Analysis of an “off-ladder” allele at the Penta D short tandem repeat locus

Y.L. Yang[1], J.G. Wang[2], D.X. Wang[1], W.Y. Zhang[1], X.J. Liu[3], J. Cao[1] and S.L. Yang[4]

[1]Clinical Laboratory Center, Traditional Chinese Medical Hospital of Langfang City, Langfang, Hebei, China
[2]Department of Internal Medicine, Traditional Chinese Medical Hospital of Langfang City, Langfang, Hebei, China
[3]Department of Forensics Laboratory, Nanning Institute of Transfusion Medicine, Nanning, Guangxi, China
[4]Department of Hematology, Traditional Chinese Medical Hospital of Langfang City, Langfang, Hebei, China

Corresponding author: S.L. Yang
E-mail: shulianyang@126.com

Genet. Mol. Res. 14 (4): 15096-15101 (2015)
Received April 23, 2015
Accepted July 22, 2015
Published November 24, 2015
DOI http://dx.doi.org/10.4238/2015.November.24.18

ABSTRACT. Kinship testing of a father and his son from Guangxi, China, the location of the Zhuang minority people, was performed using the PowerPlex® 18D System with a short tandem repeat typing kit. The results indicated that both the father and his son had an off-ladder allele at the Penta D locus, with a genetic size larger than that of the maximal standard allelic ladder. To further identify this locus, monogenic amplification, gene cloning, and genetic sequencing were performed. Sequencing analysis demonstrated that the fragment size of the Penta D-OL locus was 469 bp and the core sequence was $[AAAGA]_{21}$, also called Penta D-21. The rare Penta D-21 allele was found to be distributed among the Zhuang population from the Guangxi

Zhuang Autonomous Region of China; therefore, this study improved the range of DNA data available for this locus and enhanced our ability for individual identification of gene loci.

Key words: Cloning; Sequencing; Penta D allele; Short tandem repeat; Zhuang nationality

INTRODUCTION

Short tandem repeats (STRs), also known as microsatellites, are repeating sequences of 2-7 bp DNA, and an alteration in the number of repeating sequences can result in individual differences and genetic polymorphisms at STR loci (Dore et al., 2001). STRs have many characteristics including wide distribution across the human genome, easy detection, enormous information content, and high numbers of polymorphisms. Furthermore, STRs follow the laws of Mendelian co-dominant inheritance. Their use has been successfully applied in several fields including genetic mapping, linkage analysis, kinship testing, and individual identification (Narkuti et al., 2008; Chen et al., 2012). Commercial kits are also available and utilize multiplex polymerase chain reaction (PCR) and capillary electrophoresis analysis to allow the simultaneous detection of several STR loci (Miozzo et al., 2007). The commercial kit PowerPlex® 18D (Promega, Madison, WI, USA) offers testing of the core CODIS loci plus amelogenin and the loci D2S1338, D19S433, Penta E, and Penta D (Myers et al., 2012). Penta D is a five nucleotide repeat STR locus located on 21q22.3, with the core repeated unit AAAGA (Chen et al., 2012), and was one of the common loci used for kinship testing and to establish the human DNA database. In this study, we utilized the PowerPlex® 18D system to perform kinship testing and found an off-ladder (OL) allele at the Penta D locus in the Zhuang minority population in China, which exhibited a size larger than that of the maximal standard allelic ladder. This study used monogenic PCR amplification, gene cloning, and DNA sequencing to analyze the PentaD-OL allele for its use in kinship testing.

MATERIAL AND METHODS

Subjects

A father and his son were identified as Zhuangs, and in order to register their residence, kinship testing was performed to identify their relationships. Their blood specimens were collected and diluted with pure water before shakening vigorously for 15 minutes at room temperature; then centrifuged to collect the precipitate. This study was conducted in accordance with the Declaration of Helsinki. This study was conducted with approval from the Ethics Committee of the Traditional Chinese Medical Hospital of Langfang City. Written informed consent was obtained from all participants.

