

SHORT COMMUNICATION

## Population study of the Combined DNA Index System (CODIS) core loci D3S1358, D5S818, D8S1179 short tandem repeat (STR) polymorphisms in Sri Lanka

W. Chatumal<sup>2</sup>, Dinuka A. Markalanda<sup>1</sup>, Ruwan J. Illeperuma<sup>1</sup>, Gaya R. Ranawaka<sup>2</sup> and Neil D. Fernandopulle<sup>1</sup>

<sup>1</sup> Genetech Molecular Diagnostics, No.54, Kitulwatte Road, Colombo 08.

<sup>2</sup> Department of Zoology, Faculty of Natural Sciences, The Open University of Sri Lanka, Nawala.

Submitted: 26 June 2009; Accepted: 25 September 2009

**Abstract:** Short tandem repeats (STRs) are a group of highly polymorphic loci in the human genome. STR loci consist of tandemly repeated sequences of 2-7 base pairs in length. As they are small in size STR loci can easily be amplified by the polymerase chain reaction (PCR). Allele frequencies of short tandem repeat loci D3S1358, D5S818, D8S1179 were analyzed by PCR amplification and polyacrylamide gel electrophoresis in 50 unrelated Sri Lankan individuals. All STR loci were in Hardy-Weinberg equilibrium. Statistical parameters of forensic importance, the power of discrimination (PD), observed heterozygosity value (H), polymorphic information content (PIC), typical paternity index (PI) and power of exclusion (PE) were calculated for the loci. These parameters indicated the usefulness of the loci in paternity testing and individual identification in the Sri Lankan population. The combined power of discrimination was 0.926. This is the first time that Sri Lankan population data for D3S1358, D5S818, and D8S1179 loci of forensic importance are reported.

**Keywords:** D3S1358, D5S818, D8S1179, polymorphism, population study, short tandem repeats.

### INTRODUCTION

DNA typing is the characterization of individuals at the level of their genetic material, deoxyribonucleic acid (DNA). The DNA of every human being (with the exception of monozygotic twins) is unique. The DNA of an individual is identical whether it is extracted from hair bulbs, white blood cells, or semen. These principles of individual uniqueness and identical DNA structure within all tissues of the same body provide the basis for DNA typing<sup>1</sup>. Whereas previous methods of identification were based on gene products<sup>2</sup>, which only indirectly represented the genetic identity of an individual, DNA

typing methods seek a direct analysis of the genetic uniqueness of individuals<sup>3</sup>.

Tandemly repeated DNA sequences in the human genome are a source of genetic markers useful for genetic studies, medical diagnostics, paternity testing and identification of individuals in forensic casework<sup>3-5</sup>. Short tandem repeat (STR) loci or microsatellites are a subgroup of the highly polymorphic variable number of tandem repeats (VNTR). They are composed of tandemly repeated sequences of 2-7 base pairs in length and are abundant in the human genome. DNA typing based on amplification of STR loci is an effective means to overcome some of the problems encountered in forensic practice such as typing of minute amounts of DNA, highly degraded DNA or mixtures of DNA from different individuals<sup>6</sup>. Although most STR systems have a much lower heterozygosity than restriction fragment length polymorphisms (RFLPs) and some amplified fragment length polymorphisms (AmpFLPs), this disadvantage is reduced by increasing the number of systems tested and by the use of multiplex polymerase chain reaction (PCR)<sup>7</sup>. The STR loci D3S1358, D5S818 and D8S1179 of the present study can be amplified using PCR with locus specific primers. The PCR products can be resolved by denaturing polyacrylamide gel electrophoresis (PAGE) and allele designation can be done unambiguously by comparison to an allelic ladder.

The use of genetic markers, such as STR, in identity testing must accompany allele/genotype frequency data from the relevant population for estimating the frequency of any particular genetic profile.

\*Corresponding author (dinuka\_g@yahoo.com)

**Table 1:** Description of the three loci

STR locus	Chromosomal location	Repeat motif	Product size (bp)
D3S1358	3p	TCTA complex	99-147
D5S818	5q23.3-32	AGAT	129-177
D8S1179	8q	TCTA complex	157-209

In Sri Lanka, currently eleven autosomal STR loci are used for human identity testing. They are CSF1PO, TPOX, TH01, D16S539, D7S820, D13S317, FESFPS, F13A, F13B, vWA and LPL. These eleven autosomal STRs studied are highly polymorphic and are widely used in human identification in Sri Lanka for paternity testing as well as forensic DNA analysis. The allelic frequencies for these STR loci have been calculated and databases have been established for the Sri Lankans <sup>8,9</sup>.

