Forensic Science International: Genetics 4 (2010) 221–227

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

A universal strategy to interpret DNA profiles that does not require a definition of *low-copy-number*

Peter Gill [a,b,*], John Buckleton [c]

[a] University of Strathclyde, Glasgow, UK
[b] Institute of Legal Medicine, University of Oslo, Oslo, Norway
[c] ESR, Auckland, New Zealand

ARTICLE INFO

Article history:
Received 31 July 2009
Received in revised form 11 September 2009
Accepted 15 September 2009

Keywords:
LT-DNA
LCN
Biological model
Drop-out
Drop-in
Contamination
Likelihood ratio
RMNE

ABSTRACT

In this paper we critically examine the causes of the underlying confusion that relates to the issue of low-template (LT) DNA profile interpretation. Firstly, there is much difficulty in attempting to distinguish between LT-DNA vs. conventional DNA because there is no discrete 'cut-off' point that can be reasonably defined or evaluated. LT-DNA is loosely characterised by *drop-out* (where alleles may be missing) and *drop-in* (where additional alleles may be present). We have previously described probabilistic methods that can be used to incorporate these phenomena using likelihood ratio (LR) principles. This is preferred to the random man not excluded (RMNE) method, because we cannot identify a coherent way forward within the restrictions provided by this framework. Most LT-DNA profiles are interpreted using a 'consensus' profile method, we called this the '*biological model*', where only those alleles that are duplicated in consecutive tests are reported. We recognise that there is an increased need for probabilistic models to take precedence over the *biological model*. These models are required for all kinds of DNA profiles, not just those that are believed to be low-template. We also recognise that there is a need for education and training if the methods we recommend are to be widely introduced.

1. Introduction

Since the beginning of the DNA profiling revolution in 1985 [1] partial profiles that exhibited the phenomenon of drop-out were regularly observed. More than 20 years later, these effects are still observed. In a recent paper Budowle et al. [2] argue for caution in the application of DNA techniques when the template levels are low. We readily concur—but we also believe that the same caution needs to be applied to 'conventional' DNA techniques. We see no need to distinguish between the conventional and the low-template LT-DNA profile, primarily because no satisfactory definition can be applied to delineate between the two states. Rather than to attempt an arbitrary categorisation of methodology, we prefer to work towards a comprehensive interpretation framework that can be universally applied. Unfortunately, we cannot see a method to introduce such a framework that utilises the random man not excluded calculation (RMNE)[3]at pg 219–223. Consequently, we advocate the use of an LR framework to interpret complex evidence. This paper is essentially a review of our (and other authors) discussions on the subject, written primarily in the last decade. The concerns noted by Budowle et al. [2] have been previously described (and accommodated) by us. There remains the need to implement new software in order to facilitate statistical analysis, and the requirement to educate all of those engaged with the criminal justice system on the meaning and limitations of DNA profiling evidence.

The argument seems to revolve solely around an arbitrary definition of LT-DNA vs. conventional DNA profiling. We contend that it is unwise to attempt to distinguish between the two states. There has been much confusion surrounding the meaning of *low-copy-number* (LCN). The phrase is typically used to describe a technique that employs elevated cycle number or, to a lesser extent, increased injection time. However, we now reject this definition because the *stochastic effects* associated with the analysis of LT-DNA, including analysis by LCN, are undeniably observed with all DNA profiling technologies. We have therefore abandoned the LCN term if used to describe a sample with low levels of DNA and use the LT-DNA term instead. We recognise that it may be necessary for some providers to retain the LCN term because it is used as a product description describing a technique. We assert that the rationale applied to LT-DNA profiles should be applied equally to all DNA profiles, regardless of the method used to produce them. The Budowle et al. paper opens areas that are

* Corresponding author at: University of Strathclyde, Centre for Forensic Science, Department of Pure and Applied Chemistry, Royal College Building, Glasgow G1 1XW, UK.
E-mail address: peter.gill@strath.ac.uk (P. Gill).

1872-4973/$ – see front matter © 2009 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.fsigen.2009.09.008

worthy of clarification in order to prevent further confusion. These areas include:

1. The nature of variability in DNA replicates
2. The definition of LT-DNA
3. The development of LT-DNA interpretation
4. The risk associated with application of LT-DNA
5. The wider interpretation issues associate with DNA profiling.

2. Variability in DNA replicates

As a basic premise we note that no two replicate profiles from one sample are exactly the same. There will be differences in peak and stutter heights and in the ratios of these heights. This is true regardless of the number of cycles used in amplification or the methodology used. Extensive empirical studies [4–8] have shown that the variability increases as the peak heights decrease.

2.1. Reproducibility vs. reliability

Budowle et al. argue that the loss of reproducibility equates to a loss of reliability. In some definitions reproducibility is one of the requirements for reliability. The Concise Oxford dictionary gives, *inter alia*, "of sound and consistent quality" hence the Budowle et al. comment is not completely without traction. However it is misleading to describe reproducibility to be either a Daubert requirement or a Frye requirement,[i] neither does this conform with guidance from the UK courts. For a discussion of these standards with respect to LT-DNA see Buckleton [9]. This seems to be reasonable, as exact reproducibility cannot be expected. Variability, and indeed uncertainty, is a part of most, if not all, scientific endeavours.

