**Research Article**
Volume 16:04, 2025

Journal of Forensic Research

ISSN: 2157-7145 Open Access

# Allele Frequency Database for 21 Short Tandem Repeats in the Ugandan Population

**Jane Nabwowe*, Musa Kirya, Dianah Katiti, Tarsisius Byamugisha, Immaculate Natukwasa, Precious Nyangoma, Alex Nkusaule, Charles Kigozi, Francis Mbolwa, Amon Serunkuma, Agnes Namatovu and Kepher Kuchana Kateu**

*Department of Government Analytical Laboratory, Ministry of Internal Affairs, Kampala, Uganda*

## Abstract

Genetic studies on the Ugandan population are very limited and little has been done to characterize its population structure. Here, we used 21 autosomal STRs included in the GlobalFilerTMAmplification Kit to amplify DNA from 1301 non-related adult individuals randomly selected from the four geographical regions of the country, namely Eastern, Western, Central and Northern region. The DNA samples were extracted using PrepFiler Express extraction protocol, amplified using the GlobalFilerTMPCR amplification kit and separated using capillary electrophoresis on the 3500xL Genetic Analyzer. A total of 342 alleles were observed ranging from 4 to 43.2 repeat units. Genotype distribution agreed with Hardy-Weinberg equilibrium ($p > 0.05$) except at D3S1358 loci that did not show significant departures from Hardy-Weinberg equilibrium after Bonferroni correlation. Tri-allelic patterns were observed at TPOX and D7S820 and the most polymorphic loci with 49 and 30 different alleles were SE33 and FGA loci respectively. The most discriminating loci were SE33 (0.9900) and D2S133 (0.9788) while heterozygosity ranged from 0.700 for D13S317 to 0.9131 for SE33 in the Ugandan population. The Combined Power of Exclusion (CPE), Combined Match Probability (CMP), Combined Power of Discrimination (CPD) and Combined Typical Paternity Index were 0.999980519, 1 in $6.76703 \times 10^{27}$, 1 and 695669937.8 respectively. Results from the study showed that when combined, the 21 Short tandem repeat loci showed high gene diversity since they were highly informative, polymorphic and discriminative providing a Ugandan STR population dataset resourceful in paternity and forensic testing.

**Keywords:** Allele frequency • Autosomal short tandem repeats • Ugandan population • Forensic statistics parameters • Globalfiler

## Introduction

Short Tandem Repeats (STRs) are the most common genetic markers, well widespread in the human genome and have a broad range of applications in DNA profiling [1-3]. The relatively short length of STRs is critical because it makes them the most amenable to multiplex amplification by Polymerase Chain Reaction (PCR) and therefore ideal forensic markers [4]. Short tandem repeat loci's importance as the most informative genetic markers providing high statistical capability of discrimination and individualization is widely acknowledged, documented and the data generated is currently being used in various forensic and judicial settings [5-7]. Genetic studies on Ugandan population to date are very limited and yet knowledge of any such structure is important in the interpretation of the significance of DNA-based analysis for human identification in parentage testing as well as forensic investigations [8].

Prior to the application of any DNA based identification method, it is essential to estimate the allele frequencies and forensic statistical parameters of targeted STR loci in each population in order to provide a more precise reference database for forensic investigation. Following the expansion of the Combined DNA Index System (CODIS) core loci number from 13 to 20 [9], no population studies have been conducted on the overall Ugandan population. The paucity of data on autosomal STR markers of Forensic and Paternity interest in the Ugandan population prompts this study to be carried out following the recommendations of the DNA commission of ISFG (International Society for Forensic Genetics; 2020) [10].

Genetic studies done show allele frequencies differ worldwide, some alleles are more frequent while others are absent in certain populations. For this reason, the aim of the present study is to determine allele frequencies for 21 short tandem repeat loci in the Ugandan population. The results indicate that all loci are informative, highly polymorphic and discriminative. Combined, the 21 STRs are resourceful in paternity and forensic investigations.

## Materials and Methods

### Sample collection and DNA extraction

Ethical approval to conduct this research was granted by the Makerere University School of Biomedical Sciences Research and Ethics committee, (Research file reference: SBS-2023-478) and Uganda National Council of Science and Technology (Research file reference: NS716ES). Blood samples were collected from 1301 unrelated consenting individuals in the four major regions of the country, namely; Central, Eastern, Western and Northern region. For each sample, the identity of the participants was dissociated from the sample collected to respect confidentiality.

### DNA extraction and quantitation

Genomic DNA was extracted following standard operating procedure of extracting DNA using Prepfiler Express on Automate Express DNA extraction system [11]. Following the manufacturer's guide [12], the quantities of extracted DNA samples were determined using Quantifiler™ Trio DNA Quantification Kit on the 7500 Real Time PCR machine.

### Multiplex PCR amplification

Multiplex amplification of genomic DNA after normalisation was performed as per GlobalFiler™ and GlobalFiler™ IQC PCR Amplification Kits protocol [12] on a Proflex PCR machine.

***Address for Correspondence:** Jane Nabwowe, Department of Government Analytical Laboratory, Ministry of Internal Affairs, Kampala, Uganda, E-mail: jnabwowe@yahoo.com*



**Received:** 23 June, 2025, Manuscript No. jfr-25-167119; **Editor Assigned:** 25 June, 2025, PreQC No. P-167119; **Reviewed:** 10 July, 2025, QC No. Q-167119; **Revised:** 17 July, 2025, Manuscript No. R-167119; **Published:** 25 July, 2025, DOI: 10.5281/zenodo.17706946

### Genotyping

The amplicons were separated by capillary electrophoresis on the ABI 3500xL genetic Analyser (Thermal Fisher Scientific) using POP-4™ polymer and LIZ-600 as internal size standard (Life Technologies) and raw data was captured with 3500 Series Data Collection software 2 (Thermal Fisher Scientific) [13,14]. GeneMapper ID-X version 1.5 software (Thermal Fisher Scientific) was used to genotype the raw data and the corresponding allelic ladder employed as per the manufacturer's guide [15].

