ANIL AGGRAWAL'S
Internet Journal of Forensic Medicine and Toxicology
The world's first online journal devoted entirely to forensic medicine and toxicology · Published 1 January and 1 July

VOL. 19, NO. 2 · JULY – DECEMBER 2018
Downloaded from www.anilaggrawal.com · ISSN 0972-8074 (Online) · https://www.anilaggrawal.com/ij/vol_019_no_002/papers/paper005.html

# STR Analysis For Random Delhi-NCR Population Sample Using AmpFlSTR Identifiler Plus PCR Amplification Kit

Monika Chakravarty[1,2], Nidhi Singh[1], Dhruw Sharma[1], Naresh Kumar[1,2], Prateek Pandya[2*], Anupama Raina[3]

[1]Forensic Science Laboratory, Home Department, Govt. of NCT Delhi, Rohini, India
[2]Amity Institute of Forensic Sciences, Amity University, Noida, Uttar Pradesh, India
[3]All India Institute of Medical Sciences, New Delhi, India

* Corresponding author: Prateek Pandya · ppandya@amity.edu

CITE THIS ARTICLE

Chakravarty M, Singh N, Sharma D, Kumar N, Pandya P, Raina A. STR Analysis For Random Delhi-NCR Population Sample Using AmpFlSTR Identifiler Plus PCR Amplification Kit. *Anil Aggrawal's Internet Journal of Forensic Medicine and Toxicology* [serial online], 2018; Vol. 19, No. 2 (July – December 2018): [about 18 p]. Available from: https://www.anilaggrawal.com/ij/vol_019_no_002/papers/paper005.html

## ABSTRACT

Allelic frequencies for 15 STR's (D8S1179, D21S11, D7S820, CSF1PO, D3S1358, TH01, D13S317, D16S539, D2S1338, D19S433, vWA, TPOX, D18S51, D5S818, FGA) included in the AmplSTR Identifiler plus 15 STR system were derived from a sample of 200 unrelated random individuals. Hardy Weinberg Equilibrium was tested for each locus using Chi square test. No deviation was observed and the departure observed at these loci was not significant. The average heterozygosity level for both observed and expected values were above 70%, thus making these loci suitable for human identification for forensic purpose.

## INTRODUCTION

The repeat DNA sequence variation among individuals was identified almost 30 years back (1). The STR loci are ubiquitous in human genome, show high degree of polymorphism and can be amplified by PCR techniques, which are highly discriminatory in human identification specifically for parentage testing and in medico-legal situations (2-3).

Gene and genotype frequencies are required for forensic DNA typing for the estimation of probabilities involving matches in inclusion cases. In recent years, allelic data for STR loci have been published from a large number of populations (4-10).

## MATERIALS AND METHODS

**Population:** Randomly selected from case work routinely done in Forensic Science Laboratory, Rohini, Delhi, India.

**Extraction:** DNA was extracted using the phenol- chloroform –isoamyl alcohol procedure (11) and also using the FTA technology following the manufacturer's procedure as described in Whatman FTA protocol. DNA was quantified using Quantifiler human DNA quantification kit as per the manufacturer's recommended protocol (Applied Biosystems, Foster city USA)

**PCR**: A quantity of 0.5ng target DNA was used for each reaction prepared using the Identifiler plus PCR amplification kit. PCR amplification was carried out on Veriti Thermal cycler following the manufacturer's recommended protocol (12).

**Typing**

ABI 3500xl, positive controls, negative controls, reagent blanks and allelic ladder were used to type PCR products. The samples were heat denatured at 95 ºC for 3 minutes and chilled for 5 minutes in an ice bath before electrophoresis.

## Data Analysis

The results were analyzed using the Gene Mapper IDx software version v1.4. Allelic designations were determined by comparison with allelic ladder. The peak height threshold was 50 RFU for heterozygous and 200 RFU for homozygous alleles.

## Statistical Analysis

After the samples have been collected, DNA extracted, amplicons were genotyped for the 15 STR loci. The genotyping information was then converted into allele frequencies by counting the frequency of each allele for each locus. (Gene count method / standard counting methods described by Weir (13), Evett & Weir (14) and Sham (15). For comparing gene counts between samples and for testing Hardy Weinberg Equilibrium (HWE) within each sample, Chi square tests were used (16).

**Supplementary Table 1: The characteristics of the 15 STR loci examined in this study:**

| Locus designation | Chromosome location | STR sequence |
|---|---|---|
| D8S1179 | 8 | TCTA |
| D21S11 | 21q11.2-q21 | TCTA |
| D7S820 | 7q11.21-22 | GATA |
| CSF1PO | 5q33.3-34 | AGAT |
| D3S1358 | 3p | TCTA |
| THO1 | 11p15.5 | AATG |
| D13S317 | 13q22-31 | TATC |
| D16S539 | 16q24-qter | GATA |
| D2S1338 | 2q35-37.1 | TGCC |
| D19S433 | 19q12-13.1 | AAGG |
| vWA | 12p12-pter | TCTA |
| TPOX | 2p23-2per | AATG |
| D18S51 | 18q21.3 | AGAA |
| D5S818 | 5q21-31 | AGAT |
| FGA | 4q28 | TTTC |

***Negi D.S, Shrivastava P, Das S. P. Biochemical Characterzation of Molecular Markers for Human Genetic Identification in Paternity testing by DNA profiling. Asian Journal of Biomedical and Pharmaceutical Sciences; 04(37);2014,52-56 DOI: 10.15272/ajbps.v4i37.621***

## RESULTS AND DISCUSSION

Hardy Weinberg Equilibrium (HWE) was performed by calculating the observed allele frequencies and expected genotype frequencies based on the allele frequencies. ($p^2$ & $q^2$) for genotype frequencies of homozygotes & 2pq for genotype frequencies of heterozygotes.

