International Research Journal of Computer Science (IRJCS)
Issue 6, Volume 2 (June 2015)
ISSN: 2393-9842
www.irjcs.com

Quality Control and Duplication for Concordance in Forensic DNA Samples: Implications for Interpretation in Mixtures

Heather Miller Coyle*
Forensic Science Department, Henry C. Lee College of Criminal Justice & Forensic Sciences, University of New Haven, 300 Boston Post Road, West Haven, CT 06516 USA

Abstract— In DNA human identification, concordance between duplicate samples is required for quality assurance of results. When variation is noted, the underlying science points to factors such as sub-optimal DNA template concentration, potential contamination events due to primary or secondary transfer, and DNA degradation or breakage. All of these factors relate back to PCR amplification efficiency which also influences the error rates for the reliable detection of alleles. For low level DNA mixtures or interpretation of major donor alleles in a mixture, PCR amplification kinetics and detection of data close to the baseline of the instrumentation is critical for any attempt to establish the true number of contributors to a sample. With software, quantitative analysis of run to run variation can be established by evaluating peak height differences and is useful as an aid in DNA mixture interpretation.

Keywords— DNA, forensic science, stochastic fluctuation, scientific accuracy, PCR amplification efficiency

I. INTRODUCTION

Concordant and duplicate analyses in forensic laboratories are used to detect sample mix-up and confirm the presence of DNA mixtures. Duplicate processing is for determining internal concordance of samples in high-throughput laboratories that are concerned about sample tracking or error due to the physical switching of evidentiary DNA tubes during the genotyping procedure. The general laboratory policy is to confirm DNA results either by having concordant DNA results within a case, or by duplicating the DNA results with separate aliquot, amplification, and electrophoresis steps and comparing the data from both runs. The two most common reasons for performing concordance testing is to confirm known reference sample processing for quality control and to repeat DNA testing when a low amount of DNA is amplified to verify a result for scientific accuracy. This article will focus on the latter issue (scientific accuracy) as it relates to reproducibility in duplicate amplifications and detection of the major donor to an evidentiary sample.

II. PCR AMPLIFICATION EFFICIENCY AND END POINT MEASUREMENTS

The enzymatic copying process of different regions of genomic DNA is performed in the laboratory by a process called polymerase chain reaction (PCR) amplification. Small quantities of total genomic DNA are placed in a sterile tube and temperature cycled with short complementary PCR primers, Taq polymerase enzyme, co-factors and free nucleotides to generate a sufficient quantity of DNA from minimal template for forensic detection. This process requires a three-step reaction: denaturation of double-stranded DNA, annealing of short PCR primer sequences to complementary DNA template and synthesis of new strands of DNA incorporating label for downstream detection by capillary electrophoresis. When one generates a DNA profile from high quality and sufficient quantity of DNA template, one would expect that the same DNA profile is generated with equivalent peak heights every time the PCR reaction is performed. This is referred to as data reproducibility. In practice there is some variability in duplicate samples and contributing factors are further discussed to explain the variance in peak heights as measured by software analysis systems.

One main factor in PCR reproducibility is amplification efficiency which is correlated to the initial template concentration, presence of inhibitory compounds and enzyme reaction kinetics [1], [2], [3]. One would assume that every DNA template amplifies in the same manner in an enzymatic reaction but this is not true [4]. Preferential amplification at some loci or for one DNA template versus another can result in what is commonly called a “flip-flop” of anticipated results, even when one template has a higher concentration than the other. In general, one anticipates the higher quantity template in a DNA mixture to have the greater allelic peaks using capillary electrophoresis detection methods; however, when the input templates are close to equivalent in input quantity, the difference in preferential amplification efficiency can cause the lower input quantity template to have higher detectable peaks. Estimates of the effect on calculation of DNA quantity by RT-PCR detection have been provided in the following example; a 5% reduction in amplification efficiency can reduce the DNA concentration estimate by 46-fold in a 30 cycle PCR method [1].

III. MEASURING STOCHASTIC EVENTS

A stochastic event in PCR refers to a fluctuation in peak amplification between instrument runs (Table 1). This is evident in peak height differences between runs as the height is not exact and is expressed on the y-axis as relative fluorescence units (RFU). The variance between runs can be calculated by comparing the corresponding peak heights per locus for each run. In Table 1, when comparing run 1 and run 2, three additional alleles (10, 12, 12.2) were detected at chromosome locus

