Legal Medicine 30 (2018) 10–13

Contents lists available at ScienceDirect

Legal Medicine

journal homepage: www.elsevier.com/locate/legalmed

# Null alleles and sequence variations at primer binding sites of STR loci within multiplex typing systems

Yining Yao[a,1], Qinrui Yang[a,1], Chengchen Shao[a], Baonian Liu[a], Yuxiang Zhou[a], Hongmei Xu[a], Yueqin Zhou[a], Qiqun Tang[b], Jianhui Xie[a,*]

[a] Department of Forensic Medicine, Shanghai Medical College of Fudan University, Shanghai 200032, China
[b] Department of Biochemistry and Molecular Biology, Shanghai Medical College of Fudan University, Shanghai 200032, China

ARTICLE INFO

*Keywords:*
Short tandem repeats
Genotyping
Null alleles
Primer binding sites
Sequence variations

ABSTRACT

Rare variants are widely observed in human genome and sequence variations at primer binding sites might impair the process of PCR amplification resulting in dropouts of alleles, named as null alleles. In this study, 5 cases from routine paternity testing using PowerPlex®21 System for STR genotyping were considered to harbor null alleles at TH01, FGA, D5S818, D8S1179, and D16S539, respectively. The dropout of alleles was confirmed by using alternative commercial kits AGCU Expressmarker 22 PCR amplification kit and AmpFℓSTR®. Identifiler® Plus Kit, and sequencing results revealed a single base variation at the primer binding site of each STR locus. Results from the collection of previous reports show that null alleles at D5S818 were frequently observed in population detected by two PowerPlex® typing systems and null alleles at D19S433 were mostly observed in Japanese population detected by two AmpFℓSTR™ typing systems. Furthermore, the most popular mutation type appeared the transition from C to T with G to A, which might have a potential relationship with DNA methylation. Altogether, these results can provide helpful information in forensic practice to the elimination of genotyping discrepancy and the development of primer sets.

## 1. Introduction

Short tandem repeat (STR) analysis serves as a dominant means in individual identification and paternity testing, thus a variety of multiplex typing systems have been developed [1–8]. Efforts have been made to obtain ideal primers, which optimize the assay to promote both specificity and efficiency of PCR amplification. However, rare variants at primer binding sites inevitably occur among populations [9], which might lead to unexpected mismatches in PCR amplification. Variants from single base modifications at certain regions of primer binding sites potentially affect annealing and/or elongation of primers, which can result in the dropout of one or both of the alleles presenting at the particular locus [10]. The dropout of alleles from sequence variation at primer binding sites of STRs is also known as null alleles or silent alleles [11]. As a matter of fact, concerns have been raised about null alleles since a genotyping error from commercial STR typing systems or a discrepancy between them could bring up a false or inaccurate interpretation in forensic practice [12]. Therefore, it is important to characterize variants at primer binding sites of STRs within multiplex typing systems to eliminate genotyping error or discrepancy in individual identification and paternity testing.

Null alleles are more possibly observed in concordance studies where different PCR primer sets are introduced to STR typing, and in paternity tests where uncertain discrepancy of allele transfers occurs [13]. In this study, 5 cases with uncertain discrepancy of allele transfers were observed using PowerPlex®21 System (PP21) at TH01, FGA, D5S818, D8S1179, and D16S539, respectively. The dropout of alleles at these 5 STRs was verified using AGCU Expressmarker 22 PCR amplification kit (EX22) and AmpFℓSTR® Identifiler® Plus Kit (Identifiler Plus), and sequence variations were determined by sequencing. Furthermore, null alleles and variants at primer binding sites of STRs from commonly used multiplex typing systems were collected from previous reports for further investigation.

## 2. Materials and methods

### 2.1. Samples collection, DNA extraction, PCR amplification, and electrophoresis

Samples from 5 independent cases in routine paternity tests were

* Corresponding author.
*E-mail address:* jhxie@fudan.edu.cn (J. Xie).
[1] These authors equally contributed to this work.

http://dx.doi.org/10.1016/j.legalmed.2017.10.007
Received 15 July 2017; Received in revised form 22 October 2017; Accepted 27 October 2017
Available online 28 October 2017

collected with the obtainment of informed consent. Genomic DNA was extracted by using Chelex-100 and proteinase K [14]. The polymerase chain reaction (PCR) was performed using PP21 (Promega, USA), EX22 (AGCU, China) or Identifiler Plus (Applied Biosystems, USA) in a GeneAmp PCR System 9700 (Applied Biosystems, USA) according to the manufacturers' recommendations. PCR products were separated by capillary electrophoresis in ABI PRISM 3130xL Genetic Analyzer (Applied Biosystems, USA). Allele designation was determined according to allelic ladders by using the GeneMapper® ID software v3.2 (Applied Biosystems, USA). All procedures in this study were approved by the ethics committee of Shanghai Medical College, Fudan University.

### 2.2. DNA sequencing

Novel primer sets for TH01, FGA, D5S818, D8S1179, and D16S539 (Table S1) were used for PCR amplification according to the conditions consisting of an initial denaturation at 95 °C for 5 min followed by 30 cycles of 95 °C for 30 s, 60 °C for 30 s, 72 °C for 1 min as well as a final extension at 72 °C for 10 min. PCR products were separated by agarose gel electrophoresis and sequences were determined by Sanger sequencing (Sangon Biotech., China). Sequence variations in the flanking regions were obtained by aligning with reference sequences from UCSC Genome Browser [15] and designation was determined by calculating the number of base pairs upstream or downstream from the core region of STRs according to ISFG recommendations concerning nomenclature of STR alleles [16].

