MedCrave
Step into the World of Research
Forensic Research & Criminology International Journal

Research Article
Open Access
CrossMark

Comparison of different STR typing for a special allele at D7S820 locus by using ten different STR multiplex system

Abstract

During a DNA database construction and population study in China, we observed a special genotype 10/10.1 at D7S820 locus which was misjudged by seven out of ten different STR multiplex systems. Allele typing results acquired by using ten diverse STR multiplex kits (PowerPlex® Fushion6C, PowerPlex® 21, PowerPlex® 24, ACGU® EX25, AppliedBio® HuaXiaBaijin, SureID® PanGlobal, Goldeneye™ 20A, GoldenEye™ 25A, AppliedBio® GlobalFiler, HuaDa® YanHuagn) and then analyzed by GeneMapper®1.4 software. In this paper, we conclude the reasons for the misjudgment of this allele as well as how to reduce these kinds of errors during STR analysis.

Keywords: short tandem repeat, D7S820, float bin, virtual bin, off-ladder, marker range, bin range, OL peak

Volume 7 Issue 3 - 2019

Xiao Lei
Beijing Institute of Radiation Medicine, China

Correspondence: Xiao Lei, Beijing Institute of Radiation Medicine, Beijing, P.R. China, Email xiaolei269@hotmail.com

Received: May 08, 2019 | **Published:** June 19, 2019

Introduction

Nowadays, more and more rare alleles are studied and supplemented to common Loci system, especially for the construction of DNA database and population studies. The ability to accurately judge the rare Loci becomes more essential. D7S820 is one of the useful markers for human identification, paternity and maternity testing and sex determination in forensic sciences.[1] It has been revealed 4 microvariant alleles: 8.1, 9.1, 10.1 and 10.3.[2] Here we designed a study to evaluate the performance of allele typing at D7S820 locus, with ten different STR multiplex systems, which mostly have been used in current forensic application for a long time. In this paper, we observed an abnormal genotype shown different results when tested by ten different STR multiplex kits.

Materials and methods

Genetic characterization of one special individual sample was carried out using blood (DNA extraction: CN-QIAamp-DNA-Investigator) (Qiagen). PCR products of D7S820 were generated using following kits: PowerPlex® Fushion6C, PowerPlex® 21, PowerPlex® 24, ACGU®EX25, AppliedBio® HuaXiaBaijin, SureID® PanGlobal, Goldeneye™ 20A, GoldenEye™ 25A, AppliedBio® GlobalFiler, HuaDa® YanHuagn. Automated fragment analysis was carried out on the ABI 3500 Genetic Analyzers (Applied Biosystems, Foster City, CA, USA). Allele designations were automatically assigned by GeneMapper 1.4 Software and every kit above by size comparison between sample alleles and allelic ladder alleles, run on the same gel or set of injections. Peaks are labled with the allele category and the calculated fragment size using the internal sizing standard (Liz600) with GeneMapper 1.4. The direct Taq-cyclesequencing method was performed.[3–7]

Results and conclusion

Parallel tests with AppliedBio®GlobalFiler shows two different allele typing: homozygote-10/10, 10.1/10.1 and likely split peaks. So we performed re-examination by using PowerPlex systems (PowerPlex® Fushion6c, PowerPlex® 21, PowerPlex® 24) (Figure 1 & Figure 2).

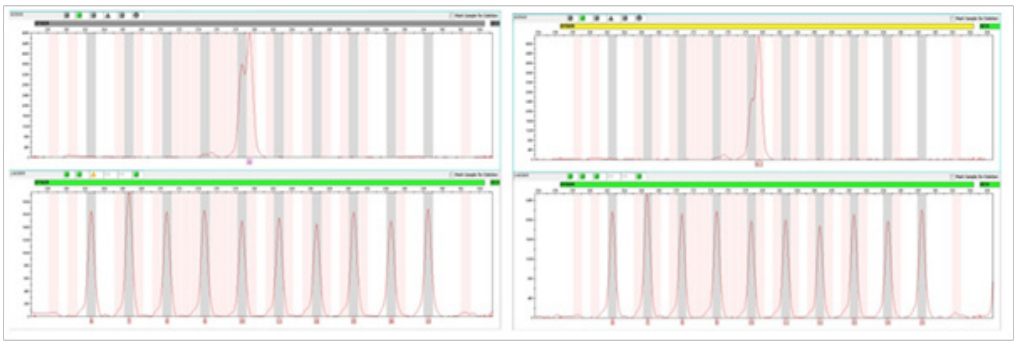


Figure 1 Allele typing by AppliedBio®GlobalFiler show different results.

PowerPlex® 21 (heterozygote 9.3/10.1) (lower left); PowerPlex® 24 (heterozygote 10/10.1) (lower right) show different results. Same sample was tested by all three kinds of PowerPlex kits: PowerPlex® Fushion6C, PowerPlex® 21 and PowerPlex® 24. Although, clearly, there is a split peak at this locus, two different heterozygous typing are observed: the first two kits have identical results:9.3/10.1. and the other kit is 10/10.1. Then we applied six other multiplex kits including ACGU® EX25, AppliedBio®HuaXiaBaijin, SureID® PanGlobal, GoldenEye™ 20a, GoldenEye™25a and HuaDa ® YanHuagn to find out the real peak-shape and the right typing of this rare allele (Figure 3).