---

**Hardy- Weinberg law states that in a large, randomly mating population both the allele frequencies & the genotype frequencies are constant from generation, provided the population lack's mutation, migration & natural selection.**

**Genotype frequencies can be predicted by the allelic frequencies ($p^2$ & $q^2$ for genotype frequencies of homozygotes & 2pq for genotype frequencies of heterozygotes).**

1. **Allelic frequencies should add upto 1**

**$p + q = 1$**

1. **Genotype frequencies can be calculated from allelic frequency. The sum of genotype frequencies should equal to1.**

**$p^2 + 2pq + q^2 = 1$**

**If the observed genotype proportion are different from those expected, one or more of the HW assumption have been violated.**

---

The observed genotype frequencies at a given locus are calculated by dividing the number of individuals with one genotype by the total number of individuals in the sampled population. The observed and expected genotype frequencies were then compared using a Chi square test. The Chi square value is used to determine a P value. If P value exceeds 0.05 the deviation of expected from observed genotype frequencies is not considered statistically significant. The null hypothesis states that all the STR loci tested are in HWE. The null hypothesis was rejected if the computed P values were below 0.05.

**Table 1 (i-xv) shows the allele frequency distribution of the 15 STR loci, n=200**

**Table 1(i) Allelic Frequency of D8S1179 locus**

| Allele | Frequency |
|---|---|
| 8 | 0.0075 |
| 9 | 0.0100 |
| 10 | 0.1825 |
| 11 | 0.0800 |
| 12 | 0.1375 |
| 13 | 0.2075 |
| 14 | 0.1275 |
| 15 | 0.1625 |
| 16 | 0.0700 |
| 17 | 0.0125 |
| 18 | 0.0025 |
| 19 | - |
| HWE | |
| df=1 | $\chi^2$ =0.174944 |
| | P =0.675758 |

**Table 1(ii) Allele Frequency of D21S11 locus**

| Allele | Frequency |
|---|---|
| 24 | - |
| 24.2 | - |
| 25 | - |
| 26 | - |
| 27 | 0.0175 |
| 28 | 0.1100 |
| 28.2 | - |
| 29 | 0.2125 |
| 29.2 | 0.0075 |
| 30 | 0.1625 |
| 30.2 | 0.0475 |
| 31 | 0.0450 |
| 31.2 | 0.1175 |
| 32 | 0.0100 |
| 32.2 | 0.1925 |
| 33 | - |
| 33.2 | 0.0675 |
| 34 | 0.0025 |
| 34.2 | 0.0075 |
| 35 | - |
| 35.2 | - |
| 36 | - |
| 37 | - |
| 38 | - |
| **HWE** | |
| df=1 | $\chi^2$ =0.37998 |
| | P =0.537657 |

**Table 1(iii) Allele Frequency of D7S820 locus**

| Allele | Frequency |
|---|---|
| 6 | - |
| 7 | 0.0175 |
| 8 | 0.2100 |
| 9 | 0.0650 |
| 10 | 0.2350 |
| 11 | 0.2650 |
| 12 | 0.1650 |
| 13 | 0.0325 |
| 14 | 0.0100 |
| 15 | - |
| **HWE** | |
| df=1 | $\chi^2$ =0.41103<br>P =0.52144 |

**Table 1(iv) Allele Frequency of CSFIPO locus**

| Allele | Frequency |
|---|---|
| 6 | - |
| 7 | 0.0025 |
| 8 | - |
| 9 | 0.0350 |
| 10 | 0.2175 |
| 11 | 0.2900 |
| 12 | 0.3725 |
| 13 | 0.0700 |
| 14 | 0.0100 |
| 15 | 0.0025 |
| **HWE** | |
| df=1 | $\chi^2$ =0.079744 |
| | P =0.77765 |

**Table 1(v) Allele Frequency of D3S1358 locus**

| Allele | Frequency |
|---|---|
| 12 | - |
| 13 | 0.0025 |
| 14 | 0.0350 |
| 15 | 0.3150 |
| 16 | 0.3225 |
| 17 | 0.2250 |
| 18 | 0.0925 |
| 19 | 0.0075 |
| **HWE** | |
| df=1 | $\chi^2$ =0.058966 |
| | P =0.808147 |

**Table 1(vi) Allele Frequency of TH01 locus**

| Allele | Frequency |
|---|---|
| 4 | - |
| 5 | - |
| 6 | 0.2225 |
| 7 | 0.1750 |
| 8 | 0.1375 |
| 9 | 0.2725 |
| 9.3 | 0.1800 |
| 10 | 0.0125 |
| 11 | - |
| 13.3 | - |
| **HWE** | |
| df=1 | $\chi^2$ =0.031922 |
| | P =0.858203 |

**Table 1(vii) Allele Frequency of D13S317 locus**

| Allele | Frequency |
|---|---|
| 8 | 0.1900 |
| 9 | 0.0725 |
| 10 | 0.0750 |
| 11 | 0.3000 |
| 12 | 0.2600 |
| 13 | 0.0700 |
| 14 | 0.0300 |
| 15 | 0.0025 |
| **HWE** | |
| df=1 | $\chi^2$ =0.10897 |
| | P =0.7414 |

**Table 1(viii) Allele Frequency of D16S539 locus**

| Allele | Frequency |
|---|---|
| 5 | - |
| 8 | 0.0625 |
| 9 | 0.1650 |
| 10 | 0.1075 |
| 11 | 0.3525 |
| 12 | 0.2125 |
| 13 | 0.0875 |
| 14 | 0.0125 |
| 15 | - |
| **HWE** | |
| df=1 | $\chi^2$ =0.10705 |
| | P =0.743528 |

