

## IDENTIFICATION AND CHARACTERIZATION OF VARIANT ALLELE 34.1 AT SHORT TANDEM REPEAT (STR) LOCUS D21S11

SHARIF AKHTERUZZAMAN\*, AHMAD FERDOUS<sup>1</sup> AND TANIA HOSSAIN<sup>1</sup>

*Department of Biochemistry & Molecular Biology, University of Dhaka,  
Dhaka-1000, Bangladesh*

*Key words:* Microsatellites, Allele, Genotype, Repeat motif

### Abstract

D21S11 is one of the core short tandem repeat (STR) locus included in most of the commercial kits used for DNA parentage testing, kinship analysis and human identification purpose. It is a complex tetranucleotide repeat located on the long arm of chromosome 21 which surrounds a constant 43 bp section. The allele designation of a particular locus is correctly assigned by comparing the alleles with an allelic ladder that consists of all possible alleles at that particular locus. In this study we observed a variant allele at D21S11 during a routine paternity test while typing the putative father. The variant allele was called by the Genotyper™ software as 34.1, though this allele was not included in the allelic ladder provided by the manufacturer. Sequencing analysis revealed that the variant allele was generated as a change in the repeat motif (TCTA) for TCA at the 3' end of the repeat region. This variant allele was first reported in Bangladeshi population from our laboratory in a previous report. This study involves further characterization of the reported allele by means of direct DNA sequencing of the D21S11 locus.

### Introduction

Short tandem repeat (STR) based DNA typing has now become the most powerful tool in human identification, parentage testing, kinship analysis and population genetics studies<sup>(1-4)</sup>. Invariably 15 - 20 autosomal STR markers are routinely examined in various combinations<sup>(5)</sup>. The current analysis method involves PCR amplification of multiple STR loci in a multiplex fashion, followed by capillary electrophoresis of the PCR products. Manufacturers of commercial kits provide allelic ladder that assists the analyst in assigning the accurate allele designation.<sup>(6)</sup> Allelic ladders are developed on the basis of variations observed in different populations, and the vast majority of alleles being typed world-wide each year are usually represented in it. However, variant or off-ladder alleles are observed occasionally in a small number of individuals<sup>(7)</sup>. This happened due to the fact that, the DNA sample under investigation belongs to a population is different from that were studied during kit development. Forensic DNA typing community should therefore be aware of the existence of rare alleles so that they can be dealt with proper precautions during casework interpretation.

---

\*Author for correspondence: <sharif\_akhteruzzaman@yahoo.com>. <sup>1</sup>National Forensic DNA Profiling Laboratory, Dhaka Medical College, Dhaka-1000, Bangladesh.

D21S11 is a core STR loci present in most of the commercial kits<sup>(8-10)</sup>. It is a complex tetranucleotide repeat found on the long arm of chromosome 21. A constant 43 bp section made up of {[TCTA]3 TA }TCTA]3 TCA [TCTA]2 TCCA TA} sequence is surrounded by a variable number of TCTA and TCTG repeats at this locus<sup>(11)</sup>. The known alleles of this locus range from 12 to 41.2. The *x.2* microvariant alleles arise due to a 2 bp [TA] insertion at the 3' end of the repeat region<sup>(12)</sup>. The *x.1* microvariants are very rare and therefore not included in the allelic ladder provided by the commercial kit manufacturer.

This study reports the identification and characterization of variant allele 34.1 at D21S11 locus, typed by both AmpF/STR<sup>®</sup> SGM plus and Powerplex<sup>®</sup> 16 PCR amplification kit. In spite of the absence of this allele in the allelic ladder the allele was called by the GeneMapper ID software. The aim of this study was therefore to characterize the allele at molecular level by sequencing analysis.

