



Forensic Science Regulator

Guidance

**The Interpretation of DNA Evidence
(Including Low-Template DNA)**

FSR-G-202

Issue 2

© Crown copyright 2020

The text in this document (excluding the Forensic Science Regulator's logo, any other logo, and material quoted from other sources) may be reproduced free of charge in any format or medium providing it is reproduced accurately and not used in a misleading context. The material must be acknowledged as Crown copyright and its title specified.

This document is not subject to the Open Government Licence.

1.	Introduction	5
1.1	Background.....	5
2.	Scope.....	5
3.	Implementation	6
4.	Modification.....	6
5.	Terms and Definitions	7
5.2	Historical Development of the Low Copy Number and Low Template DNA Terminology	7
6.	Standardisation.....	9
6.2	Assessment framework.....	10
6.3	Setting technical standards.....	10
6.4	Monitoring technical standards and competence	10
6.5	External performance testing	10
7.	Stochastic Effects	11
8.	Negative Controls and Characterisation of Drop-in.....	15
9.	Estimation of the Probability of Drop-out.....	17
10.	The Purpose of the Quantification Test.....	17
11.	The Consensus Interpretation Methodology	18
12.	Replication.....	19
13.	Population Databases.....	20
14.	Summary of the Basic Interpretation Method Principles	20
15.	Validation.....	22
16.	Determination of the Homozygote Threshold and its Impact on the National DNA Database®.....	22
17.	Acknowledgements.....	25

Guidance – Guidance – Guidance – Guidance – Guidance – Guidance – Guidance

18. Review25

19. References25

20. Abbreviations and Acronyms30

21. Glossary.....31

1. Introduction

1.1 Background

- 1.1.1 Following the publication of ‘A review of the science of Low Template DNA Analysis’ by Professor Brian Caddy (Caddy et al., 2008), the Forensic Science Regulator (FSR) commissioned Professor Peter Gill to provide advice on the level of consensus in the interpretation of DNA profiles and mixed profiles.
- 1.1.2 The FSR adopted a set of principles in a descriptive rather than prescriptive manner that described the UK consensus regarding the features that an interpretation methodology should have to interpret DNA profiles, especially those that are complex in some way because the target material is at a low level (or degraded).
- 1.1.3 Methods to interpret full single-sourced DNA profile(s), where all alleles are present are already largely standardised and non-problematic; methods such as the likelihood ratio, match probability and even random man not excluded calculations are used around the world. Interpretation methodologies in use are qualitative or probabilistic, or a combination of the two.

2. Scope

- 2.1.1 These guidelines apply to England and Wales. Scotland, Northern Ireland and the Republic of Ireland have instituted parallel arrangements for their jurisdictions.
- 2.1.2 This guidance sets out the basic principles to interpret complex DNA profiles and, where appropriate, inform scientific developments to enable their application in practice.
- a. Complex DNA profiles are subject to the same effects that are typically associated with low-level target DNA, i.e. stochastic effects (see Section 7). Profiles are often partial, which means that alleles may be missing (allele drop-out). Additional alleles may also be present – either

because they are mixtures of two or more individuals, and/or because of the allele drop-in phenomenon.

- b. In the case of the complex DNA profile there is now guidance set out in FSR-G-222 'Mixture Interpretation', which aims to standardise interpretation methodology.

3. Implementation

- 3.1.1 This guidance is available for incorporation into a provider's quality management system from the date of publication and is effective from 01 October 2020.

4. Modification

- 4.1.1 This is the second issue of this guidance; it is a major rewrite of the previous version.
- 4.1.2 The modifications made to create Issue 2 of this document were, in part, to ensure compliance with The Public Sector Bodies (Websites and Mobile Applications) (No. 2) Accessibility Regulations 2018. However, the document has also been updated to reflect changes due to the passage of time since its creation. Section 18.1.1 sets out future plans for this document.
- 4.1.3 The Regulator uses an identification system for all documents. In the normal sequence of documents this identifier is of the form 'FSR-#-###' where (a) (the first '#') indicates a letter to describe the type or document and (b) '###' indicates a numerical, or alphanumeric, code to identify the document. For example, this document is FSR-G-202 and the 'G' indicates that it is a guidance document. Combined with the issue number this ensures that each document is uniquely identified.
- 4.1.4 If it is necessary to publish a modified version of a document (for example, a version in a different language), then the modified version will have an additional letter at the end of the unique identifier. The identifier thus becoming FSR-#-####.
- 4.1.5 In all cases the normal document bearing the identifier FSR-#-### is to be taken as the definitive version. In the event of any discrepancy between the

normal version and a modified version then the text of the normal version shall prevail.

