Evaluation of five new European Standard Set (ESS) STR loci in a Bangladeshi Population Sample

Anik Samiur Rahman[a], Sudipta Arka Das[a] , Mahamud Hasan[b], Sharif Akhteruzzaman[a*]

[a] *Department of Biochemistry & Molecular Biology, University of Dhaka, Dhaka-1000, Bangladesh*
[b] *National Forensic DNA Profiling Laboratory, Dhaka Medical College, Dhaka-1000, Bangladesh*

ABSTRACT: Genetic polymorphism of five new European Standard Set (ESS) STR loci was studied in a sample population of Bangladesh using Investigator ESSplex kit (QIAGEN). Allele frequencies and forensic efficiency parameters were evaluated in a sample of 88 unrelated individuals. No significant deviation from Hardy-Weinberg equilibrium was observed for the loci. The combined probability of match (PM) and combined power of exclusion (PE) for the studied loci was calculated to be 1.22×10^{-6} and 0.98306 respectively. However, combining this data with data from other STR loci typically incorporated in most of the commercial kits using our population sample gave very high level of individualisation. Typing of these five ESS loci therefore may provide a useful addition to the previously established set of autosomal STR loci in Bangladeshi population.

Keywords: allele frequency, discrimination power, polymorphism, locus

Introduction

Short tandem repeat (STR) based DNA typing has now become the gold standard in most of the forensic DNA laboratories in the world. STRs are widespread in human genome and occur on average in every 10,000 nucleotides [1]. Due to their smaller size, high heterozygosity, low mutation rate and multiplexing capability, they have been very useful markers in personal identification, parentage testing and population genetic studies [2-4]. Since its introduction in mid 1990s, STRs remained the mainstays in most of the forensic laboratories as these markers provide high statistical capability of discrimination and individualisation [5].

Over the past two and half decades, a set of core loci has been established by the human identity testing community [6, 7]. A panel of 10-15 loci is invariably analysed in most laboratories which are available as commercial kits from different sources [8-10]. The current state-of-the-art DNA typing method using STRs has been proved to be very satisfactory and being adopted by more and more laboratories across the globe. The benefit of DNA typing combined with continued technological advances have led to the necessity of establishing DNA databases. Countries either initiated or established their DNA databases solely based on batteries of STR loci. Because of the difficulty and expense of changing a well-established technology, STR technology is therefore expected to be the predominant DNA typing method in the next decade. Other systems such as, SNPs, mtDNA or InDels will continue to develop, but their use will probably be not a replacement, rather supplement of the current STR technology.

Today, over 120 countries in the world are known to use DNA profiling in their criminal investigations and more than 50 countries that are known to have their national DNA Database. These databases have various organisational structures and defined set of loci depending on the legislative background of the implementing country [11]. Over the last few years, countries in Europe and the United States have been working to expand core set of genetic markers to be used for routine forensic investigation. This expansion will help to establish a legal basis for exchange of DNA profiles between countries and reduce the potential for adventitious matches as the database grows bigger. The FBI laboratory has already announced to expand its 13 CODIS loci to at least 18 loci in near future [12]. The ESS of STR loci has already been increased from seven to 12 in 2009 [13]. Considering this growing need, commercial STR kit manufacturers have recently released GlobalFiler™ (Life Technologies) and PowerPlex® Fusion (Promega Corporation) STR typing kit with

increased number of loci. In this study, we report the allele frequencies and other forensic parameters to evaluate the effectiveness of the five new ESS loci in Bangladeshi population.

Materials and Methods

Population

Peripheral blood samples were collected from 88 individuals randomly selected form the mainstream Bengali population. Being ethnically homogeneous, the Bengalis comprise 98% of the total Bangladesh population. The remainder is mostly indigenous ethnic minorities living in the southeastern, north-central and northeastern part of Bangladesh.

DNA Extraction

Genomic DNA was extracted using Chelex-100 method as described by Walsh *et. al.* [14]. Extracted DNA was quantified by NanoDrop-1000 (NanoDrop Technologies, Inc, Wilmington DE 19810, USA).

PCR amplification and STR typing

Approximately 1-2 ng of template DNA was used for each PCR amplification process. A total of 15 autosomal STR loci included in Investigator ESSplex kit (QIAGEN, Germany) was amplified using Veriti™ thermal cycler (Applied Biosystems). Thermal cycling parameters were employed according to the protocol provided by the manufacturer. The PCR amplified products were separated by capillary electrophoresis on ABI Prism® 3100 *avant* Genetic Analyzer (Applied Biosystems) using POP-4 polymer and data collection software version 2.0. Peak sizing and genotype assignments were done by GeneMapper ID software version 3.2.

Analysis of data

The Hardy-Weinberg equilibrium analysis, calculation of observed and expected heterozygosity was performed with Arlequine Software version 3.11 [15, 16]. Allele frequencies at each locus and other statistical parameters of forensic efficiency were calculated by using PowerStat Microsoft Excel Workbook version 1.2 [17].

Results and Discussion

The allele frequencies of the five new European Standard Set (ESS) STR loci (D1S1656, D2S441, D10S1248, D12S391 and D22S1045) are presented in Table 1.