### Materials and Methods

*Samples and DNA extraction:* Blood samples from the disputed child, mother and alleged father was collected in EDTA tube with written informed consent. Genomic DNA was extracted using Chelex-100 method<sup>(13)</sup>. 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. DNA typing was carried out by using SGM Plus (Applied Biosystems, USA) and PowerPlex<sup>®</sup> 16 (Promega Corporation, Madison, WI) following the protocol provided by the manufacturer. PCR amplification was carried out by using Veriti thermal cycler (Applied Biosystems). The PCR amplified products were separated by capillary electrophoresis on ABI Prism<sup>®</sup> 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 ver 3.2.

*STR allele sequencing:* For sequencing of the variant allele, D21S11 locus was amplified by monoplex PCR using forward and reverse primers PCR primes that anneal sufficiently away from the STR repeat in order to generate PCR products containing the primer binding regions of all the known commercial kit primers<sup>(14)</sup>. The PCR reaction was carried out in 10 µl reaction volume containing 5-7 ng of extracted genomic DNA, 1×GenAmp<sup>®</sup> PCR Gold buffer (Applied Biosystems), 2 µM MgCl<sub>2</sub>, 250 µM dNTPs, 0.2 µM forward (5'-CCAGCTTCCCTGATTCTTCA-3') and reverse primers (5'-CACTGA GAAGGGAGAAACACTG-3'), 1U AmpliTaq Gold DNA polymerase and 0.16 mg/ml BSA. The PCR was carried out for 28 cycles after an initial denaturation for 10 min at 95°C. Each cycle comprised by 1 min at 94°C, 1 min at 57°C, and 1 min at 72°C, followed by extension for 45 min at 60°C, and a final soak at 25°C. PCR products from heterozygous alleles were separated on acrylamide based gels containing 12% total acrylamide with 3% cross linked bis-acrylamide. The gel was then developed by silver

staining. The desired band was excised from the gel by using a razor blade and placed overnight in 50 - 100  $\mu$ l of TE buffer (10 mmol/l Tris-HCl, 0.1 mmol/l EDTA, pH 8.0). Homozygous allele samples did not require any gel separation. A 5  $\mu$ l aliquote of the gel separated heterozygote alleles were re-amplified using the same PCR condition as described.

Excess PCR primers and unincorporated dNTPs were functionally inactivated by the addition of 2  $\mu$ l ExoSap-IT™ (USB Corporation) to 5  $\mu$ l of PCR product. The mixture was then incubated 90 min at 37°C, 20 min at 80°C, and a final soak at 25°C. Sequencing was performed by using ABI PRISM®BigDye® Terminator cycle sequencing kit v3.1 (Applied Biosystems, USA). The sequencing reaction consisted 5  $\mu$ l purified PCR product, 6.5  $\mu$ l of 2.5 $\times$  sequencing dilution buffer, 3.8  $\mu$ l of dH<sub>2</sub>O, 2  $\mu$ l BigDye Terminator ready reaction mix, and 3.2  $\mu$ l of forward or reverse primer at a concentration of 1.0  $\mu$ M. The thermal cycling conditions were as follows: 5 min at 94°C, amplification for 25 cycles of 10 sec at 94°C, 5 sec at 50°C, and 2 min at 60°C, extension for 5 min at 60°C, and a final soak at 4°C. The resultant sequencing amplicons were purified with a 2.2% SDS solution in Centri-Sep spin columns.

*Analysis on the ABI Prism 3100 genetic analyzer:* The purified amplicons were diluted by adding 1  $\mu$ l of PCR product 0.35  $\mu$ l of GS500-LIZ size standard to 14  $\mu$ l HiDi formamide (Applied Biosystems). The samples were then electrophoresed in an ABI Prism 3100 *avant* Genetic analyzer using POP-6™ and a 60 cm capillary array with appropriate dye set.

*Data analysis:* DNA sequencing data were analyzed using sequencing analysis software v3.7. Sequences were then aligned with GenBank reference allele and edited using Sequencer Plus 4.0.5b11 (GeneCodes, Ann Arbor, MI).