D19S433. Though many of the peaks have higher RFU values in run 2, the additional alleles were detected in run 1. Barring such artefacts such as spectral pull-up, the best explanation is differential PCR amplification efficiency or alternatively, a contamination event. Stochastic events are even more pronounced than in Table 1 when DNA template concentrations are low and create a situation where the peak height may vary by greater than 30% between PCR amplifications [5], [6], [7]. At its most extreme, stochastic fluctuation can lead to absence of peaks in a DNA profile that re-appear in a separate amplification. Since this phenomenon affects detection of data, the estimate for the number of contributors (sources) to a sample can be in error. In one low copy number DNA study, 86% of high template samples (100pg or more, 14% error rate) and approximately 76% of low template samples (50-100pg, 24% error rate) were detected correctly as four person mixtures [7]. However, for samples with a DNA template of 50pg or less, the four person mixtures were detected as two and three person mixtures a large proportion of the time due to missing data points [7]. The variability was greatest for samples below the 25pg quantity [6]. For the data that is detected, DNA software packages such as GeneMarker (Soft Genetics, LLC; State College, PA) have a flagging system by which stochastic results are ranked and identified for the end-user and data can be tabulated to include peak heights for each allele automatically. With other software systems such as GeneMapper (Life Technologies; Grand Island, NY), peak heights are also tabulated automatically as output data but the mathematical evaluation of peak height differences in percent are often calculated by hand by the end-user.

IV. EFFECT OF DNA TEMPLATE CONCENTRATION

With standard 28 cycle PCR, and the optimal manufacturer's kit input DNA template of 1 ng, duplicate amplifications are very similar or even identical for peak detection. However, as in low copy number DNA testing (<100pg, 31 cycle PCR) or standard forensic DNA tests with low level mixtures, duplication of data is not always possible and leads to scientific inaccuracies in DNA detection and measurements regardless of the software analysis package. To mitigate the effect of stochastic fluctuation in low template data, additives such as betaine have been shown to improve amplification efficiency and yield by reducing secondary structure in the G:C rich regions of DNA during the template copying process [8]. One concerning effect with low level DNA template amplification is DNA contamination enhancement resulting in extraneous DNA being detected that is unaccounted for when compared to the known reference samples even in controlled studies [6], [9], [10]. Contamination events, even with a single cell, through contact transfer prior to evidentiary processing or during processing steps significantly confuse the interpretation of DNA mixtures and evaluation of true number of contributors to a sample for statistical match calculations.

Ultraviolet irradiation can help with reducing levels of DNA on surfaces before DNA processing by causing DNA degradation of fragments by random nicking events resulting in insufficient DNA template for PCR amplification [9]. For samples that have detectable alleles in the negative controls, most software packages can subtract out the alleles from the statistical calculations or they can be manually excluded from consideration. Under these circumstances, most forensic science laboratories, rerun samples with new sets of controls to maintain quality standards in test methods.

V. CONCLUSIONS

Data reproducibility is expected in most forensic test methods and as a quality assurance measure to have confidence in the results. Variable data, however, is generated under certain forensic conditions and with certain types of samples and are to be treated with caution when interpreting if a particular individual could be the source of the DNA in the sample. Allelic drop-out leading to partial DNA profiles and contamination events contributes to highly subjective interpretation of the true number of contributors to a sample and error rates. Computer analysis programs aid in data interpretation by allowing the analyst to evaluate data close to the baseline of the instrumentation (below the analytical threshold) used for DNA detection so as to recover all of the necessary information to establish the potential source as well as for providing quantitative measurement of variance between duplicate runs to define the expected variation due to differences in PCR amplification efficiency.

VI. ACKNOWLEDGEMENTS

Thank you to Attorney Troy A. Smith, 34 S. Broadway, Suite 218, White Plains, NY 10601 USA for providing a DNA case with transfer issues for review. Thank you to The Bronx Defenders, 360 East 161st Street, Bronx, NY 10451 USA for inquiry and for providing DNA concordance data for review.

REFERENCES

[1] N. Chatterjee, T Banerjee, S Datta, "Accurate Estimation of Nucleic Acids by Amplification Efficiency Dependent PCR", *PLoS ONE*, 7(8): e42063. doi:10.1371/journal.pone.0042063, 2012.
[2] D Sisti, M Guescini, M Rocchi, P Tibollo, M D'Atri, V Stocchi, "Shape Based Kinetic Outlier Detection in Real-Time PCR", *BMC Bioinformatics*, 11: pp. 186-198, 2010.
[3] R Rutledge, D Stewart, "Critical Evaluation of Methods Used to Determine Amplification Efficiency Refutes the Exponential Character of Real-Time PCR", *BMC Molecular Biology*, 9: pp. 96-108, 2008.