It is not the existence of variability but rather the magnitude and potential consequences of any variability that needs to be assessed and reported to the court. There are many examples by which this variability can be fully accounted for - for example, size bias corrections or sub-population corrections [10–14]. The assessment of LT-DNA is no different. Once all of the facts relating to a case are adduced and the science candidly reported, it is usually the courts responsibility to decide what weight to place on the evidence.

3. The definition of LT-DNA

Budowle et al. suggest that all profiles that give a quantification value of 200 pg or less should be both defined and treated as LT-DNA.

The origin of the 200 pg threshold in the Budowle et al. paper has been taken from the Caddy et al. report to the UK Home Office [15,16] (hereafter "the Caddy report). Budowle et al. claim "the maximum template value has been raised to less than 200 pg", the original level being 100 pg.

However, this implies that there is an 'official' threshold. The Caddy report simply used this level as a 'loose' arbitrary definition that was provided in discussion with UK suppliers. There was no data evaluation to inform such a tight descriptor, neither was this the intention of the report (Adrian Linacre, pers. Comm.).

Furthermore, for reasons explained in our previous papers [9,17] we suggest that any definition based on an arbitrary and generalised quantification level is unfounded. Full profiles at less than 200 pg can be generated using 28 PCR cycles. We certainly observe drop-out in profiles developed from 50 or 100 pg at 28 cycles but we have yet to observe stutter peaks that exceed the size of alleles.

There is no reason why profiles related to any DNA quantity cannot be characterised, and stutter ratios and other measures of stochastic effects assessed. "Our own validation work [6] at 34 cycles and 25 or 12 pg starting template showed that 95% of stutters were less than 0.15 of the parent allele and 99% were less than 0.25, with the maximum observed being 0.57. The drop-in rate was 13.4% per sample, and 1.34% per locus. The proportion of loci exhibiting allelic drop-out was 12.3% with 2.9% of the heterozygotes exhibiting locus drop-out."

In addition a generalised quantification value does not take into account the relative contribution of mixtures. Consequently, the minor component of a mixture may be less than 200 pg. Within a degraded profile, low molecular weight loci will be disproportionately represented. Even within a non-mixed sample, some components could be described as LT-DNA, whereas others will be conventional. We have provided warnings elsewhere [18] of the dangers of ascribing some difficult to interpret, individual loci as 'neutral' evidence.

As explained previously [17] we find it difficult to attempt a definition of LT-DNA at all. This is largely because the underlying variability is continuous. This means that there is no 'magic' cut-off point that can be elucidated. This is why definitions that attempt to relate DNA quantity with the 'state' of conventional vs. LT-DNA are ambiguous, loose guidelines rather than definitive indicators.

The efficacy of the quantification test is dependent upon the system used. Notably, commercial systems such as HY plexor (Promega) or Quantifiler (Applied Biosystems) utilise fragments that are relatively small in comparison with the target molecule. This means that the quantity of DNA measured in a degraded sample will tend to be an overestimate. Ideally, what we should be measuring is the amount of DNA that it is possible to amplify, conditioned on the target fragment of specific interest. This will vary *between* loci and alleles, dependent upon their size.

In fact the best predicator of *relative quantities* of DNA is provided by the electropherogram (epg) itself [9,17]. We prefer to infer the likely magnitude of stochastic effects from the peak heights rather than the quantification result. This means that we cannot support the definition of LT-DNA at 200 pg and we do not know of any active laboratory that would use this definition as an absolute delineator to decide whether to report a DNA profile.

In summary, we will continue to use the term LT-DNA. For the reasons explained previously, our definition of this 'state' is loose. There is no delineator that can be provided. The definition is not based on any technique or process. We only refer to the characteristics of LT-DNA. We cannot therefore define LT-DNA as a delineated 'state'. We can only refer to the consequences of decreasing template-number, independent of the test.

4. The development of LT-DNA interpretation

When the greater variability of LT-DNA profiles was first recognised in the late 90's we, and others, developed an interpretation strategy that significantly compensated for stochastic variability associated with low numbers of molecules which we called the *statistical model* [19]. This strategy was not commented upon in the Budowle et al. paper. In the Budowle et al. paper the discussion was based on the determination of the genotype via a consensus strategy, which we termed the *biological*

[i] To meet the Frye standard, scientific evidence must be demonstrated to be 'generally accepted' by the relevant scientific community. Under Daubert, the trial court assumes the role of gatekeeper for the admission of expert evidence. State courts that have adopted the Daubert rule look to a variety of factors including: (1) whether the theory or technique is subjected to peer review and publication; (2) the known or potential rate of error; (3) whether the theory can be and has been tested; (4) whether there is 'general acceptance' of the opinion or technique in the relevant scientific community (pers. Comm.. Michelle Kazuba, Queens County District Attourney's Office).

model. Most of the Budowle et al. concerns can be attributed to this confusion.

