FORENSIC SCIENCE JOURNAL SINCE 2002

Forensic Science Journal
2014;13(1):1-8
Available online at:fsjournal.cpu.edu.tw

Off-ladder Alleles and Tri-allelic Variants Observed in Short Tandem Repeats (STRs) Typing of Populations in Taiwan

Chun-Wei Tsai [1] , M.S.; Chung-Yen Pai [2] , Ph.D.; Chiou-Herr Yang [2*] , Ph.D

[1] Biological Office, Forensic Science Center, Criminal Investigation Bureau, National Police Administration, Taiwan (ROC)

[2] Department of Forensic Sciences, Central Police University, Taoyuan 33304, Taiwan (ROC)

Received: December 20, 2013 / Accepted: March 24, 2014

Abstract

Amplified STR alleles of tested individuals are determined by comparing to the allelic ladders provided by commercial kits' manufactures. However, alleles were sometimes found outside of this ladder range or slightly different length to allelic ladders, which are called off-ladder (OL) allele. A total of 10,974 STR profiles from Taiwanese were screened for variants using the AmpFlSTR® Identifiler kit. The results showed that 10 of 15 STR loci have non-ladder alleles, and the number of observed OL alleles was 307, with a percentage of 1.40%. Totally, we detected 29 different OL allele types from the 307 alleles, 10 located at outside the range of commercial allelic ladder, and 19 at inside this range. Among them, two OL types (allele 19.3 of D19S433 and allele 4.3 of D5S818) have not been reported in STRBase (http://www. cstl.nist.gov / biotech / strbase). The two most common OL variants observed in Taiwan were allele 9.1 of D7S820 (0.67%) and allele 30.3 of D21S11 (0.39%), which are different from and much higher than those in other populations except Chinese Han. In addition to the OL alleles, 3 loci in 4 samples with tri-allelic patterns (0.04%) were identified from the 10,974 individuals. The OL alleles and tri-allelic pattern data can help to increase the power of cases identification when samples are involved with those variant loci.

Keywords: forensic science, short tandem repeat (STR) DNA, off-ladder allele, tri-allelic pattern.

Introduction

Short tandem repeats (STRs) DNA are widely used in forensic and parentage identification. Samples' STR allele designation using commercial kits are usually determined by comparing a PCR-amplified DNA to their ladder markers [1, 2]. However, alleles have been occasionally found outside of the allelic ladder range or slightly different length to the ladder marker; we call them off- ladder (OL) allele [3, 4]. In rare cases, three alleles were also found at a single STR locus, called a tri-allelic pattern [5].

OL alleles have been reported in literatures [6, 7] and websites [(http://www. cstl.nist.gov / biotech / strbase)]. Some of them are shorter or longer than the allelic ladder with one or more full repeat units (deletion or insertion), and the others contain a slightly different length (microvariant, deleted or inserted 1 or 2 bp in a repeat unit) comparing to commercial ladder allele. Tri-allelic patterns are rare but could be occasionally found from a large sample size [8, 9]. These patterns were formed by the process of cell differentiation due to localized duplication of a locus or due to chromosomal trisomy. The frequencies of OL alleles and tri-allelic patterns quite vary between populations [10, 11], so it is important to establish the variants data of each population for users.

The OL alleles and tri-allelic patterns data can increase the power of discrimination in individual identification, paternity or sibling testing, since these

* Corresponding author: una033@mail.cpu.edu.tw

variant alleles are very rare [10]. These variants would also be helpful for the interpretation of practical cases involving a decomposed or mixture DNA.

Materials and Methods

Oral swab samples were taken from 10974 individuals and then most of the samples were preserved in FTA card. In this study, most DNA samples were extracted from FTA cards and a few were from cotton swab. FTA buccal spots were prepared by transferring oral cotton swab onto a 2.5 cm diameter FTA card (Whatman FTA®), the cards were then dried and stored at room temperature. A 2-mm diameter punch was taken from each subject's FTA card and prepared for PCR by washing with 200μL FTA purification reagent and twice with 200μL TE buffer. The sample is finally washed with 200μL water and air-dried at room temperature, ready for DNA amplification. For cotton swab samples, DNA extractions were carried out by using the QIAmp mini commercial DNA extraction kit (Qiagene, Valencia, CA, USA) according to the manufacturer's instruction.

