

ALLELE FREQUENCIES OF 15 AMPFISTR IDENTIFILER LOCI IN THE NEPALESE POPULATION

Dinesh Kumar Jha*, **Luis Javier Martinez Gonzalez****, **Jiwan Prasad Rijal***,
Bhinu Shova Tuladhar* and **Nirajan Thapa Chhetri***

*National Forensic Science Laboratory, Khumaltar, Nepal.

**Laboratory of Genetic Identification, Department of Legal Medicine,
Toxicology and Psychiatry, Faculty of Medicine, University of Granada.

Abstract: Allele frequencies for the 15 STR loci in the AmpFISTR® Identifiler® and statistical parameters were estimated from a sample of 233 unrelated individuals from different area of Nepal. A total of 161 alleles were found, with corresponding allelic frequencies ranging from 0.0021 to 0.4077. The MP, PD, PIC, PE, TPI, Ho and He ranged from 0.0282 to 0.1511, 0.8489 to 0.9717, 0.6478 to 0.8700, 0.4546 to 0.7631, 1.7651 to 4.3148, 0.7167 to 0.8841 and 0.7009 to 0.8833 respectively. Deviations from the Hardy–Weinberg Equilibrium were observed for D16S539, D18S51, D21S11 and TPOX Markers. One new allele has been detected in the process.

Key words: Nepalese population; Autosomal STR markers; Allele frequencies; Matching probability; Power of discrimination; Power of exclusion; Hardy-Weinberg equilibrium.

INTRODUCTION:

Short tandem repeat (STR) form approximately 3% of the total human genome and occur on average in every 10,000 nucleotides [15]. STRs remained the mainstays in most of the forensic laboratories in the world for the last two decades, as these markers provide high statistical capability of discrimination and individualization [6] and thus established as widely used genetic markers for individual identity and paternity testing in forensic applications [9]. It is suitable for analyzing degraded, outmoded, and minute amounts of human DNA samples [10,2]. The usefulness of genetic markers for identity testing and paternity analysis is based on known allele frequencies for the genetic markers analyzed. With the growing number of laboratories, which use STR markers, more and more population data are reported from all over the world. It is well described the importance of knowing the allele distribution of the genetic markers used in forensic laboratories, in the population they will be applied [1]. Diversity of STR loci in some Nepalese population is available in the literature (13, 17, 20). The present study describes the allele frequencies for 15 short tandem repeat (STR) loci and statistical parameters of forensic interest from a sample of 233 individuals undergoing paternity testing from different regions of Nepal.

MATERIALS AND METHODOLOGY:

1. Samples: A total of 233 unrelated (as stated in their

identification form) Nepalese individuals casework samples received in National Forensic Science Laboratory, Nepal from different court, immigration purpose and personal concern for the meaning of DNA based paternity issues were taken as the source of the study samples. These samples are represents from the different districts (Fig. 1) of Nepal. The surnamewise included samples (alphabetically) are; Adhikari, Ahir, Ale Magar, Ansari, Aryal, Badu, Basnet, Bhandari, Bhatt, Bhujel, Bik, Bishwakarma, Biswas, Budha, Budhathoki, Chalawne, Chaudhary, Chetri, Chonzom, Dagaura, Dahal, Dangi, Dev, Devkota, Dharel, Dubey, Gautam, GC, Gelal, Ghale, Ghimire, Giri, Gurung, Husain, Jha, Karki, KC, Kewat,

Fig. 1: District representation of samples (indicated by black colour) studied for genetic variation at 15 STR loci in Nepalese population.



Table 1: Allele frequencies and associated statistical parameters of AmpFISTR® Identifiler® PCR Amplification Kit loci in Nepalese population (n = 233).

