for the rest of the loci.

Based on statistical parameters, this study found that Libyan population of may use these 15 STRs loci as a vital tool for forensic personal identification and paternity testing.

Consistent with prior study (Immel *et al*., 2006) by using same kit that was used in our study showed that the allele frequencies ranged from 0.005 (0.5%) for the rarest alleles found at most loci studied to 0.500 (50%) for the most common allele (allele 8 at the TPOX locus). Moreover, another prior study (Houssein *et al*., 2001) by using AmpFISTR Identifiler (Applied Biosystems), showed that the allele frequencies ranged from 0.005 (0.5%) for the rarest alleles found at most loci studied to 0.434 (43.4%) for the most common allele (allele 8 at the TPOX locus).

Forensic parameters for the 15 studied loci are the PD and the PE for the 15 studied loci were higher than 0.999, observed and expected heterozygosity were ranged from 0.657 - 0.889 and 0.716 - 0.843 respectively.

In other studies, the most frequent allele for forensic STR loci can vary depending on the population, for example 12 (0.4093) at D5S818 in Slovenian population (Drobnic et al., 2005), 15 (0.4143) at D3S1358 in Bolivians (Cifuentes et al., 2008), 8 (0.4890) at TPOX among Wallachians in South Romania (Stanciu *et al*., 2009) and 8 (0.424) at TPOX in Adaima community from Egypt (Coudray et al., 2007.

**Table 1. Frequencies of 15 STR loci in 150 unrelated Libyan samples (frequent of 151 alleles for 15 loci in Libyan population).**

| Allele | D3S1358 | TH01 | D18S51 | Penta E | D5S818 | D13S317 | D7S820 | D16S539 | CSF1P0 | Penta D | VWA | D8S1179 | TPOX | FGA | D21S11 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2.2 | - | - | - | - | - | - | - | - | - | 0.076 | - | - | - | - | - |
| 3.2 | - | - | - | - | - | - | - | - | - | 0.003 | - | - | - | - | - |
| 5 | - | - | - | 0.036 | - | - | - | - | - | - | - | - | - | - | - |
| 6 | - | 0.220 | - | - | - | - | - | - | 0.0033 | - | - | - | 0.003 | - | - |
| 7 | - | 0.260 | - | 0.100 | - | 0.003 | 0.0166 | - | 0.0067 | 0.003 | - | - | 0.003 | - | - |
| 8 | - | 0.140 | - | 0.073 | 0.043 | 0.130 | 0.1133 | 0.030 | 0.016 | 0.036 | - | 0.013 | 0.463 | - | - |
| 9 | - | 0.263 | 0.003 | - | 0.06 | 0.05 | 0.133 | 0.130 | 0.016 | 0.183 | - | 0.016 | 0.156 | - | - |
| 9.3 | - | 0.093 | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 10 | - | 0.023 | - | 0.063 | 0.100 | 0.073 | 0.320 | 0.063 | 0.313 | 0.133 | - | 0.053 | 0.080 | - | - |
| 11 | - | - | 0.010 | 0.150 | 0.233 | 0.266 | 0.253 | 0.35 | 0.313 | 0.116 | - | 0.136 | 0.253 | - | - |
| 12 | - | - | 0.120 | 0.160 | 0.363 | 0.326 | 0.133 | 0.226 | 0.283 | 0.086 | - | 0.16 | 0.040 | - | - |
| 13 | - | - | 0.210 | 0.1166 | 0.18 | 0.123 | 0.023 | 0.166 | 0.043 | 0.216 | - | 0.253 | - | - | - |
| 14 | 0.090 | - | 0.110 | 0.060 | 0.013 | 0.026 | 0.006 | 0.033 | 0.003 | 0.103 | 0.100 | 0.173 | - | - | - |
| 15 | 0.263 | - | 0.133 | 0.057 | 0.006 | - | - | - | - | 0.033 | 0.153 | 0.12 | - | - | - |
| 16 | 0.257 | - | 0.140 | 0.046 | - | - | - | - | - | 0.003 | 0.223 | 0.056 | - | - | - |
| 17 | 0.260 | - | 0.083 | 0.053 | - | - | - | - | - | 0.003 | 0.233 | 0.01 | - | 0.003 | - |
| 18 | 0.113 | - | 0.090 | 0.030 | - | - | - | - | - | - | 0.190 | - | - | 0.010 | - |
| 18.2 | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.006 | - |
| 19 | 0.017 | - | 0.040 | 0.036 | - | - | - | - | - | - | 0.073 | - | - | 0.090 | - |
| 19.2 | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.003 | - |
| 20 | - | - | 0.030 | 0.006 | - | - | - | - | - | - | 0.020 | - | - | 0.066 | - |
| 21 | - | - | 0.016 | - | - | - | - | - | - | - | 0.003 | - | - | 0.130 | - |
| 21.2 | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.006 | - |
| 22 | - | - | 0.006 | 0.006 | - | - | - | - | - | - | 0.003 | - | - | 0.170 | - |
| 22.2 | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.003 | - |
| 23 | - | - | 0.0-06 | - | - | - | - | - | - | - | - | - | - | 0.226 | - |
| 23.2 | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.006 | - |
| 24 | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.126 | - |
| 25 | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.086 | - |
| 25.2 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.003 |
| 26 | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.030 | - |
| 27 | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.020 | 0.030 |
| 28 | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.010 | 0.116 |
| 29 | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.006 | 0.270 |
| 30 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.006 |
| 30.2 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.046 |
| 31 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.066 |
| 31.2 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.003 |
| 32 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.003 |
| 32.2 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.063 |
| 33 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.003 |
| 33.2 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.066 |
| 34.2 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.010 |
| 35 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.006 |
| 35.2 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.003 |