They incorrectly state that there is no method available for the interpretation of mixtures and large stutters. This is not so [3,19–21]. It is the *statistical model* that provided the confidence that *stochastic effects* can be compensated to a significant extent and that enables robust reporting to occur. We were also able to use the statistical model to determine when the *biological model* was at risk of non-conservative reporting. It is worth emphasising that this discussion has little to do with delineation between conventional vs. LT-DNA. All DNA profiles that may be subject to allele drop-out are affected.

It is worthwhile going through the two approaches in some detail as they are philosophically quite different. The *biological model* attempts to infer the genotype from the replicates by *consensus*, the *statistical model* attempts to assess the probability of the replicates from all possible genotypes.

4.1. The biological model

Let us start with the *biological model*. First we note that the moment we accept the existence of *drop-out*, *drop-in* and large stutters there is difficulty in inferring the genotype from the epg (or from replicate epgs). This is not new. In mixed stains it is often difficult to infer the minor contributor genotype and in some mixtures it can be hard to determine the major and the minor. We are therefore used to dealing with ambiguity in the genotypes of the contributors and have developed methods to deal with this ambiguity. Furthermore, it is also important to mention that the methodology was not applied in hind-sight. We argue that all of the challenges were recognised in 2000 [19]. There is still no challenge to our rationale.

Even single stains can have ambiguity. If the peak heights are low enough then it can be difficult to determine whether a locus is a homozygote or a heterozygote with *drop-out*. The 2p rule was developed to deal with this situation, but was subject to the caveat that it was conservative "provided that the band was low in peak area" [19], emphasising that we recognised the limitations *before* implementation.

4.2. Optimum number of replicates

With our rationale, there has never been the suggestion that we are attempting to reconstruct complete genotypes from replicate analyses [22]. This would certainly be necessary under a RMNE philosophy, but is not required within the LR framework. Neither is there an optimum number of replicates. Again the mathematics takes care of the strength of the evidence—the number of replicates is incorporated holistically into this process.

The *biological model* was originally developed in order to facilitate the reporting of DNA profiles that were subject to the twin phenomena of *drop-out* and *drop-in*.[ii] Clearly, methods that existed before 2000 did not specifically deal with these events, even though it is obvious that both phenomena have been prevalent throughout the history of DNA profiling technology. We, and Jonathan Whitaker [19], were the first to: (a) identify these pre-existing phenomena and (b) define a probabilistic method to interpret the phenomena.

[ii] We must distinguish between 'drop-in' and 'contamination' Drop in events are single independent events consisting of fragmented chromosomes that are all pervasive in the environment. Such events are rare, and typically result in the addition of one or two unexplained alleles in some samples. Contamination events are multiple spurious alleles (more than two) present in the profile. These 'gross contamination' events can be dealt with by calculating the LR to include an additional 'unknown' contributor in numerator and denominator.

4.3. The biological model was validated by the statistical model

We prefer probabilistic methods (the *statistical model*) as the way forward, rather than developing a consensus. The probabilistic model does give an assessment of the reliability of the full set of replicates (whatever that number is) and never proceeds via a consensus.

The *biological model* was developed in order to facilitate reporting of low-template DNA profiles in the absence of software solutions which came later [20]. However, the *statistical model* was concurrently made available to check calculations provided by the *biological model* and can be applied without software. We would have expected that court-going challenges to the *biological model* could be addressed by the *statistical model*. The *statistical model* is required to justify the *biological model*. Similarly, it was not the intention that the *biological model* should be preferred to a full *statistical model*. Although cumbersome, the mathematics that we developed could be used to check the results of any low-template result originally interpreted using the biological model. It is of course disappointing that nearly a decade later, vendors still have not developed commercial solutions based on our statistical thinking. Recently, Balding and Buckleton [18] have developed a freeware solution. If validated, this could form the basis for implementation of widely used statistical models that would be used to replace the *biological model*.

4.4. The statistical model

It may seem unusual to state that one can, and should, interpret LT-DNA profiles without ever trying to infer what genotype(s) the epg(s) represent. However this is exactly what we advocate and it is justified mathematically. We will not restate the full logic here having previously published it extensively [3,19–21,23–28]. However consider a locus with a single peak at position A in the one replicate attempted. Let the height of the A peak be low. What is required is to assess the probability of seeing this single A peak **IF** the contributor is an AA homozygote and to assess the probability of this single A peak **IF** the contributor is an AX heterozygote (where X means any other allele). Clearly if the A peak is high then the chance of observing a single A peak from an AX heterozygote is low and so forth. These assessments should be founded on empirical data [24,29]. The mathematics of the extrapolation to multiple replicates and to mixtures follows in a straight-forward way. Note that at no point do we ever pronounce that the contributor or contributors are a certain genotype. The final weight of the evidence involves a summation over all plausible contributor genotypes under two hypotheses. The first hypothesis will be that of the prosecution, usually termed Hp, and will typically be the suggestion that the suspect of genotype, say, AB is a contributor (or one of several contributors). The summation is either simply across the single possibility that the contributor is AB, for simple stains, or across AB and all other possibilities for a mixture. The second hypothesis will be that of the defence, usually termed Hd, and will typically be that the suspect is not a contributor. The summation is across all possible single or multiple contributors depending on whether the stain is treated as a mixture or not.