The commercial STR kit used for typing was AmpFlSTR® IdentifilerTM (Applied Biosystems, Foster City, USA), including 15 STR loci, D8S1179, D21S11, D7S820, CSF1PO, D3S1358, TH01, D13S317, D16S539, D2S1338, D19S433, vWA, TPOX, D18S51, D5S818, FGA , and a gender marker of Amelogenin.

PCR amplification was performed in 25μl of reaction mixture in a 9700 thermal cycler (Applied Biosystems) with the following conditions: denatured at 95°C for 10 min, followed by 28 cycles of 95°C for 10 sec, 60°C for 30 sec and 72°C for 30 sec [Applied Biosystems, AmpFlSTR Identifiler PCR Amplification Kit User's Manual]. The PCR products were checked on a 3 % agarose gel, and then stored at 4°C before analyses.

The STR profiles were identified with an Applied Biosystems 3130 XL Genetic Analyzer or ABI Prism® 3500-24 according to the protocol of the manufacturer (Applied Biosystems). The genotypes were determined by the software of geneMapper ID v3.1 [Applied Biosystems, AmpFlSTR Identifiler PCR Amplification Kit User's Manual].

DNA extraction, PCR amplification, and genotype determination were performed at least twice when samples had OL alleles or tri- allele patterns. Off-ladder allele designations were calculated using the following formula [12].

$\delta_1 = S_Y - L_Y$, $S_2 = S_{OL} - L_X$, $c = | \delta_1 - \delta_2 |$

δ_1 represents the size difference (in base pairs) between sister allele Y (S_Y) and ladder allele Y (L_Y). δ_2 represents the size difference (in base pairs) between off-allele S_{OL} and the ladder allele X close to the OL. The c value is the size difference between the OL and the ladder allele. Allele designations for off-allele lying outside of the ladder ranges were extrapolated by comparing the OL sizes to the smallest or largest ladder alleles [10]. OL frequencies were calculated by dividing the number of observations by the product of 2N (N, the number of individuals studied).

Likelihood ratios in the form of sibling indices (SIs), as well as the probability of sibling statistics, were calculated using allele frequencies from an in-house database and formulae derived elsewhere [13].

Results and discussion

Off-ladder (OL) alleles

STR profiles were screened from 10,974 individuals by using the AmpFlSTR® Identifiler kit. The results revealed that off-ladder alleles were observed in 10 of 15 STR loci in the kit. A total of 29 OL types were found from the 307 OL alleles out of 10,974 samples (1.40%, 307/(10,974×2)). Among the 29 OL types, 10 are outside the ladder alleles range and appeared at the 6 loci: D3S1358, D13S317, D19S433, D18S51, D5S818 and FGA. The other 19 OL types are found to be inside the ladder range and detected at 10 loci (Table 1). Seven of the 29 OL types were shorter or longer than the sizes of the shortest or longest ladder allele with one or more full repeat unit (deletion or insertion), and the other 22 OL types contained a slightly different length (microvariant, 1 or 2 bp of deletion or insertion). The OL allele lying outside of ladder range can only be represented as larger than the largest allele or as smaller than the smallest allele, since their DNA sequences were not determined in this study. Thus, these OL alleles were just called, in this study, closely similar off-allele. For all the OL alleles in a single STR locus, their designations were estimated by comparing the OL allele size to the smallest or largest ladder allele. This length estimation by extrapolation method had been used by other authors [10], who proved that the result of extrapolation method is close to that by

DNA sequence analysis.

There were 7 different kinds of OL alleles detected at D21S11 and 5 OL alleles detected at FGA. Comparing the OL alleles detected in this study with those of the known data, we found that 27 of 29 OL had been reported in STRBase (http://www. cstl.nist.gov / biotech / strbase) ; the other two unreported alleles found in this study are allele 19.3 of D19S433 and allele 4.3 of D5S818.