**Table 1(ix) Allele Frequency of D2S1338 locus**

| Allele | Frequency |
|---|---|
| 15 | - |
| 16 | 0.0175 |
| 17 | 0.0475 |
| 18 | 0.1725 |
| 19 | 0.1400 |
| 20 | 0.1300 |
| 21 | 0.0475 |
| 22 | 0.0900 |
| 23 | 0.1975 |
| 24 | 0.0750 |
| 25 | 0.0650 |
| 26 | 0.0125 |
| 27 | 0.0025 |
| 28 | 0.0025 |
| **HWE** | |
| df=1 | $\chi^2$ =0.40638 |
| | P =0.523813 |

**Table 1(x) Allele Frequency of D19S433 locus**

| Allele | Frequency |
|---|---|
| 9 | - |
| 10 | 0.0050 |
| 11 | 0.0025 |
| 12 | 0.0750 |
| 12.2 | 0.0075 |
| 13 | 0.3275 |
| 13.2 | 0.0200 |
| 14 | 0.2425 |
| 14.2 | 0.0600 |
| 15 | 0.1150 |
| 15.2 | 0.0500 |
| 16 | 0.0575 |
| 16.2 | 0.0250 |
| 17 | 0.0125 |
| 17.2 | - |
| **HWE** | |
| df=1 | $\chi^2$ =0.656484 |
| | P =0.417409 |

**Table 1(xi) Allele Frequency of vWA locus**

| Allele | Frequency |
|---|---|
| 11 | - |
| 12 | - |
| 13 | 0.0050 |
| 14 | 0.1275 |
| 15 | 0.0800 |
| 16 | 0.2100 |
| 17 | 0.2950 |
| 18 | 0.1975 |
| 19 | 0.0675 |
| 20 | 0.0175 |
| 21 | - |
| 22 | - |
| 23 | - |
| 24 | - |
| **HWE** | |
| df=1 | $\chi^2$ =0.092262 |
| | P =0.761323 |

**Table 1(xii) Allele Frequency of TPOX locus**

| Allele | Frequency |
|---|---|
| 6 | 0.0025 |
| 7 | 0.0050 |
| 8 | 0.4075 |
| 9 | 0.1475 |
| 10 | 0.0875 |
| 11 | 0.3150 |
| 12 | 0.0350 |
| 13 | - |
| **HWE** | |
| df=1 | $\chi^2$ =0.026518 |
| | P =0.870661 |

**Table 1(xiii) Allele Frequency of D18S51 locus**

| Allele | Frequency |
|---|---|
| 7 | - |
| 9 | 0.0025 |
| 10 | 0.0025 |
| 10.2 | - |
| 11 | 0.0275 |
| 12 | 0.1000 |
| 13 | 0.1200 |
| 13.2 | - |
| 14 | 0.2625 |
| 14.2 | 0.0025 |
| 15 | 0.1825 |
| 16 | 0.1375 |
| 17 | 0.0700 |
| 18 | 0.0425 |
| 19 | 0.0225 |
| 20 | 0.0225 |
| 21 | 0.0025 |
| 22 | 0.0025 |
| 23 | - |
| 24 | - |
| 25 | - |
| 26 | - |
| 27 | - |
| **HWE** | |
| df=1 | $\chi^2$ =0.807213 |
| | P =0.368946 |

**Table 1(xiv) Allele Frequency of D5S818 locus**

| Allele | Frequency |
|---|---|
| 7 | - |
| 8 | - |
| 9 | 0.0400 |
| 10 | 0.1025 |
| 11 | 0.4000 |
| 12 | 0.2750 |
| 13 | 0.1700 |
| 14 | 0.0125 |
| 15 | - |
| 16 | - |
| **HWE** | |
| df=1 | $\chi^2$ =0.089409 |
| | P =0.764941 |

**Table 1(xv) Allele Frequency of FGA locus**

| Allele | Frequency |
|---|---|
| 17 | - |
| 18 | 0.0050 |
| 19 | 0.0550 |
| 20 | 0.1325 |
| 21 | 0.1000 |
| 22 | 0.1400 |
| 23 | 0.2100 |
| 24 | 0.1725 |
| 25 | 0.1250 |
| 26 | 0.0475 |
| 26.2 | - |
| 27 | 0.01 |
| 28 | - |
| 29 | - |
| 30 | - |
| 30.2 | - |
| 31.2 | - |
| 32.2 | 0.0025 |
| 33.2 | - |
| **HWE** | |
| df=1 | $\chi^2$ =0.351716 |
| | P =0.553147 |

**Table 2 (i-xv) shows the Genotype Frequency distribution of the 15 STR loci, n=200**

**Table 2 (i) Genotype frequency table of D8S1179 locus**

| Genotype | Observed Frequency (O) |
|---|---|
| 8,10 | 0.005 |
| 8,13 | 0.005 |
| 8,14 | 0.005 |
| 9,11 | 0.005 |
| 9,12 | 0.005 |
| 9,13 | 0.005 |
| 9,15 | 0.005 |
| 10,10 | 0.025 |
| 10,11 | 0.030 |
| 10,12 | 0.035 |
| 10,13 | 0.080 |
| 10,14 | 0.070 |
| 10,15 | 0.050 |
| 10,16 | 0.025 |
| 10,17 | 0.010 |
| 11,11 | 0.020 |
| 11,12 | 0.025 |
| 11,13 | 0.040 |
| 11,14 | 0.005 |
| 11,15 | 0.020 |
| 11,16 | 0.005 |
| 12,12 | 0.015 |
| 12,13 | 0.065 |
| 12,14 | 0.035 |
| 12,15 | 0.055 |
| 12,16 | 0.015 |
| 12,17 | 0.005 |
| 13,13 | 0.020 |
| 13,14 | 0.065 |
| 13,15 | 0.055 |
| 13,16 | 0.055 |

| Genotype | Observed Frequency (O) |
|---|---|
| 13,17 | 0.010 |
| 14,14 | 0.010 |
| 14,15 | 0.035 |
| 14,16 | 0.015 |
| 15,15 | 0.035 |
| 15,16 | 0.030 |
| 15,18 | 0.005 |