Consider the situation discussed above. We have a single replicate showing a single A allele, and an AB suspect. We assume that Hp is that the suspect is the donor. For simplicity we treat this as a simple stain. If the suspect is the donor then we require no drop-out of the A allele (with probability $\bar{D}$), drop-out of the B allele (with probability D) and no drop-in (with probability $\bar{C}$).

Hence we model the probability of this profile under Hp as $D\bar{D}\bar{C}$. Note that we at no point state that drop-out has occurred, or not, for any given allele, neither do we state that drop-in has or has not occurred.

Under Hd we assume that the suspect is not the donor. Reasonably we assume that the true donor is an AA homozygote or an AX heterozygote where X stands for any other allele. These have estimated frequencies $p(\mathrm{AA})$ and $p(\mathrm{AX})$ in some population. To obtain a single A peak from an AA homozygote requires no drop-out of a homozygote (with probability $\bar{D}_2$) and no drop-in (with probability $\bar{C}$). To obtain the single A peak from an AX heterozygote requires no drop-out of the A allele (with probability D), drop-out of the X allele (with probability $\bar{D}$) and no drop-in (with probability $\bar{C}$). Hence we model the probability of the single A peak under Hd as $\bar{D}_2\bar{C}\,p(\mathrm{AA}) + D\bar{D}\bar{C}\,p(\mathrm{AX})$ suggesting

$$\mathrm{LR} = \frac{D\bar{D}\bar{C}}{\bar{D}_2\bar{C}\,p(\mathrm{AA}) + D\bar{D}\bar{C}\,p(\mathrm{AX})} = \frac{D\bar{D}}{\bar{D}_2\,p(\mathrm{AA}) + D\bar{D}\,p(\mathrm{AX})}$$

It is often reasonable to assume[iii] that $\bar{D}_2 = \bar{D}(1+D)$ and hence

$$\mathrm{LR} \approx \frac{D}{(1+D)\,p(\mathrm{AA}) + D\,p(\mathrm{AX})} \quad (1)$$

We will return to this equation later as it is the basis of our concerns about the 2p rule.

The test of the strength of evidence is assessed on a continuous basis to formulate the likelihood ratio. If alleles don't appear, or are visualised just a few times in multiple replicates, then the LR is low. This is the requirement advocated by Budowle et al. who appear to have missed the solution in our published work.

In all DNA work there is ambiguity in the number of contributors. This is true even for simple stains [10]. It is claimed, often by advocates of the RMNE approach, that this represents a problem for the LR approach [30]. Again this is not so. The LR can be developed by summing across all possible numbers of contributors weighted by their prior. We accept that in practice, this method is unlikely to be used in the court-room. But the DNA commission on mixture interpretation [26] agreed that it seemed reasonable to allow the prosecution to set the number to that represented by their hypothesis and to optimise the number for the defence. This optimum is usually at the minimum number required to explain the number of peaks [10]. Of course, there is no reason why the defence may not propose a different number of contributors. The LR provides a convenient framework to allow exploratory calculations to be carried out. We have argued elsewhere that it is the RMNE approach that requires the number of contributors in order to declare inclusion or exclusion [31]. Consider the simple situation of a locus showing the alleles AB. Do we exclude an AA homozygote? The answer depends entirely on the number of contributors. Ignoring the plausible number of contributors will lead to false inclusions.

The *statistical model* can be used in two different ways. It can be used to develop a likelihood ratio *per se*, or it can be used to determine whether the consensus approach is "safe" under the circumstances described. If challenged, then there is no reason why the *biological model* cannot be tested directly against the *statistical model*. We have previously tabulated a number of safe and unsafe situations. This list can be extended, but at this stage we prefer to advocate a move by the community towards formal probabilistic *statistical models*.

Budowle et al. have advocated the use of the 2p rule as conservative [2,30]. Consider a single allele peak A in one replicate. With a stochastic threshold (Budowle et al.'s MIT [30]) of 200RFU. If the peak height of A is 201rfu then this is an exclusion against a heterozygote AB suspect whereas, if the 2p rule is used, a peak at 199rfu is strong evidence against the same suspect. Intuitively, this is unreasonable and can be shown to be unreasonable mathematically [26,28]. Consider equation 1. If D is low then the LR is also low whereas the 2p rule is not conservative in all situations [19,26,28] and it has been necessary for us to publish some warnings regarding its use.

The risk area is just below the stochastic threshold and only when the suspect is AB not when he is genetically homozygous AA.

Budowle et al. to some extent warn against bias. We support this stance.

However there have been misquotes, stating: "If a locus shows ab alleles in the crime stain and the suspect is an ab genotype. . .no contamination has occurred" The key to understanding the point lies in the missing parts of the edited quote. The entire (original) quote reads: "For example if a locus shows *ab* alleles in the crime stain and the suspect is an *ab* genotype then we write $p(\mathrm{notC})$ meaning that no contamination has occurred".[iv] Whilst this may have been written more clearly, all that it states is that contamination is not required to explain an *ab* profile under the prosecution hypothesis that it comes from an *ab* suspect. It is implicit that the results can be explained by two (albeit unlikely) drop-in events under the alternative defence hypothesis.