Alleles	D8S1179	D21S11	D7S820	CSFIPO	D3S1358	TH01	D13S317	D16S539	D2S1338	D19S433	vWA	TPOX	D18S51	D5S818	FGA
6	-	-	-	-	-	0.2017	-	-	-	-	-	-	-	-	-
7	-	-	0.0107	-	-	0.2017	0.0086	-	-	-	-	-	-	-	-
8	-	-	0.2425	0.0021	-	0.1073	0.1609	0.0408	-	-	-	0.4077	-	-	-
9	0.0214	-	0.0794	0.0386	-	0.3626	0.1309	0.2704	-	-	-	0.1566	0.0021	0.0515	-
9.3	-	-	-	-	-	0.1223	-	-	-	-	-	-	-	-	-
10	0.1481	-	0.1759	0.1888	0.0086	0.0042	0.1116	0.103	-	-	0.0043	0.0815	0.0064	0.1524	-
10.1	-	-	<u>0.0021</u>	-	-	-	-	-	-	-	-	-	-	-	-
11	0.0601	-	0.2489	0.2939	-	-	0.2618	0.2704	-	0.0043	0.0021	0.3197	0.0086	0.3326	-
11.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
12	0.1223	-	0.2060	0.3841	-	-	0.2425	0.2038	-	0.0644	-	0.0300	0.0558	0.2554	-
12.2	-	-	-	-	-	-	-	-	-	0.0064	-	-	-	-	-
13	0.1888	-	0.0322	0.0879	-	-	0.0686	0.0965	-	0.2661	0.0021	0.0043	0.2038	0.1901	-
13.2	-	-	-	-	-	-	-	-	-	0.0558	-	-	-	-	-
14	0.1867	-	0.0021	0.0021	0.0322	-	0.0150	0.0086	0.0043	0.2403	0.1309	-	0.2296	0.0172	-
14.2	-	-	-	-	-	-	-	-	0.0021	0.0987	-	-	-	-	-
15	0.1974	-	-	0.0021	0.3176	-	-	0.0021	0.0043	0.1287	0.0407	-	0.1523	-	-
15.2	-	-	-	-	-	-	-	-	-	0.0815	-	-	-	-	-
16	0.0601	-	-	-	0.3541	-	-	-	0.0086	0.0193	0.2511	-	0.1051	-	-
16.2	-	-	-	-	-	-	-	-	-	0.0193	-	-	-	-	-
17	0.0150	-	-	-	0.1845	-	-	-	0.0558	0.0043	0.2618	-	0.0708	-	-
17.2	-	-	-	-	-	-	-	-	-	0.0064	-	-	-	-	-
18	-	-	-	-	0.0987	-	-	-	0.1351	-	0.2103	-	0.0279	-	0.0214
19	-	-	-	-	0.0043	-	-	0.0021	0.1716	0.0021	0.0794	-	0.0579	-	0.0708
19.2	-	-	-	-	-	-	-	0.0021	-	-	-	-	-	-	-
20	-	-	-	-	-	-	-	-	0.0901	-	0.0150	-	0.0365	-	0.0858
20.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0043
21	-	-	-	-	-	-	-	-	0.0429	-	0.0021	-	0.0172	-	0.1202
21.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0064
22	-	-	-	-	-	-	-	-	0.0515	-	-	-	0.0150	-	0.1459
22.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0064
23	-	-	-	-	-	-	-	-	0.1824	-	-	-	0.0043	-	0.1974
23.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0086
24	-	0.0021	-	-	-	-	-	-	0.1437	-	-	-	0.0043	-	0.1266
24.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0086
25	-	-	-	-	-	-	-	-	0.0837	0.0021	-	-	0.0021	-	0.0965
25.2	-	-	-	-	-	-	-	-	0.0193	-	-	-	-	-	0.0064
26	-	0.0021	-	-	-	-	-	-	0.0043	-	-	-	-	-	0.0708
27	-	0.0172	-	-	-	-	-	-	-	-	-	-	-	-	0.0172
28	-	0.1073	-	-	-	-	-	-	-	-	-	-	-	-	0.0064
28.2	-	0.0107	-	-	-	-	-	-	-	-	-	-	-	-	-
29	-	0.2339	-	-	-	-	-	-	-	-	-	-	-	-	-
29.2	-	0.0064	-	-	-	-	-	-	-	-	-	-	-	-	-
30	-	0.2360	-	-	-	-	-	-	-	-	-	-	-	-	-
30.2	-	0.0300	-	-	-	-	-	-	-	-	-	-	-	-	-
31	-	0.0601	-	-	-	-	-	-	-	-	-	-	-	-	-
31.2	-	0.0837	-	-	-	-	-	-	-	-	-	-	-	-	-
32	-	0.0129	-	-	-	-	-	-	-	-	-	-	-	-	-
32.2	-	0.1223	-	-	-	-	-	-	-	-	-	-	-	-	-
33	-	0.0021	-	-	-	-	-	-	-	-	-	-	-	-	-
33.2	-	0.0558	-	-	-	-	-	-	-	-	-	-	-	-	-
34.2	-	0.0107	-	-	-	-	-	-	-	-	-	-	-	-	-
35.2	-	0.0064	-	-	-	-	-	-	-	-	-	-	-	-	-
MP	0.0456	0.0418	0.0763	0.1270	0.1258	0.0936	0.0645	0.0848	0.0308	0.0529	0.0694	0.1511	0.0395	0.0994	0.0282
PD	0.9544	0.9581	0.9237	0.8729	0.8741	0.9063	0.9355	0.9152	0.9691	0.9471	0.9305	0.8489	0.9604	0.9005	0.9717
PIC	0.8266	0.8310	0.7678	0.6739	0.6824	0.7242	0.7862	0.7592	0.8623	0.8103	0.7693	0.6478	0.8419	0.7230	0.8700
PE	0.6280	0.5491	0.6199	0.4685	0.4826	0.4546	0.6280	0.5879	0.7544	0.6444	0.4970	0.4615	0.5958	0.5339	0.7631
TPI	2.7093	2.1981	2.6477	1.8203	1.8790	1.7651	2.7093	2.4270	4.1607	2.8414	1.9416	1.7923	2.4787	2.1182	4.3148
Ho	0.8154	0.7725	0.8111	0.7253	0.7339	0.7167	0.8154	0.7940	0.8798	0.8240	0.7425	0.7210	0.7983	0.7639	0.8841
He	0.8475	0.8495	0.8001	0.7226	0.7304	0.7622	0.8139	0.7922	0.8769	0.8321	0.8005	0.7009	0.8590	0.7631	0.8833
p	0.2930	0.0075	0.6868	0.1845	0.0669	0.1417	0.8063	0.0015	0.6531	0.1697	0.2584	0.0104	0.0035	0.4042	0.8172