The most common OL allele in this study is allele 9.1of D7S820, representing 47.88% (147/307) of OL alleles, with a population frequency of 0.67% (147/(10,974×2), Table 1). The second one is allele 30.3 of D21S11 with a population frequency of 0.39% (85/(10,974×2)). These two alleles are also the highest frequent alleles detected in Chinese with a frequency of 0.40% (allele 9.1 of D7S820, 81/(10,071×2)) and 0.16% (allele 30.3 of D21S11, 32/(10,071×2)), but they are not common in other populations, only 0.01% (4/(16,185×2)) of allele 9.1 of D7S820 detected in the Caucasian [10] and 0.004% (1/(14,015×2)) of allele 30.3 of D21S11 in African American [10]. On the other hand, the two highest OLs in African American are allele 33.1 of D21S11 (0.16%, 45/(14,015×2)) and allele 9 of D3S1358 (0.11%, 32/(14,015×2)) [10]. The two highest OLs in Hispanic are allele 10.3 of D7S820 (0.21%, 8/(1,872×2)) and allele 20.1 of D3S1358 (0.08%, 3/(1,872×2)) [10]. The highest OL in Serbia and Bosia is allele 7.3 at D7S820 (0.11%, 24/(11,000×2)) [11]. The most frequent OL allele in Caucasian was found to be allele 20.1 of D3S1358 (0.06%, 19/(16,185×2)) [10]. The frequency of the most common OL allele in this study is much higher than that of other populations (Table 1), however, close to that of Chinese [14]. It also indicated that there is large difference in off-alleles frequency between different populations [10, 11, 15]. Therefore, it is important to have off-alleles data for each population, since forensic, paternity or sibling identification involving non-ladder alleles can increase the power of discrimination.

Table 1 Off-ladder (OL) allele frequency detected in different populations. N, the total number of samples tested; the corresponding off-ladder allele frequency in parentheses (%).

locus	OL allele	OL allele number observed					
		Taiwanese [1]	Chinese (Han) [2]	Caucasian [3]	African American [3]	Hispanic [3]	Bosnia & Serbia [4]
		N = 10,974	N = 10,071	N = 16,185	N = 14,015	N = 1,872	N = 11,000
D8S1179	7				1(0.0036)		
D21S11	21.1		1(0.0050)				
	24.3				16(0.0571)	1(0.0267)	
	25.3				1(0.0036)		
	27.1				4(0.0143)		
	28.3	2(0.0091)					
	29.1			1(0.0031)			
	29.3	1(0.0046)		7(0.0216)	4(0.0143)		
	30.3	85(0.3873)	32(0.1589)		1(0.0036)	1(0.0267)	
	31.1	1(0.0046)				1(0.0267)	
	32.1	1(0.0046)		1(0.0031)	1(0.0036)		
	33.1	1(0.0046)		4(0.0124)	45(0.1605)	2(0.0534)	4(0.0182)
	34.1	3(0.0137)		2(0.0062)	10(0.0357)	1(0.0267)	3(0.0136)
	35.1				9(0.0321)		
	36.1				1(0.0036)		
D7S820	6.3			3(0.0093)			
	7.3						24(0.1091)
	8.1					1(0.0267)	
	8.3				1(0.0036)		
	9.1	147(0.6698)	81(0.4021)	4(0.0124)	2(0.0071)		15(0.0681)
	9.3				2(0.0071)	1(0.0267)	1(0.0045)
	10.1	13(0.0592)	3(0.0149)	2(0.0062)	9(0.0321)	1(0.0267)	