After several repeated experiments were settled, the possibility of contamination was eliminated. From Figure 3 we can see that results of six kits are also varied from both peak shape and allele typing. Single-peak and double-peak pattern were observed and the typing include both homozygote and heterozygote. Some of the heterozygote typing have one OL (off ladder) peak within the white longitudinal stripe Figure 3(A–D), which means they are out of marker/bin range and cannot be marked by the system ladder automatically. Although we can label the off-ladder allele 10.1 manually, we still leave them unmarked Figure 3(A–D) for parallel comparison. In order to find out the real typing of this rare allele, we also applied Sanger Sequencing test.[8,9]

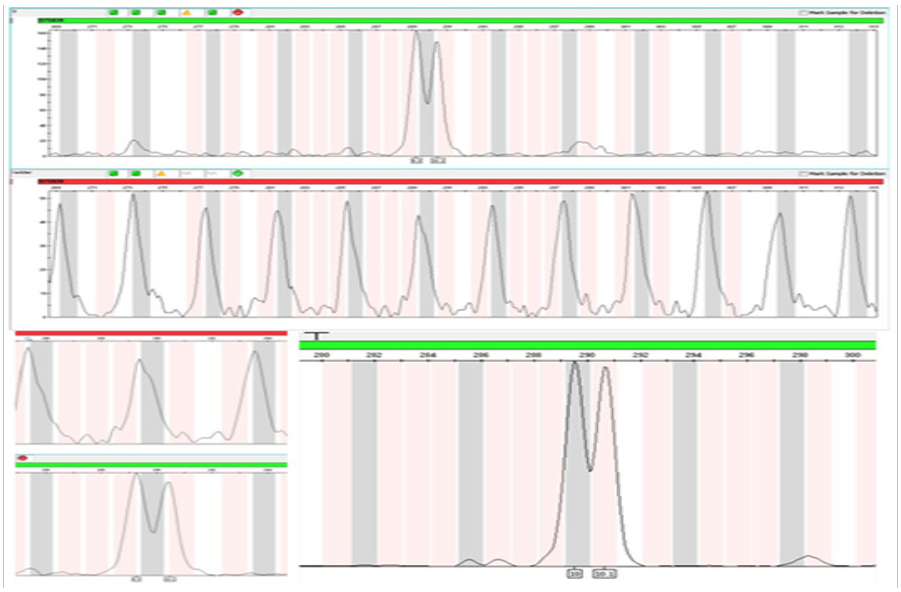


Figure 2 Allele typing by Promega PowerPlex®Fushion 6c kit (heterozygote 9.3/10.1) (uper).

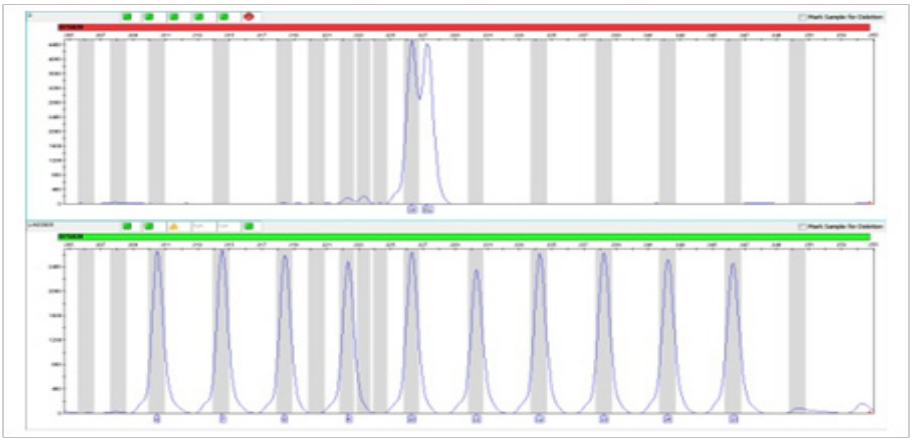


Figure 3(A) Allele typing by ACGU EX25 kit (heterozygote 10/OL).

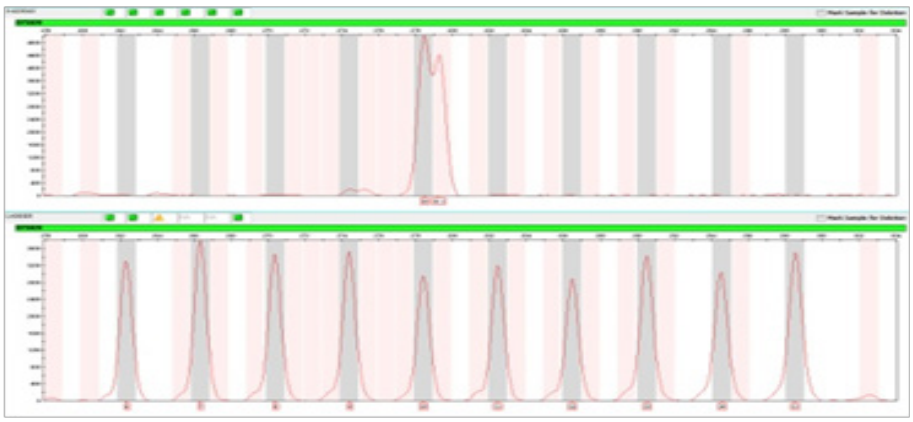


Figure 3(B) AppliedBio®HuaXiaBaijin kit (heterozygote 10/10.1).

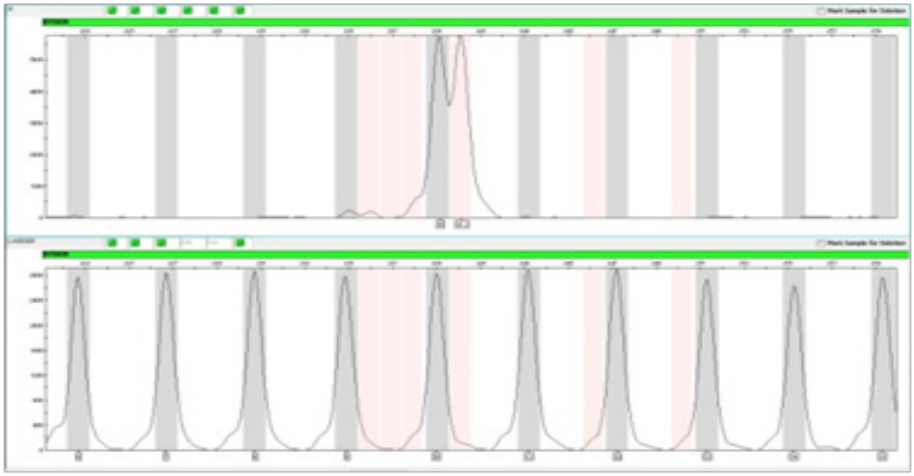


Figure 3(C) SureID®PanGlobal kit (heterozygote 10/10.1).