**Table 2. Statistical parameters of the 15 STR loci for forensic interest (n = 150)**

| Allele | D3S1358 | TH01 | D18S51 | Penta E | D5S818 | D13S317 | D7S820 | D16S539 | CSF1P0 | Penta D | VWA | D8S1179 | TPOX | FGA | D21S11 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ***H*obs** | 0.7467 | 0.773 | **0.893** | 0.86 | 0.7666 | 0.76 | 0.76 | 0.72 | 0.74 | 0.81333 | 0.7866 | 0.7933 | 0.7133 | 0.8266 | 0.8467 |
| ***H*exp** | 0.776 | 0.785 | 0.874 | **0.9030** | 0.7654 | 0.7814 | 0.78416 | 0.77542 | 0.7208 | 0.86071 | 0.8195 | 0.8396 | 0.6885 | 0.8668 | 0.8036 |
| ***h*** | 0.224 | 0.214 | 0.126 | 0.0969 | 0.2345 | 0.2185 | 0.21584 | 0.22458 | **0.2791** | 0.13929 | 0.1804 | 0.1603 | 0.3114 | 0.1331 | 0.1964 |
| **PM** | 0.0856 | 0.085 | 0.034 | **0.0216** | 0.0927 | 0.0846 | 0.08124 | 0.0824 | **0.1366** | 0.03671 | 0.0630 | 0.0460 | 0.1422 | 0.0336 | 0.0793 |
| **PD** | 0.9144 | 0.914 | 0.965 | **0.9783** | 0.9072 | 0.9153 | 0.91876 | 0.9176 | 0.8633 | 0.96329 | 0.9369 | 0.9539 | **0.8577** | 0.9663 | 0.9207 |
| **PIC** | 0.7397 | 0.752 | 0.861 | **0.8952** | 0.7319 | 0.7507 | 0.75297 | 0.74371 | 0.6686 | 0.84548 | 0.7945 | 0.8202 | **0.6424** | 0.8530 | 0.7795 |
| **PE** | 0.5157 | 0.555 | **0.775** | 0.7276 | 0.5381 | 0.5356 | 0.5367 | 0.48075 | 0.4844 | 0.64063 | 0.5871 | 0.6037 | **0.4384** | 0.6633 | 0.67 |
| **PIT** | 2.231 | 2.334 | 3.968 | **5.158** | 2.131 | 2.287 | 2.316 | 2.226 | 1.791 | 3.589 | 2.770 | 3.118 | **1.605** | 3.7550 | 2.5464 |

Hobs, observed heterozygosity; Hexp, expected heterozygosity; h, homozygosity; PM, matching probability; PD, power of discrimination; PIC, polymorphism information content; PE, power of exclusion; PIT, typical paternity index.