Table 1: Allele frequencies and forensic efficiency parameters for D1S1656, D10S1248, D22S1045, D12S391, D2S441 STR loci in a population sample from Bangladesh (n = 88)

Allele	D1S1656	D10S1248	D22S1045	D12S391	D2S441
6	-	-	-	-	-
7	-	-	-	-	-
8	0.0682	-	-	-	-
9	-	-	-	-	0.0114
9.3	-	-	-	-	-
10	0.0057	-	0.0057	-	0.3125
11	0.1477	0.0227	0.2500	-	0.3636
11.3	-	-	-	-	0.0568
12	0.1080	0.0341	-	-	0.0682
12.3	-	-	-	-	
13	0.1136	0.1989	-	-	0.0057
13.3	-	-	-	-	-
14	0.0966	0.2500	0.0568	-	0.1591
14.3	-	-	-	-	-
15	0.1648	0.2727	0.3807	0.0057	0.0170
15.3	0.0114	-	-	-	-
16	0.1534	0.1761	0.1591	0.0057	0.0057
16.3	0.0170	-	-	-	-
17	0.0739	0.0455	0.1364	0.1534	-
17.3	0.0114	-	-	0.0170	-
18	0.0170	-	0.0114	0.2386	-
18.3	0.0114	-	-	0.0114	-
19	-	-	-	0.1875	-
19.3	-	-	-	0.0057	-
20	-	-	-	0.1023	-
21	-	-	-	0.0966	-
22	-	-	-	0.0795	-
23	-	-	-	0.0682	-
24	-	-	-	0.0227	-
25	-	-	-	-	-
26	-	-	-	0.0057	-

The overall distribution of the allele frequencies for each system was found to be in fair accordance with Hardy-Weinberg equilibrium rule by compliance with the exact test [15]. Two rare alleles (allele 8 at D1S1656, allele 19.3 at D12S391) were observed, off which allele 8 at D1S1656 is reported only once in STRbase [18]. Allele 15 was found to be the most frequent allele at D1S1656, D10S1248 and D22S1045 locus. Whereas at D12S391 and D2S441 loci, the most frequent allele was 18 and 11 respectively.

Forensic efficiency parameters such as, Power of discrimination (PD), Observed and expected heterozygosity (Ho & He), Polymorphism information content (PIC), Probability of match (PM) Power of exclusion (PE), and Typical paternity index (TPI) were calculated for the studied loci. The results are presented in Table 2. Out of the five loci, D12S391 and D1S1656 were the most informative locus based on observed heterozygosity (0.8068 and 0.8977 respectively). The combined probability of match (PM) and combined power of exclusion for the studied loci were 1.22×10^{-6} and 0.98306, respectively. However, combining this data with data from other STR loci typically incorporated in most of the commercial kits using our population sample gave very high level of individualization, Table 3.

Table 2: Forensic efficiency parameters for D1S1656, D10S1248, D22S1045, D12S391, D2S441 STR loci in a population sample from Bangladesh

Allele	D1S1656	D10S1248	D22S1045	D12S391	D2S441
Ho	0.8068	0.7273	0.6932	0.8977	0.6250
He	0.8875	0.7933	0.7496	0.8574	0.7407
PM	0.0300	0.0775	0.0994	0.0475	0.1116
PD	0.9700	0.9225	0.9006	0.9525	0.8884
PIC	0.8709	0.7562	0.7063	0.8361	0.6947
PE	0.6118	0.4717	0.4178	0.7908	0.3220
TPI	2.5882	1.8333	1.6296	4.8889	1.3333
p	0.6163	0.3073	0.7364	0.5328	0.1152

Ho, observed heterozygosity; *He*, expected heterozygosity; PM, probability of match; PD, power of discrimination; PIC, polymorphism information content; PE, power of exclusion of paternity; TPI, typical paternity index; *p*, Hardy-Weinberg equilibrium exact test.

Table 3: Calculated combined PM and PE using data obtained from this study and other published data for this population

Loci set	Total no of loci	Combined PM	Combined PE	Reference
Five new ESS loci	5	1.22×10^{-6}	0.983059614	This study
Five new ESS loci + ISSOL	12	2.27×10^{-15}	0.999989725	[20]
Five new ESS loci + SGM Plus loci	15	7.18×10^{-19}	0.999999706	[19]
Five new ESS loci + Identifiler™ loci	20	1.53×10^{-24}	0.999999987	Unpublished data of this laboratory
Five new ESS loci + PowerPlex® 16 loci	20	2.51×10^{-24}	0.999999994	[20]

Most of the forensic DNA laboratories engaged in human DNA typing around the world invariably use 10-15 STR loci, plus amelogenin as the sex typing marker. All these loci are incorporated in commercial kits available from many different sources. Of these commercial kits, seven loci (THO1, vWA, FGA, D21S11, D3S1358, D8S1179, and D18S51) are typically incorporated in most of the commercial kits available today. The first ESS selected the above seven loci which are the same as the INTERPOL standard set of loci (ISSOL). In view of the tremendous growth of many national DNA database, concern over adventitious matches between profiles and difficulties in international data sharing led to the European Network of Forensic Science Institutes (ENFSI) and European DNA Profiling Group (EDNAP) to extend the ESS by adopting five additional loci [18, 19]. These five loci are currently available in ESSplex (QIAGEN), NGM (Applied Biosystems), Powerplex ESX and ESI 16 (Promega) kit. The FBI laboratories in US and INTERPOL have also announced to expand their existing loci in near

future [12]. This study, evaluated the allele frequencies and forensic efficiency parameters for the five new ESS STR loci in a Bangladeshi population sample using ESSplex kit. The obtained data from the present study demonstrated that the five new ESS STR loci may also provide a very useful addition to the previously established set of autosomal STR loci in Bangladeshi population.