- Homozygotes = 25
  Heterozygotes =175

**Table 2 (ii) Genotype Frequency of D21S11 locus**

| Genotype | Observed Frequency (O) |
|---|---|
| 27,29 | 0.015 |
| 27,30 | 0.01 |
| 27,30.2 | 0.005 |
| 27,31 | 0.005 |
| 28,28 | 0.25 |
| 28,29 | 0.01 |
| 28,29.2 | 0.005 |
| 28,30 | 0.04 |
| 28,30.2 | 0.005 |
| 28,31 | 0.02 |
| 28,31.2 | 0.035 |
| 28,32.2 | 0.05 |
| 28,33.2 | 0.005 |
| 29,29 | 0.045 |
| 29,29.2 | 0.005 |
| 29,30 | 0.085 |
| 29,30.2 | 0.015 |
| 29,31 | 0.015 |
| 29,31.2 | 0.04 |
| 29,32 | 0.005 |
| 29,32.2 | 0.115 |
| 29,33.2 | 0.025 |
| 29,34.2 | 0.005 |
| 29.2,32.2 | 0.005 |
| 30,30 | 0.025 |
| 30,30.2 | 0.01 |
| 30,31 | 0.015 |
| 30,31.2 | 0.05 |
| 30,32.2 | 0.045 |
| 30,33.2 | 0.02 |
| 30.2,30.2 | 0.005 |
| 30.2,31.2 | 0.01 |

| Genotype | Observed Frequency (O) |
|---|---|
| 30.2,32.2 | 0.025 |
| 30.2,33.2 | 0.01 |
| 30.2,34 | 0.005 |
| 31,31 | 0.005 |
| 31,31.2 | 0.005 |
| 31,32.2 | 0.01 |
| 31,33.2 | 0.01 |
| 31.2,31.2 | 0.015 |
| 31.2,32 | 0.015 |
| 31.2,32.2 | 0.025 |
| 31.2,33.2 | 0.025 |
| 32.2,32.2 | 0.035 |
| 32.2,33.2 | 0.035 |
| 32.2,34.2 | 0.005 |
| 33.2,34.2 | 0.005 |

Homozygotes =31
Heterozygotes =169

**Table 2(iii) Genotype Frequency of D7S820 locus**

| Genotype | Observed Frequency (O) |
|---|---|
| 7,8 | 0.005 |
| 7,10 | 0.015 |
| 7,11 | 0.005 |
| 7,14 | 0.01 |
| 8,8 | 0.06 |
| 8,9 | 0.025 |
| 8,10 | 0.085 |
| 8,11 | 0.1 |
| 8,12 | 0.06 |
| 8,13 | 0.025 |
| 9,9 | 0.005 |
| 9,10 | 0.025 |
| 9,11 | 0.04 |
| 9,12 | 0.03 |
| 10.10 | 0.025 |
| 10,11 | 0.16 |
| 10,12 | 0.125 |
| 10,13 | 0.01 |
| 11,11 | 0.085 |
| 11,12 | 0.045 |
| 11,13 | 0.005 |
| 11,14 | 0.005 |
| 12,12 | 0.02 |
| 12,13 | 0.025 |
| 12,14 | 0.005 |

Homozygotes =39
Heterozygotes =161

**Table 2(iv) Genotype Frequency of CSF1PO locus**

| Genotype | Observed Frequency (O) |
|---|---|
| 7,11 | 0.005 |
| 9,10 | 0.015 |
| 9,11 | 0.02 |
| 9,12 | 0.035 |
| 10,10 | 0.045 |
| 10,11 | 0.17 |
| 10,12 | 0.13 |
| 10,13 | 0.025 |
| 10,15 | 0.005 |
| 11,11 | 0.055 |
| 11,12 | 0.21 |
| 11,13 | 0.06 |
| 11,14 | 0.005 |
| 12,12 | 0.16 |
| 12,13 | 0.045 |
| 12,14 | 0.015 |
| 13,13 | 0.005 |

Homozygotes = 53
Heterozygotes =147

**Table 2(v) Genotype Frequency of D3S1358**

| Genotype | Observed Frequency (O) |
|---|---|
| 13,16 | 0.005 |
| 14,15 | 0.01 |
| 14,16 | 0.035 |
| 14,17 | 0.015 |
| 14,18 | 0.01 |
| 15,15 | 0.1 |
| 15,16 | 0.205 |
| 15,17 | 0.145 |
| 15,18 | 0.06 |
| 15,19 | 0.01 |
| 16,16 | 0.1 |
| 16,17 | 0.15 |
| 16,18 | 0.045 |
| 16,19 | 0.005 |
| 17,17 | 0.04 |
| 17,18 | 0.06 |
| 18,18 | 0.005 |

Homozygotes =48
Heterozygotes =152

**Table 2(vi) Genotype Frequency of TH01 locus**

| Genotype | Observed Frequency (O) |
|---|---|
| 6,6 | 0.045 |
| 6,7 | 0.095 |
| 6,8 | 0.065 |
| 6,9 | 0.135 |
| 6,9.3 | 0.05 |
| 6,10 | 0.01 |
| 7,7 | 0.025 |
| 7,8 | 0.04 |
| 7,9 | 0.1 |
| 7,9.3 | 0.065 |
| 8,8 | 0.02 |
| 8,9 | 0.07 |
| 8,9.3 | 0.055 |
| 8,10 | 0.005 |
| 9,9 | 0.07 |
| 9,9.3 | 0.095 |
| 9,10 | 0.01 |
| 9.3,9.3 | 0.045 |