### Quality control

The negative control and the target were checked for peaks in order to discount contamination. The sizes of the observed peaks (above 150 RFU), quality, presence or absence of interferences and possible null alleles were assessed. After successfully passing the quality control tests, these profiles were used for statistical analysis. Quality control guidelines outlined by Schneider PM [16] were followed while performing the laboratory work. Our laboratory participates in the quality control proficiency tests annually organised by Forensic Assurance Proficiency Testing Provider (www.forensicassurance.com).

### Statistical analysis

Genetic diversity, forensic and population statistical parameters including allele frequency, Observed Heterozygosity (Ho), expected Heterozygosity (He), Observed Homozygosity (H obs), Power of Discrimination (PD), Polymophic Information Content (PIC), Match Probability (MP), Power of Exclusion (PE), were estimated using STRAF 2.0 [17,18]. Minimum allele frequency (MAF) was computed according to the NRC recommendations as 5/2N [19]. Exact tests were carried out with the Arlequin 3.5 software [20] to check Hardy-Weinberg Expectations (HWE) by locus. Bonferroni correction by Weir BS [21] was applied to the probability of HWE to determine significant deviations.

## Results

### Allele frequency, rare alleles, microvariant alleles and tri-allelic patterns

A total of 1301 unrelated adult individuals were successfully genotyped for the STR loci CSFIPO, D10S1248, D12S391, D13S317, D16S539, D18S51, D19S433, D1S1656, D21S11, D22S1045, D2S1338, D2S441, D3S1358, D5S818, D7S820, D8S1179, FGA, SE33, TH01, TPOX, vWA, DYS391 and the Amelogenin loci yielding 2602 gene copies. A total of 342 alleles were detected and these ranged from allele 4.2 at SE33 to allele 43.2 at FGA as showed in Table 1.

The allele frequencies obtained from 21 autosomal STR markers in the Ugandan population are indicated in Table 1 with the least allele frequency of 0.0004 for rare alleles observed at many loci except at D8S1179 and THO1 and the highest frequency of 0.3797 at allele 12 of D5S818.

The common microvariants observed were of two incomplete repeats on loci namely SE33, D19S433, FGA, D21S11, D18S51 and D3S1358 though microvariants with three incomplete repeats were also observed at D1S1656, THO1, D2S441, D7S820 and FGA. Rare microvariants of one incomplete repeat were as well observed in the Ugandan population at D21S11, D12S391, D7S820 and D5S818. Rare microvariants and tri-allelic patterns were genotyped twice and the same allele calls were generated. The minimum allele frequency for the Ugandan population is 0.0019 to be used were rare alleles were observed that did not meet 5 occurances.

Tri-allelic patterns, a rare occurance, were observed at TPOX and D7S820 loci and the genotypes of the 10 individuals who displayed these patterns are shown in Table 2. Type 2 tri-allelic pattern were the only variants in this study.