To summarise, in our preferred approach, no allele needs to be designated as allelic, or as stutter, or as drop-in. What can be stated is that if the profile is ab and the suspect is ab, then drop-in does not need to be postulated under Hp. Contrast this with an ab single stain profile and an aa suspect, then drop-in does need to be postulated under Hp. Our process also accommodates Budowle et al.'s concerns regarding replicate stutters, i.e. 'the likelihood of stutter being observed twice in replicate analyses' This is, in fact, part of the elegance of the statistical model, in that it is not necessary to assign peaks in a definitive manner. This means that there isn't an absolute requirement to assign alleles, drop-in/out events, stutters, etc. We always take account of the possibility that peaks are extraneous to the suspect by this method. This is something that cannot be envisaged within the RMNE framework.

Note that modern software solutions such as that described by the *LoComatioN* software [20] includes an assessment of the probability that alleles matching the suspect are all drop-in events, in the LR.

Our purpose was to define a simple method (the *biological model*) that did not misstate the strength of the evidence, along with suitable warnings and caveats about the limitations, which were to be concurrently applied. Thus when Budowle et al. state: "limitations should be explained" we can agree and also point out that this was stated a long time ago and we would concurrently hope that it is universally happening. But we cannot agree that the existence of these limitations should be based on an arbitrary DNA quantity of 200 pg or less. Budowle's list of 10 considerations also apply equally to DNA profiles that are generated from >200 pg.

For example the German 'Phantom' [32] was a widespread contamination incident that occurred in relation to '*conventional*' DNA profiling. A false sense of security is a likely consequence when there are artificial divisions of techniques, where the results obtained from '*conventional*' profiles are interpreted using methods that don't follow the same cautions that are applied to LT-DNA profiles.

4.5. General points on contamination

We welcome the points made by Budowle et al. in relation to contamination issues. The same points had already been made by us previously but are not referenced in the Budowle et al. paper.

[iii] This is obtained by assuming the two alleles of a homozygote act independently. If this is so then in order to see an A allele we need neither to *drop-out* with probability $\bar{D}^2$, or one but not the other to *drop-out* with probability $2D\bar{D}$. Adding these gives $\bar{D}^2 + 2D\bar{D} = \bar{D}(\bar{D}+2D) = \bar{D}(1+D)$ since $D+\bar{D}=1$.

[iv] It is important to reiterate that we are discussing drop-in events (as defined previously) and not a gross contamination event. When $p(C)$ is used it always refers to independent drop-in events.

We hope that the reference list provided in this paper provides clarity.

Budowle et al. suggests that we confine our definition of *drop-in* to laboratory processes. However Gill and Kirkham [33] have explained the mechanisms in great detail and this has been expounded in training workshops internationally. It is not necessary to provide a detailed account here, since we simply refer the interested reader to this work. We can summarise that our definitions/analysis encompasses transfer from sources at the crime scene; at the evidence recovery unit; as well as from the DNA unit itself. Furthermore we provide the methods to (a) assess levels of contamination and (b) assess the impact of *contamination* and *drop-in* by computer simulation models. If known, it is straightforward to assimilate these probabilities into LR calculations (provided levels of contamination are low, the impact on the LR is very small).

4.6. Drop-in vs. contamination

We must not confuse '*drop-in*' with '*gross contamination*'. They are two different concepts (*we accept that there is some confusion on this and we provide more clarity here*). The former relates to appearance of *one* or *two* alleles per sample that arise from *independent* sources. *[Gross] contamination* refers to multiple alleles from a single unknown source. In the latter case these extraneous alleles are *dependent* events (and are therefore not accommodated by the *drop-in* model). Buckleton et al. [3] have carried out a formal assessment of the robustness of the drop-in model of multiple events. However, if multiple alleles (more than two) are present, these are unlikely to be drop-in events and we prefer to invoke an additional contributor to calculate the LR instead. There is no reason why both models may not be used simultaneously to determine the practical implications of using models with different underlying assumptions. The origin of an unknown profile is not relevant to the calculation of the LR. Nothing special (or new) is required to take account of the gross contamination event in mathematical terms.

It cannot be claimed that the mathematical methods/theory that we developed are in routine use in all laboratories. But our main point is that the theory has been developed, and is available for use. The theory is not specific to vague concepts of LCN or LT-DNA

It is worth re-iterating our most important conclusion: "The primary risk of [random] contamination is wrongful exclusion, particularly if the contaminant masks the perpetrator's profile". The actual mechanism of DNA transfer is a separate issue that we discuss below.

4.7. Relevance of evidence

The previous discussion leads naturally onto a consideration of the 'relevance of evidence' [34,35]. This led to the theory of the hierarchy of propositions [36–41] and the ideas were generalised by Gill [42] specifically in relation to DNA profiling, including a consideration of LT-DNA.

Evidence can arise in three broad ways: (a) by 'innocent means', (b) as a result of the crime event itself and (c) as a result of 'contamination', or inadvertent transfer [43].

