GENOTYPING IRREGULARITIES OBSERVED IN STR MARKERS USED IN FORENSIC DNA ANALYSIS IN THE LEBANESE POPULATION

Ansar El Andari[1,2], Issam Mansour[1]
[1]*Molecular Biology Laboratory,* Faculty of Health Sciences, American University of Science and Technology, Lebanon.
[2]*School of Criminal Justice,* University of Lausanne, Switzerland.

ABSTRACT

El Andari, Ansar and Mansour, Issam. 2018. Genotyping irregularities observed in STR markers used in forensic DNA analysis in the Lebanese population. Lebanese Science Journal, 19(2): 200-209.

Forensic DNA analysis is based on the evaluation of DNA profiles obtained from reference samples or crime scene traces. Profiles are determined using different commercial STR multiplex kits, which include allelic ladders for the correct designation of alleles of a given STR system. However, some new STR variants could be observed and their identification is essential for proper allele size calling in casework interpretation and cases of kinship matching or relationship testing. A total of 6392 samples were analyzed as part of the DNA work at our DNA testing facility from 2004 until 2016, where one hundred and twelve new different STR variant alleles of 20 STR loci were observed in the Lebanese population. Forty-five of these variant alleles were internationally reported in the STR-base whereas the remaining sixty-seven alleles were observed for the first time. The loci included D3S1358, D19S433, D21S11, D13S317, TPOX, FGA, D7S820, D18S51, CSF1PO, D2S1338, D10S1248, D12S391, D22S1045, D2S441, Penta-D, Penta-E, SE33, D1S1656, vWA, and THO1. Also, four triallelic patterns were observed at D21S11, D12S391, D13S317 and Penta-D loci. Primer binding site mismatch that resulted in a discordance between the amplification of different kits was observed once at the D19S433 locus.

Keywords: variant alleles, microvariants, off-ladder, DNA analysis, human identification, forensic STR testing.

http://dx.doi.org/10.22453/LSJ-019.2.200209

INTRODUCTION

Currently, DNA forensic analysis and human identification rely on the construction of a DNA profile determined by the amplification of a given number of STR markers, usually 12 to 17 markers (Butler, 2005; Roewer, 2013). STR regions are amplified using commercially available kits. These kits provide allelic ladders and multiplex primers of a given set of STR markers. Ladders include all of the already known alleles for the given amplified STR systems and help in the designation of the alleles by adjusting different sizing measurements obtained from different instruments. While the majority of the STR alleles fall usually within the allelic ladder sizing bins, sometimes variant or "off-ladder" alleles could be detected. Variant alleles are called either new alleles or microvariants and are driven by addition(s) or deletion(s) of a complete repeat or part of a complete repeat respectively (Rene, 2007). Sequence variation or point mutations may also affect the primer binding site and may lead to either a complete amplification failure or a partial amplification, which result in null alleles or peak height imbalance respectively (Kline *et al*., 2011). Triallelic patterns or sequence mutations could also be rarely observed. Triallelic patterns are generally due to chromosomal trisomy or duplications at a localized locus during the embryonic development. (Brinkman *et al*., 1998; Buckleton, 2005).

The detection and proper allele designation of these variations is crucial for sound interpretation and conclusions of the cases under investigation. As a matter of fact, scientists dealing with DNA STR profiling need to have the population database of variant alleles to properly assign them to the profile under study and avoid a potential misinterpretation. In addition, variant alleles could increase the power of discrimination in DNA profiling and DNA comparisons (Allor et al., 2005). For these reasons, international databases, in which the forensic science community shares new findings and information concerning STR markers, have been developed (Butler *et al*., 2006; Sho, 1997; Bodner *et al*., 2016). These databases are of great assistance to practitioners in updating data and saving time in proper casework analysis. Variant alleles and genotyping findings of STR markers in different populations have been reported earlier (Mizuno *et al.*, 2003; Heinrich *et al*., 2005; Grubwieser *et al*., 2005; Hyunwoong *et al*., 2012; Onofri *et al*., 2008). However, none have been reported from the Lebanese population.

The present study aims at screening and reporting results for off-ladder alleles and tri-allelic patterns observed during ten years of DNA typing of samples from the Lebanese population.