**Table 1.** Uganda allele frequencies database and forensic statistics parameters.

| Allele | CSF1PO | D10S1248 | D12S391 | D13S317 | D16S539 | D18S51 | D19S433 | D1S1656 | D21S11 | D22S1045 | D2S1338 | D2S441 | D3S1358 | D5S818 | D7S820 | D8S1179 | FGA | SE33 | TH01 | TPOX | vWA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 4.2 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.0008 | - | - | - |
| 5 | - | - | - | - | 0.0019 | - | - | - | - | - | - | - | - | - | - | - | - | 0.0004 | - | - | - |
| 6 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | | 0.1745 | 0.0634 | - |
| 6.3 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.0004 | - | - | - |
| 7 | 0.0400 | - | - | - | - | - | - | - | - | - | - | - | - | 0.0008 | 0.0158 | 0.0008 | - | - | 0.3713 | 0.0088 | - |
| 7.3 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.0035 | - | - | - | - | - | - |
| 8 | 0.0461 | - | - | 0.0231 | 0.0446 | - | - | - | - | - | - | - | - | 0.0676 | 0.2333 | - | - | - | 0.2448 | 0.2890 | - |
| 8.1 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.0004 | - | - | - | - | - | - |
| 8.3 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.0004 | - | - | - | - | - | - |
| 9 | 0.0945 | 0.0015 | - | 0.0204 | 0.2052 | | 0.0038 | | | 0.0004 | | 0.0035 | 0.0004 | 0.0177 | 0.1230 | 0.0012 | - | - | 0.1410 | 0.3209 | - |
| 9.2 | - | - | - | - | - | 0.0004 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 9.3 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.0515 | | - |
| 10 | 0.2967 | 0.0042 | - | 0.0350 | 0.1176 | 0.0008 | 0.0200 | 0.0077 | | 0.0450 | | 0.0534 | | 0.0527 | 0.3732 | 0.0077 | - | 0.0004 | 0.0146 | 0.0776 | - |
| 10.1 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.0004 | | - | - | - | - | - |
| 10.2 | - | - | - | - | - | 0.0019 | 0.0004 | - | - | - | - | - | - | - | - | - | - | 0.0015 | | | - |
| 11 | 0.1868 | 0.0653 | - | 0.3436 | 0.2963 | 0.0019 | 0.0542 | 0.0503 | | 0.1875 | | 0.2832 | 0.0008 | 0.2095 | 0.1414 | 0.0388 | - | 0.0023 | 0.0008 | 0.2118 | 0.0046 |
| 11.2 | - | - | - | - | - | - | 0.0019 | - | - | - | - | | - | - | - | - | - | 0.0015 | - | - | - |
| 11.3 | - | - | - | - | - | - | - | - | - | - | - | 0.0480 | - | - | - | - | - | | - | - | - |
| 12 | 0.2744 | 0.1326 | - | 0.3782 | 0.1964 | 0.0242 | 0.1103 | 0.0603 | | 0.0357 | | 0.1526 | 0.0046 | 0.3797 | 0.0865 | 0.1530 | - | 0.0019 | 0.0015 | 0.0281 | 0.0004 |
| 12.1 | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.0004 | - | - | - | | - | - | - |
| 12.2 | - | - | - | - | - | - | 0.0238 | - | - | - | - | 0.0004 | - | - | - | - | - | 0.0085 | - | - | - |
| 12.3 | - | - | - | - | - | - | | - | - | - | - | 0.0177 | - | - | - | - | - | | - | - | - |
| 13 | 0.0527 | 0.2921 | 0.0004 | 0.1441 | 0.1245 | 0.0430 | 0.2563 | 0.1057 | - | 0.0077 | | 0.0427 | 0.0012 | 0.2567 | 0.0181 | 0.1645 | - | 0.0061 | - | 0.0004 | 0.0104 |
| 13.2 | - | - | - | - | - | 0.0015 | 0.0615 | | - | - | - | - | - | - | - | - | - | 0.0073 | - | - | |
| 13.3 | - | - | - | - | - | - | 0.0004 | - | - | - | - | 0.0015 | - | - | - | - | - | | - | - | |
| 14 | 0.0077 | 0.2471 | 0.0050 | 0.0530 | 0.0119 | 0.0546 | 0.2463 | 0.2610 | - | 0.0807 | 0.0004 | 0.3505 | 0.0884 | 0.0135 | 0.0035 | 0.3228 | - | 0.0342 | - | - | 0.0911 |
| 14.2 | - | - | - | - | - | 0.0027 | 0.0749 | | - | | | | | | | | - | 0.0065 | - | - | |
| 14.3 | - | - | - | - | - | | 0.0004 | 0.0177 | - | | | 0.0015 | | | | | - | | - | - | |

| | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 15 | 0.0008 | 0.1768 | 0.0942 | 0.0015 | 0.0015 | 0.1368 | 0.0580 | 0.1776 | - | 0.2375 | 0.0019 | 0.0430 | 0.3236 | 0.0015 | 0.0008 | 0.2252 | 0.0004 | 0.0334 | - | - | 0.1760 |
| 15.2 | - | - | - | | | 0.0061 | 0.0600 | | - | | | | 0.0004 | - | - | | | 0.0058 | - | - | |
| 15.3 | - | - | - | | | | | 0.0200 | - | | | | | - | - | | | | - | - | |
| 16 | - | 0.0626 | 0.1003 | 0.0012 | | 0.1833 | 0.0108 | 0.1357 | - | 0.2414 | 0.0592 | 0.0015 | 0.3228 | - | - | 0.0742 | 0.0008 | 0.0669 | - | - | 0.2763 |
| 16.2 | - | - | - | - | - | 0.0050 | 0.0150 | | - | | | | 0.0004 | - | - | | | 0.0035 | - | - | |
| 16.3 | - | - | - | - | - | | | 0.0968 | - | | | | | - | - | | | | - | - | |
| 17 | - | 0.0161 | 0.1845 | - | - | 0.1910 | | 0.0288 | - | 0.1549 | 0.0842 | 0.0004 | 0.2129 | - | - | 0.0081 | 0.0019 | 0.0853 | - | - | 0.1818 |
| 17.1 | - | - | 0.0008 | - | - | | | - | - | - | - | - | - | - | - | - | - | | - | - | - |
| 17.2 | - | - | - | - | - | 0.0015 | 0.0015 | - | - | - | - | - | - | - | - | - | - | 0.0027 | - | - | - |
| 17.3 | - | - | - | - | - | | | 0.0273 | | | | | | - | - | | | | - | - | - |
| 18 | 0.0004 | 0.0015 | 0.3244 | - | - | 0.1480 | | 0.0061 | | 0.0077 | 0.0780 | | 0.0415 | - | - | 0.0038 | 0.0061 | 0.1341 | - | - | 0.1453 |
| 18.1 | - | - | 0.0004 | - | - | | | | | | | | | - | - | | | | - | - | - |
| 18.2 | - | - | - | - | - | | 0.0004 | 0.0004 | | | | | | - | - | | 0.0054 | 0.0015 | - | - | - |
| 18.3 | - | - | - | - | - | | | 0.0023 | | | | | | - | - | | | | - | - | - |
| 19 | - | - | 0.1264 | - | - | 0.1049 | | | | 0.0004 | 0.1630 | | 0.0031 | - | - | | 0.0384 | 0.1372 | - | - | 0.0753 |
| 19.1 | - | - | 0.0027 | - | - | - | - | - | - | - | - | - | - | - | - | - | | | - | - | - |
| 19.2 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.0035 | 0.0012 | - | - | - |
| 19.3 | - | - | - | - | - | - | - | 0.0023 | - | - | - | - | - | - | - | - | | | - | - | - |
| 20 | - | - | 0.0696 | - | - | 0.0527 | - | - | - | 0.0004 | 0.0895 | - | - | - | - | - | 0.0477 | 0.0980 | - | - | 0.0319 |
| 20.1 | - | - | 0.0004 | - | - | | - | - | - | - | - | - | - | - | - | - | | | - | - | - |
| 20.2 | - | - | - | - | - | 0.0004 | - | - | - | - | - | - | - | - | - | - | 0.0004 | 0.0058 | - | - | - |
| 21 | - | - | 0.0361 | - | - | 0.0204 | - | - | - | | 0.1191 | - | - | - | - | - | 0.0723 | 0.0623 | - | - | 0.0065 |
| 21.1 | - | - | - | - | - | | - | - | - | | | - | - | - | - | - | - | 0.0008 | - | - | - |
| 21.2 | - | - | - | - | - | 0.0015 | - | - | - | | | - | - | - | - | - | - | 0.0115 | - | - | - |
| 22 | - | - | 0.0307 | - | - | 0.0119 | - | - | - | 0.0004 | 0.1533 | - | - | - | - | - | 0.2145 | 0.0223 | - | - | 0.0004 |
| 22.2 | - | - | - | - | - | | - | - | - | - | | - | - | - | - | - | - | 0.0081 | - | - | - |
| 23 | - | - | 0.0138 | - | - | 0.0038 | - | - | - | - | 0.0919 | - | - | - | - | - | 0.1649 | 0.0038 | - | - | - |
| 23.2 | - | - | - | - | - | | - | - | - | - | - | - | - | - | - | - | 0.0012 | 0.0077 | - | - | - |
| 23.3 | - | - | - | - | - | | - | - | - | - | - | - | - | - | - | - | 0.0004 | | - | - | - |
| 24 | - | - | 0.0054 | - | - | 0.0012 | - | - | - | - | 0.0822 | - | - | - | - | - | 0.1495 | 0.0008 | - | - | - |
| 24.2 | - | - | - | - | - | | - | - | 0.0004 | - | 0.0004 | - | - | - | - | - | - | 0.0188 | - | - | - |
| 24.3 | - | - | - | - | - | | - | - | 0.0038 | - | | - | - | - | - | - | 0.0008 | | - | - | - |
| 25 | - | - | 0.0031 | - | - | 0.0004 | - | - | - | - | 0.0365 | - | - | - | - | - | 0.1007 | | - | - | - |
| 25.2 | - | - | | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.0315 | - | - | - |
| 25.3 | - | - | | - | - | - | - | - | 0.0012 | - | - | - | - | - | - | - | 0.0004 | | - | - | - |
| 26 | - | - | 0.0008 | - | - | - | - | - | 0.0042 | 0.0004 | 0.0227 | - | - | - | - | - | 0.0565 | 0.0004 | - | - | - |
| 26.1 | - | - | | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.0004 | - | - | - | - |
| 26.2 | - | - | | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.0503 | - | - | - |
| 27 | - | - | 0.0012 | - | - | - | - | - | 0.0434 | - | 0.0150 | - | - | - | - | - | 0.0369 | 0.0008 | - | - | - |
| 27.2 | - | - | - | - | - | - | - | - | | - | - | - | - | - | - | - | | 0.0427 | - | - | - |
| 28 | - | - | - | - | - | - | - | - | 0.2571 | - | 0.0008 | - | - | - | - | - | 0.0400 | | - | - | - |
| 28.2 | - | - | - | - | - | - | - | - | | - | - | - | - | - | - | - | | 0.0415 | - | - | - |
| 29 | - | - | - | - | - | - | - | - | 0.1891 | - | 0.0019 | - | - | - | - | - | 0.0281 | 0.0019 | - | - | - |
| 29.2 | - | - | - | - | - | - | - | - | 0.0008 | - | - | - | - | - | - | - | | 0.0192 | - | - | - |
| 30 | - | - | - | - | - | - | - | - | 0.1872 | - | - | - | - | - | - | - | 0.0096 | | - | - | - |
| 30.2 | - | - | - | - | - | - | - | - | 0.0092 | - | - | - | - | - | - | - | 0.0069 | 0.0108 | - | - | - |
| 31 | - | - | - | - | - | - | - | - | 0.0765 | - | - | - | - | - | - | - | 0.0023 | | - | - | - |
| 31.2 | - | - | - | - | - | - | - | - | 0.0338 | - | - | - | - | - | - | - | 0.0073 | 0.0088 | - | - | - |
| 32 | - | - | - | - | - | - | - | - | 0.0196 | - | - | - | - | - | - | - | | | - | - | - |
| 32.2 | - | - | - | - | - | - | - | - | 0.0588 | - | - | - | - | - | - | - | 0.0019 | 0.0027 | - | - | - |
| 33 | - | - | - | - | - | - | - | - | 0.0123 | - | - | - | - | - | - | - | - | - | - | - | - |
| 33.1 | - | - | - | - | - | - | - | - | 0.0054 | - | - | - | - | - | - | - | - | - | - | - | - |
| 33.2 | - | - | - | - | - | - | - | - | 0.0242 | - | - | - | - | - | - | - | 0.0004 | 0.0015 | - | - | - |
| 34 | - | - | - | - | - | - | - | - | 0.0154 | - | - | - | - | - | - | - | - | | - | - | - |
| 34.2 | - | - | - | - | - | - | - | - | 0.0023 | - | - | - | - | - | - | - | - | 0.0012 | - | - | - |
| 35 | - | - | - | - | - | - | - | - | 0.0377 | - | - | - | - | - | - | - | - | 0.0031 | - | - | - |
| 35.2 | - | - | - | - | - | - | - | - | 0.0012 | - | - | - | - | - | - | - | - | | - | - | - |
| 36 | - | - | - | - | - | - | - | - | 0.0108 | - | - | - | - | - | - | - | - | 0.0004 | - | - | - |
| 37 | - | - | - | - | - | - | - | - | 0.0050 | - | - | - | - | - | - | - | - | - | - | - | - |
| 39 | - | - | - | - | - | - | - | - | 0.0008 | - | - | - | - | - | - | - | - | - | - | - | - |
| 43.2 | - | - | - | - | - | - | - | - | | - | - | - | - | - | - | - | 0.0008 | - | - | - | - |