DNA extraction

For this study, we collected 1.5 µL ethylenediaminetetracetic acid (EDTA) whole-blood samples, from which the DNA was extracted using the Chelex 100 (Bio-Rad, Hercules, CA, USA) procedure as previously described by Walsh et al. (1991).

PCR-STR

According to the PowerPlex® 18D System manufacturer instructions, PCR amplifications were performed in a GeneAmp PCR System 9700 (Applied Biosystems, Foster City, CA, USA) and an ABI3130 Genetic Analyzer instrument (Applied Biosystems) was used to perform capillary electrophoresis for sequence analysis. The Genemapper v3.2 software (Applied Biosystems) was used to perform data analysis.

The Penta D locus was amplified by monoplex PCR using the primers as listed in STRBase (http://www.cstl.nist.gov/biotech/strbase/str_Penta_D.htm): Penta-D forward: 5' - GAA GGT CGA AGC TGA AGT G - 3', and Penta-D reverse: 5' - ATT AGA ATT CTT TAA TCT GGA CAC AAG - 3'. PCR amplification was performed in a total volume of 50 µL containing 5 µL 10X Buffer (Mg^{2+}), 4 µL 2.5 mM of each dNTPs (Qiagen, Hilden, Germany), 0.3 µL exTaq (5U) (Qiagen), 2.5 µL 10 µM primers (Promega Corporation), 3 µL template DNA, and 35.3 µL ddH_2O. Amplification cycling conditions were as described in the PowerPlex® 18D System technical manual.

Sequence analysis

The PCR products were purified using the PCR product recycling kit (Huashun Ltd., Shanghai, China), following the manufacturer protocol. Recycled PCR products were ligated to the pGEM-T easy vector (Promega) and transformed into *Escherichia coli* DH5α (Sambrook and Russell, 2001). Positive clones were sequenced on a 3130 Genetic Analyzer using the Big Dye v3.1 sequencing kit (Life Technologies, Carlsbad, CA, USA).

RESULTS

PCR-STR

We detected 17 STR loci using PowerPlex® 18D System fluorescence detection kits, and except for the Penta D locus, 16 of these followed the laws of Mendelian inheritance (Table 1). The father's Penta D locus was 9,OL (Figure 1A) and his son's Penta D locus was 10,OL (Figure 1B); the OL locus was found at approximately 20 bp after the Allelic Ladder (2.2-17). In order to confirm the variant, amplifications were repeated on the same specimens, and the results were consistent with those obtained previously.

Table 1. Results of PCR-STR analysis of the father and his son.

Locus	Father	Son	Locus	Father	Son	Locus	Father	Son
D3S1358	15,16	15,15	D13S317	9,12	11,12	D8S1179	10,15	10,16
TH01	7,9	9,9	D7S820	9,12	10,12	TPOX	11,12	9,11
D21S11	30	30,32.2	D16S539	12,12	11,12	FGA	22,27	22,27
D18S51	13,15	13,15	CSF1PO	10,12	10,11	D19S433	14.2,15.2	15.2,16.2
Penta E	11,16	11,21	AMEL	X,Y	X,Y	D2S1338	19,20	17,20
D5S818	10,12	10,12	vWA	14,16	14,18	Penta D	9,?	10,?

PCR-STR = polymerase chain reaction-short tandem repeat.

Sequencing analysis

The Penta D locus from father and his son was subjected to monogenic amplification,

cloning, and sequencing. Sequencing analysis found that the suspicious locus was identical between the father and his son, and that the fragment size was 469 bp. The core sequence of the OL allele at the Penta D locus was $[AAAGA]_{21}$ (Figure 2) and consisted of 21 complete repeated units. Furthermore, this allele had 4 [AAAGA] repeat units more than the maximal allele of Penta D-17. Therefore, we concluded that both the father and his son carried the Penta D-21 allele through inheritance.