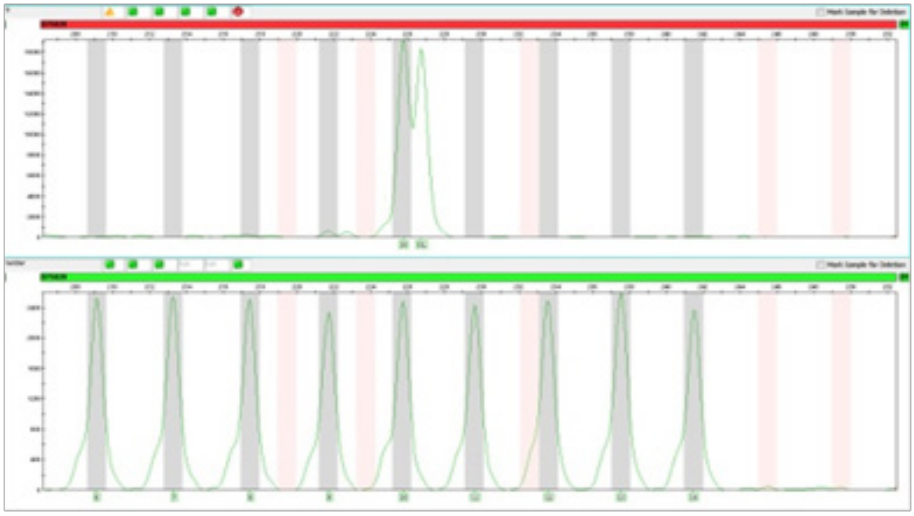


Figure 3(D) Goldeneye™20A kit (heterozygote 10/OL).

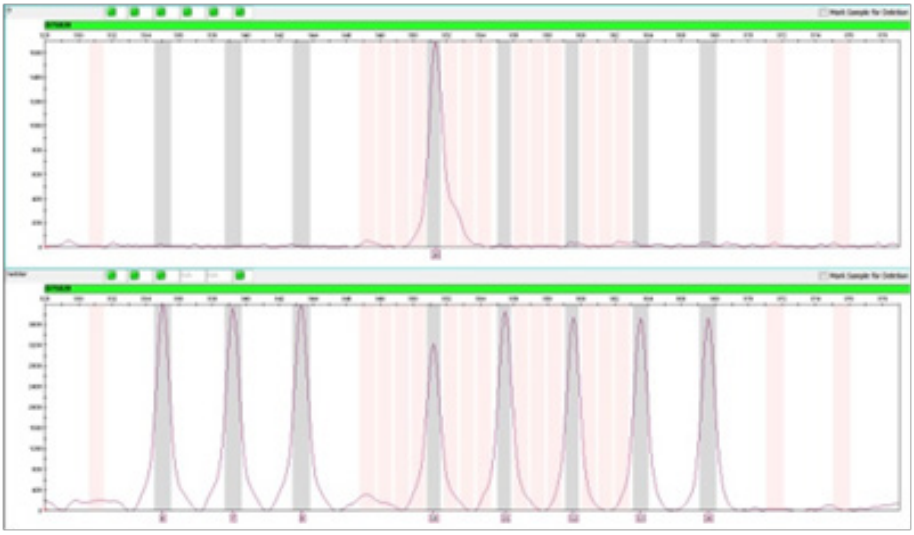


Figure 3(E) Goldeneye™25A kit (homozygote 10).

Citation: Lei X. Comparison of different STR typing for a special allele at D7S820 locus by using ten different STR multiplex system. *Forensic Res Criminol Int J.* 2019;7(3):135–142. DOI: 10.15406/frcij.2019.07.00277

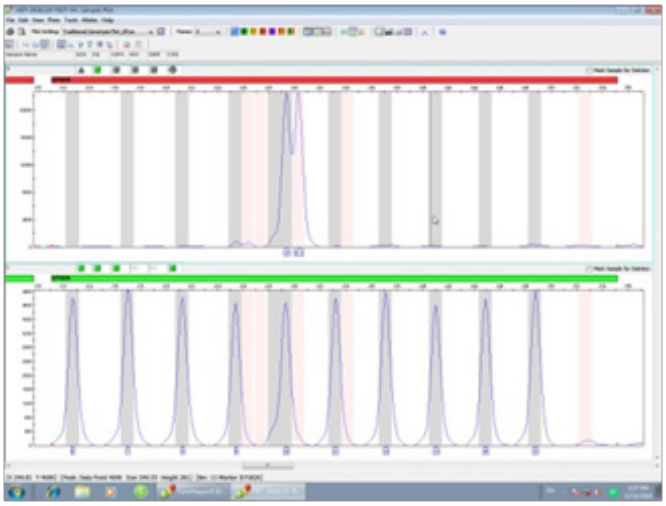


Figure 3(H) HuaDa®YanHuang kit (heterozygote 10/10.1).