**Hobs** ranged from **0.71** to **0.89** (Pointing to random mating population).
**h** ranged from **0.12** to **0.27** (inbreeding is insignificant in population and excess of heterozygous genotype).
**PM** ranged from **0.021** to **0.136** (the chance to find two persons with the same DNA profile is very low).
**PD** is very high, ranged from **0.85** to **0.97** (chances of picking out two people randomly that are genetically different are high).
**PIC** ranged from **0.64** to **0.89** (more allele diversity).
**PE** with average **0.58** (rare to find a random man who could not be excluded as the biological father of the child).
**PIT** ranged from **1.6** to **5.15** (testing DNA is more likely if tested man is the biological father of the child than if the biological father is another man)

**Acknowledgements:**

The authors would like to thank the ministry of martyrs and missing persons, Tripoli Libya for providing DNA data, International Commission of Missing Persons (ICMP) labs, Sarajevo, Bosnia and Herzegovina for analysis all of samples, exhumation team in Technical administration for identification of missing persons, and thanks goes to the Zoology Department, Faculty of science, Tripoli university.

**Recommendations:**

Use of more kits for multiplying STR loci.

Calculate of allele frequency in different Libyan ethnicity for comparison.

# حساب معدّل تكرار ألّيلات المواقع القصيرة المتعاقبة في المجتمع الليبي

أشرف بلعيد

**المستخلص:**

تهدف الدراسة الى وصف تعدّد التكرارات القصيرة المتعاقبة (STR) في المجتمع الليبي، وكذلك نسبة تكرار كل موقع من الـ15 موقعاً على حدى, وذلك للمساهمة بتحليلها في الأغراض الجنائية وأغراض تحديد الهويّة. تصنيف النمط الجيني (Genotype) ومعدل تكرار الألّيلات في المجتمع الليبي للمواقع التي تم تضخيمها بواسطة تفاعل انزيم البلمرة (PCR) اعتمد على خمسة عشر موقعاً من مواقع التكرارات القصيرة المتعاقبة, وهي (D3S1358, THO1, D21S11, D18S51, Penta E, D5S818, D13S317, D7S820, D16S539, CSF1PO, Penta D, VWA, D8S1179, TPOX, and FGA)

تمّ تضخيم هذه المواقع باستخدام PowerPlex® 16 HS System (Promega).

أظهرت النتائج أن تكرار المواقع القصيرة المتعاقبة (STR allele frequencies) في هذه الدراسة للمجتمع الليبي تراوحت بين 36.3% إلى 46.3%. كما أظهرت الدراسة وجود 151 ألّيلاً مختلفاً في المجتمع الليبي لـ 15 موقعاً من التكرارات القصيرة المتعاقبة.

**References:**


1. Butler, John M. (John Marshall), (2012). Advanced topics in forensic DNA typing: methodology 99.
2. Cifuentes L, Jorquera H, Acuña M, Ordóñez J, Sierra AL (2008) Allele frequencies for 12 autosomal short tandem repeat loci in two Bolivian populations. Genet Mol Res 7: 271-275.
3. Coudray C, Guitard E, el-Chennawi F, Larrouy G, Dugoujon JM (2007) Allele frequencies of 15 short tandem repeats (STRs) in three Egyptian populations of different ethnic groups. Forensic Sci Int 169: 260-265.
4. Drobnic K PN, Bakal N and Marjanovic D (2005) Allele Frequencies for 15 Short Tandem Repeat Loci in Representative Sample of Croatian Population. Croat Med J 48(4): 473–477
5. Gill, P., Urquhart, A., Millican, E.S., Oldroyd, N.J., Watson, S., Sparkes, R. and Kimpton, C.P. (1996) International Journal of Legal Medicine, 109, 14–22.
6. Goodwin. W, Adrian Linacre, Sibte Hadi (2007) An introduction of forensic genetics. First edition, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England,105-112.
7. Houssein Khodjet-el-Khil1, Karima Fadhlaoui-Zid1, Leonor Gusma, Cıntia Alves, Amel Benammar-Elgaaied1 & Anto nio Amorim (2012). Allele frequencies for 15 autosomal STR markers in the Libyan population. Journal of Annals of Human Biology, January – February 2012; 39 (1): 80–83.
8. Immel U.-D, M. Erhuma, T. Mustafa, M. Kleiber, M. Klintschar (2006). Population genetic analysis in a Libyan population using the PowerPlex 16 system. Journal of International Congress Series 1288 (2006) 421– 423.

9. Jeffreys AJ, Wilson V, Thein SL (1985). Individual-specific fingerprints of human DNA. Nature, 316: 76-79.
10. Linacre AMT, Phillips P, Wally A, Goodwin W (2001). *STR* data for the SGM PlusTM loci from three Indonesian populations. Forensic Sci. Int. 123: 232-234.
11. Stanciu F, Stoian IM, Popescu OR (2009) Population data for 15 short tandem repeat loci from Wallachia Region, South Romania. Croat Med J 50: 321-325.