Conclusion

Typing of the five new ESS loci may provide a useful adjunct to the previously studied autosomal STR loci evaluated for this population in the past.

References

1. Lander ES, Linton LM, Birren B. *et al*. (2001)Initial sequencing analysis of the human genome. *Nature*, 409(6822), 860-921.
2. Hammond HA, Jin L, Zhong CT, Caskey CT, Chakraborty R. (1994). Evaluation of 13 short tandem repeats loci for use in personal identification applications. *American Journal of Human Genetics*, 55, 175-189.
3. Edward A, Civitello A, Hammond CT, Caskey CT. (1991). DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *American Journal of Human Genetics*, 49, 746-756.
4. Rosenberg NA, Pritchard JK, Weber JL, Cann HK, Kidd KK, Zhivorovsky LA, Feldman MW. (2002). Genetic Structure of human populations. *Science*, 298, 2381-2385.
5. Myers BA, King JL, Budowle B. (2012). Evaluation and comparative analysis of direct amplification of STRs using PowerPlex® 18D and Identifiler® Direct systems. *Forensic Science International: Genetics*, 6(5), 640-5.
6. Butler JM. (2007). Short tandem repeat typing technologies used in human identity testing. *Biotechniques*, 43(4), 2-5.
7. Butler JM. (2006). Genetics and genomics of core short tandem repeats loci used in human identity testing. *Journal of Forensic Science*, 51(2), 253-265.
8. Cotton EA, Allsop RF, Guest JL, Frazier RRE, Koumi P, Callow IP, Seager A, Sparkes RL. (2000) Validation of the AMPFISTR® SGM Plus system for use in forensic casework. *Forensic Science International*, 112, 151-161.
9. Krenke BE, Tereba A, Anderson SJ, Buel E, Culhane S, Finis CJ, Tomsey CS, Zachetti JM, Masibay A, Rabbach DR, Amiott EA, Sprecher CJ. (2002) Validation of 16-locus fluorescent multiplex system. *Journal of Forensic Scienc*, 47, 773–785.
10. Zupanic PI, Gornjak PB, Balazic J, Zupanc T, Stefanic B. (2012). Highly efficient nuclear DNA typing of the World War II skeletal remains using three new autosomal short tandem repeat amplification kits with the extended European Standard Set of loci. *Croatian Medical Journal*, 53 (1), 17-23.
11. Schneider PM, Martin PD. (2001). Criminal DNA databases: The European situation. *Forensic Science International*, 119, 232-238.
12. Hares DR (2012). Expanding the CODIS core loci in the United States. *Forensic Science International:* Genetics, 6, e52-e54.
13. Schneider PM. (2009). Expansion of the European Standard Set of DNA database loci –The current situation. *Profiles in DNA*. 12(1), 6-7.
14. Walsh PS, Metzger DA, Higuchi R. (1991). Chelex 100 as a simple medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques*, 10, 413-506.
15. Guo SW, Tompson EA. (1992). Performing the exact test of Hardy-Weinberg equilibrium proportion for multiple alleles. *Biometrics*, 48, 361-372.
16. Powerstats version 1.2, Promega Corporation, assessed at: http://www.promega.com/geneticidtools/powerstats/
17. Laval LG, Schneider S. (2005). Arlequin ver. 3.0: An integrated software package for population genetic data analysis. *Evolutionary Bioinformatics Online*. 1, 47-50.
18. Short Tandem Repeat DNA Internet Database (NIST Standard Reference Database SRD 130): http://www.cstl.nist.gov/biotech/strbase.
19. Ferdous A, Ali ME, Alam S, Hossain T, Hany U, Dissing J, Akhteruzzaman S. (2006). Genetic data on 10 autosomal STR loci in Bangladeshi population. *Legal Medicine*, 8, 297- 299.
20. Ferdous A, Ali ME, Alam S, Hasan M, Hossain T, Akhteruzzaman S. (2009). Forensic evaluation of STR data for the PowerPlex™ 16 loci in a Bangladeshi population. *Legal Medicine*, 11, 198-199.

Additional information and reprint request:
Sharif Akhteruzzaman, Ph.D
Email:sharif_akhteruzzaman@yahoo.com
Department of Biochemistry & Molecular Biology
University of Dhaka, Dhaka-1000, Bangladesh.
Tel: +88 02 9661900-50/7641,
+88 01714 111 267
Fax: +88 02 8615583