Figure 3

10.1: GGGGCTAACGCAGTGCAGCTTGCATG CCTGCAGGTCGACGATTGACCCCCTATGGAATTTTTT TGTTTGTTTGTTTTTATTTATTTCTTTATCTTGA GATGGAGTCTCACTCTGTCACCCAGGCTGGAGTG CAGTGGTGCGATCTCGGCTCACTGCAACCTCCGCTT CTTGGGTCAAGTGGTTCTCCTGCCCCAGCCT CCTGAGTAGCTGGGACTACAGGCATGTGCTACTGC ATCCAGCTAATTTTTGTATTTTTTTTTAGAGAC GGGGTTTCACCATGTTGGTCAGGCTGACTATGGAGTT ATTTTAAGGTTAATATATATAAAGGGTAT GATAGAACACTTGTCATAGTTTAGAACGAACTAACA $(GATA)_{10}$GACAGATTGATAG$(T)_9$ATCTCACT AAATAGTCTATAGTAAACATTTAATTACCAATATGT GGTGCAATTCTGTCAATGAGGATAAATGTG GAATCGTTATAATTCTTAAAAATATATATTCCCTCTGA GTTTTTGATACCTCAGATTTTAAGACCTC ACAATTATCTCACAAGGCTTAAAATCAATCATATTTTGA GGATCACCTTATGGTATTTTTGCCTGTT TTTATTCCTTCTGGTGTGAAAACTGATGCCTTCCATCGT GTAACTCTTGTTCACACTGGTTTCAGTA TTTTGTTTTGAATCTCTAGAGGATCCCCGGGTACCGAG CTCGAATTCGTAATCATGGTCATAGCTGT TTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACA ACATACGAGCCGGAAGCATAAAGTGTA AAGCCTGGGGTGCCTAATGAGTGAGCTAACTC ACATTAATTGCGTTGCGCTCACTGCCCGCTTTCC AGTCGGGAAACCTGTCGTGCCAGCTGCATTAATG AATCGGCCAACGCGCGGGGAGAGGCGGTTTG CGTATTGGGCGCTCTTCCGCTTCCTCGCTCAC

As we known, the rare allele 10.1 have two microvariations: TAACA $(GATA)_{10}$ GACAGATTGATAG $(T)_9$ and TAAC-$(GATA)_{10}$GACAGATTGATAG$(T)_{10}$.[2] In our case, the sequence of this rare allele is TAACA $(GATA)_{10}$ GACAGATTGATAG$(T)_9$ (Figure 4), which is in concordance with the data reported by Tsuji et al.[10] Hence, In previous tests of ten STR multiplex kits, only three kits achieved the correct typing (Here we only present sequencing result for the rare allele 10.1, no allele 10, since allele 10 was for sure in this locus), they are: HealthGene PanGlobal, HuaDa® YanHuagn, and AppliedBio® HuaXiaBaijin. (Although ACGU EX25 and Goldeneye™ 20A have identical split peaks, one of the peak is still marked 'OL' and cannot be labeled by GeneMapper software, hence, they cannot be counted for correct results).

Discussion

Generally speaking, STR multiplex system ladders were divided into three colored longitudinal stripe: Grey, Pink, and White. Grey means float bin, Pink for virtual bin, and White stripe is off-ladder area or out of marker range Figure 5.

Sample peaks within Grey or Pink stripes can be automatically marked by STR analysis software. In our tests, two 'OL' peaks fall outside of the bin range, means they are out of marker/bin range. For PowerPlex® Fushion6C and PowerPlex® 21), both ladders have a little deviation to the left, hence, their sample peaks are all near stripe border between virtual bin and off-ladder area which tend to be another source of error and could easily misjudged by GeneMapper software. Sample peaks can be marked within a bin range (ether float bin or virtual bin), but when outside these bins, they cannot be marked by GeneMapper or other STR analysis software automatically. However, an off-ladder peak can be marked manually by comparing the position of the peak on the panel with nearest bin on the ladder. Virtual bin centers were created using the offset value from a neighboring allele and the reference (sequence length) size of the virtual allele. For example, some virtual alleles either size within 4bps of the smallest or largest allelic ladder allele, or contain 2bps partial repeat units. Virtual bins for alleles containing 1bp or 3bps partial repeat units were not included. Integer designations for these variant alleles must be assigned manually. In addition to the substantial expansion of nearly 300 configured markers, support for novel microvariants has been included for all loci with expanded 'virtual bin sets' comprising each potential base call within the allelic range rather than only observed nominal allele bins.[11]

For some kits like ACGU EX25 and Goldeneye™ 20A, their virtual bins for rare alleles like 10.1, are reduced or removed fromD7F820 locus, as for some rare allele no longer in bin/marker range and cannot be recognized by GeneMapper® or other STR analysis software automatically.

Manually analysis of an off-ladder allele at D7S820 locus with AGCU EX25 (Figure 7) is shown in GeneMapper Software plots (bp size versus relative fluorescence units). To do so, first we should know the repeat number of the certain locus. Let's take the system panel of AGCU EX25 For example, the repeat base value of D7S820 labeled as 'marker' is '4', which can be found in the upper middle 'panel manager' tab, marked by blue bar (Figure 6). This value also means microvarient alleles can only include 10.1,10.2 and 10.3. We can also check the label 'Ladder Alleles' to the right on the same blue bar of he tab. Allele 10.1 is not in the list, which means we must label this allele manually.

Sample allele was identified by recognition within floating bins (vertical gray stripes) around D7S820 alleles of the AGCU EX25 allelic Ladder Figure 7. One of the alleles that fell outside allelic ladder bins was flagged as 'OL' allele is shown. In this example, peak position is compared to the nearest allelic ladder using GeneMapper software. To be specific, as shown in Figures 7A & B, we put the mouse at off-ladder peak 'OL' and find the value of the 'Size' label in the lower left

of the tab shows: 227.25bps. Then we put the mouse at the nearest ladder peak '10', the value is 226.33bps.The apparent size difference relative to the D7S820 allele 10, verified by DNA sequencing as a 1bp insertion, defined a D7S820 10.1 allele. Sometimes, easier way is to use the abscissa value in lower left corner Figure 7C of the tab when pointing at the bin's peak near the off-ladder peak instead of the 'Size' value of the ladder.

