Forensic Science International: Genetics 12 (2014) 208–214

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

The effect of the uncertainty in the number of contributors to mixed DNA profiles on profile interpretation

Jo-Anne Bright [a,b,*], James M. Curran [b], John S. Buckleton [a]

[a] ESR Ltd., Private Bag 92021, Auckland 1142, New Zealand
[b] University of Auckland, Private Bag 92019, Auckland 1142, New Zealand

ARTICLE INFO

Article history:
Received 1 April 2014
Received in revised form 8 June 2014
Accepted 12 June 2014

Keywords:
Forensic DNA
Interpretation
Mixed DNA profiles
Identifiler
STRmix™

ABSTRACT

The effect of uncertainty in the number of contributors to a profile is a matter of some contention in forensic DNA interpretation. Interpretation methods are moving towards continuous models. Within this paper the effect of misspecification of the number of contributors to a profile caused by one artefactual peak, either a large back stutter or a forward stutter, was investigated using a continuous model. The misassignment of the number of contributors to a profile either has no significant effect or decreases the *LR* for the true contributors. It often increases the chance of an adventitious link.

1. Introduction

Forensic DNA profiles are often presented as electropherograms (epgs). Each distinct peak at a locus may correspond to an allele or an artefact. The height of peaks within the epg is measured in relative fluorescent units (rfu) and are roughly proportional to the amount of DNA added to the PCR reaction. The heights of peaks from an individual tend to decrease as the molecular weight of the alleles increases [1,2].

Mixed profiles arise when DNA from two or more individuals is present in a DNA extract. Interpretation of mixed DNA profiles is complicated by the occurrence of artefacts. The most prevalent artefact is back stutter [3], a peak one repeat unit less than the parent allele generated during the PCR amplification of STR loci [4]. Di- and trinucleotide repeats stutter more than tetranucleotide repeats which in turn stutter more than pentanucleotide repeats [5]. Forward stutter, a peak one repeat larger than the parent allele, occurs less frequently but can also complicate profile interpretation. Stutter, either forward or back, is most problematic when the height of stutter peaks is similar to minor contributors within a mixed profile. Back stutter is typically quantified by a stutter ratio, *SR*:

$$SR = \frac{O_{a-1}}{O_a}$$

where O_{a-1} refers to the observed height of the stutter peak, and O_a, the parent peak. Traditionally alleles are assigned on the basis of a threshold applied per locus, or per multiplex, such as all peaks with $SR \geq 0.15$ being assigned as allelic and those with $SR \leq 0.15$ being putative stutter peaks [6].

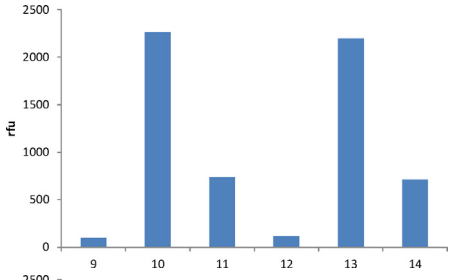
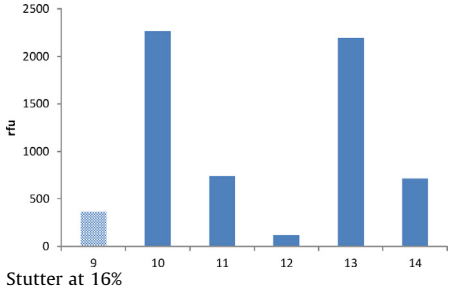
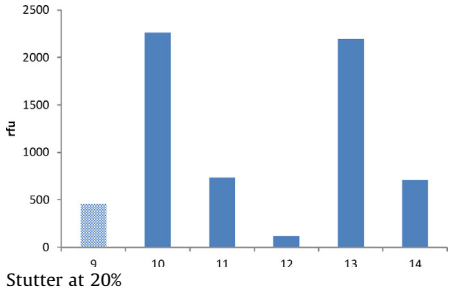
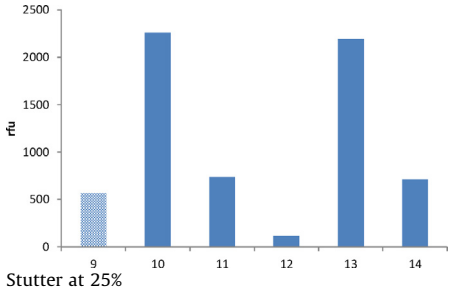
Interpretation methodologies have not kept pace with advances in 'front-end' areas such as DNA extraction and amplification, testing chemistries, and robotics [7]. As the sensitivity of forensic DNA typing procedures increases, more and more mixed DNA profiles are encountered. There has been a push from many jurisdictions for standardisation and more research due to the complexity of low level and mixed profile interpretation [8–10]. The creation of statistical software packages to advance the development, and implementation, of generally accepted standards for forensic genetics has been encouraged [7].