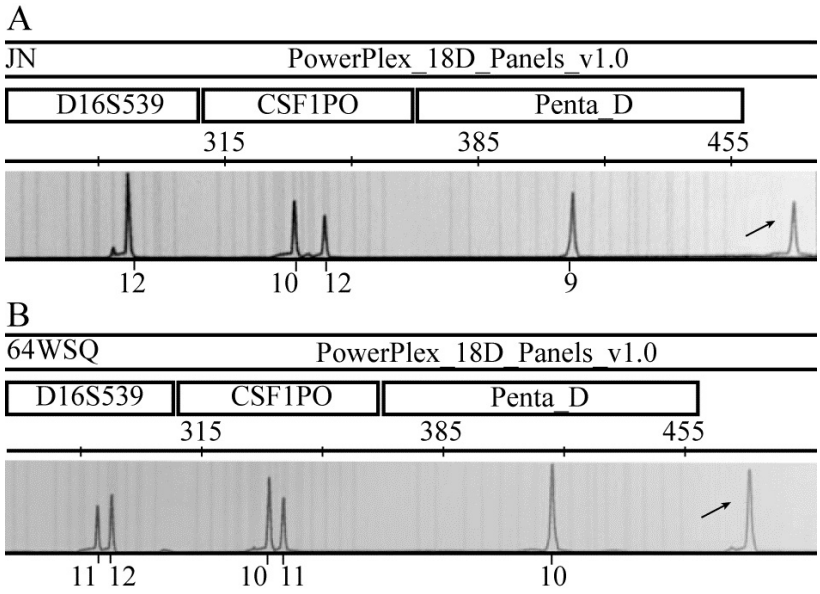

Figure 1. Penta D allele typing map. The arrow indicates the suspicious locus. **A.** Father (9, ?); **B.** Son (10, ?).

Figure 2. Sequencing result of the OL allele at the Penta D locus (the core sequence is underlined). OL = off-ladder.

DISCUSSION

Autosomal STR loci have many advantages for molecular research; for example, they are present in a great number of species, are widely distributed, have high polymorphic content, are easily detectable, currently are the type of genetic marker widely used for DNA kinship testing in the laboratory, they play an important role kinship testing, and are predicted to have extensive future use in this field (Ji and Shen, 2013). Although the ladders included in commercial STR kits contain the majority of common allele fragments, newly identified rare alleles might also be discovered during large scale STR typing that would not have corresponding fragments represented by the ladder alleles (Li et al., 2008). Rare alleles have low distribution frequencies in humans. Several reports have shown and characterized variant alleles in STR systems, including alleles shortened by one base in the Penta E locus, caused by a partial repeat motif (Mizuno et al., 2003; Park et al., 2012), inter-alleles in the D21S11 locus with a TA insertion in the variable region3 (Brinkmann et al., 1996), and a dinucleotide insertion within the 3' flanking region of the D18S51 locus (Barber and Parkin, 1996). In 2013, Ozeki and Tamaki (2013) performed an STR analysis in a Japanese population and found rare alleles at the FGA, D5S818, and D19S433 loci. Davis et al. (2012) also found rare alleles at the SE33 locus by DNA sequencing analysis. Furthermore, the variant allele 9.2 of the Penta D locus was first reported in 2007 (Miozzo et al., 2007). Therefore, until now, there were 38 rare alleles reported at this locus (www.cstl.nist.gov/biotech/strbase/var_Penta_D.htm); however the Penta D-21 allele identified in this study has not previously been reported. In comparison, two types of loci beyond the allele ladder have been reported; one was out of the ladder range and the other was between two alleles (Crouse et al., 1999). Here, we identify Penta D-21 as belonging to the former category. The Penta D-21 allele was beyond the ladder range and had 4 complete repeat units more than the maximal allele; in addition, this locus was shown to been transmitted to the next generation.

There are approximately 365 population studies on allele frequencies at STR loci from around the world (www.cstl.nist.gov/div831/strbase). Alleles varied between different population and allelic frequency distributions exhibited obvious geographic differences, and some studies found that there were significant differences between Chinese and American populations (Semikhodskii et al., 2012). This study is the first to report the Penta D-21 allele in the Zhuang Ethnic Group of Guangxi Province, China.