| | CSF1PO | D10S1248 | D12S391 | D13S317 | D16S539 | D18S51 | D19S433 | D1S1656 | D21S11 | D22S1045 | D2S1338 | D2S441 | D3S1358 | D5S818 | D7S820 | D8S1179 | FGA | SE33 | TH01 | TPOX | vWA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| N | 1301 | 1301 | 1301 | 1301 | 1301 | 1301 | 1301 | 1301 | 1301 | 1301 | 1301 | 1301 | 1301 | 1301 | 1301 | 1301 | 1301 | 1301 | 1301 | 1301 | 1301 |
| Nall | 10 | 10 | 19 | 9 | 9 | 25 | 19 | 16 | 24 | 14 | 17 | 14 | 12 | 10 | 13 | 11 | 30 | 49 | 8 | 8 | 12 |
| PIC | 0.7554 | 0.7671 | 0.7982 | 0.6657 | 0.7715 | 0.8558 | 0.8237 | 0.8374 | 0.8314 | 0.7909 | 0.8842 | 0.7311 | 0.6905 | 0.6967 | 0.7300 | 0.7570 | 0.8668 | 0.9221 | 0.7100 | 0.7187 | 0.8006 |
| PM | 0.0759 | 0.0739 | 0.0533 | 0.1275 | 0.0684 | 0.0313 | 0.0407 | 0.0370 | 0.0405 | 0.0608 | 0.0211 | 0.0903 | 0.1161 | 0.1098 | 0.0898 | 0.0760 | 0.0260 | 0.0100 | 0.1007 | 0.0958 | 0.0537 |
| PD | 0.9241 | 0.9261 | 0.9467 | 0.8725 | 0.9316 | 0.9687 | 0.9593 | 0.9630 | 0.9595 | 0.9392 | 0.9789 | 0.9097 | 0.8839 | 0.8902 | 0.9102 | 0.9240 | 0.9740 | 0.9900 | 0.8993 | 0.9042 | 0.9463 |
| PE | 0.5739 | 0.6267 | 0.6036 | 0.4286 | 0.5894 | 0.7241 | 0.6428 | 0.6936 | 0.6830 | 0.6122 | 0.7643 | 0.5274 | 0.4690 | 0.4944 | 0.5193 | 0.5739 | 0.7674 | 0.8223 | 0.4880 | 0.5009 | 0.6108 |
| TPI | 2.3399 | 2.6992 | 2.5311 | 1.6679 | 2.4363 | 3.6960 | 2.8283 | 3.3189 | 3.2044 | 2.5916 | 4.3367 | 2.0849 | 1.8221 | 1.9303 | 2.0456 | 2.3399 | 4.3953 | 5.7566 | 1.9020 | 1.9593 | 2.5813 |
| Hobs | 0.7863 | 0.8148 | 0.8025 | 0.7002 | 0.7948 | 0.8647 | 0.8232 | 0.8493 | 0.8440 | 0.8071 | 0.8847 | 0.7602 | 0.7256 | 0.7410 | 0.7556 | 0.7863 | 0.8862 | 0.9131 | 0.7371 | 0.7448 | 0.8063 |
| GD (Hexp) | 0.7866 | 0.7966 | 0.8188 | 0.7135 | 0.8004 | 0.8698 | 0.8411 | 0.8532 | 0.8480 | 0.8166 | 0.8941 | 0.7648 | 0.7364 | 0.7385 | 0.7634 | 0.7878 | 0.8785 | 0.9269 | 0.7493 | 0.7580 | 0.8237 |
| Fis | -0.0016 | -0.0229 | 0.019 | 0.0169 | 0.0067 | 0.0052 | 0.0202 | 0.0032 | 0.0036 | 0.0082 | 0.0093 | 0.0023 | 0.0144 | -0.0043 | 0.0076 | 0.0017 | -0.0096 | 0.0146 | 0.0124 | 0.0159 | 0.0197 |
| p-Value | 0.23501 | 0.0172 | 0.5658 | 0.6543 | 0.9834 | 0.4542 | 0.6894 | 0.1326 | 0.1687 | 0.8086 | 0.88245 | 0.9777 | 0.0020 | 0.9923 | 0.5813 | 0.4483 | 0.9728 | 0.1844 | 0.8656 | 0.2825 | 0.6850 |

N: Population; Nall: Number of alleles observed per locus; PIC: Polymorphism Information Content; PM: Random Match Probabilty; PD: Power of Discrimination; PE: Power of Exclusion; TPI: Typical Paternity Index; Hobs: Observed Heterozygosity; GD(Hexp): Gene Diversity (Expected Heterozygosity; Fis: Inbreeding Coefficient; p-Value: Hardy Weinberg equilibrium