The Federal Bureau of Investigation (FBI) of the United States of America has been a leader in developing DNA typing technology for use in the identification of perpetrators of violent crime. In 1997, the FBI announced the selection of 13 STR loci to constitute the core of the United States national database, CODIS which stands for Combined DNA Index System. All CODIS STRs are tetrameric repeat sequences and are namely, CSF1PO, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, FGA TH01, TPOX vWA and AMEL <sup>10</sup>. Only seven of the above loci are in the Sri Lankan allelic database and thus being used for DNA based paternity testing and human identification.

The present study is the first application of STR typing of three CODIS loci D3S1358, D5S818, D8S1179 in a sample of Sri Lankans. These new STR sites will further enhance the accuracy of human identification and paternity testing in Sri Lanka. The present study has evaluated the informativeness of these genetic markers and their suitability to identify Sri Lankan individuals by DNA typing technology and has formulated a preliminary allelic frequency database which can be used to express a chance match in DNA based forensic case work.

## METHODS AND MATERIALS

Human blood samples were taken after obtaining informed written consent from 50 unrelated Sri Lankan donors. Whole blood was obtained by venipuncture and collected into acid citrate dextrose anticoagulant. DNA was extracted by the Chelex-100 method<sup>11</sup>.

Amplification of D3S1358, D5S818 and D8S1179 loci (Table 1) was carried out in a GeneAmp 2400 (Applied

Biosystems, USA) Thermal Cycler in 25  $\mu$ L reaction volumes containing 10 mM Tris HCl, 50 mM KCl, 1.5 mM MgCl<sub>2</sub>, 0.1% Triton-100, 0.2 mg/mL BSA or gelatin, 0.2 mM of each dNTP, 200  $\mu$ M of each primer and 0.75 u of Taq DNA polymerase. The reactions were subjected to initial denaturation at 96 °C for 2 min; 10 cycles of denaturation at 94 °C for 1min; primer annealing at 64 °C for 1min; and primer extension at 70 °C for 1.5 min. PCR was continued with a further 20 cycles of denaturation at 90 °C for 1min; primer annealing at 64 °C for 1 min; and primer extension at 70 °C for 1.5 min (Promega Technical manual, 1998).

The PCR product (12.5  $\mu$ L) was mixed with 2.5  $\mu$ L of loading dye (10 mM NaOH, 95% formamide, 0.05% bromophenol blue and 0.05% xylene cyanol) and denatured at 95 °C for 2 min and loaded into denaturing polyacrylamide gels (6%, 37 cm long and 0.4 mm thick) containing 7 M urea and 0.5X Tris-Borate-EDTA buffer. Electrophoresis was carried out on a GibcoBRL S2001 electrophoresis apparatus (Life Technologies). The conditions of electrophoresis were set at a constant power of 50 W at ambient temperature <sup>12</sup>.

Silver staining was carried out using the DNA silver staining system according to the manufacturer's recommendations (Promega Technical manual 1998). Allelic ladders were constructed by mixing alleles with varying lengths obtained from a randomly selected set of individuals. The exact size length of the alleles in

**Table 2:** Allele frequency distribution

Allele	Frequency		
	D3S1358	D5S818	D8S1179
8			1.1%
9		3.1%	1.1%
10		10.4%	17.0%
11		33.3%	9.1%
12		37.5%	8.0%
13		12.5%	9.1%
14	7.1%	3.1%	29.5%
15	31.6%		17.0%
16	30.6%		8.0%
17	22.4%		
18	8.2%		

**Table 3:** Summary of statistically important data

	D3S1358	D5S818	D8S1179
Forensic			
Matching probability	0.115	0.139	0.074
Expressed as 1 in .....	8.7	7.2	13.4
Power of discrimination	0.885	0.861	0.926
PIC	0.70	0.67	0.80
Power of exclusion	0.306	0.546	0.768
Typical paternity index	1.29	2.18	4.40
Allele frequencies			
Homozygotes	38.8%	22.9%	11.4%
Heterozygotes	61.2%	77.1%	88.6%
Total alleles	98	96	88

the ladders were determined upon validation against a universal known size marker (K562) (Promega Corporation, USA), and allele designations were determined by comparison of sample fragments with those of the constructed allelic ladder. Allele designations of each samples were made according to recommendations of the DNA Commission of the International Society of Forensic Haemogenetics<sup>13</sup>.

Statistical parameters of forensic importance, the power of discrimination (PD)<sup>2</sup> polymorphism information content (PIC)<sup>14</sup>, typical paternity index (PI)<sup>15</sup> and power of exclusion (PE)<sup>16</sup> for each locus were determined using the PowerStats computer programme from the Promega Corporation<sup>17</sup>.