MATERIALS AND METHODS

A total of 6392 samples were tested as part of the DNA work at our DNA testing facility from 2004 until 2016, including the research study where the Lebanese STR allele frequencies were estimated (Al-Andari *et al*., 2013). Only reference samples known to be of Lebanese origins were selected for the study. Samples consisted of whole blood or buccal swabs. Genomic DNA was extracted from whole blood using the salting out method, whereas organic DNA extraction was used with buccal swabs. Approximately 1 ng of the target DNA was amplified using commercial typing kits. 3180 out of the 6392 samples were amplified using Identifiler kit from Applied Biosystems, whereas the rest 3212 samples were tested using PowerPlex® 16 HS and PowerPlex® ESI 17, totaling 23 autosomal STR markers (Promega Corporation, 1978), and all samples were typed following manufacturer's instructions. Amplification was performed using the AB 9700 thermal cycler (Applied Biosystems, 1981). Genotypic data was obtained using ABI 3130 Genetic Analyzer using Performance Optimized Polymer (POP) 7 and 4 x 36 cm capillary arrays (Applied Biosystems, 1981). GeneScan™ 500 LIZ® Size Standard was used for sizing DNA fragments amplified using Identifiler kit. Internal Lane Standards (ILS) 600 and 500 were used for sizing the amplified fragments using PowerPlex® 16 HS and PowerPlex® ESI 17, respectively. Data collection was achieved by using Data Collection Software v 3.0 from Applied Biosystems, whereas fragment analysis was done using Gene Mapper Software v 4.1 from the same manufacturer. Size calling was obtained by using allelic ladders provided in the kits. The settings used for the electrophoresis were those set by the manufacture's protocol. Samples showing genotypic irregularities were run in duplicates to confirm the results.

RESULTS

A total of 112 off-ladder alleles in 20 STR loci were observed, of which, 46 were microvariant alleles and 66 were new alleles with a males to females ratio of approximately 2:1. The 20 loci included D3S1358, D19S433, D21S11, D13S317, TPOX, FGA, D7S820, D18S51, CSF1PO, D2S1338, D10S1248, D12S391, D22S1045, D2S441, Penta-D, Penta-E, SE33, D1S1656, vWA, and THO1. Out of these 112 detected off ladder alleles, 46 were previously reported in the STR-base (Short Tandem Repeat DNA Internet Data Base, 1997), whereas the remaining 67 variant alleles were observed for the first time in this international database (El-Andari *et al*., 2013; Chouery *et al*., 2010). Table 1 shows these novel STR off-ladder STR alleles and their estimated allele frequencies calculated using Arlequin software (Excoffier and Lischer, 2010). Triallelic patterns were observed at D21S11, D12S391, D13S317 and Penta D loci (Table 2). A genotypic discordance due to a null allele was observed once between Identifiler and PowerPlex® ESI 17 at the D19S433 locus (Table 3). All these genotypic findings were

confirmed by re-extraction and amplification. None of the D16S539, D8S1179, and D5S818 showed any allele variation. Examples of some of the findings observed in this study are shown in Figure 1.

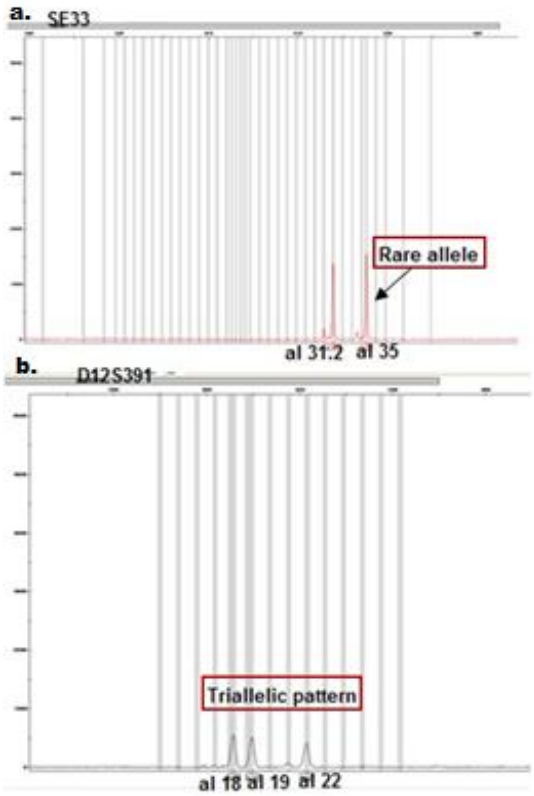

Figure 1. Example of an observed (a.) rare allele, and (b.) a triallelic pattern

Table 1. Rare microvariant and new alleles in the Lebanese population.