## Results and Discussion

In this study, we characterized a microvariant allele 34.1 at D21S11 short tandem locus. D21S11 is a complex tetranucleotide repeat located on the long arm of chromosome 21 in which a constant 43 bp section is surrounded by a variable number of TCTA and TCTG repeats (GenBank Accession No: AP000433). Most of the DNA profiling commercial kits<sup>(8-10)</sup> used for parentage testing, personal identification or kinship analyses, contain this locus due to its highly polymorphic nature. More than 80 alleles ranging from 12 to 41.2 have been reported at this locus. The current method of DNA analysis involves amplification of multiple STR loci in a multiplex fashion. Correct allele designation is achieved by comparison with an allelic ladder which is developed on the basis of the variations observed in different populations. Alleles within STR loci vary greatly between individuals and kit ladder do not represent all possible alleles. It therefore poses a great challenge for the analyst whenever a rare microvariant allele is observed in course of the DNA investigation.

A microvariant allele represents an incomplete repeat of a given allele. They are usually reported as the number of complete repeat unit and are designated as integer. Any partial repeat is designated as the decimal, followed by the number of bases in the initial repeat (e.g.  $x.1$  or  $x.2$ ). The  $x.1$  microvariants are very rare and therefore not usually included in the allelic ladder provided by the manufacturer. In this study we report a microvariant allele 34.1 first time observed in our population by using SGM plus PCR amplification kit (Fig. 1A). Fig. 1B shows that the microvariant allele was not included in the allelic ladder which was provided by the manufacturer. In order to check the accuracy of the allele designation, we got the analysis done by using a second kit (PowerPlex 16) from a different manufacturer. In both the cases the allele was called as 34.1 with its sister allele 34 (Fig. 1C) by the Geotyper™ software used in the instrument platform (ABI prism 3100 *avant* Genetic Analyzer).

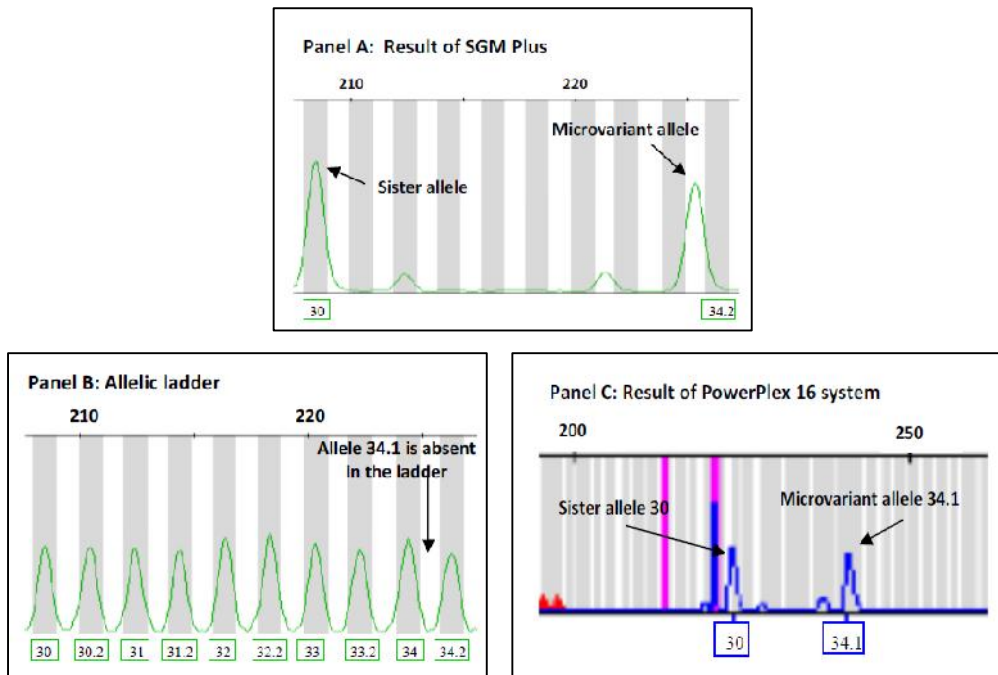


Fig. 1. Panel A: Variant allele 34.1 with its sister allele (30) while using SGM Plus PCR amplification kit. Panel B: Allelic ladder where the microvariant allele 34.1 is not included. Panel C: Variant allele 34.1 with its sister allele (30) while using PowerPlex 16 PCR amplification kit.