The introduction of probabilistic, or continuous, models removes some, but not all, of the subjectivity in profile analysis [11–13] and moves towards consistency in DNA interpretation and reporting across different laboratories. Continuous models of DNA

* Corresponding author at: ESR Ltd., Private Bag 92021, Auckland 1142, New Zealand. Tel.: +64 8153940; fax: +64 98496046.
E-mail address: Jo.bright@esr.cri.nz (J.-A. Bright).

http://dx.doi.org/10.1016/j.fsigen.2014.06.009
1872-4973/

Table 1
Summary of two person artificial mixed profiles, $t_1 = 3000$ and $t_2 = 1000$. The added ambiguous peaks are plotted as chequered bars in the artificial epgs for marker 1 (D8S1179).

Sample	Marker 1 epg	*N*	*LR* of known contributor		Maximum adventitious *LR*
			LR_1 major	LR_2 minor	
$t_1 = 3000$, $t_2 = 1000$		2	5.57E+19	3.28E+20	0
		3	9.23E+18	5.30E+10	0
		2	5.57E+19	3.28E+20	0
		3	5.57E+19	3.23E+20	25
	Stutter at 16%				
		2	5.57E+19	3.28E+20	0
		3	5.57E+19	3.36E+20	41
	Stutter at 20%				
		2	5.57E+19	3.28E+20	0
		3	8.74E+18	8.54E+10	0
	Stutter at 25%				

profile interpretation have the advantage of modelling stutter directly by assigning a probability to the profile given a proposed genotype [14].

The accurate assignment of the number of contributors in conjunction with the adoption of continuous models has become one of the most contentious issues in forensic DNA profile interpretation. The most difficult profiles to specify the number of contributors are those with peaks that may be either allelic, or artefactual, or both, and which term *ambiguous* in this paper. In our case working experience, trace DNA contributions, profiles with high stutter above an assigned threshold and stutter in a forward $(a + 1)$ position introduce uncertainty and often result in the inflation of the assumed number of contributors to a profile.

The actual number of contributors to a profile is always unknown. In many cases it may be assigned with some confidence with information from the profile itself, and with case and sample information. In the presence of ambiguous peaks it may be tempting to increase the number of assumed contributors. The probability of a given number of contributors is influenced by how likely this number is given the case circumstances and how well this number of contributors explains the profile [15]. A number of contributors that is either very unlikely given the case circumstances or very poor at explaining the profile could be considered unreasonable. Proposing an unreasonable number of contributors under the defence hypothesis, H_d, and holding the number under the prosecution hypothesis at a reasonable assignment will increase the likelihood ratio (*LR*), favouring the prosecution hypothesis, H_p [16]. In addition, Bright et al. [17] demonstrated that the assumption of additional contributors under both the prosecution and defence hypotheses over and above the number suggested by allele count tended to lower the *LR* for the true contributors. It also had the effect of increasing the number of