[4] J Ruitjer, C Ramakers, W Hoogaars, Y Karlen, O Bakker, M van der Hoff, A Moorman, "Amplification Efficiency: Linking Baseline and Bias in the Analysis of Quantitative PCR Data", *Nucleic Acids Research*, 37(6): pp. 1-12, 2009.
[5] A Lievens, S Van Aelst, M Van den Bulcke, E Goetghebeur, "Simulation of Between Repeat Variability in Real Time PCR Reactions", *PLoS ONE*, 7(11): e47112. doi:10.1371/journal.pone.0047112, 2012.
[6] T Caragine, R Mikulasovich, J Tamariz, E Bajda, J Sebestyen, H Baum, M Prinz, "Validation of Testing and Interpretation Protocols for Low Template DNA Samples Using AmpF*l*STR Identifiler", *Croat Med J*, 50(3): 250-267, 2009.
[7] J Perez, AA Mitchell, N Ducasse, J Tamariz, T Caragine, "Estimating the Number of Contributors to Two-, Three-, and Four-Person Mixtures Containing DNA in High Template and Low Template Amounts", *Croat Med J*, 52(3): 314-326, 2011.
[8] P Marshall, J King, B Budowle, "Utility of Amplification Enhancers in Low Copy Number DNA Analysis", *Int J Legal Med*, 129(1):43-52. doi: 10.1007/s00414-014-1021-1, 2015.
[9] S Champlot, C Berthelot, M Pruvost, EA Bennett, T Grange, et al., "An Efficient Multistrategy DNA Decontamination Procedure of PCR Reagents for Hypersensitive PCR Applications', *PLoS ONE*, 5(9): e13042. doi:10.1371/journal.pone.0013042, 2010.
[10]J Tamariz, K Voynarovska, M Prinz, T Caragine, "The Application of Ultraviolet Irridation to Exogenous Sources of DNA in Plasticware and Water for the Amplification of Low Copy Number DNA", *J Forensic Sci*, 51(4): 790-794, 2006.

TABLE 1. GENEMARKER OUTPUT DATA FOR A PARTIAL DNA PROFILE
(RUN 1 AND RUN 2; DUPLICATE PCR AMPLIFICATIONS FOR COMPARISON)

Run 1:	dye	size	height	height ratio	marker	allele	difference	score
1	Blue	148.7	1354	0.72	D8S1179	14	0.1	Pass
2	Blue	153.0	466	0.25	D8S1179	15	0.1	Undetermined
3	Blue	157.3	1886	1.00	D8S1179	16	0.1	Pass
4	Blue	204.2	374	1.00	D21S11	29	0.0	Pass
1	Green	123.6	3523	1.00	D3S1358	15	0.0	Pass
2	Green	175.2	456	1.00	TH01	7	0.1	Pass
3	Green	228.7	4349	1.00	D13S317	11	0.1	Pass
1	Yellow	98.8	607	0.43	D19S433		1.0	Undetermined
2	Yellow	105.7	473	0.34	D19S433	10	0.1	Undetermined
3	Yellow	109.6	1409	1.00	D19S433	11	0.1	Pass
4	Yellow	113.5	333	0.24	D19S433	12	0.1	Undetermined
5	Yellow	115.4	283	0.20	D19S433	12.2	0.0	Undetermined
6	Yellow	117.4	1364	0.97	D19S433	13	0.0	Pass
7	Yellow	119.4	957	0.68	D19S433	13.2	0.0	Undetermined
8	Yellow	178.8	2320	1.00	vWA	17	0.0	Pass
9	Yellow	229.2	144	1.00	TPOX		1.0	Undetermined
1	Red	106.6	1786	0.93	AMEL	X	0.1	Pass
2	Red	112.2	1923	1.00	AMEL	Y	0.1	Pass
3	Red	146.9	1302	1.00	D5S818	10	0.0	Pass
4	Red	151.3	698	0.54	D5S818	11	0.0	Check
5	Red	238.9	233	1.00	FGA	23	0.2	Pass
Run 2:								
1	Blue	148.5	1872	0.72	D8S1179	14	0.1	Pass
2	Blue	152.9	607	0.23	D8S1179	15	0.2	Undetermined
3	Blue	157.2	2599	1.00	D8S1179	16	0.2	Pass
4	Blue	204.1	499	1.00	D21S11	29	0.1	Pass
1	Green	123.5	5018	1.00	D3S1358	15	0.1	Pass
2	Green	175.1	622	1.00	TH01	7	0.1	Pass
3	Green	228.5	5639	1.00	D13S317	11	0.0	Pass
1	Yellow	109.5	1575	0.94	D19S433	11	0.1	Pass
2	Yellow	117.3	1673	1.00	D19S433	13	0.0	Pass
3	Yellow	119.4	1121	0.67	D19S433	13.2	0.1	Undetermined
4	Yellow	178.7	3225	1.00	vWA	17	0.0	Pass
5	Yellow	229.1	137	1.00	TPOX		1.0	Undetermined
1	Red	106.3	2486	0.93	AMEL	X	0.0	Pass
2	Red	111.9	2671	1.00	AMEL	Y	0.0	Pass
3	Red	146.8	1847	1.00	D5S818	10	0.0	Pass
4	Red	151.2	991	0.54	D5S818	11	0.0	Check
5	Red	238.7	306	1.00	FGA	23	0.2	Pass