To confirm our prediction we conducted a sequencing analysis of the locus D21S11 using the primer pair described above. The primers were designed so that they bind reasonably far away from the primer binding sites of the commercial kits. The sequencing analysis revealed that the allele 34 of this locus harbors a 43 bp constant region surrounded by variable number of TCTA and TCTG repeat. The 34.1 microvariant

arises as a result of a 3 bpTCA insertion between [TCTA]<sub>2</sub> and [TCTA]<sub>4</sub>at 3' end of the repeat region (Fig. 2). This microvariant allele has previously been reported in Caucasians, African Americans as well as in Hispanic populations<sup>(15)</sup>. The Short Tandem Repeat International DNA Database (<http://www.cstl.nist.gov/div831/strbase>) lists 17 different reports of this allele but no published report is available. This allele did not also appear in reports published describing the allele frequency of Bangladeshi population<sup>(16-17)</sup>. This microvariant allele 34.1 from D21S11 STR loci was first reported from our laboratory in Bangladeshi population in a previous report<sup>(18)</sup>. In this current study, this allele was further characterized by direct DNA sequencing. The scientific and forensic community should therefore be aware of the existence of this rare allele so that they can be handled with proper precautions during parentage test, personal identification and kinship analysis.

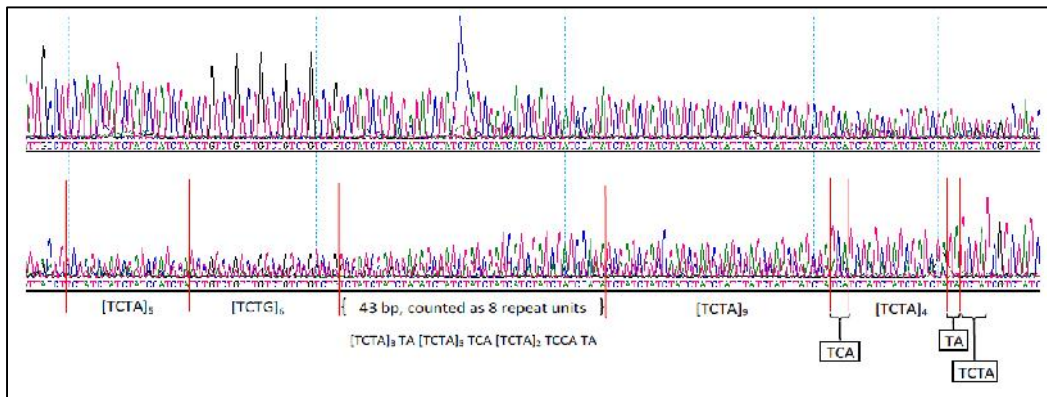


Fig. 2. Sequencing analysis showing that a 3 bp insertion between [TCTA]<sub>9</sub> and [TCTA]<sub>4</sub> at 3' end of the repeat region generates the microvariant allele 34.1.