### 2.3. Collection of reported null alleles

STR loci with null alleles in various commercial typing kits were well collected from STRBase [17]. Following, additional searching of null alleles was carried out in PubMed database using key words without any limitation including 'null allele', 'silent allele', 'dropout of allele', 'STR', 'primer binding site mutation', 'concordance', 'discordant', and/or 'discrepancy'. Detailed information of STR loci with null alleles and primer binding site mutations was collected.

## 3. Results and discussion

### 3.1. Sequence variations of primer binding sites at 5 STR loci

After the genotyping of 5846 paternity testing cases (including 721 fathermother-child trios, 4232 father-child duos, and 893 mother-child duos) using PP21 [18], 5 uncertain discrepancy of allele transfers from parents to children at TH01, FGA, D5S818, D8S1179, and D16S539, respectively (Table 1). The combined parentage index (CPI) is less than 10000 in each case according to ISFG's recommendation on biostatistics in paternity testing [19], which leaves real paternity questioned. STR genotyping was subsequently performed using both EX22 and Identifiler Plus, and results showed an extra allele at the mentioned STR loci in individuals, which verified the dropout of alleles (Table 1). Sequence variations at primer binding sites of 5 STR loci were determined by sequencing, indicating a single point mutation at each locus (Fig. 1), which were correspondingly an alteration from G to A at 27 bases upstream from the repeat region $[TCAT]_n$ of TH01 (GenBank sequence D00269), an alteration from C to T at 81 bases downstream from the repeat region $[TTTC]_3TTTTTTCT[CTTT]_nCTCC[TTCC]_2$ of FGA (GenBank sequence M64982), an alteration from C to A at 90 bases upstream from the repeat region $[AGAT]_n$ of D5S818 (GenBank sequence AC008512.4), an alteration from A to C at 89 bases downstream from the repeat region $[TATC]_n$ of D8S1179 (GenBank sequence AF216671), and an alteration from C to A at 34 bases upstream from the repeat region $[GATA]_n$ of D16S539 (GenBank sequence AC008512.4). As null alleles of the same position and mutation type at D5S818 have been already reported [20], there might be a higher frequency of transition at this position in Chinese Han population. Although the primer sequence for STR loci of PP21 is not available, the approximate primer binding site could be inferred based on previous reports. Therefore, a single base variation at inferred primer binding site should be responsible for the dropout of alleles in STR genotyping.

When both alleged father (mother) and child harbor a single allele at one STR locus with discrepant transfer of alleles, null alleles should be considered [21,22]. Generally, a single point modification at primer binding site might have not a sound impact on the annealing of primers although primer binding site mutations might impair the allelic balance. However, as commercial STR typing kits require a rigorous condition for PCR amplification, the binding of primers to primer binding sites when a single base modification occurs might fail, therefore giving rise to null alleles. Adjustment of PCR amplification conditions, such as the temperature of annealing, might recover null alleles [23]. Furthermore, mismatches at 3′ end of the primers because of a single base modification could greatly impair the elongation of primers, which should be another factor for null alleles [10]. Since commercial amplification kits usually use different primer sets for the same locus, genotypes of STRs with null alleles could be possibly corrected by using other multiplex kits. In this study, genotypes of the 5 cases with null alleles obtained using PP21 were corrected using both kits. However, it should be noted that not a single STR typing system could avoid all potential null alleles since rare variants occur at random. In fact, cases with null alleles have been reported when genotyping using EX22 [24], or Identifiler Plus kit and AmpFℓSTR® Identifiler® PCR Amplification Kit (Identifiler) [25–32], which employ the same loci and primer sets. Rare circumstances additionally happened under which two or more STR loci with null alleles were observed in a single case detected by the same commercial multiplex kit [33].

### 3.2. STRs with null alleles within commonly used multiplex typing systems

Detailed information about STR loci with null alleles from commercial typing systems is collected in this study (Table S2) [3,10,17,20,21,23–47]. Although null alleles at most CODIS STR loci have been observed to date, there is no report about null alleles at D7S820 caused by primer binding site mutations.

In contrast to other commercial amplification systems, previous studies reported a majority of null alleles from PowerPlex®16 System (PP16). It might be in that the population detected by PP16 is larger than that by other multiplex typing systems. Compared to other STR loci within PowerPlex® typing systems, null alleles at D5S818 were

**Table 1**
Typing discrepancies between PP21 and additional typing systems.

| Loci | PP21 | | | EX22 | | | Identifiler Plus | | |
|---|---|---|---|---|---|---|---|---|---|
| | Alleged Father | Child | Mother | Alleged Father | Child | Mother | Alleged Father | Child | Mother |
| TH01 | 9 | 7 | / | 6/9 | 6/7 | / | 6/9 | 6/7 | / |
| FGA | 20/23 | 23 | 24 | 20/23 | 18/23 | 18/24 | 20/23 | 18/23 | 18/24 |
| D5S818 | 12 | 13 | 12/13 | 12 | 12/13 | 12/13 | 12 | 12/13 | 12/13 |
| D8S1179 | 15 | 12 | / | 14/15 | 12/14 | / | 14/15 | 12/14 | / |
| D16S539 | 11 | 9 | / | 9/11 | 9 | / | 9/11 | 9 | / |