**Table 2.** Tri-allelic patterns observed in the Ugandan population.

| Locus | No. of Tri-alleles Observed | Alleles | Type |
|---|---|---|---|
| TPOX | 6 | 8,10,11 | Type 2 |
| | | 8,9,10 | |
| | | 8,9,10 | |
| | | 9,10,11 | |
| | | 8,9,10 | |
| | | 6,8,10 | |
| D7S820 | 4 | 10,12,14 | Type 2 |
| | | 8,11,11.3 | |
| | | 9,11,12.3 | |
| | | 9,11,14 | |

TPOX had the highest number of tri alleles observed

### Population genetic indices of the Ugandan population parameters

As shown in Table 1, all loci were in Hardy-Weinberg equilibrium except at D3S1358 (p=0.0020). However, after applying the Bonferroni correction, the departures observed were not significant. The expected heterozygosity values were very close to the observed heterozygosity value with the least observed at D13S317 (Hexp.=0.7134, Hobs=0.7002) and the highest at SE33 locus (Hexp.=0.9269, Hobs=0.9131). Total heterozygosity ranged from 0.7134 at D13S317 to 0.9270 at SE33. Low values of inbreeding coefficient (Fis), a measure of genetic relatedness, in the Ugandan population ranged from -0.0229 at D10S1248 to 0.0197 at vWA were observed. Values of genetic differentiation (Fst) were low ranging from 0.00096 at D16S539 to 0.00456 at D2S441.

### Genetic diversity and forensic efficiency parameters

The computed forensic metrics are shown in Table 1 with the most and least polymorphic loci being SE33 and D13S317 with PIC values of 0.9221 and 0.6657 respectively. Matching Probability (PM) values ranged from 0.010 (SE33) to 0.096 (TPOX) with the probability of obtaining a random match between individuals; Combined Match Probability (CMP) of $6.76703 \times 10^{27}$. All loci were highly discriminatory with high Power of Discrimination (PD) values ranging from 0.8725 at D13S317 to 0.9899 at SE33 locus.

The Ugandan population is a more diverse one based on the observed proportions of heterozygous and homozygous individuals (Power of Exclusion; PE). High PE values were associated with SE33 (PE=0.8223) and the lowest with D13S317 (PE=0.4286). Typical Paternity Index (TPI) values measuring how many times more likely that a possible father is the actual father than a randomly selected man in the population ranged from 1.6679 at D13S317 locus to 4.3953 at FGA locus.

All the 21 autosomal markers were highly informative and genetically diverse with the most informative locus in our data set being SE33 (Ho=0.9134, Hexp=0.9269, PIC=0.9221, PE=0.8223 and PD=0.9899) and the least informative locus was D13S317 (Ho=0.7002, Hexp=0.7134, PIC=0.6657, PE=0.4286 and PD=0.8725). When combined, the Combined Power of Exclusion (CPE), Combined Match Probability (CMP), Combined Power Of Discrimination (CPD) and Combined Typical Paternity Index were 0.999980519, 1 in $6.76703 \times 10^{27}$, 1 and 695669937.8 respectively. The small CMP values and large CPE values support the use of these 21 autosomal markers in individual identification and discrimination within the Ugandan population.

## Discussion and Data Analysis

This is the first study in the Ugandan population on the 21 STR markers with frequencies generated at all loci below 50% as shown by the predominant alleles at each locus tested. These low frequency values reflect the usefulness and validity of these loci in discriminating individuals [21].

The 21 STR markers in the Ugandan population are in Hardy-Weinberg equilibrium (p- values > 0.05) since no deviation was significant after applying Benferroni correction (0.05/21=0.0024) as is in previous studies from other populations such as Arabian populations from Morocco and Syria [22], African-American, US Caucasian and US Hispanic [23], NorthWest Spain (Gacilia) [24], Angola [25], Tigray population of Ethiopia [26] and Nigeria [27] which showed no deviations from Hardy-Weinberg equilibrium from all tested loci. Observed heterozygosity values were very close to the expected heterozygosity values at all loci indicating that the Ugandan population is in Hardy-Weinberg equilibrium and both heterozygosities were above 70% at all loci as in other published data [28-30]. This suggested that all these populations are not small, not closed and that there is minimal inbreeding if any [31].

TPOX loci had the highest number of tri-allelic patterns observed in the Ugandan population and these were of Type 2. Tri-allelic patterns are extensively studied [32-34] and may be as a result of duplication located on the same chromosome, translocation or chromosome aneuploidy (trisomy). Since TPOX occurs near the end of chromosome 2, it is more likely to duplicate for telomeres maintenance in order to keep intact the chromosome end [35].

The most informative and discriminatory locus in our data set was SE33 (PD =0.9900, PIC=0.9221, Ho =0.9131, He=0.9266, 49 alleles) while the lowest D13S317 (PD =0.8723, PIC=0.6656, Ho =0.7000, He=0.7131, 9 alleles)

similar to other studied populations [27,36-41]. This dataset indicates a high degree of gene diversity detected and that Uganda's population maintains a high degree of genetic variation with a history of high miscegenation. Similar results were obtained for the Power of Exclusion (PE) and Typical Paternity Index (TPI) with the highest PE and TPI values being associated with SE33 (PE=0.8223, TPI=5.7566) proving to be a powerful marker for human identification and paternity testing within the Ugandan population.