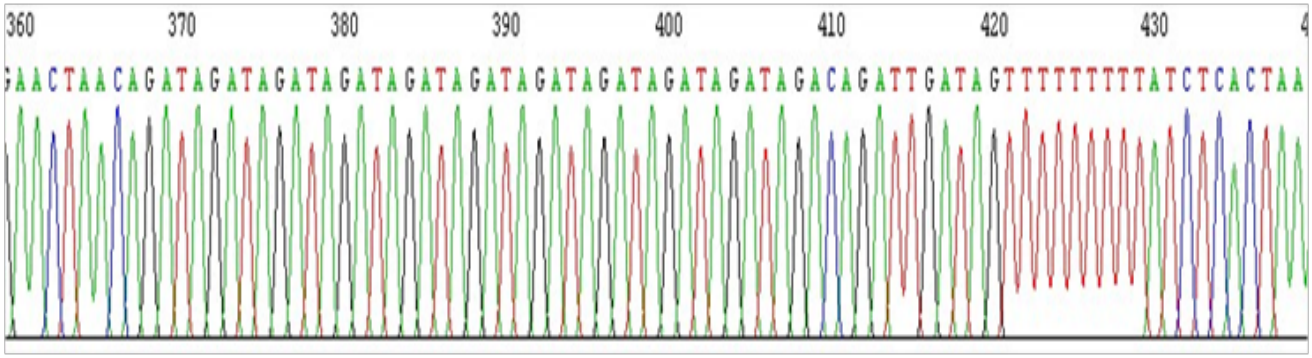


Figure 4 DNA sequencing for rare alleles 10.1.

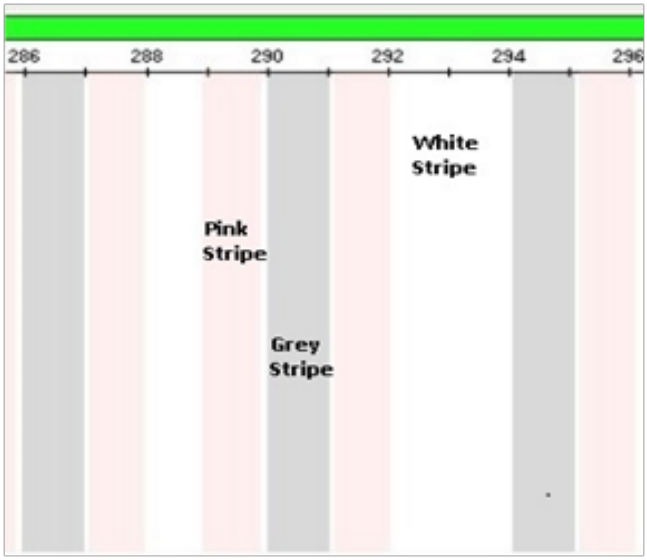


Figure 5 Grey (float bin), Pink (virtual bin), and White (no bin, out of marker range) longitudinal stripe on system marker.

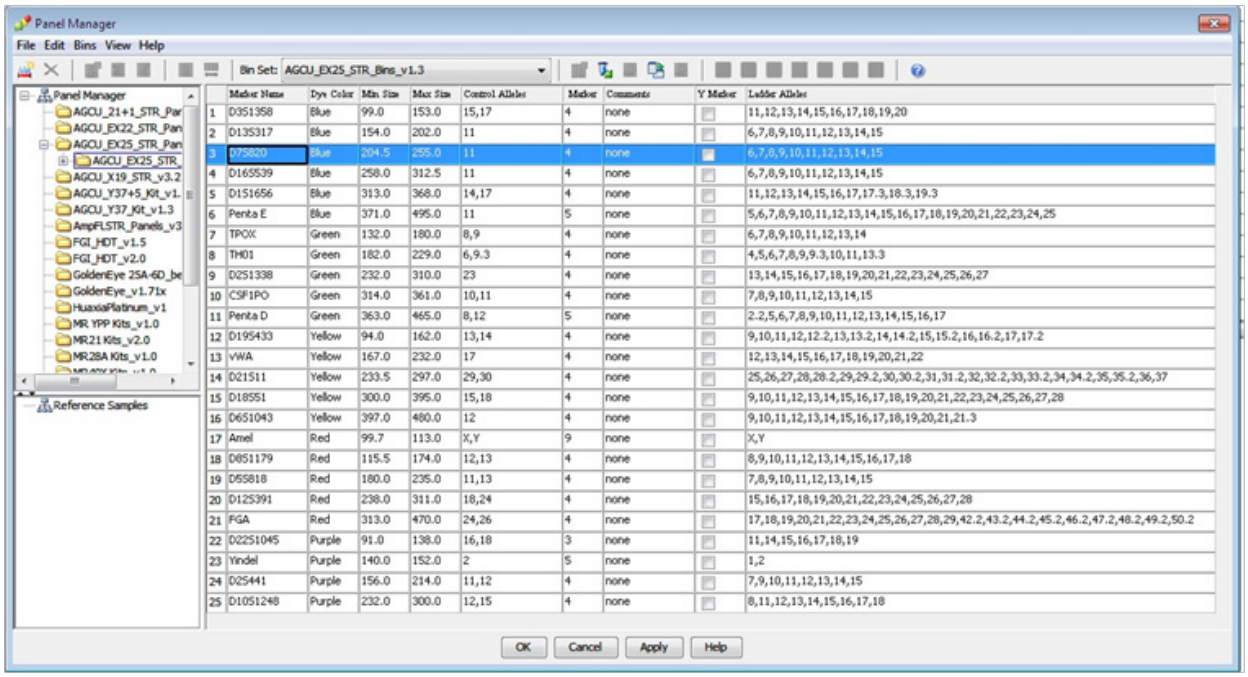
Sometimes, off-ladder allele may fall out of the locus range on system panel Figure 8. Under these circumstances, we should judge which locus the off-ladder allele should belong to at first, by comparing the distances to each locus respectively. Normally, the closer the peak to the certain locus, the more chance it should belong to. Then we can compare the position of the off-ladder peak with the nearest bin with the method described above. In this case, the difference value is 8bps, which means there are 2 repeats between the off-ladder peak and the bin, thus, the off-ladder allele should be labeled as '8'. But the most reliable way to get this kind of off-ladder allele typing is to perform ether Sanger or massive parallel sequencing test.