Systems	Alleles	Lebanese reported allele frequency	NIST International database
D19S433	6.2	0.003	Reported
	10.2	0.005	Reported
	11.2	0.003	Reported
	12.2	0.005	Not reported
	17	0.005	Not reported
	17.2	0.005	Not reported
	18	0.003	Reported

	18.2	0.005	Reported
D21S11	26	0.005	Not reported
	29.2	0.005	Not reported
	32	0.005	Not reported
	33.1	0.003	Reported
	33.3	0.003	Reported
	34.2	0.005	Reported
	35	0.005	Not reported
D13S317	6	0.003	Reported
	15	0.005	Not reported
D8S1179	18	0.005	Reported
TPOX	5	0.003	Reported
	6	0.005	Not reported
	13	0.005	Not reported
	14	0.003	Reported
FGA	16	0.003	Reported
	17	0.005	Reported
	20.2	0.003	Reported
	21	0.005	Not reported
	21.2	0.003	Reported
	22.3	0.003	Reported
	23.2	0.005	Reported
	25.2	0.005	Reported
	29	0.005	Reported
D7S820	9.1	0.003	Reported
	9.3	0.003	Reported
	12.1	0.003	Reported
	14	0.005	Not reported
D18S51	7	0.005	Reported
	8	0.003	Reported
	12.2	0.003	Reported
	13.2	0.005	Reported
	16.2	0.003	Reported

Table 1 cont'd: Rare microvariant and new alleles in the Lebanese population.

Systems	Alleles	Lebanese reported allele frequency	NIST International database
D18S51 cont'd	21	0.005	Not reported
	22	0.005	Not reported
	24	0.005	Not reported
CSF1PO	7	0.005	Not reported
	8	0.005	Not reported
	9.1	0.003	Reported
	14	0.005	Not reported
D2S1338	14	0.003	Reported
	26	0.005	Not reported
D3S1358	11	0.003	Reported
	13	0.005	Not reported
	16.3	0.003	Not reported
D12S391	17.3	0.005	Reported
	26	0.003	Not reported
D10S1248	6.2	0.006	Not reported
	9	0.003	Not reported
	10	0.005	Not reported
D22S1045	10	0.007	Not reported
D2S441	6	0.001	Not reported
	9	0.003	Not reported
	13.3	0.005	Not reported
Penta-D	5.2	0.005	Not reported
	7	0.003	Not reported
	12.2	0.006	Reported
	14.2	0.006	Not reported
	16	0.001	Not reported
	17	0.005	Not reported
	18	0.006	Reported
Penta-E	17.4	0.006	Reported
	21	0.003	Not reported
	22	0.001	Not reported
SE33	8	0.003	Not reported
	8.3	0.003	Not reported
	9	0.003	Not reported

	10	0.005	Not reported
	11	0.003	Not reported
	11.2	0.005	Reported
	12	0.005	Not reported
	13.3	0.003	Not reported

Table 1 cont'd: **Rare microvariant and new alleles in the Lebanese population.**

Systems	Alleles	Lebanese reported allele frequency	NIST International database
SE33 cont'd	14.3	0.003	Not reported
	15.3	0.006	Reported
	16.3	0.006	Reported
	20.2	0.003	Not reported
	22	0.005	Reported
	24	0.003	Not reported
	25	0.003	Not reported
	30	0.003	Not reported
	34.2	0.005	Not reported
	35	0.003	Not reported
	35.2	0.006	Not reported
	36	0.005	Reported
	40	0.006	Not reported
	46.2	0.003	Not reported
D1S1656	10	0.005	Not reported
	18	0.005	Reported
vWA	13	0.005	Not reported
	15.2	0.006	Not reported
	21	0.005	Not reported
THO1	8.3	0.006	Reported
	11	0.005	Reported

Table 2. Observed triallelic patterns.

Systems	Nb. observed	Alleles	NIST International database
D21S11	1	30-31-31.2	Not Reported
D12S391	1	18-19-22	Reported[a]
D13S317	1	8-11-12	Reported
Penta D	1	9-11-12	Reported

[a]reported by our laboratory

Table 3. Discordance observed at D19S433 System due to a null allele.

Sample Number	STR Multiplex Kit	Genotype
FE12137	**Identifiler**	16-16
FE12137	**ESI17**	12-16

DISCUSSION

The present study reports a compendium of unusual variant STR alleles detected in the Lebanese population using a combination of 3 multiplex kits, Identifiler, PowerPlex® 16 HS and PowerPlex® ESI 17. This dataset will assist practitioners in the field in acknowledging the presence of new variant and microvariant alleles present in the Lebanese population, and therefore, help in decreasing wrong designation of alleles that could affect human identification, forensic and relationship studies. All new alleles reported in this paper will be uploaded in the STRbase in order to help disseminating new information that could be of great assistance worldwide.

