Article:

Bright, J.-A., Taylor, D., Curran, J. M., & Buckleton, J. S. (2013). **Degradation of forensic DNA profiles**. *Australian Journal of Forensic Sciences*, *45*(4), 445–449.

INSTITUTE OF ENVIRONMENTAL SCIENCE AND RESEARCH LIMITED
Kenepuru Science Centre: 34 Kenepuru Drive, Kenepuru, Porirua 5022 | PO Box 50348, Porirua 5240, New Zealand
T: +64 4 914 0700 **F:** +64 4 914 0770
www.esr.cri.nz

Degradation of forensic DNA profiles

Jo-Anne Bright[1,2*], Duncan Taylor[3], James Curran[2], and John Buckleton[1]

[1] *ESR, Private Bag 92021 Auckland 1025, New Zealand*

[2] *Department of Statistics, University of Auckland, Private Bag 92019, Auckland, New Zealand*

[3] *Forensic Science South Australia, 21 Divett Place, SA 5000, Australia*

* Corresponding author at: Institute of Environmental Science and Research Limited, Private Bag 92021, Auckland, 1142, New Zealand. Email address: Jo.bright@esr.cri.nz .

Selected profiles typed at the Promega PowerPlex® 21 (PP21) loci were examined to determine if a linear or exponential model best described the relationship between peak height and molecular weight. There were fewer large departures from observed and expected peak heights using the exponential model. The larger differences that were observed were exclusively at the high molecular weight loci. We conclude that the data supports the use of an exponential curve to model peak heights versus molecular weight in PP21 profiles. We believe this observation will improve our ability to model expected peak heights for use in DNA interpretation software.

Keywords: Forensic DNA, PowerPlex® 21, degradation

Introduction

In the interpretation of forensic DNA evidence a sample associated with a crime is compared with genotype information from one or more persons. Typically the samples will be amplified using commercially manufactured short tandem repeat (STR) multiplexes that analyse many loci simultaneously, with subsequent PCR product generated on a capillary electrophoresis instrument. The resulting DNA profile is an electropherogram (epg). The heights (or areas) of the peaks within the epg are approximately proportional to the amount of undegraded template DNA [1-4]. However this relationship is affected by a number of systematic factors. Notable amongst these factors is the molecular weight (m_a) of allele, a.

A typical epg has a downward trend with increasing molecular weight. This is variously described as the degradation slope or the 'ski slope' [5-7]. The term degradation slope alludes to a suggested cause, degradation of the DNA. There are many chemical, physical and biological insults which are believed to contribute to DNA degradation or inhibition of a profile. Environmental factors such as humidity [8], bacteria [7] or other forces such as

ultraviolet light breakdown the DNA destroying some fraction of the initial template [9]. Although the cause of slope may not be known, we will refer to this ski slope effect as degradation to comport with common usage.

The modelling of expected peak heights is important in the interpretation of forensic mixtures. The authors have previously described a series of models that can be used to calculate expected values for allele and stutter peak heights, and their ratio, *SR* [10]. Known shortcomings of the binary model [11, 12] have led to the development of new and improved models that factor in the probability of dropout [13-16]. Subsequently, fully continuous interpretation models have been developed [17, 18]. These models take the quantitative information from the electropherogram (for example peak heights) and use them to calculate the probability of the peak heights given all the possible genotype combinations for the individual contributors. This approach removes some of the criticism regarding subjectivity [19, 20] in profile analysis and attempt to ensure consistency in DNA interpretation and reporting across different laboratories.

It is important to understand how degradation affects these models. The simplest model is linear. That is, the expected peak height declines constantly with respect to molecular weight. This can be demonstrated crudely by taking a paper epg and drawing a downward sloping straight line across the apex of the heterozygote peaks from the lowest molecular weight locus to the highest molecular weight locus. A linear model has previously been suggested by the current authors [10]. Tvedebrink et al [21] have proposed an exponential relationship in relation to models for allelic dropout.

If the breakdown of the DNA strand was random with respect to location, then we would expect that the observed height of peak *a*, O_a, would be exponentially related to molecular weight. In this work we investigate linear and exponential equations for modelling degradation within single source Promega PowerPlex® 21 profiles.

Methods

We analysed data from all Australian state and territory laboratories generated using the Promega PowerPlex® 21 multiplex as part of a large data analysis project to implement a continuous model of DNA interpretation in Australasia.

Single source PowerPlex® 21 (Promega Corporation, Madison WI) DNA profiles were submitted for analysis from eight laboratories either as previously analysed outputs or as raw, unanalysed data files. All raw data was analysed using Applied Biosystems' GeneMapper™ ID v 3.2.1 with an analysis threshold of 30 relative fluorescent units (rfu). Previously analysed data sets provided by the laboratories were analysed with a maximum analysis threshold of 30 rfu, with some examples at thresholds below this. All profiles were amplified at 30 cycles (as per the manufacturer's recommendations). Amplified products from two laboratories were separated using Applied Biosystem's 3500 capillary electrophoresis instruments with the remaining laboratories using Applied Biosystem's 3130 instruments.

