Characterization of Novel Mini STR Loci to Aid the Analysis of Degraded DNA Evidence in Sri Lankan Forensic Casework

NDS Goonawardhana[1,2#], GSKW Jayasekera[1,2], V Elanahai[1], RJ Illeperuma[1],
PV Udagama[2] and ND Fernandopulle[3]
[1]*Genetech Molecular Diagnostics, Colombo 8, Sri Lanka*
[2]*Department of Zoology, University of Colombo, Colombo 3, Sri Lanka*
[3]*Centre of Forensic Sciences, Ministry of Community Safety and Correctional Services, Toronto, Canada*
[#]darshanasuraj@gmail.com

Abstract *— Miniaturized Short Tandem Repeat (miniSTR) markers have become increasingly useful in the analysis of degraded DNA specimens from mass disaster or forensic evidence during the past few years. Improved success of these markers are based on theirability to generate reduced sized fragments in Polymerase Chain Reaction (PCR) .Tropical climatic conditions present in Sri Lanka promote rapid DNA degradation in biological samples shed to a crime scene by perpetrator/s. Thus, obtaining a meaningful DNA profile using conventional analysis techniques under such circumstances is a challenge and often results in failed DNA analysis and result interpretation. This study screened literature for a new set of Mini STR loci that can generate reduced sized PCR products to aid in the analysis of degraded DNA evidence in Sri Lanka. New PCR primers for the novel STR loci D4S2632, D6S2436and D19S589 were designed. Maximum length of amplified fragments for theminiSTR loci D6S2436 , D19S589 was found to be less than 95 bp and 124 bp, respectively, whileD4S2632 extended to 180 bp. Allele frequency database for these mini STR loci was established by analyzing blood samples from443 unrelated individuals comprising Sinhalese, Sri Lankan Tamils and Sri Lankan Moors. Forensically important parameters i.e. polymorphic information content (PIC), heterozygosity (OH) and typical paternity index (PI) were determined. D4S2632, D6S2436and D19S589 showed a relatively high degree of polymorphism with observed heterozygosities>0.7 for the three ethnic groups analyzed. Two of the three mini-STR loci did not show any departure from the Hardy–Weinberg equilibrium for all the three populations tested whereas D4S2632 conformed to HW expectations only after the application of a Bonferroni correction. The population genetic evaluation on novel mini-STR loci for the Sri Lankan population confirmed their potential use as additional STR loci to provide greater discrimination for the analysis of highly degraded DNA evidence in Sri Lanka.*

Keywords *— degraded DNA, Reduced sized PCR products, miniSTR*

I. INTRODUCTION

Analysis of Short Tandem Repeat (STR) markers on the human genome have become an important tool in the identification of individuals based on their genetic makeup. DNA extracted from the biological evidence is subjected to Polymerase Chain Reaction (PCR) amplification to generate DNA profiles unique to individuals. However, a number of factors account for the failure in PCR amplification of such biological DNA evidence making it challenging to the forensic analysts to obtain a complete DNA profile (Butler et al. 2003). For instance, the tropical environmental conditions of high humidity and high temperature in Sri Lanka promote the degradation of such biological materials and therefore the possibility of typing nuclear DNA is rapidly decreased under these conditions. Those conditions and environmental contaminants collectively promote the activity of bacterial, biochemical or oxidative processes in DNA degradation processes resulting in fragmentation of DNA molecules in to smaller pieces (Bar et al. 1988; Coble & Butler 2005). The current conventional DNA analyzing technology used in Sri Lanka analyzes nuclear DNA using PCR technology to generate amplicons in size ranging between 150 to 450 nucleotide bases. In such situations when DNA in a biological sample has fragmented in to smaller sizes, amplification of larger nuclear DNA targets by PCR could not be achieved. Therefore, we have screened for three new STR loci unlinked to the existing STR loci system in Sri Lanka. These showed the ability to generate PCR amplicons less than 125bp except one STR loci which generated a maximum amplicon size of 180bp upon analyzing for the Sri Lankan population. No previous records exist in the application of these three STR loci for forensic DNA analysis.

II. MATERIALS AND METHODS

A. Selection of STR loci

The selection of the three MiniSTR markers was achieved by screening the STR marker sets published in Marshfield Clinic Center for Medical Genetics (http://research.marshfieldclinic.org/genetics/). The selected STRs namely D4S2632, D6S2436and D19S589 were not adapted to human identity testing up to date. The sequences of the STR loci screened were obtained from a BLAST – nucleotide search on www.ncbi.com. The chromosomal positions were also determined by BLAT (http://genome.ucsc.edu/cgi-bin/hgBlat) and assembly of the human genome version of Dec.2013 (GRCh38/Hg38).

B. Primer designing

The primers were designed for the selected miniSTR loci using stand alone Primer premier 6 (Premier Biosoft Interpairs, Palo Alto, CA). The manual primer designing tool of the software used, enabled us to effectively locate primers adjacent to the repeated region. Primers were designed maintaining the G and C content within the range of 40 % to 60 % and basic melting temperatures from 51^{o}C-54^{o}C (Table: 2). the successfully designed primer sets were evaluated for probable inter-primer interactions using the options in Primer premier 6.

