Challenges and Perspectives

Peter Gill[1,2], John Buckleton[3]
[1]University of Strathclyde, Glasgow, UK
[2]Institute of Legal Medicine, University of Oslo, Oslo, Norway
[3]ESR, Auckland, New Zealand

Abstract
Recently, there has been much discussion of the practice of categorising DNA profiles according to whether they are deemed to be *low-copy-number* (LCN) or '*conventional*'. Various authors have previously attempted to provide definitions based on modifications designed to enhance the sensitivity of the existing techniques (e.g. elevated cycle number, post PCR manipulation) whilst others have attempted a classificatory system based on the measured amount of template DNA input to the PCR. We believe that such an approach is neither helpful nor indeed possible. This is because there is no possible natural delineator that can be used. The transition between the two 'states' is gradual rather than sudden and is independent of the methodology utilized to prepare the DNA profile. Rather it is preferable to work towards a single integrated interpretative statistical approach that can be applied universally.

Key words: LCN, LT-DNA, RMNE, LR, consensus model, Omagh Bombing

1. Differences between the RMNE and LR

The concepts of RMNE and LR are fundamentally different. The RMNE approach is not *suspect-anchored* in the same way as the *LR*. The RMNE calculation is a two-step *consecutive* process. Firstly, a determination is made to '*include*' or '*not exclude*' an individual as a contributor to the crimestain profile. Secondly, given that an "inclusion" has been declared, an assessment is made to determine what fraction of people in a defined population would also be "included" [1]. Because of the difficulties of implementing a definitive threshold to distinguish between included/excluded profiles, a third state is usually invoked, namely: "inconclusive".

By contrast, the likelihood ratio (*LR*) approach proceeds by making an assessment of the probability of the crimestain evidence under two alternative hypotheses. There is no need to declare an *inclusion* or *exclusion* because the conditional probability that is used in the calculation obviates that requirement. In some important situations (mainly those in the inconclusive category) the RMNE ignores data which may be in favour of the defence alternative. The principal advantage of the LR approach is that it weighs the evidence in a logically consistent manner and can be applied to almost any situation, provided that the relevant probability data exists.

2. Historical development of the interpretation strategy

In the early-mid 1990's, relatively large inputs of DNA (c.1ng) were required to achieve a full profile. It was recognized that greater sensitivity could be achieved by raising the number of PCR cycles. With the advent of capillary gel eletrophoresis the same effect could be achieved by alternative methods. The advantage of the approach was to increase the range of evidence types amenable to DNA analysis. The lower levels of DNA analysed, concurrently resulted in more partial profiles. Obviously, the partial profile was not a new phenomenon, but we were the first to properly characterize the effect[2]. A partial profile is the result of stochastic effects: a) allelic drop-out, b) allelic drop-in and c) heterozygote imbalance. Traditionally, interpretation of the partial profile followed conventional *'short cut'* methods (e.g. the 2p rule), but these methods are not always conservative[3], hence we devised new ways to assess the strength of evidence. A *statistical model*[2] was developed to accommodate this.. In computational terms, the model was complex and could not have been deployed without software. An alternative method of approximation was developed. This was called the *biological model* – it depends upon the derivation of a consensus profile. The *biological model* was underpinned by the *statistical model* The framework of the *statistical model* resides within the LR approach. The *statistical model* was also used to determine when the *biological model* was at risk of non-conservative reporting.

3. What is LT-DNA?

LT-DNA was a term suggested by Caddy et al [4] to capture the plethora of possible technical modifications designed to render DNA tests more sensitive. It has been defined in various ways:

1) A modification that increases the sensitivity of an existing DNA profiling technique.
2) An amount of DNA below a specific threshold level (previous authors have suggested this level to be 100-200pg).