According to Kimura M [42], alleles with a frequency of less than 0.01 in a population database are regarded as rare alleles. The Ugandan population database as well had rare alleles with the lowest frequency of 0.004 and these were mainly microvariants; alleles with incomplete repeats [31]. Locus that displayed the highest number of microvariants was SE33 and these had two incomplete repeats only while at TH01 a 9.3 microvariant was observed and this was found to have been caused by a missing adenine in the seventh repeat [43]. The same study points out that the cause of a .2 microvariant at D21S11 locus is a -TA- dinucleotide partial repeat before the last full TCTA repeat and a .2 microvariant at FGA locus is caused by a -TT- dinucleotide partial repeat after the fifth full repeat and before the variable CTTT repeat motif.

The large Combined Power of Discrimination (CDP), Combined Power of Exclusion (CPE) and Combined Paternity Index (CPI) values and low value of Combined Matching Probability (CMP) support the use of the 21 STR loci in individual identification and paternity testing within the Ugandan population. These values were remarkably higher than that of the 15 autosomal loci obtained for the studied populations [44,45].

## Conclusion

This study confirms that the 21 Short Tandem Repeat loci is a suitable tool for human identification in forensic, parentage and kinship analysis in the Ugandan population due to their genetic variability and combined discriminatory power. Since these loci studied represent a combination of CODIS core and expanded European loci, this dataset and the allelic frequencies calculated from it can help international authorities with Ugandan victims or suspects.

## Acknowledgement

We gratefully acknowledge the study participant who actively and willingly contributed to this research. We also would like to thank Mr. Kepher Kuchan Kateu the Chief Government Chemist and Mr. Tarsisius Byamugisha, the Commissioner at the Government Analytical Laboratory, Ministry of Internal Affairs for their continued technical guidance and professional suggestions all the way through the demanding moments of this research. We thank all Forensic Biology laboratory staff for their continued generosity in offering the information required as well as the enthusiasm they had for this research. Funding for this research was provided by the Government of Uganda.

## Conflict of Interest

The authors declare no potential conflicts of interest.

## References

1. Edwards, Albert Andrew Civitello, Holly A. Hammond and C. Thomas Caskey. "DNA typing and genetic mapping with trimeric and tetrameric tandem repeats." *Am J Hum Genet* 49 (1991): 746.
2. Zietkiewicz, Ewa, Magdalena Witt, Patrycja Daca and Jadwiga Zebracka-Gala, et al. "Current genetic methodologies in the identification of disaster victims and in forensic analysis." *J Appl Genet* 53 (2012): 41–60.
3. Schuler, G. D., M. S. Boguski, E. A. Stewart and L. D. Stein, et al. "A gene map of the human genome." *Science* 274 (1996): 540–546
4. Butler, John M. "Advanced topics in forensic DNA typing: Methodology." Academic Press (2011)
5. Budowle, Bruce and Robert C. Allen. "Analysis of amplified fragment-length polymorphisms (VNTR/STR loci) for human identity testing." In Forensic DNA Profiling Protocols Totowa, NJ: Humana Press (1998).
6. Butler, John M. "Forensic DNA Typing: Biology and Technology behind STR Markers." San Diego CA: Academic Press (2001): 15-21.
7. Butler, John M. "Advanced topics in forensic DNA typing: Methodology." Academic Press (2012).
8. Butler, John M. "Fundamentals of forensic DNA typing." Academic Press (2009).
9. Hares, Douglas R. "Selection and implementation of expanded CODIS core loci in the United States." *Forensic Sci Int Genet* 17 (2015): 33–34
10. Gill, Peter, Tacha Hicks, John M. Butler and E. D. Connolly, et al. "DNA commission of the International Society for Forensic Genetics: assessing the value of forensic biological evidence—guidelines highlighting the importance of propositions. Part II: Evaluation of biological traces considering activity level propositions." *Forensic Sci Int Genet* 44 (2020): 102186
11. Applied Biosystems "PrepFiler Express™ and PrepFiler Express BTA™ Forensic DNA Extraction Kits User Guide (2010) (Part No: 4442699)
12. Scientific, Thermo Fisher. "GlobalFiler and GlobalFiler IQC PCR amplification kits user guide." Thermo Fisher Scientific (2019).
13. Applied Biosystems "3500/3500xL Genetic Analyzer User guide" (2011) (Part No: 4401661).
14. Applied Biosystems "3500/3500xL Genetic Analyzer with 3500 Series Data Collection Software 3." (2014).
15. Applied Biosystems "GeneMapper®ID-X Software Version 1.5" (2015).
16. Schneider, Peter M. "Scientific standards for studies in forensic genetics." *Forensic Sci Int* 165 (2007): 238–243.
17. Gouy, Alexandre and Martin Zieger. "STRAF—a convenient online tool for STR data evaluation in forensic genetics." *Forensic Sci Int Genet* 30 (2017): 148–151.
18. Gouy, Alexandre and Martin Zieger. "The STRAF book."
19. National Research Council (NRC) Committee on DNA Forensic Science. "The evaluation of forensic DNA evidence." Washington, D.C.: National Academy Press (1996).
20. Excoffier, Laurent, Guillaume Laval and Stefan Schneider. "Arlequin (version 3.0): An integrated software package for population genetics data analysis." *Evol Bioinform* 1 (2005): 47-50.
21. Weir, Bruce S. "Genetic data analysis II." Sunderland, MA: Sinauer Associates (1996)
22. Abdin, Louai, Ichiroh Shimada, Bernd Brinkmann and Carsten Hohoff. "Analysis of 15 short tandem repeats reveals significant differences between the Arabian populations from Morocco and Syria." *Leg Med* 5 (2003): S150–S155
23. Redman, Janette W. and Margaret C. Kline. "Allele frequencies for 15 autosomal STR loci in U.S. Caucasian, African American and Hispanic populations." *J Forensic Sci* 48 (2003).
24. Fernandez-Formoso, L., C. Phillips, A. Rodriguez and R. Calvo, et al. "Allele frequencies of 20 STRs from northwest Spain (Galicia)." *Forensic Sci Int Genet* 6 (2012): e149–e150
25. Melo, Miguel Manuel, Mónica Carvalho, Virgínia Lopes and Maria João Anjos, et al. "Genetic study of 15 STR loci of Identifiler system in the Angolan population." *Forensic Sci Int Genet* 4 (2010): e153–e157
26. Haddish, K., E. Chierto, G. Di Vella and D. Lacerenza, et al. "A reference database of forensic autosomal and gonosomal STR markers in the Tigray population of Ethiopia." *Forensic Sci Int Genet* 56 (2022): 102618
27. Okolie, Victoria O., Selena Cisana, Moses S. Schanfield and Khalid O. Adekoya, et al. "Population data of 21 autosomal STR loci in the Hausa, Igbo and Yoruba people of Nigeria." *Int J Legal Med* 132(2018): 735–737.
28. Morillo, Dairis, Francheska Acosta, Santa Jiménez and Víctor Calderón, et al. "Population genetic data of 22 autosomal STR loci from the Dominican Republic." *Forensic Sci Int Rep* 4 (2021): 100228.
29. Santos, Natalia Bahia Pinheiro dos, Márcio Fabrício Falcão de Paula Filho, Abigail Marcelino dos Santos Silva and Enio Paulo Telo, et al. "Allele frequencies and forensic data of 25 STR markers for individuals in northeast Brazil." *Genes* 14 (2023): 1185