locus	OL allele	OL allele number observed					
		Taiwanese [1]	Chinese (Han) [2]	Caucasian [3]	African American [3]	Hispanic [3]	Bosnia & Serbia [4]
		N = 10,974	N = 10,071	N = 16,185	N = 14,015	N = 1,872	N = 11,000
	10.3	1(0.0046)		1(0.0031)		8(0.2137)	
	11.1	5(0.0228)	4(0.0199)	4(0.0124)	1(0.0036)		3(0.0136)
	11.3				2(0.0071)		
	12.1			1(0.0031)			
	13.1			1(0.0031)	1(0.0036)	1(0.0267)	
	15		3(0.0149)				11(0.0500)
CSF1PO	10.1						1(0.0045)
	12.2	1(0.0046)					
	17						1(0.0045)
D3S1358	9			1 (0.0031)	32 (0.1142)		
	< 11, (10) [b]	1(0.0046)					
	15.1				1 (0.0036)		
	15.3	4(0.0182)					
	17.1			2(0.0062)			
	20.1			19(0.0587)	2 (0.0071)	3 (0.0801)	
	21		4(0.0199)				
TH01	5.2	1(0.0046)					
D13S317	< 7, (5) [b]	8(0.0364)	2(0.0099)	1(0.0031)	1(0.0036)		
	< 7, (6) [b]	3(0.0137)	2(0.0099)	2(0.0062)			22(0.0999)
	7.1					2(0.0534)	
	10.3	1(0.0046)					
	17						1(0.0045)
D16S539	6		2(0.0099)				
	12.1						2(0.0091)
	16		2(0.0099)				
D19S433	< 9, (8.2) [b]	1(0.0046)					
	19.2		1(0.0292)				
	> 19, (19.3) [a, b]	1(0.0046)					
vWA	18.3			1(0.0031)			
TPOX	7.3					1(0.0267)	
	14		3(0.0149)				
	15		4(0.0199)				
D18S51	13.3				3(0.0107)		
	16.1	1(0.0046)					
	17.3					1(0.0267)	
	18.1			2(0.0062)			
	> 27, (28) [b]	3(0.0137)					
	28.1					1(0.0267)	
D5S818	< 6, (4) [b]	1(0.0046)					
	< 6, (4.3) [a, b]	1(0.0046)					
	12.3				2(0.0071)		
	18				1(0.0036)		
FGA	< 16, (13) [b]	14(0.0638)	28(0.1390)				
	< 16, (14) [b]	3(0.0137)					

locus	OL allele	OL allele number observed					
		Taiwanese [1]	Chinese (Han) [2]	Caucasian [3]	African American [3]	Hispanic [3]	Bosnia & Serbia [4]
		N = 10,974	N = 10,071	N = 16,185	N = 14,015	N = 1,872	N = 11,000
	15			2(0.0062)			
	16.1				13(0.0464)		
	19.3			1(0.0031)			
	20.1			1(0.0031)			11(0.0500)
	20.3				1(0.0036)	1(0.0267)	
	21.1			1(0.0031)			
	22.3	1(0.0046)			3(0.0107)	1(0.0267)	4(0.0182)
	23.1	1(0.0046)					
	23.3				8(0.0285)	1(0.0267)	
	24.1				3(0.0107)		
	24.3			1(0.0031)	3(0.0107)		
	25.1					1(0.0267)	
	25.3				4(0.0143)		
	28.1	1(0.0046)					
	33.1				1(0.0036)		
	34.1				2(0.0071)		
	41.2				2(0.0071)	1(0.0267)	

a：The allele 4.3 of D5S818 and allele 19.3 of D19S433 were not previously reported in STRbase.

b：Closely similar off-alleles were shown in parentheses.

1：This study,

2, 3, and 4：Data from reference 13, 10, and 11 respectively.

Here, we reported an identification case, in which two male individuals' sibling relationship was determined. The sibling data (Table 2) showed that if the variant allele 30.3 of D21S11 was omitted, the combined sibling index (CSI) would be 5.52 or when it was mistyped as allele 31, the combined sibling index (CSI) came out as 8.04. Alternatively, when the allele 30.3 of D21S11 was included, the power of CSI would raise to 178.33, and increased the probability of sibling from 84.66% to 99.45%. It indicated that if the non-ladder allele was involved, no or only few additional testing would be needed for further STRs typing.