A total of 1,295 profiles were available from single source samples prepared at optimal conditions from pristine DNA. These profiles demonstrated a range of degradation slopes despite being pristine DNA at optimal amplification conditions. Fifty of the most degraded profiles were selected by taking those with the biggest difference in the ratio of the peak heights at Penta D (a high molecular weight locus) to D16S539 (a low molecular weight locus). These represent some of the profiles with the steepest downward slopes, and thus would be described as degraded using common terminology, whether caused by

degradation or other phenomena such as inhibition. As such these profiles are more likely to provide the desired information on the nature of the relationship between peak height and molecular weight.

One consequence of a linear model is that there exists the possibility that predicted peak heights are negative. This is unreasonable. To avoid this possibility the linear function was modified as shown in equation 1. We write the expected peak height as E_a.

$$E_a^{'} = \begin{cases} E_a, & E_a \geq {}^{Z}/_{2} \\ {}^{Z}/_{2}, & E_a < {}^{Z}/_{2} \end{cases} \quad (1)$$

where Z is the analytical threshold.

The models of interest are linear,

$$E_a' = \max\left(\frac{Z}{2}, \quad t + d \times m_a\right) \quad (2)$$

and exponential

$$E_a' = t \times e^{d \times m_a} \quad (3)$$

Where:

t is the intercept of the line or the constant of proportionality for expected height vs. molecular weight

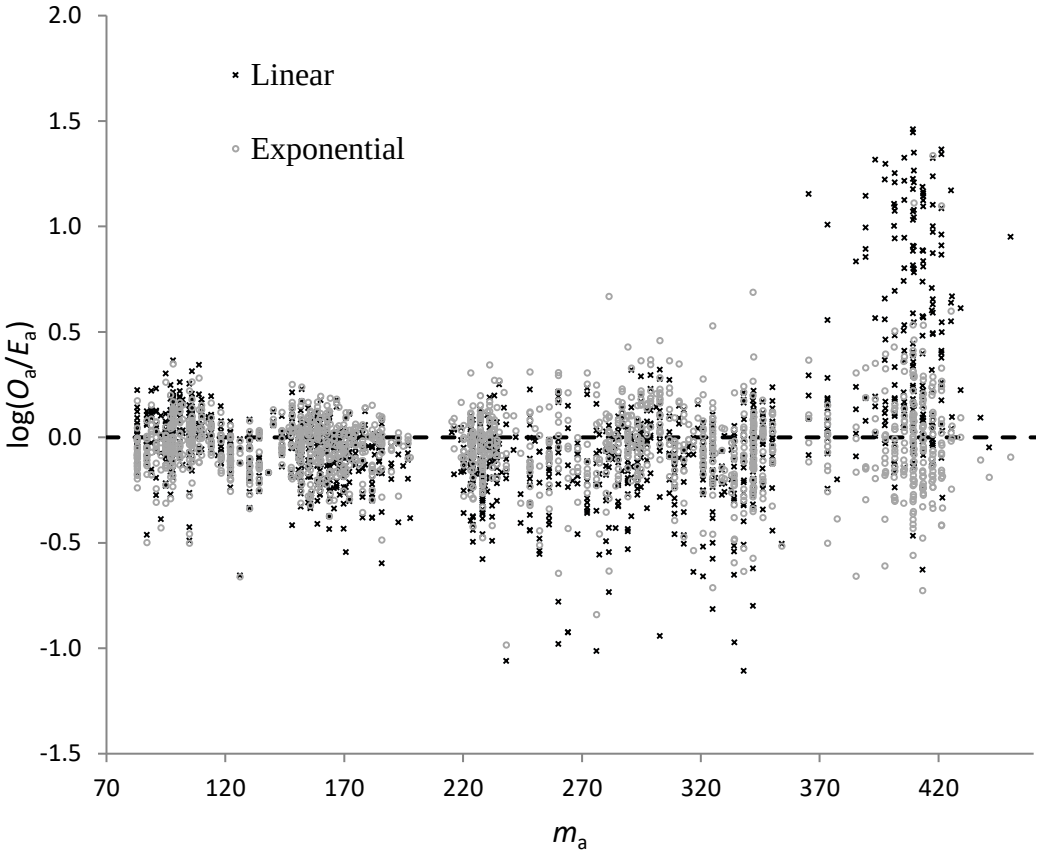
d is the slope of the line or exponent for expected height vs. molecular weight

The values for t and d for each model were determined using maximum likelihood estimation in MS Excel. The exponential and linear models(equations 2 and 3) were then fitted for each profile by least squares in MS Excel.

Results and conclusions

In Figure 1 we show a plot of $\log(\frac{O_a}{E_a'})$ versus molecular weight in base pairs using the exponential and modified linear models. We can see that more extreme positive departures from expectation occur at the high molecular weight end, approximately 350bp and above, and extreme negative departures occur in the mid-zone. This is expected if we force a straight line on an exponential curve. The exponential fit reduces the number of these types of departure. The R^2 values are 0.82 and 0.86 for linear and exponential, respectively. We conclude that this evidence supports the use of an exponential curve to model peak heights versus molecular weight in PowerPlex® 21 profiles.