Homozygotes =41
Heterozygotes =159

**Table 2(vii) Genotype Frequency of D13S317 locus**

| Genotype | Observed Frequency (O) |
|---|---|
| 8,8 | 0.05 |
| 8,9 | 0.005 |
| 8,10 | 0.04 |
| 8,11 | 0.12 |
| 8,12 | 0.09 |
| 8,13 | 0.02 |
| 8,14 | 0.005 |
| 9,9 | 0.005 |
| 9,10 | 0.01 |
| 9,11 | 0.035 |
| 9,12 | 0.05 |
| 9,13 | 0.03 |
| 9,14 | 0.005 |
| 10,11 | 0.05 |
| 10,12 | 0.03 |
| 10,13 | 0.015 |
| 10,14 | 0.005 |
| 11,11 | 0.1 |
| 11,12 | 0.155 |
| 11,13 | 0.025 |
| 11,14 | 0.015 |
| 12,12 | 0.065 |
| 12,13 | 0.035 |
| 12,14 | 0.025 |
| 12,15 | 0.005 |
| 13,13 | 0.005 |
| 13,14 | 0.005 |

Homozygotes =45
Heterozygotes =155

**Table 2(viii) Genotype Frequency of D16S539 locus**

| Genotype | Observed Frequency (O) |
|---|---|
| 8,9 | 0.005 |
| 8,10 | 0.015 |
| 8,11 | 0.04 |
| 8,12 | 0.03 |
| 8,13 | 0.03 |
| 8,14 | 0.005 |
| 9,9 | 0.025 |
| 9,10 | 0.025 |
| 9,11 | 0.155 |
| 9,12 | 0.065 |
| 9,13 | 0.025 |
| 9,14 | 0.005 |
| 10,10 | 0.015 |
| 10,11 | 0.095 |
| 10,12 | 0.04 |
| 10,13 | 0.01 |
| 11,11 | 0.12 |
| 11,12 | 0.13 |
| 11,13 | 0.04 |
| 11,14 | 0.005 |
| 12,12 | 0.055 |
| 12,13 | 0.04 |
| 12,14 | 0,01 |
| 13,13 | 0.015 |

Homozygotes =46
Heterozygotes =154

**Table 2(ix) Genotype Frequency of D2S1338 locus**

| Genotype | Observed Frequency (O) |
|---|---|
| 16,17 | 0.005 |
| 16,18 | 0.01 |
| 16,23 | 0.02 |
| 17,18 | 0.015 |
| 17,19 | 0.005 |
| 17,20 | 0.02 |
| 17,21 | 0.005 |
| 17,22 | 0.015 |
| 17,23 | 0.025 |
| 17,26 | 0.005 |
| 18,18 | 0.035 |
| 18,19 | 0.055 |
| 18,20 | 0.03 |
| 18,21 | 0.015 |
| 18,22 | 0.03 |
| 18.23 | 0.055 |
| 18,24 | 0.04 |
| 18,25 | 0.025 |
| 19,19 | 0.015 |
| 19,20 | 0.045 |
| 19,21 | 0.025 |
| 19,22 | 0.035 |
| 19,23 | 0.045 |
| 19,24 | 0.01 |
| 19,25 | 0.02 |
| 19,26 | 0.01 |
| 20,20 | 0.025 |
| 20,21 | 0.01 |
| 20,22 | 0.015 |
| 20,23 | 0.035 |
| 20,24 | 0.025 |

| Genotype | Observed Frequency (O) |
|---|---|
| 20,25 | 0.025 |
| 20,26 | 0.005 |
| 21,22 | 0.01 |
| 21,23 | 0.02 |
| 21,24 | 0.015 |
| 21,25 | 0.005 |
| 21,28 | 0.005 |
| 22,23 | 0.04 |
| 22,24 | 0.005 |
| 22,25 | 0.01 |
| 22,27 | 0.005 |
| 23,23 | 0.05 |
| 23,24 | 0.035 |
| 23,25 | 0.015 |
| 23,26 | 0.005 |
| 24,25 | 0.02 |
| 25,25 | 0.005 |

Homozygotes =26
Heterozygotes =174

**Table 2(x) Genotype Frequency of D19S433 locus**

| Genotype | Observed Frequency (O) |
|---|---|
| 10,13 | 0.005 |
| 10,14 | 0.005 |
| 11,15 | 0.005 |
| 12,12 | 0.01 |
| 12,13 | 0.035 |
| 12,14 | 0.035 |
| 12,14.2 | 0.01 |
| 12,15 | 0.015 |
| 12,15.2 | 0.01 |
| 12,16 | 0.02 |
| 12,17 | 0.005 |
| 12.2,12.2 | 0.005 |
| 12.2,13 | 0.005 |
| 13,13 | 0.11 |
| 13,13.2 | 0.025 |
| 13,14 | 0.18 |
| 13,14.2 | 0.025 |
| 13,15 | 0.075 |
| 13,15.2 | 0.035 |
| 13,16 | 0.02 |
| 13,16.2 | 0.03 |
| 13.2,14.2 | 0.005 |
| 13.2,15 | 0.005 |
| 13.2,16 | 0.005 |
| 14,14 | 0.055 |
| 14,14.2 | 0.025 |
| 14,15 | 0.06 |
| 14,15.2 | 0.0.25 |
| 14,16 | 0.025 |
| 14,16.2 | 0.005 |
| 14,17 | 0.015 |

| Genotype | Observed Frequency (O) |
|---|---|
| 14.2,14.2 | 0.01 |
| 14.2,15 | 0.005 |
| 14.2,15.2 | 0.015 |
| 14.2,16 | 0.015 |
| 15,15 | 0.025 |
| 15,16 | 0.01 |
| 15,16.2 | 0.005 |
| 15.2,16 | 0.01 |
| 15.2,17 | 0.005 |
| 16,16 | 0.005 |
| 16.2,16.2 | 0.005 |