30. Babiker, Hiba M. A., Carina M. Schlebusch, Hisham Y. Hassan and Mattias Jakobsson. "Genetic variation and population structure of Sudanese populations as indicated by 15 Identifiler sequence-tagged repeat (STR) loci." *Invest Genet* 2 (2011): 12

31. Butler, J. M. "Forensic DNA typing: Biology, technology and genetics of STR markers." Elsevier Academic Press 660 (2005).

32. Butler, John M. "Genetics and genomics of core short tandem repeat loci used in human identity testing." *J Forensic Sci* 51 (2006): 253–265

33. Clayton, T. M., J. L. Guest, A. J. Urquhart and P. D. Gill. "A basis for anomalous band patterns encountered during DNA STR profiling." *J Forensic Sci* 49 (2004): JFS2003145–8.

34. Lane, A. B. "The nature of tri-allelic TPOX genotypes in African populations." *Forensic Sci Int Genet* (2008): 134–13

35. Louis, Edward J. and Alexander V. Vershinin. "Chromosome ends: Different sequences may provide conserved functions." *Bioessays* 27 (2005): 685–697

36. Moyses, Cinthia Bachir, Weslley Mottoi Tsutsumida, Paulo Eduardo Raimann and Carlos Henrique Ares Silveira da Motta, et al. "Population data of the 21 autosomal STRs included in the GlobalFiler® kits in population samples from five Brazilian regions." *Forensic Sci Int Genet* 26 (2017): e28–e30.

37. Paul, Gasana, JunLin Liu, Pan Ma and Abiy Wendifraw Assefa, et al. "Genetic polymorphism of 24 autosomal STR in the population of Rwanda." *Biochem Genet* 60 (2022): 80–93.

38. Al-Snan, Noora R., Safia Messaoudi, Saranya R. Babu and Moiz Bakhiet. "Population genetic data of the 21 autosomal STRs included in GlobalFiler kit of a population sample from the Kingdom of Bahrain." *PLoS One* 14 (2019): e0220620

39. Kofi, Abban Edward, Hashom Mohd Hakim, Hussein Omar Khan and Siti Afifah Ismail, et al. "Population dataset for 21 simple tandem repeat loci in the Akan population of Ghana." *Data Brief* 31 (2020): 105746.

40. Muinde, Jane Mbithe, Devi R. Chandra Bhanu, Rita Neumann and Richard Okoth Oduor, et al. "Geographical and linguistic structure in the people of Kenya demonstrated using 21 autosomal STRs." *Forensic Sci Int Genet* 53 (2021): 102535.

41. Ristow, Peter Gustav, Kevin Wesley Cloete and Maria Eugenia D'Amato. "GlobalFiler® Express DNA amplification kit in South Africa: Extracting the past from the present." *Forensic Sci Int Genet* 24 (2016): 194–201

42. Kimura, Motoo. "Rare variant alleles in the light of the neutral theory." *Mol Biol Evol* 1 (1983): 84–93

43. McBeth, Molly. "A study of the molecular basis of microvariants at the FGA and D21S11 loci used in forensic DNA testing." 2006

44. Gomes, Verónica, Paula Sánchez-Diz, Cíntia Alves and Iva Gomes, et al. "Population data defined by 15 autosomal STR loci in the Karamoja population (Uganda) using AmpF/STR Identifiler kit." *Forensic Sci Int Genet* 3 (2009): e55–e58

45. Nabwowe, Jane, Musa Kirya, Emmanuel Okello and Ann Nanteza. "Allele frequency of 15 short tandem repeats (STRs) in a Buganda population (central Uganda): Forensic utility and parentage testing." *J Forensic Res* 5 (2014): 1

**How to cite this article:** Nabwowe, Jane, Musa Kirya, Dianah Katiti and Tarsisius Byamugisha, et al. "Allele Frequency Database for 21 Short Tandem Repeats in the Ugandan Population." *J Forensic Res* 16 (2025): 668.