C. Nomenclature of STR Loci

Nomenclature of the STR loci D4S2632, D6S2436 and D19S589 was done in accordance with the International Society for Forensic Genetics (ISFG) guidelines on nomenclature of STR loci (Bär et al. 1997).

D. Collection of blood samples

Blood samples were collected from 443 healthy adult individuals belonging to three ethnic groups; Sinhalese, Sri Lankan Tamils and Sri Lankan Moors after obtaining their written informed consent to participate in the study. Ethical clearance for sample collection was granted by the Medical faculty of Sri Jayawardhanapura University, Sri Lanka.

Genomic DNA was isolated from blood-stained filter papers using Chelex-100 protocol (Walsh et al. 1991) and quantified using the NanoDrop instrument, prior to analysis.

E. PCR amplification and Genotyping

From an extracted genomic DNA sample, 5ng of each amplified in PCR with the following conditions: an initial denaturation of 94^{o}C for 2 min; 94 ^{o}C for 30 sec; 55.5 ^{o}C for 45 sec and 72^{o}C for 60 sec repeated for 32 cycles and a final extension at 60^{o}C for 30 min. Sizing of the amplified products were performed using 6% polyacrylamide gel electrophoresis followed by DNA sliver staining procedure described in Promega silver sequence™ DNA Sequencing System. Calling of alleles was done by comparing the PCR fragments with the allelic ladder prepared in-house and validated by sequencing of at least two alleles for each STR loci.

F. Analysis of data

Forensically important statistical parameters were analyzed using the software PowerStatsV12 spreadsheet (http://www.promega.com/geneticidtools/powerststs/). Observed (HO) and expected heterozygosities (HE), AMOVA and the exact test for Hardy–Weinberg equilibrium was analyzed using Arlequin Version 3.5 (Excoffier et al. 2005).

III. RESULTS AND DISCUSSION

Forensically important parameters for the three miniSTR markers, sizes of observed alleles, allele ranges and the allele frequency distribution obtained for the Sri Lankan population are given in Tables 1,3, and 4.

Table 1. Information on three novel STRs evaluated in this study. Chromosomal location and base pair (bp) position of each marker was determined by using BLAT (http://genome.ucsc.edu/cgi-bin/hgBlat) and the Dec.2013 (GRCh38/Hg38) assembly of the human genome. The miniSTR amplicon length is based on the GenBank allele observed.

STR Locus	Chromosomal location	Chromosomal pb Position	Repeat Motif	GenBank accession	GenBank allele	MiniSTRamplicon length (bp)
D4S2632	4p15-p14	35704165	(AGAT)n	G08391.1	13	105
D6S2436	6q24.1	154136091	(AGAT)n	G27284.1	9	91
D19S589	19q13.42	58498394	(AGAT)n	G08026.1	13	98

Table 2. miniSTR pimers designed for the study. All Anti-Sense primers are designed with a Guanine residue to promote adeneylation (Butler et al. 2003). Negative number at 3' end of primers denotes the number of bases that are designed in to the core repeat region. Basic melting temperature for each primer is denoted as **T_m**

STR Locus	Primer sequence (5'-3')		Distance from repeat motif (bp)	T_m (^{0}C)	GC %
D4S2632	Sense	ATTGAAGTTGTGACTCAGGTC	16	52.5	42.9
	Anti-Sense	**G**CAGGCTGTAACCAATCTAGGA	18	53.9	47.6
D6S2436	Sense	AAGGCACCACCAATAAGATA	-4	51.2	40
	Anti-Sense	**G**CAATTACCCTAATAAGGTCCC	6	52.1	42.9
D19S589	Sense	GCTGAGTAGTATTCCATCAGATAG	-5	52.2	41.7
	Anti-Sense	**G**ATCAATGAGTGGATAAAGGGAC	13	52.2	40.9

Table 3. Forensically important parameters on three novel STRs evaluated in this study.

STR Locus	N	Observed Allele Range	Observed Size Range (bp)	H(ob)	*P*-Value	PIC	PI	PE	PD
D4S2632	443	8-26	108- 180	0.8239	0.2870	0.84	2.84	0.644	0.962
D6S2436	443	6-15	59-95	0.8025	0.5533	0.79	2.53	0.604	0.941
D19S589	443	11-18	96- 124	0.7166	0.1126	0.67	1.75	0.451	0.873

Key: N – Number of samples, H(ob) – Observed Heterozygosity, PIC – Polymorphism Information Content, PD- Power of discrimination. PI- Typical Paternity Index, PE- Power of Exclusion

These potential miniSTR loci were screened from Previously well characterized STR loci (Ghebranious et al. 2003) including the whole genome screening sets of STRs developed by the Marshfield Clinic Center for Medical Genetics. These screening sets were extensively reported in genetic linkage studies across the entire nuclear genome to determine specific genes that cause or have linked to human diseases. A subset of the screening markers were also used to characterize the genetic diversity in global populations(Rosenberg et al. 2002; Rosenberg et al. 2003; Ayub et al. 2003).