Table 2 Off-ladder (OL) alleles of D21S11 in the sibling identification case

genotype tested			sibling index		
STR Locus	elder brother	alleged brother	OL allele omitted	OL allele misdetermined	OL allele included
D8S1179	11/14	14/14	1.586	1.586	1.586
D21S11	29/30.3	30/30.3	-- [a]	1.457 [b]	32.525
D7S820	11/11	11/12	0.966	0.966	0.966
CSF1PO	10/13	10/13	8.180	8.180	8.180
D3S1358	15/16	15/15	0.970	0.970	0.970
TH01	6/9	6/9	3.958	3.958	3.958
D13S317	10/12	10/12	7.662	7.662	7.662
D16S539	11/12	9/13	0.250	0.250	0.250
D2S1338	18/24	19/21	0.250	0.250	0.250
D19S433	14/14	13/13	0.250	0.250	0.250
vWA	15/16	15/20	3.610	3.610	3.610

genotype tested			sibling index		
STR Locus	elder brother	alleged brother	OL allele omitted	OL allele misdetermined	OL allele included
TPOX	8/11	8/11	1.697	1.697	1.697
D18S51	17/18	14/16	0.250	0.250	0.250
D5S818	11/12	11/13	0.645	0.645	0.645
FGA	18/24	22/24	0.969	0.969	0.969
Combined sibling Index			5.521	8.041	178.328
Probability of sibling			84.664%	88.939%	99.446%

a：“--” The allele 30.3 omitted, the sibling index comes out as 1

b：The allele 30.3 mistyped as allele 31, the combined sibling index would be 8.041.

Triallelic patterns

From the 10974 individuals, three loci in 4 samples with tri-allelic pattern (0.04%, 4/10,974) were observed at D21S11, D13S317 and D19S433. Table 3 reveals that the frequency of our tri-allelic variants is lower than those in Chinese (0.08%, 8/10,500) [16, 17] and in Bosnia and Serbia (0.12%, 13/11,000) [11].

Table 3 Triallelic patterns observed in different populations

locus	No. of observation (%)		
	Taiwan [1] N=10,974	Chinese Han [2] N=10,500	Bosnia & Serbia [3] N=11,000
D8S1179			1(0.0091)
D3S1358		1(0.0095)	
D21S11	1(0.0091)	1(0.0095)	3(0.0273)
D5S818		1(0.0095)	
CSF1PO			1(0.0091)
TH01			1(0.0091)
D13S317	1(0.0091)	2(0.0190)	
D19S433	2(0.0182)		
vWA			1(0.0091)
TPOX			1(0.0091)
D18S51		2(0.0190)	3(0.0273)
FGA		1(0.0095)	2(0.0182)
Total	4(0.0364)	8(0.0762)	13(0.1182)
Triallelic variants observed in Taiwan, and all are type 1 pattern.			
D21S11	29/30/31		
D13S317	8/9/10		
D19S433	12/14/15, 12.2/14/15.2		

1：this study.

2：data from reference16,17.

3：data from reference 11.

In this study, the tri-allelic patterns were originally observed in 9 samples from routine analysis. Four individuals were detected to carry the tri-allelic variants after at least triple confirmation. All of tri-allelic variants are of the type I pattern, which is characterized by uneven peak heights for two mutated alleles that sum to the height of the unmutated allele [11]. Our result is similar to the literature reported by Huel RL [11], who found 12 type I of 15 tri-allelic patterns.

Conclusion

This study offered the data of off-ladder alleles and tri-allelic pattern detected in a large database of STR profiles from Taiwanese by using the AmpFlSTR® Identifiler kit. The two most common OL alleles in this study were allele 9.1 of D7S820 (0.67%) and allele 30.3 of D21S11 (0.39%). In addition to the OL alleles, 3 loci in 4 samples with tri-allelic pattern (0.04%) were identified from 10,974 individuals. The frequencies of our OL alleles are higher and tri-allelic patterns are lower comparing to the data of other populations. The off-ladder and tri-allelic variants would increase the power of discrimination in sibling identification. We suggest that these data should be established individually for different populations.