However, even within bin range, results could be inaccurate (Figure 2). This mainly due to the peak deviation. Sometimes, ladders could have little deviation to the left or right from bins, hence, sample peaks from same batch could also have displacement in the same direction. When these peaks near stripe border between bins (gray or pink stripe) and off-ladder area (white stripe), they could be misjudged by GeneMapper software, error occurred. Peak deviation caused by many reasons, apart from experimental errors, one of the most important reasons is that the primer and position design of the locus on the panel was not optimized. Optimization for rare allele can either by adding virtual bins to the flank regions of neighboring loci or put loci with rare alleles within ranging of 75-250 bps in the system panel when designing multiplex primers.[11] As we known, the best position for detecting rare alleles is within 75-250bps on the STR multiplex kit's system panel. This is because the further away from this range, the speed of electrophoresis is getting more unstable, the peaks became wider and lower, and the system perform lesser resolution, especially for rare alleles, therefore, peak deviation appears.

Table 1 shows that except for AppliedBio® HuaXiaBaijin, two of the three kits* which achieved the correct typing are within preferred range of 75-250bps. Since AppliedBio® HuaXiaBaijin is specially designed for Chinese population, although it is not in the optimized range, there are certain virtual bins in locus with rare alleles like D7S820 which also contribute to the detection of rear alleles. Other two kits**, ACGU EX25 and Goldeneye™ 20A, although have split-peak shape and within the preferred range, the 10.1 allele didn't marked by GeneMapper software in short of certain virtual bins. But we can still achieve the correct typing by manual analysis as mentioned above. Goldeneye™25A only have a single peak, not even a trend of split-peak as GlobalFiler does. The main reason is because the position for locus D7S820 on the panel is around 350bps, far beyond the preferred detection range of 75-250bps. However, the genotype 10, 10.1 is tricky since if panel and locus bins for D7S820 are not optimized, presence of two alleles can be mistaken for inefficient non-template addition and split peaks. It is similar situation as in genotype with 9.3/10 allele of TH01 locus. Fortunately, we also detected this genotype with SureID®PanGlobal (Figure 9).

During STR test, when encountered with certain rare alleles, peaks could be easily misjudged as OLs (off ladders), since GeneMapper software cannot find corresponding data from the system ladder. To reduce these kinds of errors, companies should make necessary optimization including both putting loci with rare alleles into 75-250bps on the panel and adding certain virtual bins for rare alleles at the same time when designing STR multiplex kits. For certain

population group, virtual bins for some specific loci should also be taken into account. Traditional capillary electrophoresis (CE) is still widely used for forensic DNA typing, mainly due to its time- and cost-effectiveness. Nowadays, native Bio-companies in China tend to add more and more numerous loci into 5- or 6-color fluorescence STR multiplex systems. On one hand, this may statistically increase the whole system accuracy. But on other hand, more rare alleles may appear and need to be optimized. The lack of optimization for rare alleles may inevitably cause more off-ladder peaks. Technical speaking, In order to avoid wrong judgement for rare alleles, kit designers should expand the assigned length on the panel and increase the number of virtual bins for every locus with rare allele in the system to cover all necessary known rare alleles (or rare alleles of certain population) to make those off-ladder rare allele peaks back into marker/bin range. For analyzers, the best way to avoid the error is manually analyze the typing of an off-ladder rare allele with the methods we mentioned above. An experienced lab staff could easily tell the real typing of an off-ladder rare allele by comprehensive analysis of different data like allele fragment length and peak position/shape via STR analysis software or by Sanger/Massive parallel sequencing test. In DNA database construction, rare allele types can greatly increase the power of discrimination. However, particular care should be taken in kinship matching and forensic cases, since incorrect designation of any deviations from allelic ladders could lead to a false conclusion.[2] Therefore, it is necessary to increase the number of useful references on non-standard allele patterns.[2] At the same time, parallel comparison is necessary to choose a suitable brand of STR multiplex kit as well as one or two standby kits for rare allele test during DNA database construction.