The occurrence of low frequency rare alleles easily leads to misjudgements and affects the establishment of DNA databases and the exchange of information between laboratories. Identification of rare alleles not only improves individual reorganization of STR loci and the elimination rate of incorrect alleles, but also benefits data collection in order to generate more appropriate ladders for the Chinese population. Furthermore, individual identification and kinship testing by DNA typing are more proper to Chinese population. Timely identification of STR loci in clinical practice and the performance of deep genetic analysis play essential roles in kinship testing. In conclusion, the results of this study have important significance for establishment of a complete DNA database and improvement of loci that might be used for individual identification.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

Research supported by the Science and Technology Research and Development Program of Lang Fang City (#2014013124) and the Shulian Yang Prestigious Doctor Studio.

REFERENCES

Barber MD and Parkin BH (1996). Sequence analysis and allelic designation of the two short tandem repeat loci D18S51 and D8S1179. *Int. J. Legal Med.* 109: 62-65.

Brinkmann B, Meyer E and Junge A (1996). Complex mutational events at the HumD21S11 locus. *Hum. Genet.* 98: 60-64.

Chen JG, Pu HW, Chen Y, Chen HJ, et al. (2012). Population genetic data of 15 autosomal STR loci in Uygur ethnic group of China. *Forensic Sci. Int. Genet.* 6: e178-179.

Crouse CA, Rogers S, Amiott E, Gibson S, et al. (1999). Analysis and interpretation of short tandem repeat microvariants and three-banded allele patterns using multiple allele detection systems. *J. Forensic Sci.* 44: 87-94.

Davis C, Ge J, King J, Malik N, et al. (2012). Variants observed for STR locus SE33: a concordance study. *Forensic Sci. Int. Genet.* 6: 494-497.

Dore MP, Realdi G, Mura D, Onida A, et al. (2001). Genetic instability in chronic viral hepatitis and hepatocellular carcinoma. *Hum. Pathol.* 32: 698-703.

Ji FQ and Shen XL (2013). Research progress of autosomal STR loci genetic paternity testing. *Int. J. Lab. Medicine* 34: 2871-2873.

Li CT, Guo H, Zhao ZM and Li L (2008). Analysis of genetics and genomics of short tandem repeat loci commonly used in kinship testing. *Fa Yi Xue Za Zhi* 24: 214-220.

Miozzo MC, Maxzud MK, Pacharoni CM, Mutal SA, et al. (2007). Characterization of the variant allele 9.2 of Penta D locus. *J. Forensic Sci.* 52: 1073-1076.

Mizuno N, Sekiguchi K, Sato H and Kasai K (2003). Variant alleles on the Penta E locus in the PowerPlex 16 kit. *J. Forensic Sci.* 48: 358-361.

Myers BA, King JL and Budowle B (2012). Evaluation and comparative analysis of direct amplification of STRs using PowerPlex® 18D and Identifiler® Direct systems. *Forensic Sci. Int. Genet.* 6: 640-645.

Narkuti V, Vellanki RN, Anubrolu N, Doddapaneni KK, et al. (2008). Single and double incompatibility at vWA and D8S1179/D21S11 loci between mother and child: implications in kinship analysis. *Clin. Chim. Acta* 395: 162-165.

Ozeki M and Tamaki K (2013). Allele frequencies of 37 short tandem repeat loci in a Japanese population. *Leg. Med.* 15: 342-346.

Park H, Ra EK, Park SS and Seong MW (2012). Detection of very large off-ladder alleles at the PentaE locus in a 15 locus autosomal STR database of 199 Korean individuals. *Forensic Sci. Int. Genet.* 6: e189-e191.

Sambrook J and Russel DW (2001). Molecular Cloning: A Laboratory Manual. 3rd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 96-99.

Semikhodskii A, Kozub NA and Sozinov IA (2012). Genetic data of 15 STR loci in the Caucasian population of the Russian Federation. *Tsitol. Genet.* 46: 53-58.

Walsh PS, Metzger DA and Higuchi R (1991). Chelex 100 as a medium for simple extraction of DNA for PCR based typing from forensic material. *Biotechniques* 10: 506-513.