The mechanism of transfer of a DNA profile is a consideration for every case reported. As previously alluded to, the presence of the DNA profile tells us nothing about how it became evidential. But these considerations are not specific to LT-DNA samples—they are also a serious consideration for '*conventional*' profiles too. Recently, the cases of the German 'phantom'—inadvertent transfer of high-levels of DNA attributed to numerous evidential materials via swabs - were analysed using *conventional* methods recommended by manufacturers (Fig. 1).

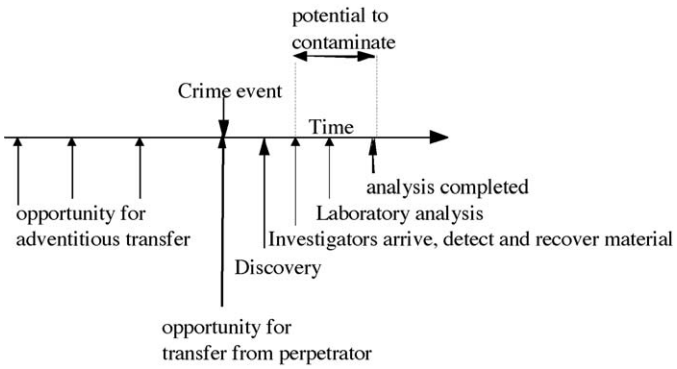

Fig. 1. A generalised timeline that illustrates the potential means by which a DNA profile may be propagated.

A very similar issue arose during the Omagh trial [15]. Not all of the DNA profiles were LT-DNA. Some profiles were complete and matched the suspect's profile. The scientists for the prosecution considered propositions such as:

a. Hp: The DNA came from the suspect
Hd: The DNA came from a random man

However, the issue that concerned the courts was the relevance of the evidence. This suggests propositions of the type:

b. Hp: The DNA came from the suspect when he made the devices
Hd: The DNA came from the suspect by deliberate or inadvertent transfer

Whereas set (a) are questions that can be dealt with by the scientist, set (b) are questions for the jury to consider. Probabilistic determinations can be made using graphical models (or Bayes nets) [44–47] but these require the utilisation of prior probabilities, which is problematic for scientists to use within the UK courts [48,49].

Set (b) and related issues are currently not for the scientist to consider. We agree with Budowle et al., that there is a responsibility for the scientist to place the evidence in context and to point out the limitations of interpretation as described above. However, it is a fallacy to assume that these limitations apply only to LT-DNA.

Unfortunately, there is a *mystique* that surrounds DNA. There is a general public perception that says: '*if there is DNA evidence that matches the suspect then he must be guilty of the offence*'. This perception also extends to some scientists, judges and lawyers. It is highly dangerous thinking, however. Furthermore, it would be very misleading to suppose that this 'problem' was confined solely to the vaguely defined LT-DNA. The *hierarchy of propositions* framework provides a universal method to place the evidence into context, without falling into the trap of straying into areas that are close to the 'ultimate issue' of guilt vs. innocence.

The confusion that arose in the Omagh trial had nothing to do with the DNA profiling evidence *per se*. The difficulty that arose in the case was purely a result of the court's pre-conceptions that assumed the presence of a DNA profile was related to an activity, i.e. the main issues were not within the realm of the scientist to consider. It was the relevance of the evidence (i.e. various modes of transfer) that was the issue—not the process of achieving and interpreting the profile itself. There has been considerable misunderstanding on this point, and we welcome the opportunity to clarify this. However, there remains a perception that failure to convict somehow translates into a failure of science. This would be a very dangerous concept to be given any credence. Whether a suspect is convicted or not is irrelevant—it is the responsibility of the scientist to properly explain the evidence in the context of the case.

Thus we agree with Budowle et al. that limitations of evidence must be described (as they were in the Omagh trial). We assert that the framework provided by the hierarchy of propositions is already used for this purpose. Nothing else is required, other than to educate scientists, judges and lawyers on its uses and practicalities. The numerator of the LR statistic simply considers questions of the style given in set (a). There is no mathematical requirement for a statement on how transfer of the evidence occurred. The LR remains valid regardless. The question of how the DNA was transferred is one for the jury to consider—the scientists' main role is to outline the various modes of transfer that exist and to advise on the relative risks associated with these.

We strongly disagree with Budowle et al. that the limitations described above are confined to LT-DNA. They clearly very much apply to every kind of DNA profiling method. The rationale also applies in a much more general sense (glass, fibres, footprints, etc.).

The uncertainties about the mode of transfer are certainly increased with so-called 'touch DNA' evidence, i.e. evidence that cannot be associated with a particular body-fluid [9]. Budowle et al. write that "some touch DNA samples do not qualify as LCN samples" because they are present in high quantities (>200 pg)—presumably this means that they can be interpreted as *conventional* profiles? It clearly illustrates the difficulties of utilising a definition based on quantification as we would not recommend a different mode of reporting for 'touch-DNA'.

We prefer to consider the position that *all* DNA profiles could be the subject of phenomena that were originally attributed to low-template DNA.