	Marker Name	Dye Color	Min Size	Max Size	Control Alleles	Marker	Comments	Y Marker	Ladder Alleles
1	D3S1358	Blue	99.0	153.0	15,17	4	none		11,12,13,14,15,16,17,18,19,20
2	D13S317	Blue	154.0	202.0	11	4	none		6,7,8,9,10,11,12,13,14,15
3	D7S820	Blue	204.5	255.0	11	4	none		6,7,8,9,10,11,12,13,14,15
4	D16S539	Blue	258.0	312.5	11	4	none		6,7,8,9,10,11,12,13,14,15
5	D15S1656	Blue	313.0	368.0	14,17	4	none		11,12,13,14,15,16,17,17.3,18.3,19.3
6	Penta E	Blue	371.0	495.0	11	5	none		5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25
7	TPOX	Green	132.0	180.0	8,9	4	none		6,7,8,9,10,11,12,13,14
8	TH01	Green	182.0	229.0	6,9.3	4	none		4,5,6,7,8,9,9.3,10,11,13.3
9	D2S1338	Green	232.0	310.0	23	4	none		13,14,15,16,17,18,19,20,21,22,23,24,25,26,27
10	CSF1PO	Green	314.0	361.0	10,11	4	none		7,8,9,10,11,12,13,14,15
11	Penta D	Green	363.0	465.0	8,12	5	none		2.2,5,6,7,8,9,10,11,12,13,14,15,16,17
12	D19S433	Yellow	94.0	162.0	13,14	4	none		9,10,11,12,12.2,13,13.2,14,14.2,15,15.2,16,16.2,17,17.2
13	vWA	Yellow	167.0	232.0	17	4	none		12,13,14,15,16,17,18,19,20,21,22
14	D21S11	Yellow	233.5	297.0	29,30	4	none		25,26,27,28,28.2,29,29.2,30,30.2,31,31.2,32,32.2,33,33.2,34,34.2,35,35.2,36,37
15	D18S51	Yellow	300.0	395.0	15,18	4	none		9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28
16	D6S1043	Yellow	397.0	480.0	12	4	none		9,10,11,12,13,14,15,16,17,18,19,20,21,21.3
17	Amel	Red	99.7	113.0	X,Y	9	none		X,Y
18	D8S1179	Red	115.5	174.0	12,13	4	none		8,9,10,11,12,13,14,15,16,17,18
19	D5S818	Red	180.0	235.0	11,13	4	none		7,8,9,10,11,12,13,14,15
20	D12S391	Red	238.0	311.0	18,24	4	none		15,16,17,18,19,20,21,22,23,24,25,26,27,28
21	FGA	Red	313.0	470.0	24,26	4	none		17,18,19,20,21,22,23,24,25,26,27,28,29,42.2,43.2,44.2,45.2,46.2,47.2,48.2,49.2,50.2
22	D22S1045	Purple	91.0	138.0	16,18	3	none		11,14,15,16,17,18,19
23	Yindel	Purple	140.0	152.0	2	5	none		1,2
24	D2S441	Purple	156.0	214.0	11,12	4	none		7,9,10,11,12,13,14,15
25	D10S1248	Purple	232.0	300.0	12,15	4	none		8,11,12,13,14,15,16,17,18

Figure 6 Panel manager tab in GeneMapper 1.4.

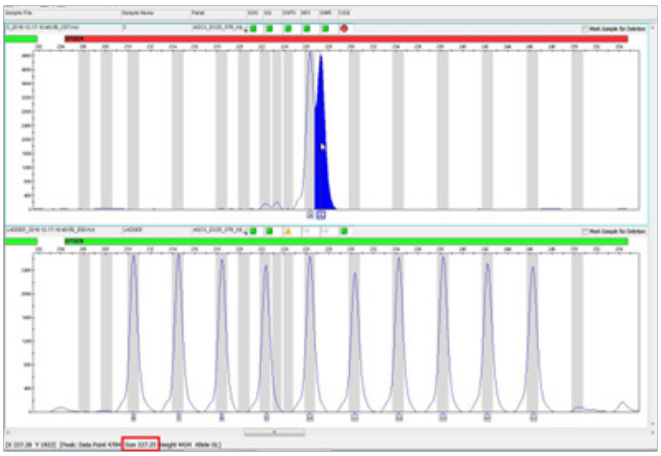


Figure 7(A)

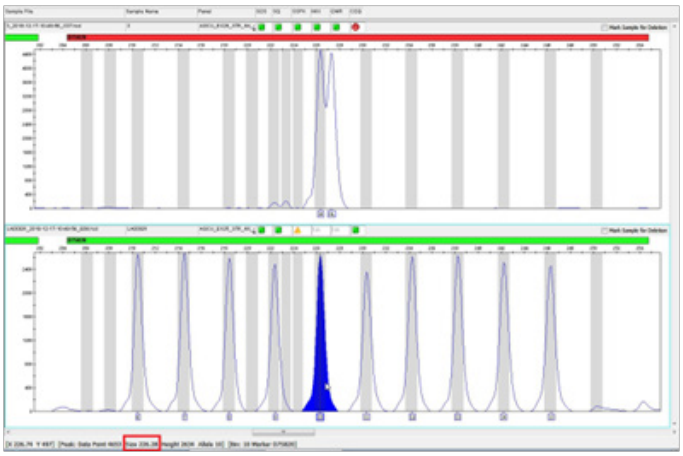


Figure 7(B)

Citation: Lei X. Comparison of different STR typing for a special allele at D7S820 locus by using ten different STR multiplex system. *Forensic Res Criminol Int J.* 2019;7(3):135–142. DOI: 10.15406/frcij.2019.07.00277

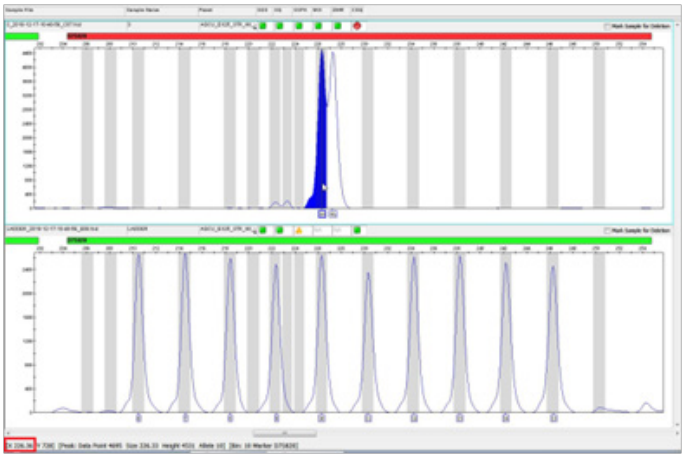
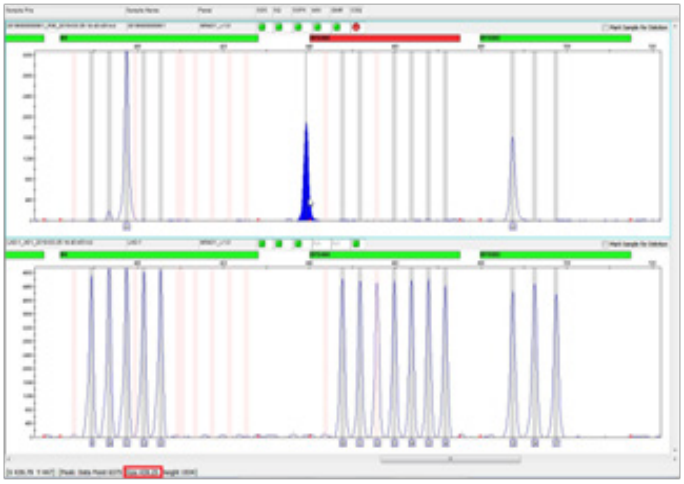


Figure 7(C)

Figure 7 Identify 'OL' peak by compare the position of the floating bins on the panel.