3.1 Difficulties of absolute definitions:
Our position is that there is no definition for LT-DNA that can be usefully applied because:

1) The stochastic effects associated with LT-DNA profiling (drop-out, drop-in, increased heterozygote imbalance and increased stutter proportions) are observed with all DNA profiling methods. The extent is dependent upon the level of input DNA and not the technical approach taken.
2) A definition based on quantification value fails for the following reasons:
    a. The separate components of mixtures will be different quantities, yet a single approach is taken using classical interpretive methods. With a classic major/minor example, one contributor may be classed as "LCN" and another may be classed as "conventional".
    b. If a sample is degraded, then the low molecular weight DNA fragments may be more prevalent than the high molecular weight fragments – i.e. *within* one contributor some loci may be "LCN" and others may be classed as "conventional".
    c. Attempts to apply a definitive quantity, e.g. 200pg as a delineator, leads to illogical consequences – e.g. a sample that is 199pg is classed as *'LCN'* whereas a 200pg sample is classed as *'conventional'* under this rationale., regardless of the quality of the profile.

For these reasons we have therefore abandoned the LCN (*low-copy-number* term) and prefer to use the LT-DNA term instead. It is not the case that we merely replace one term with another – rather we assert that the interpretative strategy applied to LT-DNA profiles should be applied equally to all DNA profiles, regardless of the method used to produce them, and regardless of the quantity of DNA present. A unified statistical approach is required, that does not rely on the arbitrary classification of what constitutes 'LCN' or 'low-template'.

4. The use of thresholds:
Thresholds are used to make decisions. An example is the *limit of detection* (LOD) threshold – if an allele is below a certain level, typically 50rfu, then it cannot be distinguished from the background noise. The *stochastic* threshold (*T)* is used to decide whether it is 'safe' to designate a locus that shows a single allele as homozygous, rather than heterozygous (with allele drop-out).

As the underlying model is continuous, a threshold can never be set where it is absolutely certain that all events will be captured. Thresholds can only ever be set relative to a *risk analysis* where a predetermined or acceptable level of risk has been defined by someone. If the level is set too high then fewer samples meet the guideline and so are deemed inconclusive. If it is set too low the risk of mistaken interpretation becomes unacceptably high. Hence it is not simply a case of *being conservative.* It is a balancing of bidirectional risk. Moreover, if those cases that clearly do not support the prosecution hypothesis of 'inclusion' are treated as 'inconclusive' when the results favour the defendant, this may be described as an example of prosecution bias.
The determination of thresholds in relation to risk analysis is considered in much greater detail by Gill et al [5].

A Likelihood ratio framework offers several distinct advantages over RMNE :

1) There is no requirement to apply interpretative thresholds
2) The framework can be easily extended to accommodate any set of probabilities
3) The ISFG DNA commission recommended the use of likelihood ratios as the preferred method to evaluate complex mixtures

5. The effects of low level DNA profiling
Whilst it is not possible to provide a working definition of what constitutes a LT-DNA profile, it is much easier to describe the *effects* that characterize such a profile[2]:

1) Drop-out: where, under the prosecution hypothesis, an allele is missing.

2) Drop-in: where, under the prosecution hypothesis, an additional allele is present in the DNA profile.
3) Heterozygotes that exhibit imbalance.

The phenomenon of *'drop-in'* should not be confused with *'gross contamination'*. They are different. Drop-in refers to the occasional and seemingly spontaneous, appearance of alleles that are not part of the sample under test that cannot be reproduced by successive amplifications. They are thought to be the result of the presence of single amplicons (fragmented chromosomal material) present on plasticware or in reagents and are independent random events. As such events are random and independent of eachother. Gross contamination is a partial or full profile that is from an individual – clearly these alleles are not independent events. This occurrence is easily accommodated within the existing LR framework by including an additional (unknown) person in the calculation [6].

6. The hierarchy of propositions [9] applied to DNA profiling

In the criminal arena, the magistrates, judges or jurors are tasked with evaluating the DNA evidence. There are four basic questions that may need to be addressed:
1. To **whom** does the DNA belong?
2. From **what** type of biological material did that DNA originate?
3. **How** did the biological material get to where it was found?
4. **When** did the biological material get to where it was found

Questions 1 and 2 relate to the source of the DNA and in the hierarchy of propositions are referred to as level 1 propositions. Question 2 is sometimes termed 'body fluid attribution' and it has been suggested [7] that in DNA profiling cases, level one propositions should be split into two sub-sets (level one and sub-level one) to reflect that whilst it might be possible to address the issue of who left the DNA, it might not be possible to ascribe that DNA to any particular type of bodily material.