MP: Matching probability, PD: Power of Discrimination, PIC: Polymorphism information content, PE: Probability of paternity exclusion, TPI: Typical Paternity Index, Ho: Observed Heterozygosity, He: Expected Heterozygosity, p: Hardy-Weinberg equilibrium. The most common alleles are indicated in bold and new allele is underlined

Table 2: Comparison of allele (the most common allele/the least allele) within Nepalese and some other populations of Asia.

Locus	Nepalese (this study)	Nepalese (13)	Nepalese (17)		Bhutanese (14)	Lasa Tibetan (16)	Philippino (5)	Korean (12)	Japanese (19)
			Sherpa	Kathmandu					
D8S1179	15/17	14/9	15/11	14/8,9	13/18	13/17,18	13/12,17	13/9	13/8
D21S11	30/24, 26,33	29,30/35.2	30/29.2	30/26,35.2	29/33.2	29/34.2,35	30/26,30.2, 34.2,36.2	30/34,35	30/24
D7S820	11/10.1,14	11/8.1	11/13	11/7	11/14	11/14	11/13	11/7	11/7
CSF1PO	12/8,14,15	12/7	12/15	12/15	12/7,8	11/7,8	12/7,8,14	12/7	12/6
D3S1358	16/19	16/20	16/19	16/19	16/13	15/13	17/19	15/12	15/13,20
THO1	9/10	9/10	9/10	9/10	9/10	9/9.3	9/5	9/8	9/4
D13S317	11/7	11/7	10/14	11/7,10.1	11/7,15	12/14	11/14	11/7	8/15
D16S539	9,11/15, 19, 19.2	9/15	12/14	11/6	11/15	12/14	9/7	9/15	9/15
D2S1338	23/14.2	19/27	23/26	23/21	23/19.3	23/26	19/26	19/19.2	19/15
D19S433	13/19	13/11.2,11.3, 13.1	13/16.2	13/12.2,17	14/11.2	14/12.2,17	13/10	13/11,17.2	14/9,11.2, 17.2,19.2
vWA	17/11, 13, 21	17/12, 13,22	16/15,20	17/12	17/21	17/20	18/13	17/20	17/12,13, 21,22
TPOX	8/13	8/14	8/10	8/7	8/13	17/20	8/10	8/13	8/13
D18S51	14/9, 25	14/28	13/11, 21,22	14/22,23	13/10,24	13/13.2,17.2, 22	14/21,24	14/8,10	14/14.2,25
D5S818	11/14	11/8	11/7,8	11/14	11/15	11/14	10/14	11/8	11/7
FGA	23/20.2	24/17	24/23.2	22/24.2	23/17, 27.2	23/20.2,21.2, 28.2	22/23.2,27	23/20,23	23/21.2,23.2, 26.2,28,29

Khadgi, Khatri, Kohar, Koirala, Kumal, Kurmi, Lama, Lonia, Madhu, Magar, Majhi, Malla, Manandhar, Mishra, Miya, Murao, Nakarmi, Negi, Ojha, Palkyi, Pandey, Pant, Pariyar, Pasarwal, Paudel, Pokhrel, Pradhan, Pudasaini, Pun, Purja, Rai, Rajbanshi, Rana, Raskoti, Regmi, Rijal, Rimal, Sah, Sapkota, Sarki, Shah (mountain), Shahi, Shakya, Sharma (mountain), Sharma (Terai), Shrestha, Subedi, Sudi, Tamang, Thakur (terai), Thapa, Thapa Magar, Tsomo, Udas, Upadhyay, Uprety and Yadav.