## RESULTS AND DISCUSSION

The distribution of observed allele frequencies for the three loci is shown in Table 2. PD, PIC, PI and PE are listed in Table 3.

In the tested Sri Lankan population the most common alleles were allele 15 for D3S1358 ( $f=0.316$ ), allele 12 for D5S818 ( $f=0.375$ ) and allele 14 for D8S1179 ( $f=0.295$ ).

The PIC values show that all three loci were highly informative ( $PIC>0.5$ ). The PD values for the present loci showed a high power of discrimination ranging from 0.885 to 0.926. The combined power of discrimination for the three loci was 0.9988. Several more loci that show random association with the present STR loci need to be typed to increase the significance of any match.

The PI is a means of presenting the genetic odds in favour of paternity given the genotypes for the mother, the child and the putative father<sup>15</sup>. The combined typical paternity index for the three loci was 12.37. An

alternative value used in paternity analyses is the PE<sup>16</sup>. The combined power of exclusion for the three loci was 0.1283.

## CONCLUSION

A preliminary Sri Lankan population database has been established for the STR loci D3S1358, D5S818 and D8S1179. The observed allele frequencies may be used to estimate genotype frequencies and the product rule used to calculate combined genotype frequencies across loci in this population. It is clear that after appropriate validation work, the use of these highly polymorphic STR loci will provide a useful tool for forensic identification, and paternity testing in Sri Lanka, and confirm the versatility and robustness of STRs. All three loci studied were highly informative ( $PIC>0.5$ ) and therefore these loci can be used for DNA based identification of individuals of the Sri Lankan population.

## References

1. Edwards A., Civitello A., Hammond H.A. & Caskey C.T. (1991). DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *American Journal of Human Genetics* **49**(4):746-756.
2. Jones. D.A. (1972). Blood samples: probabilities of discrimination. *Journal of the Forensic Science Society*. **12**(2):355-359.
3. Jobling M.A. & Gill P. (2004). Encoded evidence: DNA in forensic analysis. *Nature Reviews Genetics* **5**(10):739-51.
4. Hammond H.A., Jin L., Zhong Y., Caskey C.T. & Chakraborty R. (1994). Evaluation of 13 short tandem repeat loci for use in personal identification applications. *American Journal of Human Genetics* **55**(1):175-89.
5. Busque L., Desmarais D., Provost S., Schumm J.W., Zhong Y. & Chakraborty R. (1997). Analysis of allele distribution for six short tandem repeat loci in the French Canadian population of Quebec. *Journal of Forensic Science* **42**(6) :1147-1153.

6. Butler J.M., Schoske R. & Vallone P.M. (2003). Allele frequencies for 15 autosomal STR loci on U.S. Caucasian, African American, and Hispanic populations. *Journal Forensic Science* **48**(4): 908–911.
7. Klintschar M. & Crevenna R. (1996). Validation of the STR system FXIII B for forensic purposes in an Austrian population sample. *Forensic Science International* **81**(1): 35-42.
8. Fernanandopulle N.D. (2001). Development of DNA markers for the identification of individuals in the Sri Lankan human population and its applications in forensic casework. *PhD Thesis*. University of Colombo, Colombo 03.
9. Illeperuma R.J., Mohotti S.M., De Silva T.M., Fernandopulle N.D. & Ratnasooriya W.D. (2008). Genetic profile of 11 autosomal STR loci among the four major ethnic groups in Sri Lanka. *Forensic Science International Genetics* **3**(3):105-106.
10. Butler J.M. (2006). Genetics and genomics of core STR loci used in human identity testing. *Journal of Forensic Sciences* **51**(2):253-265.
11. 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. *Bio Techniques* **10**(4): 506.
12. Sambrook J., Fritsch E.F. & Maniatis T. (1992). *Molecular Cloning: A Laboratory Manual*. p.545. Cold Spring Harbor, New York, USA.
13. Bar W. et al. (1997). DNA recommendations: further report of the DNA commission of the ISFH. regarding the use of Short Tandem Repeat Systems. *International Journal of Legal Medicine* **110**: 175.
14. Bostein D., White R.L., Skolnick M. & Davis R.W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* **32**(3):314-331.
15. Brenner C. & Morris. J. (1989). Paternity index calculations in single locus hypervariable DNA probes: validation and other studies. *Proceedings of the International Symposium on Human Identification*. pp.21-53. Promega Corporation. Madison, Wisconsin, USA.
16. Powerstats. Profiles in DNA 3, Tools for Analysis of Population statistics. <http://www.promega.com/geneticidtools/powerstats/>.