Reference

1. Jobling MA, Gill P. Encoded evidence: DNA in forensic analysis. Nat Rev (Genetics) 2004; 5(10): 739-51.
2. Griffiths RA, Barber MD, Johnson PE, Gillbard SM, Haywood MD, Smith CD, et al. New reference allelic ladders to improve allelic designation in a multiplex STR system. Int J Legal Med 1998; 111(5): 267-72.
3. Mizuno N, Sekiguchi K, Sato H, Kasai K. Variant alleles on the penta E locus in the PowerPlex 16 kit. J Forensic Sci 2003; 48(2): 358-61.
4. Margolis-Nunno H, Brenner L, Cascardi J, Kobilinsky L.A new allele of the short tandem repeat (STR) locus, CSF1PO. J Forensic Sci 2001; 46(6): 1480-3.
5. Lukka M, Tasa G, Ellonen P, Moilanen K, Vassiljev V, Ulmanen I. Triallelic patterns in STR loci used for paternity analysis: Evidence for a duplication in chromosome 2 containing the TPOX STR locus. Forensic Sci Int 2006; 164: 3-9.
6. Crouse CA, Rogers S, Amiott E, Gibson S, Masibay A. Analysis and interpretation of short tandem repeat microvariants and three-banded allele patterns using multiple allele detection systems. J Forensic Sci 1999; 44(1): 87-94.
7. Heinrich M, Felske-Zech H, Brinkmann B, Hohoff C. Characterization of variant alleles in the STR systems D2S1338, D3S1358 and D19S433. Int J Legal Med 2005; 119(5): 310-3.
8. Clayton TM, Guest JL, Urquhart AJ, Gill PD. A genetic basis for anomalous band patterns encountered during DNA STR profiling. J Forensic Sci 2004; 49: 1207-14.
9. Rolf B, Wiegand P, Brinkmann B. Somatic mutations at STR loci – a reason for three-allele pattern and mosaicism. Forensic Sci Int 2002; 126: 200-2.
10. Allor C, Einum DD, Scarpetta M. Identification and characterization of variant alleles at CODIS STR loci. J Forensic Sci 2005; 50: 1128-33.
11. Huel RL, Basić L, Madacki-Todorović K, Smajlović L, Eminović I, Berbić I, Milos A, Parsons TJ. Variant alleles, triallelic patterns, and point mutations observed in nuclear short tandem repeat typing of populations in Bosnia and Serbia. Croat Med J 2007; 48(4): 494-502.
12. Gill P, Urquhart A, Millican ES, Oldroyd NJ, Watson S, Sparkes R, Kimpton CP. A new method of STR interpretation using inferential logic-development of a criminal intelligence database. Int J Leg Med 1996; 109: 14-22.
13. Wenk R E, Traver M, Chiafari F A. Determination of sibship in any two persons. Transfusion, 1996; 36: 259-62.
14. Lu Hui-Ling, Tai Yun-chun, Liu Chao, Li Han-yan. Sequences of off-ladder alleles of PowerPlexTM16 kit in Chinese Han Population. J Forensic Med 2006; 22(3): 186-189. (in Chinese)
15. M.d. Eunus Ali, Ahmad Ferdous, Shafiul Alam, Ummey Hany, Tania Hossain, Mahamud Hasan and Sharif Akhteruzzaman. Identification of variant alleles at AmpFlSTR SGM Plus STR loci in a sample population of Bangladesh. African J Biotech 2008; 7(20): 3603-5.
16. Han Li-li, Pan Leng, Shen Xiao-li, Lin Sai-mei, LIN Li-fang, Tang Hai-yan, Hu Jie. Genetic analysis of 3 tri-allelic variant cases in paternity identification. Hainan Med J 2011; 22(15): 1-3. (in Chinese)
17. Zeng Yan-hong, Sun Hong-yu, Wu Xin-yao, Cai Gui-qing, Chen Yong. Five tri-banded allele patterns cases in analyzing STR Loci. Chinese J Forensic Med 2003; 18(2): 112-3. (in Chinese)