Figure 1. A plot of $\log(\frac{O_a}{E_a'})$ versus m_a in base pairs using the exponential and linear fitting



Acknowledgements

We warmly acknowledge the comments of Catherine McGovern and Stuart Cooper that have greatly improved this paper. We acknowledge the gracious provision of data from Victoria Police Forensic Services Department. This work was supported in part by grant 2011-DN-BX-K541 from the US National Institute of Justice. Points of view in this document are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Justice.

References

[1] A. Edwards, A. Civitello, H. A. Hammond, C. T. Caskey, DNA typing and genetic mapping with trimeric and tetrameric tandem repeats, American Journal of Human Genetics. 49(4) (1991) 746-756.

[2] J. Ballantyne, E. K. Hanson, M. W. Perlin, DNA mixture genotyping by probabilistic computer interpretation of binomially-sampled laser captured cell populations: Combining quantitative data for greater identification information, Science & Justice.http://dx.doi.org/10.1016/j.scijus.2012.04.004(0).

[3] M. Bill, P. Gill, J. Curran, T. Clayton, R. Pinchin, M. Healy, et al., PENDULUM—a guideline-based approach to the interpretation of STR mixtures, Forensic Science International. 148(2–3) (2005) 181-189.

[4] H. Kelly, J.-A. Bright, J. M. Curran, J. Buckleton, Modelling heterozygote balance in forensic DNA profiles, Forensic Science International: Genetics. 6(6) (2012) 729-734.

[5] J. A. Nicklas, T. Noreault-Conti, E. Buel, Development of a Real-Time Method to Detect DNA Degradation in Forensic Samples*, Journal of Forensic Sciences. 57(2) (2012) 466-471.

[6] D. T. Chung, J. Drábek, K. L. Opel, J. M. Butler, B. R. McCord, A study on the effects of degradation and template concentration on the amplification efficiency of the STR Miniplex primer sets, Journal of Forensic Sciences. 49(4) (2004) 733-740.

[7] B. McCord, K. Opel, M. Funes, S. Zoppis, L. M. Jantz, An Investigation of the Effect of DNA Degradation and Inhibition on PCR Amplification of Single Source and Mixed Forensic Samples.https://www.ncjrs.gov/App/Publications/abstract.aspx?ID=258707 (2011).

[8] E. A. Cotton, R. F. Allsop, J. L. Guest, R. R. E. Frazier, P. Koumi, I. P. Callow, et al., Validation of the AMPF*l*STR® SGM Plus™ system for use in forensic casework, Forensic Science International. 112(2-3) (2000) 151-161.

[9] T. M. Diegoli, M. Farr, C. Cromartie, M. D. Coble, T. W. Bille, An optimized protocol for forensic application of the PreCR™ Repair Mix to multiplex STR amplification of UV-damaged DNA, Forensic Science International: Genetics. 6(4) (2012) 498-503.

[10] J.-A. Bright, D. Taylor, J. Curran, J. Buckleton, Developing allelic and stutter peak height models for a continuous method of DNA interpretation, Forensic Science International: Genetics. http://dx.doi.org/10.1016/j.fsigen.2012.11.013 (2013).

[11] J. Buckleton, C. M. Triggs, Is the 2p rule always conservative?, Forensic Science International. 159 (2006) 206-209.

[12] J. S. Buckleton, C. M. Triggs, S. J. Walsh. Forensic DNA Evidence Interpretation. Boca Raton, Florida: CRC Press; 2004.

[13] H. Haned, Forensim: An open-source initiative for the evaluation of statistical methods in forensic genetics, Forensic Science International: Genetics. 5(4) (2011) 265-268.

[14] H. Haned, P. Gill, Analysis of complex DNA mixtures using the Forensim package, Forensic Science International: Genetics Supplement Series. 3(1) (2011) e79-e80.

[15] D. J. Balding, J. Buckleton, Interpreting low template DNA profiles, Forensic Science International: Genetics. 4(1) (2009) 1-10.

[16] K. E. Lohmueller, N. Rudin, Calculating the Weight of Evidence in Low-Template Forensic DNA Casework, Journal of Forensic Sciences.10.1111/1556-4029.12017 (2012) n/a-n/a.

[17] M. W. Perlin, M. M. Legler, C. E. Spencer, J. L. Smith, W. P. Allan, J. L. Belrose, et al., Validating TrueAllele® DNA Mixture Interpretation, Journal of Forensic Sciences. 56(6) (2011) 1430-1447.

[18] I. W. Evett, P. D. Gill, J. A. Lambert, Taking account of peak areas when interpreting mixed DNA profiles, Journal of the Forensic Sciences. 43(1) (1998) 62-69.

[19] L. Geddes, How DNA evidence creates victims of chance, The New Scientist. 207(2774) (2010) 8-10.

[20] L. Geddes, C. King, C. Stier, Between Prison and Freedom, The New Scientist. 207(2773) (2010) 8-11.

[21] T. Tvedebrink, P. S. Eriksen, H. S. Mogensen, N. Morling, Statistical model for degraded DNA samples and adjusted probabilities for allelic drop-out, Forensic Science International: Genetics. 6(1) (2012) 97-101.