When screening the probable STR loci that can be made in to miniSTR markers, consideration was given to several characteristics of STR loci in order to obtain the best possible loci. Firstly, the STR loci consisting of tetrameric repeat motif were selected. The selection of tetrameric repeat motif is highly advisable in forensic DNA analysis, since these markers are highly polymorphic and results in reduced stutter effect in the PCR reaction (Kimpton et al. 1993). Secondly, the STR loci containing tetrameric repeat motifs were evaluated for clean flanking regions up and downstream to the repeat motif (approximately 75bp). This approach is very important when the PCR primers are located near to the repetitive region since a clear flanking region can give rise to good primer hybridization in the PCR reaction (Coble & Butler 2005). Thirdly, the STR loci that can produce PCR fragments less than 150bp were selected and PCR primers were designed. Finally, the selected three novel STR loci D4S2632, D6S2436 and D19S589 were evaluated for the allele distribution in the Sri Lankan population comprising Singhalese, Sri Lankan Tamils and Sri Lankan moors. The observed allele range of the three miniSTR marker panel for the Sri Lankan population ranged from 59 base pairs to 180 base pairs collectively. The maximum length of amplified fragments for the miniSTR loci D6S2436 and D19S589 was found to be less than 124 base pairs except for D4S2632 which extended to 180 base pairs. All the three miniSTR markers showed a relatively high degree of polymorphism with observed heterozygosities H(ob)>0.8 whereas D19S589 showed a moderate degree of polymorphism with the observed heterozygosities H(ob) >0.7 for the three populations analyzed.

Table 4. Allele frequency distribution of miniSTR loci D4S2632, D6S2436 and D19S589 for the Sri Lankan population comprising Singhalese, Sri Lankan Tamils, Sri Lankan Moors.

	STR Loci		
Allele	D4S2632	D6S2436	D19S589
5	-	-	-
6	-	0.009	-
7	-	0.154	-
8	0.081	0.025	-
9	0.005	0.163	-
10	0.000	0.090	-
11	0.001	0.090	0.002
12	0.016	0.326	0.031
13	0.095	0.121	0.101
14	0.214	0.021	0.421
15	0.236	0.001	0.288
16	0.111	-	0.113
17	0.094	-	0.039
18	0.067	-	0.006
19	0.017	-	-
20	0.003	-	-
21	0.006	-	-
22	0.009	-	-
23	0.023	-	-
24	0.008	-	-
25	0.011	-	-
26	0.005	-	-

IV. CONCLUSION

This approach outlines the initial efforts to develop miniSTR markers that can be used as supplement to the loci included in Sri Lankan forensic analysis system as well as to CODIS (Combined DNA Index System) core loci system to increase the power of discrimination in analyzing highly degraded DNA evidence in forensic casework. Apart from using these markers for forensic purposes, they can be successfully incorporated in cases where additional markers are needed to conclude the analysis such as human parentage testing or the analysis between closely related individuals.

ACKNOWLEDGMENT

We acknowledge the immense contribution of Genetech Molecular Diagnostics and School of Gene technology, Sri Lanka with all the staff members, the blood donors and the National Science Foundation (NSF) for funding the research.

REFERENCES

Ayub, Q. et al., 2003. Reconstruction of human evolutionary tree using polymorphic autosomal microsatellites. *American journal of physical anthropology*, 122(3), pp.259–268.

Bar, W. et al., 1988. Postmortem stability of DNA. *Forensic science international*, 39(1), pp.59–70.

Bär, W. et al., 1997. DNA recommendations. Further report of the DNA Commission of the ISFG regarding the use of short tandem repeat systems. *Forensic Science International*, 87(3), pp.179–184.

Butler, J.M., Shen, Y. & McCord, B.R., 2003. The development of reduced size STR amplicons as tools for analysis of degraded DNA. *Journal of forensic sciences*, 48(5), pp.1054–1064.

Coble, M.D. & Butler, J.M., 2005. Characterization of new miniSTR loci to aid analysis of degraded DNA. *Journal of forensic sciences*, 50(1), pp.43–53.

Excoffier, L., Laval, G. & Schneider, S., 2005. Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, 1, pp.47–50.

Ghebranious, N. et al., 2003. STRP screening sets for the human genome at 5 cM density. *BMC genomics*, 4(1), p.6.

Kimpton, C.P. et al., 1993. Automated DNA profiling employing multiplex amplification of short tandem repeat loci. *PCR methods and applications*, 3(1), pp.13–22.

Rosenberg, N.A. et al., 2002. Genetic structure of human populations. *Science (New York, N.Y.)*, 298(5602), pp.2381–2385.

Rosenberg, N.A. et al., 2003. Informativeness of genetic markers for inference of ancestry. *American journal of human genetics*, 73(6), pp.1402–1422.

Walsh, P.S., Metzger, D.A. & Higuchi, R., 1991. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *BioTechniques*, 10(4), pp.506–513.