Homozygotes =45
Heterozygotes =155

**Table 2(xi) Genotype Frequency of vWA locus**

| Genotype | Observed Frequency (O) |
|---|---|
| 13,14 | 0.005 |
| 13,18 | 0.005 |
| 14,14 | 0.025 |
| 14,15 | 0.015 |
| 14,16 | 0.075 |
| 14,17 | 0.05 |
| 14,18 | 0.04 |
| 14,19 | 0.01 |
| 14,20 | 0.01 |
| 15,15 | 0.01 |
| 15,16 | 0.03 |
| 15,17 | 0.06 |
| 15,18 | 0.01 |
| 15,19 | 0.02 |
| 15,20 | 0.005 |
| 16,16 | 0.05 |
| 16,17 | 0.12 |
| 16,18 | 0.07 |
| 16,19 | 0.025 |
| 17,17 | 0.09 |
| 17,18 | 0.135 |
| 17,19 | 0.035 |
| 17,20 | 0.01 |
| 18,18 | 0.045 |
| 18,19 | 0.035 |
| 18,20 | 0.01 |
| 19,19 | 0.005 |

Homozygotes =45
Heterozygotes =155

**Table 2(xii) Genotype Frequency of TPOX locus**

| Genotype | Observed Frequency (O) |
|---|---|
| 6,8 | 0.005 |
| 7,8 | 0.005 |
| 7,11 | 0.005 |
| 8,8 | 0.18 |
| 8,9 | 0.105 |
| 8,10 | 0.07 |
| 8,11 | 0.24 |
| 8,12 | 0.03 |
| 9,9 | 0.035 |
| 9,10 | 0.03 |
| 9,11 | 0.08 |
| 9,12 | 0.01 |
| 10,10 | 0.005 |
| 10,11 | 0.065 |
| 11,11 | 0.105 |
| 11,12 | 0.03 |

Homozygotes =65
Heterozygotes =135

**Table 2(xiii) Genotype Frequency of D18S51 locus**

| Genotype | Observed Frequency (O) |
|---|---|
| 9,17 | 0.005 |
| 10,12 | 0.005 |
| 11,11 | 0.005 |
| 11,13 | 0.01 |
| 11,15 | 0.01 |
| 11,16 | 0.015 |
| 11,17 | 0.005 |
| 11,18 | 0.005 |
| 12,12 | 0.005 |
| 12,13 | 0.015 |
| 12,14 | 0.055 |
| 12,15 | 0.045 |
| 12,16 | 0.035 |
| 12,17 | 0.02 |
| 12,18 | 0.005 |
| 12,19 | 0.005 |
| 12,20 | 0.005 |
| 13,13 | 0.01 |
| 13,14 | 0.09 |
| 13,15 | 0.035 |
| 13,16 | 0.025 |
| 13,17 | 0.015 |
| 13,18 | 0.015 |
| 13,19 | 0.005 |
| 13,20 | 0.005 |
| 13,21 | 0.005 |
| 14,14 | 0.075 |
| 14,15 | 0.1 |
| 14,16 | 0.06 |
| 14,17 | 0.03 |
| 14,18 | 0.005 |

| Genotype | Observed Frequency (O) |
|---|---|
| 14,19 | 0.005 |
| 14,20 | 0.025 |
| 14,22 | 0.005 |
| 14.2,15 | 0.005 |
| 15,15 | 0.035 |
| 15,16 | 0.04 |
| 15,17 | 0.005 |
| 15,18 | 0.03 |
| 15,19 | 0.02 |
| 15,20 | 0.005 |
| 16,16 | 0.02 |
| 16,17 | 0.04 |
| 16,18 | 0.015 |
| 16,19 | 0.005 |
| 17,17 | 0.005 |
| 17,19 | 0.005 |
| 17,20 | 0.005 |
| 18,18 | 0.005 |

Homozygotes =32
Heterozygotes =168

**Table 2(xiv) Genotype Frequency of D5S818 locus**

| Genotype | Observed Frequency (O) |
|---|---|
| 9,11 | 0.065 |
| 9,12 | 0.01 |
| 9,14 | 0.005 |
| 10,10 | 0.01 |
| 10,11 | 0.065 |
| 10,12 | 0.065 |
| 10,13 | 0.05 |
| 10,14 | 0.005 |
| 11,11 | 0.175 |
| 11,12 | 0.215 |
| 11,13 | 0.105 |
| 12,12 | 0.065 |
| 12,13 | 0.12 |
| 12,14 | 0.01 |
| 13,13 | 0.03 |
| 13,14 | 0.005 |