2. DNA extraction, PCR amplification and genotypic determination: DNA samples were collected and stored as bloodstains on FTA classic card (Whatman). DNA amplification were carried out by using the AmpFISTR® Identifiler® PCR Amplification Kit (Applied BioSystems, Foster City, CA, USA) containing the 15 Autosomal STR markers (D8S1179, D21S11, D7S820, CSF1PO, D3S1358, THO1, D13S317, D16S539, D2S1338, D19S433, vWA, TPOX, D18S51, D5S818 and FGA) with amelogenin marker for gender determination. The standard protocol was followed for DNA purification, PCR amplification and Genotype determination^[11].

3. Data analysis: The Statistical parameters of population genetics and forensic utility of these markers was evaluated by calculating the allele frequencies, Matching probability (MP), polymorphism information content (PIC), Typical paternity index (TPI), power of discrimination (PD), and power of exclusion (PE) using PowerStats software v. 1.2 (Promega Corporation, USA) ^[18]. The population's genetic structure deviation from Hardy–Weinberg equilibrium (*p*), observed heterozygosity (Ho) and expected heterozygosity (He) was calculated using methods implemented in the Arlequin V3.5 software [8]

4. Quality control: Positive control DNA and allelic ladder provided in AmpFISTR® Identifiler® PCR Amplification Kit were used for the genotyping analyses. Some samples run in duplicate (whole process, DNA extraction, PCR and

genotyping). The National Forensic Science Laboratory, Nepal has been successfully participated in the GITAD – AB Indo-Pacific DNA Quality control exercise (2006, 2008 and 2009) annual proficiency test.

RESULTS AND DISCUSSION:

The observed allele frequencies and statistical parameters based on the 15 STR loci in Nepalese population are summarized in Table 1. A total of 161 alleles at these 15 STR loci were found with corresponding allelic frequencies ranging from 0.0021 to 0.4077 in the Nepalese population. The number of alleles varied from 6 (THO1, TPOX and D5S818) to 17 (D21S11, D18S51 and FGA). A new allele (allele 10.1 of D7S820 locus) that was not included in the allelic ladder provided in AmpF–/STR® Identifiler® kit was observed. These 15 STR markers are found to be highly polymorphic in the populations.

MP: Matching probability, PD: Power of Discrimination, PIC: Polymorphism information content, PE: Probability of paternity exclusion, TPI: Typical Paternity Index, Ho: Observed Heterozygosity, He: Expected Heterozygosity, *p*: Hardy-Weinberg equilibrium. The most common alleles are indicated in bold and new allele is underlined

The values of the matching probability (MP), the power of discrimination (PD), the polymorphism information content (PIC), the power of exclusion (PE), Typical Paternity Index (TPI), the observed heterozygosity (Ho) and the expected heterozygosity (He) are from 0.0282 (FGA) to 0.1511 (TPOX), 0.8489 (TPOX) to 0.9717 (FGA), 0.6478 (TPOX) to 0.8700 (FGA), 0.4546 (THO1) to 0.7631 (FGA), 1.7651 (THO1) to 4.3148 (FGA), 0.7167 (THO1) to 0.8841 (FGA) and 0.7009 (TPOX) to 0.8833 (FGA) respectively.

Deviations from the Hardy–Weinberg Equilibrium (*p* < 0.05) were observed for D16S539 (*p* = 0.0015), D18S51 (*p* = 0.0035), D21S11 (*p* = 0.0075) and TPOX (*p* = 0.0104) Markers. The values of combined power of discrimination and

probability of matching for all 15 STR loci are $1-2,27078089046415E-18$ and $2,27078E-18$ (1 in $4,40377E+17$) respectively. The high value of combined matching probability of 15 STR loci examined in the present work, specifically designed for population genetics studies, also turned out to be suitable for general forensic applications. The comparison of our results within some Nepalese and other populations of Asia at the same loci are mentioned in Table 2.

Previous studies (4, 7) suggested that 100-150 tested individuals per population strength likelihood calculation. Collecting information from more samples usually adds to the precision of the obtained results (3) and considering the size of recent Nepalese population, the total number of unrelated individuals ($n = 233$) tested in this case is adequate for reliable estimation of the frequencies of major alleles for any DNA locus.

CONCLUSION:

The allele frequencies presented for the 15 STRs included in this study will be a useful tool for forensic identification and parentage testing in the Nepalese population.