We propose that both the technology and the interpretation science of LT-DNA analysis is much more advanced than has been given credit in the Budowle et al. paper. We also propose that this stance is motivated because there is no solution that can be applied within the inherent limitations of the RMNE method. Our methodology is firmly within the LR framework. We can readily support a call for caution, for attention to potential bias, for more experimental research, for further Police training and awareness and for candour in reporting. We do not see that there is any merit in proposing cautions or definitive rules based on arbitrary delineations of LT-DNA vs. conventional DNA. We believe that if due care is taken and the court is candidly appraised of the limitations of the technique then it is the court's purpose to weigh the strength of the evidence. We do not believe it is the role of the scientist to act as gate-keeper to decide whether evidence should or should not be reported based on arbitrary criteria.

Finally, we sincerely hope that our response has now clarified and will now end the so-called LCN 'debate'. As we have previously asserted, it is a debate about something we cannot define. A universal approach is required. There is a desperate need for educational programs to inform practitioners in all parts of the criminal justice system. There is also a need for specialist software to enable the probabilistic solutions to be fully implemented. We have demonstrated that loose definitions based on arbitrary criteria will always be problematic. Attempts to define and delineate between LCN and conventional DNA profiling has caused much confusion in court.

Let us conclude as follows: LCN or LT-DNA is not a method or technique, it is a way of thinking.

Acknowledgements

We wish to acknowledge the comments of Sue Vintiner and Stuart Cooper that have greatly improved this paper.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.fsigen.2009.09.008.

References

[1] A.J. Jeffreys, V. Wilson, S.L. Thein, Individual-specific 'fingerprints' of human DNA, Nature 316 (1985) 76–79.
[2] B. Budowle, A.J. Eisenberg, A. van Daal, Validity of low copy number typing and applications to forensic science, Croatian Medical Journal 50 (2009) 207–217.
[3] J.S. Buckleton, C.M. Triggs, S.J. Walsh, DNA Evidence, CRC Press, Boca Raton, Florida, 2004.
[4] T. Caragine, et al., Validation of testing and interpretation protocols for low template DNA samples using AmpFlSTR® Identifiler®, Croatian Medical Journal 50 (2009) 250–267.
[5] L. Forster, J. Thomson, S. Kutranov, Direct comparison of post-28-cycle PCR purification and modified capillary electrophoresis methods with the 34-cycle 'low copy number' (LCN) method for analysis of trace forensic DNA samples, Forensic Science International: Genetics 2 (4) (2008) 318–328.
[6] S. Petricevic, et al., Low Copy Number DNA profiling a valid forensic technique? Forensic Science International: Genetics, in draft.
[7] J.-A. Bright, J. Turkington, J.S. Buckleton, Examination of the variability in mixed DNA profile parameters for the Identifiler™ multiplex. Forensic Science International: Genetics, (in press).
[8] J.P. Whitaker, E.A. Cotton, P. Gill, A comparison of the characteristics of profiles produced with the AMPFlSTR® SGM Plus™ multiplex system for both standard and low copy number (LCN) STR DNA analysis, Forensic Science International 123 (2001) 215–223.
[9] J. Buckleton, Validation issues around DNA typing of low level DNA, Forensic Science International: Genetics 3 (4) (2009) 225–260.
[10] J.S. Buckleton, J.M. Curran, P. Gill, Towards understanding the effect of uncertainty in the number of contributors to DNA stains, FSI Genetics 1 (1) (2007) 20–28.
[11] J. Buckleton, J. Curran, S. Walsh, How reliable is the sub-population model in DNA testimony? Forensic Science International 157 (1–2) (2006) 144–148.
[12] P.D. Gill, et al., A comparison of adjustment methods to test the robustness of an STR DNA database comprised of 24 European populations, Forensic Science International 131 (2003) 184–196.
[13] J.M. Curran, J.S. Buckleton, C.M. Triggs, What is the magnitude of the subpopulation effect? Forensic Science International 135 (1) (2003) 1–8.
[14] S.A. Harbison, J.S. Buckleton, Applications and extensions of subpopulation theory: a caseworkers guide, Science & Justice 38 (1998) 249–254.
[15] B. Caddy, G. Taylor, A. Linacre, A review of the science of low template DNA analysis. 2008. Available from: http://police.homeoffice.gov.uk/news-and-publications/publication/operational-policing/Review_of_Low_Template_DNA_1.pdf?view=Binary.
[16] T.F.S. Regulator, Response to Professor Brian Caddy's Review of the Science of Low Template DNA Analysis. 2008.Available from: http://police.homeoffice.gov.uk/news-and-publications/publication/operational-policing/Review_of_Low_Template_DNA_1.pdf?view=Binary.
[17] P. Gill, et al., National recommendations of the Technical UK DNA working group on mixture interpretation for the NDNAD and for court going purposes, Forensic Science International: Genetics 2 (1) (2008) 76–82.
[18] D.J. Balding, J.S. Buckleton, Interpreting low template DNA profiles. Forensic Science International: Genetics, in press.
[19] P. Gill, et al., An investigation of the rigor of interpretation rules for STR's derived from less that 100 pg of DNA, Forensic Science International 112 (1) (2000) 17–40.
[20] P. Gill, A. Kirkham, J. Curran, LoComatioN: A software tool for the analysis of low copy number DNA profiles, Forensic Science International 166 (2–3) (2007) 128–138.
[21] J.M. Curran, P. Gill, M.R. Bill, Interpretation of repeat measurement DNA evidence allowing for multiple contributors and population substructure, Forensic Science International 160 (2005) 47–55.
[22] P. Taberlet, et al., Reliable genotyping of samples with very low DNA quantities using PCR, Nucleic Acids Research 24 (16) (1996) 3189–3194.
[23] P. Gill, et al., Interpretation of complex DNA profiles using empirical models and a method to measure their robustness, Forensic Science International: Genetics 2 (2) (2008) 91–103.
[24] P. Gill, R. Puch-Solis, J. Curran, The low-template DNA (stochastic) threshold - its determination relative to risk analysis for national DNA databases, Forensic Science International: Genetics 3 (2) (2009) 104–111.
[25] P. Gill, J. Curran, C. Neumann, Interpretation of complex DNA profiles using Tippett plots, Forensic Science International: Genetics Supplement Series 1 (1) (2008) 646–648.
[26] P. Gill, et al., DNA commission of the International Society of Forensic Genetics: Recommendations on the interpretation of mixtures, Forensic Science International 160 (2006) 90–101.
[27] J. Buckleton, C. Triggs, Dealing with allelic dropout when reporting the evidential value in DNA relatedness analysis, Forensic Science International 160 (2006) 134–139.
[28] J. Buckleton, C.M. Triggs, Is the 2p rule always conservative? Forensic Science International 159 (2006) 206–209.