Homozygotes =56
Heterozygotes =144

**Table 2(xv) Genotype Frequency of FGA locus**

| Genotype | Observed Frequency (O) |
|---|---|
| 18,23 | 0.005 |
| 18,24 | 0.005 |
| 19,19 | 0.015 |
| 19,20 | 0.035 |
| 19,22 | 0.005 |
| 19,23 | 0.005 |
| 19,24 | 0.03 |
| 19,26 | 0.005 |
| 20,20 | 0.015 |
| 20,21 | 0.02 |
| 20,22 | 0.03 |
| 20,23 | 0.055 |
| 20,24 | 0.06 |
| 20,25 | 0.03 |
| 20,26 | 0.005 |
| 21,21 | 0.005 |
| 21,22 | 0.075 |
| 21,23 | 0.05 |
| 21,24 | 0.015 |
| 21,25 | 0.02 |
| 21,26 | 0.005 |
| 21,27 | 0.005 |
| 22,22 | 0.005 |
| 22,23 | 0.055 |
| 22,24 | 0.04 |
| 22,25 | 0.04 |
| 22,26 | 0.01 |
| 22,27 | 0.005 |
| 22,32.2 | 0.005 |
| 23,23 | 0.065 |
| 23,24 | 0.06 |

| Genotype | Observed Frequency (O) |
|---|---|
| 23,25 | 0.03 |
| 23,26 | 0.025 |
| 23,27 | 0.01 |
| 24,24 | 0.015 |
| 24,25 | 0.075 |
| 24,26 | 0.03 |
| 25,25 | 0.02 |
| 25,26 | 0.015 |

Homozygotes =28
Heterozygotes =172

The most frequent allele types for each locus were

1. Allele **13** for D8S1179
2. Allele **29** for D21S11
3. Allele **11** for D7S820
4. Allele **12** for CSF1PO
5. Allele **16** for D3S1358
6. Allele **9** for TH01
7. Allele **11** for D13S317
8. Allele **11** for D16S539
9. Allele **23** for D2S1338
10. Allele **13** for D19S433
11. Allele **17** for vWA
12. Allele **8** for TPOX
13. Allele **14** for D18S51
14. Allele **11** for D5S818
15. Allele **23** for FGA

The most frequent genotypes for each locus were

1. Genotype **10,13** for D8S1179
2. Genotype **29,32.2** for D21S11
3. Genotype **10,11** for D7S820
4. Genotype for **11,12** CSF1PO

5. Genotype **15,16** for D3S1358
6. Genotype **6,9** for TH01
7. Genotype **11,12** for D13S317
8. Genotype **9,11** for D16S539
9. Genotype **18,19** and **18,23** for D2S1338
10. Genotype **13,14** for D19S433
11. Genotype **17,18** for vWA
12. Genotype **8,11** for TPOX
13. Genotype **14,15** for D18S51
14. Genotype **11,12** for D5S818
15. Genotype **21,22** and **24,25** for FGA

**Table 4: The observed ($H_{obs}$) heterozygosity and expected ($H_{exp}$) heterozygosity values for 15 loci**

| Allele | D8S1179 | D21S11 | D7S820 | CSF1P0 | D3S1358 | TH01 | D13S317 | D16S539 | D2S1338 | D195433 | VWA | TPOX | D18S51 | D5S818 | FGA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $H_{obs}$ | 0.875 | 0.62 | 0.805 | 0.735 | 0.755 | 0.795 | 0.775 | 0.77 | 0.87 | 0.775 | 0.775 | 0.675 | 0.84 | 0.72 | 0.86 |
| $H_{exp}$ | 0.855 | 0.816 | 0.799 | 0.725 | 0.737 | 0.794 | 0.796 | 0.784 | 0.89 | 0.807 | 0.804 | 0.705 | 0.847 | 0.725 | 0.86 |

No deviations from Hardy Weinberg Equilibrium were observed. There was no need to apply Bonferroni Correction (BC) value of 0.05/15 = 0.0033. This procedure is used to adjust probability (P) values when making multiple statistical tests. The BC was proposed to circumvent the problem that as the number of tests increase, the likelihood of a type I error i.e. concluding that a significant difference is present when it is not (17).

The average heterozygosity level for both observed and expected values were above the required 70%. It is, therefore concluded that these loci are suitable for human identification (18). The observed heterozygosity values were ranging from (0.62) D21S11 & (0.675) TPOX to (0.875) D8S1179 & (0.87) D2S1338. The expected heterozygosity values were ranging from (.706) TPOX to (.890) D2S1338. Bhuvnesh et al (19) carried out a study on Saraswat Brahmins community of Northern India and observed populations were in Hardy-Weinberg equilibrium and found that D2S1338, D18S51 and FGA are considered most informative markers for forensic testing. Further, D Dogra et al (20) observed in their study that all loci were in HWE except at locus FGA in Ramgariha Sikhs and locus D16S539 in Arora Sikhs.

The amount of heterozygosity at a locus in a sampled population is a measure of genetic variation. The higher the heterozygosity, the more variation at a given locus. The percentage of individuals who are heterozygous is higher at all the loci as compared to the percentage of individuals who are homozygous, which shows that there is greater heterozygosity and therefore less chance of a random match sampling.

## CONCLUSION

Hardy Weinberg Equilibrium describes the assumptions that must be present to consider a population in equilibrium. The law describes a situation in which a population is under no evolutionary change and the variability present in the population is preserved. The relationship between allele frequency and genotype frequency can be expressed as:

$P^2 + 2pq + q^2 = 1$ or $(p+ q)^2 = 1$ also $p + q = 1$

From this binomial expression it is clear that in a large random mating population, gene and genotype frequencies will remain constant; in other words, the population is genetically stable. This is evident from this data as no deviations from Hardy-Weinberg Equilibrium were observed. Also, all the 15 STR loci are highly polymorphic and suitable for human identification.

A database has thus been established for the loci D8S1179, D21S11, D7S820, CSF1PO, D3S1358, TH01, D13S317, D16S539, D2S1338, D19S433, vWA, TPOX, D18S51, D5S818, FGA of the 15 STR systems that have been shown to be useful in forensic DNA analysis.