Question 3 relates to transfer and question 4 to persistence of the biological material and collectively they are referred to as 'activity level' propositions. Not all of the questions may be of relevance in any particular case. All parties may, for instance, agree that the DNA is that of the victim and can be attributed to blood, but they may differ in their view concerning the method of transfer to the item.

Critics of LT DNA have frequently confused these separate questions. Issues relating to how and when DNA was transferred (questions 3 and 4) have been generally conflated with questions relating to the source of the DNA (questions 1 and 2). Some have argued that the inability to address questions 2 or 3 means that question 1 cannot be evaluated.
We reject these arguments.

7. Issues of secondary transfer and contamination with DNA profiling

Some biological materials are no different from other trace evidence types (e.g. textile fibres, GSR, glass, paint, explosives) in that when wet, or in a particulate form, they can be transferred secondarily or can be vulnerable to cross-contamination during the forensic process unless suitable anti-contamination precautions are employed. The more sensitive the DNA test the more likely that it can detect such events as the amount of material required is reduced. Much has been made of the fact that the investigators and scientists themselves are potential sources of DNA, Careless procedures used during collection at the crimescene may compromise the *relevance* of the evidence but it should be remembered that the result of an investigator or scientist contaminating a sample with his or her own DNA will either result in a false exclusion (if it passes undetected). If compared to an elimination database of investigators or scientists, then the contamination event can be verified to the source.

8. The Omagh Case (R v Hoey)

Much confusion has resulted from the judgment following this case. However the salient features were as follows:

1) The interpretation of the DNA profile evidence *per se* (question 1*)* was not the main issue. Rather, it was the mode of transfer (question 3) and when the transfer occurred (question 4).
2) No recognizable body fluid (question 2) was targeted but this was irrelevant to the case.
3) Some of the profiles reported were unambiguous (they did not show any drop in or drop out under the prosecution hypothesis) and clearly matched the suspect. The use of elevated PCR cycles was therefore not relevant to the interpretation of the evidence.
4) Criticisms were made about the way in which the evidence was collected by crime scene investigators. This information can be properly assimilated by the scientist reporting the DNA evidence within the framework described above. Because similar considerations must affect all casework, there is no question that the DNA method is somehow compromised. The role of the scientist is to ensure that the evidence is placed into correct perspective.

It is incorrect to suggest that the LT DNA technique was discredited.

10. The future development of the interpretation strategy

1) Because there is no satisfactory definition to LT-DNA (or LCN) we have abandoned the concept in favour of development of a universal strategy that can be used for all DNA profiles regardless of technique used.
2) If thresholds are used to make decisions, then there needs to be an improved level of understanding about how to assess the risks associated with any given level. This is also closely tied to a discussion on 'acceptable levels of risk' measured by rates of '*false exclusion*' and '*false inclusion*'.
3) We hope that a policy of providing freeware solutions under initiatives currently under discussion within ISFG/EDNAP/ENFSI will provide the way forward. Of course it will be essential to concurrently provide education and training in order to implement the new solutions. Furthermore it will be necessary to educate lawyers and others in the criminal justice system. The relevance of the evidence (mode of transfer) is often the main issue in trials where LT-DNA is assessed

9. References

[1] J. Buckleton,J. Curran, A discussion of the merits of random man not excluded and likelihood ratios, Forensic Sci Int Genet. 2 (2008) 343-8.

[2] P. Gill, J.P. Whitaker, C. Flaxman, N. Brown,J. Buckleton, An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA, Forensic Science International. 112 (2000) 17-40.

[3] J. Buckleton,C. Triggs, Is the 2p rule always conservative?, Forensic Sci Int. 159 (2006) 206-9.

[4] B. Caddy, A. Lineacre,G. Taylor, A review of the science of low template DNA analysis, http://police.homeoffice.gov.uk/publications/operational-policing/Review_of_Low_Template_DNA_1.pdf?view=Binary. (2008).

[5] P. Gill, R. Puch-Solis,J. Curran, The low-template-DNA (stochastic) threshold--its determination relative to risk analysis for national DNA databases, Forensic Sci Int Genet. 3 (2009) 104-11.

[6] P. Gill,A. Kirkham, Development of a simulation model to assess the impact of contamination in casework using STRs, J Forensic Sci. 49 (2004) 485-91.

[7] P. Gill, Application of low copy number DNA profiling, Croat Med J. 42 (2001) 229-32.