ACKNOWLEDGMENTS

We are grateful to Promega Corporation for granting this project with the necessary multiplex kits. Special thanks to Ms. Myriam Helou and Marc Obeid for their valuable help in the logistics.

REFERENCES

Buttler, J. M. 2005. Forensic DNA Typing, second edition ED., USA: Elsevier Academic Press.

Roewer, L. 2013. DNA fingerprinting in forensics: past, present, future. *Investigative Genetics,* 4: 22.

Huel, R. L. Bašić, L. Madacki-Todorović, K. Smajlović, L. Eminović, I. Berbić, I. Miloš, A. and Parson, T. 2007. Variant alleles, triallelic patterns, and point mutations observed in nuclear short tandem repeat typing of populations in Bosnia and Serbia. C*roatian Medical Journal,* 48(4): 494-502.

Kline, M. C. Hill, C. R., Decker, A. E. and Butler, J. M. 2011. STR sequence analysis for characterizing normal, variant, and null alleles. *Forensic Science International: Genetics,* 5: 329–332.

Brinkmann, B. Klintschar, M. Neuhuber, F. Huhne, J. and Rolf, B. 1998. Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. *American Journal of Human Genetics,* 62: 1408-1415.

Buckleton, J. 2005. Forensic DNA evidence interpretation, BocaRaton, fl: CRC Press, PP. 399-400.

Allor, C., Einum, D. and Scarpetta M. 2005. Identification and characterization of variant alleles at CODIS STR loci. *Journal of Forensic Sciences,* 50: 1128-1133.

Buttler, J. M., Coble, M. D. Decker, A. E. Duewer, D. L. Hill, C. R. Kline, M. C. Redman, J. W. and Vallone, P. M. 2006. Setting standards and developing technology to aid the human identity testing community. *International Congress Series,* 1288: 628-635.

Short tandem repeat DNA internet database, 1997. [online]. Available: http://www.cstl.nist.gov/strbase/.

Bodner, M. Bastisch, I. Butler, J. Fimmers, R. Gill, P. Gusmao, L. Morling, N. Phillips, C. Prinz, M. Schneider, P. and Parson, W. 2016. Recommendations of the DNA commission of the International Society for Forensic Genetics (ISFG) on quality control of autosomal short tandem repeat allele frequency databasing. *Forensic Science International: Genetics,* 24: 97-102.

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

Heinrich, M. Felske-Zech, H. Brikimann, B. and Hohoff, C. 2005. Characterization of variant alleles in the STR systems D2S1338, D3S1358 AND D19S433. *International Journal of Legal Medicine,* 119: 310-13.

Grubwieser P. Mühlmann, R. Niederstätter, H. Pavlic, M. and Parson, W. 2005. Unusual variant alleles in a commonly used short tandem repeat loci. *International Journal of Legal Medicine,* 119: 164-66.

Hyunwoong, P. Eun Kyung, R. Sung Sup, P. and Moon-Woo, S. 2012. Detection of very large off-ladder alleles at the Penta E locus in a 15 locus autosomal STR database of 199 korean individuals. *Forensic Science Internationl Genetics*, 6(6): 189-191.

Onofri, V. Pesaresi, M. Alessandrini, F. Turchi, C. and Tagliabracci, A. 2008. D16S539 Microvariant or D2S1338 Off-ladder allele? a case report about a range overlapping between two loci. *Forensic Science International: Genetics Supplement Series,* 1(1): 123-124.

EL Andari, A. Othman, H. Taroni, F. and Mansour, I. 2013. Population genetic data for 23 STR markers from Lebanon. *Forensic Science International: Genetics*, 7: 108-113.

Promega Corporation. 1978. Madison, USA.

Applied Biosystems. 1981. California, USA.

Chouery, E. Coble, M. D. Strouss, K. M., Saunier J. L., Jalkh N., Medlej Hashim M., Ayoub F., and Megarbane A. 2010. Population genetic data for 17 STR markers from Lebanon. *Legal Medicine*, 12(6): 324-326.

Excoffier, L. and Lischer, H. 2010. Arlequin suite ver 3.5: A new series of programs to perform. *Molecular Ecology Resources,* 564-567.