## References


1. Jeffreys AJ, Wilson V and Thein SL Hypervariable minisatellite regions in human DNA, Nature 1985; 314: 67-73
2. Edwards A, Civileto A, Hammond HA, Caskey CT, DNA typing and genetic mapping with trimeric and tetrameric tandem repeats, Am J Hum Genet 1991; 49: 746-56
3. Edwards A, Hammond HA, Jin L, Caskey CT, Chkraborty R, Genetic variation at five trimeric and tetrameric repeat loci in four human population groups, Genomics 1992; 12: 241-53
4. Budow.le B, Moretti TR, Baumstark AL, Defenbaugh DA, Keys KM. Population data on the thirteen CODIS core short tandem repeat loci in African Americans US, Caucasians, Hispanics, Bahamians, Jamaicans and Trinidadians. J Forensic Science 1999; 44: 1277-86
5. Santos SMM, Budowle B, Smerick JB, Keys KM, Moretti TR. Portuguese population data on the sir short tandem repeat loci- CSF1PO, TPOX, TH01, D3S1358, vWA and FGA Forensic Sci Int 1996; 83: 229-35
6. Huang NE, Schhumm J, Budowle B. Chinese population data on three tetrameric short tandem repeat loci- HUMTH01, TPOX, and CSF1PO- derived using multiplex PCR and manual typing. Forensic Sci Int 1995; 71: 131-36
7. Turowska B, Sanak M. Frequency data on the loci vWA, FES, FPS, F13A01, TH01, TPOX and CSF1PO in a population from poland. Int J Legal Med 2000; 113: 123-25
8. Neuhuber F, Radacher M, Meisner N, Tutsch- Bauer E. Nine STR markers plus amelogenin (AmpFl STR. STR Profiler Plus): a forensic study in an Austrian population. Int J Legal Med 1999; 113: 60-62
9. Hantschel M, Hausmann R, Lederer T, Martus P , Betz P. Population genetics of nine short tandem repeat (STR): DNA typing using the AmpFl STR Profiler PCR amplification kit. Int J Legal Med 1999; 112: 393-95
10. Beleza S, ALves C, Reis F, Amorim A, Carracedo A, Gusmao L, 17 STR Data (AmpFL STR Identifiler and Powerplex 16 system) from Cabinda (Angola) Forensic Sci. Int 141 2004; 193-96
11. Auxbel FM, Brent R, Kingston, RE, Moore, DD, Seidman JG, Smith JA et. al., Current protocol in molecular biology, Wiley, New York- 1987
12. AmpFl STR Identifiler Plus PCR Amplification kit, User's manual, Applied Biosystems, Foster City, USA.
13. Weir BS. Genetic Data analysis II, Sunderland, MA: Sirauer, 1996.
14. Evett IW, Weir BS. Interpreting DNA Evidence. Sunderland, MA: Sunderland 1998.
15. Sham P. Statistics in human genetics. London: Arnold, 1998.
16. Richard Li, Forensic Biology, 2nd Edition- CRC Press.
17. Armstrong RA, when to use the Bonferroni Correction. Opthalmic Physiol Opt 2014; 34 502-08.
18. Demeter Jozef-Szabolcs, Allelic frequency of STR loci in some populations from Transylvania Summary of PhD Thesis; Cluj-Napoca, 2010.
19. Yadav B, Raina A, Dogra TD, Genetic proximity of Saraswat Brahmin community of Northern India based on Autosomal STR markers, J Bio Innov 2016; 5 764-85.
20. Dogra D, Shrivastava P, Chaudhary R, Gupta U and Jain T, Population Genetics for Autosomal STR Loci in Sikh Population of Central India, Hereditary Genet 4:142.

## What's New in This paper

### What is already known on this topic?

Genomic diversity at fluorescent labelled STR loci in few important populations of state of Uttar Pradesh & Maharashtra, India has been studied by Kashyap et al (2002) using Powerplex 16 system. They found the homozygosity values for all the 15 loci in studied population range between 0.053 & 0.999 following Hardy-Weinberg Equilibrium & based on the distribution of allele pattern these markers were found to be highly polymorphic & informative.

Similar studies were done on the populations from Central India by Shrivastava P. et al (2012) also shows that the 15 STR loci were highly informative & discriminating & all the studied populations were in Hardy-Weinberg Equilibrium.

Bhuvnesh *et.al.* carried out a study on Saraswat Brahmin community of Northern India & observed populations were in Hardy-Weinberg Equilibrium. Further D. Dogra *et. al.* observed in their study that all loci were in Hardy-Weinberg Equilibrium except at locus FGA in Ramgariha Sikh & locus D16S539 in Arora Sikh.

### What this study adds?

In this study Short Tandem Repeat analysis was done in Random Delhi NCR population sample using AmpFlSTR Identifiler plus PCR amplification kit.

No deviations from Hardy-Weinberg Equilibrium were observed which means that the variability present in the populations is preserved. The average heterozygosity level for both observed & expected values were above 70%. So, there loci were found to be suitable for human identification. The present study adds to the existing database & is in concordance with the studies already done.

- **Suggestions for further development**

.

The incredible power of DNA technology as an identification tool had brought a tremendous change in criminal justice system. DNA database is an informative resource for the Forensic DNA typing community with details on commonly used STR DNA markers.

Time & expense limit an examination of an individual's entire genome. So, DNA FP involving the study of the loci which are polymorphic or in other words different among the individuals. Population data must be compiled to an estimate of the frequency of each possible allele & genotype.

Hardy-Weinberg Equilibrium is the foundation of population genetics. Hardy-Weinberg Equilibrium is maintained as long as population size is large, mating is random & mutation selection & migration is negligible. If not allele frequencies & genotype frequencies may change from one generation to the next. By examining this genetic diversity evolutionary patterns can be inferred & variants contributing to the cause of common diseases can be identified.

Almost every nation now maintains a database Proposed DNA bill by Ministry of law, India will allow the govt. to establish a National DNA data Bank & use the data for various specified Forensic purposes.