ACKNOWLEDGEMENT:

We are thankful to Mr. Basant Raj Pokhrel, Assistant Scientific Officer of National Forensic Science Laboratory for his valuable help during the laboratory work.

REFERENCES:

1. Abovich, M., Arellano, A., Szocs, A., Alcazar, D., Cabeller S., Rodriguez Cardozo MB. 2009. Allele frequencies of 15 STRs loci in an Argentine population sample. *Forensic Science International: Genetics Supplement Series* **2**: 369-370
2. Brinkmann, B. 1992. The use of STRs in stain analysis. In: *Proceedings from the third international symposium on human identification*. Promega Corporation, Madison, USA. 357-373
3. Butler, JM. 2005. *Forensic DNA typing: Biology, technology and genetics of STR markers*. 2nd ed. London: Elsevier Academic Press.
4. Chakraborty, R. 1992. Sample size requirements for addressing the population genetic issues of forensic use of DNA typing. *Hum Biol.* **64**:141-59. Medline:1559686.
5. De Ungria, MC., Roby, RK., Tabbada, KA., Rao-Coticone, S., Tan, MMM., Hernandez, KN. 2005. Allele frequencies of 19 STR loci in a Philippine population generated using AmpFISTR multiplex and ALF singleplex systems *Forensic Science International.* **152**: 281-284.
6. Edwards, A., Hammond, HA., Jin, L., Caskey, CT., Chakraborty, R. 1992. Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics.* **12**: 241-253.
7. Evett, IW., Gill, P. 1991. A discussion of the robustness of methods for assessing the evidential value of DNA single locus profiles in crime investigations. *Electrophoresis.* **12**:226-30. Medline: 2040269.
8. Excoffier, L., H.E. L., Lischer. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour.* 2010 May. **10** (3):564-7.
9. Gill, P., Werrett, DJ., Budowle, B., Guerrieri, R. 2004. An assessment of whether SNPs will replace STRs in national DNA databases—joint considerations of the DNAworking group of the European Network of Forensic Science Institutes (ENFSI) and the ScientificWorking Group on DNA Analysis Methods (SWGAM). *Sci Justice.* **44**:51-53.
10. Hammond, HA., Jin, L., Zhong, Y., Caskey, CT., Chakraborty, R. (1994) Evaluation of 13 short tandem repeat loci for use in personal identification applications. *Am J Hum Genet.* **55**:175-189.
11. Jha, DK., Rijal, JP., Chhetri, NT. 2010. Nepalese null AMELY males and their Y-Haplotypes. *Scientific World.* **8**: 97 - 101.
12. Kim, YL., Hwang, JY., KimYJ. et al. 2003. Allele frequencies of 15 STR loci using AmpF/STR Identifiler kit in a Korean population, *Forensic Sci. Int.* **136**: 92-95.
13. Kraaijenbrink, T., van Driem, GL., Opgenort, JRML., Tuladhar, NM., de Knijff, P. 2007. Allele frequency distribution for 21 autosomal STR loci in Nepal *Forensic Science International.* **168**: 227-231.
14. Kraaijenbrink, T., van Driem, GL., Tshering of Gaselo, K., de Knijff, P. 2007. Allele frequency distribution for 21 autosomal STR loci in Bhutan. *Forensic Science International.* **170**: 68-72.
15. Lander, ES., Linton, LM., Birren, B. et al 2001. Initial sequencing and analysis of the human genome. *Nature.* **409**: 860-921.
16. Li, KL., Jianmin, Z., Kai, L., Xiaosong, Li. 2007. Allele frequencies of 15 STR loci of Tibetan lived in Tibet Lassa. *Forensic Science International* **168**: 236-240.
17. Ota, M., Droma, Y., Basnyat, B., Katsuyama, Y., Asamura, H., Sakai, H., Fukuhsima, H. 2007. Allele frequencies for 15 STR loci in Tibetan populations from Nepal. *Forensic Science International.* **169**: 234-238.
18. Tereba, A. 1999. Tools for analysis of population statistics, *Profiles in DNA.* **3**: 14-16. <http://www.promega.com-geneticidtools/>.
19. Tie, J., Wang, X., Oxida, S. 2006. Genetic Polymorphisms of 15 STR Loci in a Japanese Population. *J Forensic Sci.* **51**:188-189.
20. Tuladhar, BS., Haslindawaty, N., Nada, B., Panneerchelvam, S., Norazmi, MN. 2006. Allele Frequency Data for Nine STRs Polymorphism in a Gurkha Population of Malaysia. *J Forensic Sci.* **51**: 1205 - 1206.