## References

1. Gill P, PL Ivanov, C Kimpton, R Piercy, N Benson, G Tully, I Evett, E Hagelberg and K Sullivan 1994. Identification of the remains of Romanov family by DNA analysis. *Nat. Genet.* 6:130-135.
2. Matsumura S, T Nakazono, A Ishiko, M Atsuchi, K Hiroshige, H Yamaguchi and S Kubo 2011. Kinship analysis using DNA typing from five skeletal remains with an unusual postmortem course. *Med. Sci. Law* 51(4): 240-3
3. Brinkman B, H Pfeiffer, M Schurenkamp and CHohoff 2001. The evidentiary value of STRs. An analysis of exclusion case. *Int. J. Leg. Med.* 114: 173-177.
4. Bowcock AM, A Ruiz-Linares, J Tomfohrde, E Minch, JR Kidd and LL Cavalli-Sforza 1994. High resolution of human evolutionary tree with polymorphic microsatellites. *Nature* 368: 455-457.
5. Butler JM 2006. Genetics and genomics of core STR loci used in human identity testing. *J. Forensic Sci.* 51: 253-256.

6. Griffiths RAL, MD Baber, PE Johnson, SM Gilbard, MD Haywood, CD Smith, J Arnold, T Burke, A Urquhart and P Gill 1998. New reference allelic ladders to improve allelic designation in a multiplex STR system. *Int J. Leg. Med.* **111**(5): 267-272.
7. Dauber EM, G Dorner, S Wenda, EM Schwart-jungi, B Glock, W Bär and FWR Mayr 2008. Unusual FGA and D19S433 off-ladder alleles and other allelic variants at the STR loci D8S1132, vWA, D18S51 and ACTBP2 (SE33). *Forensic Sci. Int. Genet. Suppl.* **1**: 109-111.
8. Cotton EA, RF Allsop, JL Guest, RR Frazier, P Koumi, IP Callow, A Seager and RL Sparkes 2000. Validation of AmpFISTR SGM Plus system for use in forensic casework. *Forensic Sci. Int.* **112**: 151-161.
9. Wang DY, CW Chang, RE Lagacé, LM Calandro and LK Hennessy 2012. Developmental validation of the AmpFISTR®Identifiler® Plus PCR amplification kit: an established multiplex assay with improved performance. *J. Forensic Sci.***57**(2): 453-65.
10. Tucker VC, AJ Hopwood, CJ Sprecher, RS McLaren, DR Rabbach, MG Ensenberger, JM Thompson and DR Storts 2012. Development validation of PowerPlex® 16 and PowerPlex®ESX 17 Systems. *Forensic Sci. Int. Genet.* **6**(1): 124-131.
11. Sharma V and M Litt 1992. Tetranucleotide repeat polymorphism at the D21S11 locus. *Human Mol. Genet.* **1**(1): 67.
12. Brinkman B, E Meyer and A Jung 1996. Complex mutational events at the HumD21S11 locus. *Hum. Genet.* **98**: 60-64.
13. Walsh P, DA Metzger and R Higuchi 1991. Chelex 100 as a simple medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques* **10**: 413-506
14. Kline MC, RH Carolyn, ED Amy and JM Butler 2011. STR sequence analysis for characterizing normal, variant and null alleles. *Forensic Sci. Int. Genet.* **5**(4): 329-332.
15. Catherine A, DE David and S Marcos 2005. Identification and characterization of variant alleles at CODIS STR loci. *J. Forensic Sci.* **50**(5): 1128-1133.
16. Ferdous A, ME Ali, S Alam, T Hossain, U Hany, J Dissing and S Akhteruzzaman 2006. Genetic data of 10 autosomal STR loci in Bangladeshi population. *Leg. Med.* **8**: 297-299.
17. Ferdous A, ME Ali, S Alam, M Hasan, T Hossain and S Akhteruzzaman 2009. Forensic evaluation of PowerPlex™ 16 loci in a Bangladeshi population. *Leg. Med.* **11**: 198-199.
18. Ali ME, A Ferdous, S Alam, U Hany, T Hossain, M Hasan and S Akhteruzzaman 2008. Identification of variant alleles at AmpFISTR SGM Plus STR loci in a sample population of Bangladesh. *African J. Biotechnol.* **7**: 3603-3605.

*(Manuscript received on 13 April, 2013; revised on 16 June, 2013)*