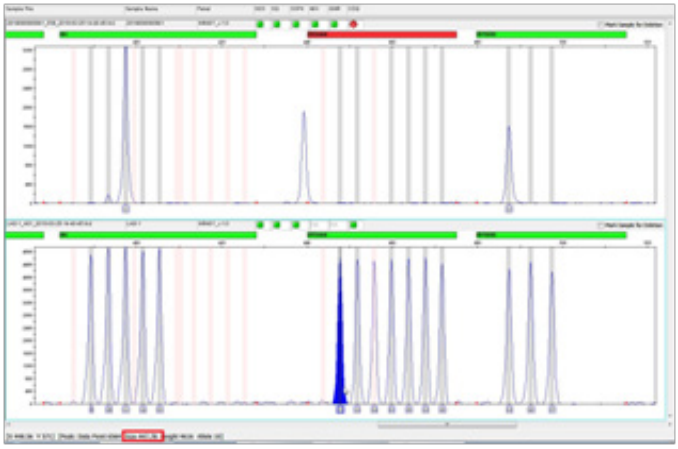


Figure 8 Off-ladder allele fall out of the locus range.

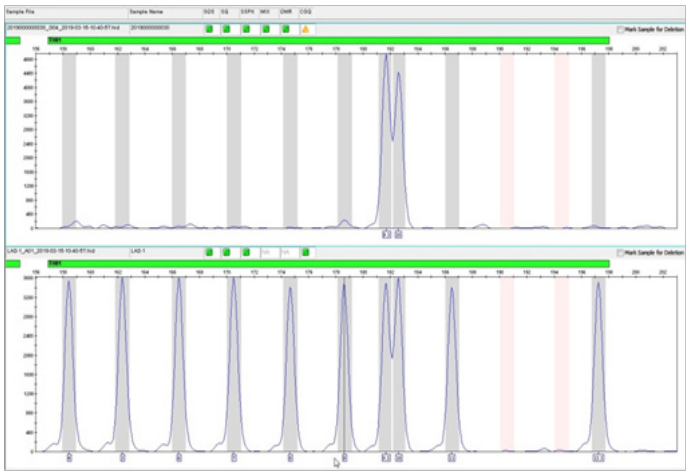


Figure 9 Genotype with 9.3/10 alleles of TH01 locus, tested by SureID®PanGlobal.

Table 1 Position designed for D7F820 locus on system panels of 10 different kits.

Multiplex systems	D7F820 locus position on panel (bps)
AppliedBio®GlobalFiler	279
Promega PowerPlex®Fushion 6c	290
PowerPlex®21	290
PowerPlex®24	290
ACGU EX25**	227
AppliedBio®HuaXiaBaijin*	279
SureID®PanGlobal*	240
GoldeneyeTM20A	226
GoldeneyeTM25A**	351
HuaDa ®YanHuang*	230

Acknowledgments

None.

Conflicts of interest

The author declares that there are no conflicts of interest.

References

1. Song XB, Zhou Y, Ying BW, et al. Short-tandem repeat analysis in seven Chinese regional populations. *Genet Mol Biol*. 2010;33(4):605–609.
2. Yoo SY, Cho NS, Park MJ, et al. A large population genetic study of 15 autosomal short tandem repeat loci for establishment of Korean DNA profile database. *Mol Cells*. 2011;32(1):15–19.
3. Deng YJ, Yan JW, Yu XG, et al. Genetic analysis of 15 STR loci in Chinese Han population from West China. *Genomics Proteomics Bioinformatics*. 2007;5(1):66–69.

4. Li S, Yan C, Deng Y, et al. Polymorphism profile of nine short tandem repeat Loci in the Han chinese. *Genomics Proteomics Bioinformatics*. 2016;1(2):166–170.

5. Liu Y, Guo L, Jin H, et al. Developmental validation of a 6-dye typing system with 27 loci and 6.application in Han population of China. *Sci Rep*. 2017;7(1):4706.

6. Grover R, Jiang H, Turingan RS, et al. FlexPlex27-highly multiplexed rapid DNA identification for law enforcement, kinship, and military applications. *Int J Legal Med*. 2017;131(6):1489–1501.

7. Hering S, Edelmann J, Dressler J. Sequence variations in the primer binding regions of the highly polymorphic STR system SE33. *Int J Leg Med*. 116 (2002) 356–367.

8. Hu N, Cong B, Gao T, et al. Application of mixsep software package: Performance verification of male-mixed DNA analysis. *Mol Med Rep*. 2015;12(2):2431–2442.

9. Zahra A, Hussain B, Jamil A, et al. Forensic STR profiling based smart barcode, a highly efficient and cost effective human identification system. *Saudi J Biol Sci*. 2018;25(8):1720–1723.

10. Tsuji A, Ishiko A, Ikeda N. The structure of a variant allele which is considered to be 30.3 in the STR locus D21S11. *Legal Med*. 2006;8(3):182–183.

11. King JL, Wendt FR, Sun J, et al. STRait Razor v2s: Advancing sequence-based STR allele reporting and beyond to other marker systems. *Forensic Sci Int Genet*. 2017;29:21–28.

Citation: Lei X. Comparison of different STR typing for a special allele at D7S820 locus by using ten different STR multiplex system. *Forensic Res Criminol Int J*. 2019;7(3):135–142. DOI: 10.15406/frcij.2019.07.00277