[29] T. Tvedebrink, et al., Estimating the probability of allelic drop-out of STR alleles in forensic genetics, Forensic Science International: Genetics 3 (4) (2009) 222–226.
[30] B. Budowle, et al., Mixture Interpretation: defining the relevant features for guidelines for the assessment of mixed DNA profiles in forensic casework, Journal of Forensic Sciences 54 (3) (2009).
[31] J. Buckleton, J. Curran, A discussion of the merits of random man not excluded and likelihood ratios, Forensic Science International: Genetics 2 (4) (2008) 343–348.
[32] C. Himmelreich, Germany's Phantom Serial Killer: A DNA Blunder, in Time. 2009: Berlin. Available from: http://www.time.com/time/world/article/0,8599,1888126,00.html accessed 14th July, 2009.
[33] P. Gill, A. Kirkham, Development of a simulation model to assess the impact of contamination in casework using STRs, Journal of Forensic Sciences 49 (2004) 485–491.
[34] D.A. Stoney, Relaxation of the assumption of relevance and an application to one-trace and two-trace problems, Journal of the Forensic Science Society 34 (1994) 17–21.
[35] C.M. Triggs, J. Buckleton, The two trace transfer problem revisited, Science and Justice 43 (3) (2003) 127–134.
[36] R. Cook, et al., A hierarchy of propositions: Deciding which level to address in casework, Science and Justice 38 (4) (1998) 231–240.
[37] R. Cook, et al., A model for case assessment and interpretation, Science and Justice 38 (3) (1998) 151–156.
[38] I.W. Evett, Towards a uniform framework for reporting opinions in forensic science casework, Science and Justice 38 (3) (1998) 198–202.
[39] I.W. Evett, G. Jackson, J.A. Lambert, More on the hierarchy of propositions: exploring the distinction between explanations and propositions, Science & Justice 40 (1) (2000) 3–10.
[40] I.W. Evett, et al., The impact of the principles of evidence interpretation on the structure and content of statements, Science and Justice 40 (4) (2000) 233–239.
[41] I.W. Evett, et al., Interpreting small quantities of DNA: the hierarchy of propositions and the use of Bayesian networks, Journal of Forensic Sciences 47 (3) (2002) 520–530.
[42] P. Gill, Low copy number DNA profiling, Croatian Medical Journal 42 (2001) 229–232.
[43] P. Gill, Role of short tandem repeat DNA in forensic casework in the UK - past, present and future perspectives, BioTechniques 32 (2) (2002) 366–385.
[44] F. Taroni, et al., A general approach to Bayesian networks for the interpretation of evidence, Forensic Science International 139 (2004) 5–16.
[45] C.G.G. Aitken, F. Taroni, Statistics and the evaluation of evidence for forensic scientists, 2nd ed, John Wiley & Sons, Ltd., 2004.
[46] C.G.G. Aitken, F. Taroni, P. Garbolino, A graphical model for the evaluation of cross-transfer evidence in DNA profiles, Theoretical Population Biology 63 (2003) 179–190.
[47] P. Garbolino, F. Taroni, Evaluation of scientific evidence using Bayesian networks, Forensic Science International 125 (2002) 149–155.
[48] Lord_Justice_Phillips, Mr_Justice_Jowitt, and Mr_Justice_Keene., R. v. Adams and R. v. Doheny, 1996.
[49] Court_of_Appeal, Regina v Doheny and Adams, Royal Courts of Justice – The Strand – London WC2, 1996.