5. Terms and Definitions

5.1.1 The terms and definitions set out in the Forensic Science Regulator’s Codes of Practice and Conduct for Forensic Science Providers and Practitioners’ (the Codes), FSR-C-108 ‘DNA Analysis’, FSR-G-222 ‘DNA Mixture Interpretation’ guidance and the Glossary (Section 21) apply.

5.1.2 The term ‘should’ has been used to indicate generally accepted good practice where the reason for not complying or any deviation is required to be recorded.

5.2 Historical Development Of The Low Copy Number And Low Template DNA Terminology

5.2.1 The term low copy number (LCN) was a commercial term, originally coined more than ten years ago (Gill et al., 2000) in relation to an enhanced method (34 polymerase chain reaction [PCR]/amplification cycles) to increase the sensitivity of a DNA profiling test. The LCN term caused confusion as other methods were developed over time, so the Caddy review (Caddy et al., 2008) used the more generic term low template (LT) DNA analysis to encompass all methods used to enhance a DNA profile. These include the use of increased PCR cycle number, longer ‘injection time’ for capillary electrophoresis and sample concentration methods. The Caddy review descriptor of LT-DNA was mainly intended to cover techniques using ‘non-standard’ protocols, or protocols specifically designed to increase sensitivity.

5.2.2 An ad hoc UK DNA technical working group (Gill et al., 2008a) considered the issue in 2008, concluding:

“We do not consider the LCN label for 34 cycles work to be useful, or particularly helpful, and propose to abandon it as a scientific concept, because a clear definition cannot be formulated. Rather, our aim is to recommend generic guidelines that can be universally applied to all DNA profiles that are independent of the method utilised.”

- 5.2.3 This position appears to be implicitly supported by the UK R. v. Reed and Reed appeal court ruling (R. v. Reed and Reed, 2009) The People v. Meganth Frye ruling (The People v. Hemant Meganth, 2010) in the USA, and R. v. Wallace appeal court ruling from New Zealand (The Queen v. Michael Scott Wallace, 2010).
- 5.2.4 The total amount of human, or primate, DNA in a sample can usually be measured. DNA concentration levels such as 100 or 200 picograms (pg) per PCR reaction have previously been suggested as arbitrary thresholds used to describe the delineation between a conventional DNA profile and a low-level target profile, i.e. where there is limited or sub-optimal amounts of DNA material available for testing (Caddy et al., 2008; R. v. Reed and Reed, 2009). Thresholds are often difficult to apply in a meaningful way, for example, in a sample that comprises DNA from two or more individuals, the total quantifiable does not reflect the individual contributions.
- 5.2.5 While the level of DNA may be above an arbitrary threshold, this does not mean that DNA from individual contributors is free from stochastic effects. This is because the quantification process generally does not evaluate the amount of DNA per contributor.¹ In consequence, DNA forming the 'minor' component of a mixture will often exhibit the stochastic effects characteristic of limited or sub-optimal amounts of target DNA material, whereas the 'major' contributor is less likely to exhibit these effects. The original guidelines for dealing with this issue was published by the technical DNA working group (Gill et al., 2008a) and there has been significant progress over recent years in the development of probabilistic theory applied to the interpretation of complex DNA profiles.
- 5.2.6 Since 2012 there has been widespread adoption of probabilistic genotyping software by forensic science providers, such that the use of 'manual' methods is limited to very simplistic calculations and some software algorithm verification. Furthermore, there is the adoption of software that

¹ An exception is where an evaluation is made of a simple male/female mixture where the Y-chromosome content is evaluated.

takes account of peak height. The original version of this guidance document referred to the development of the new theory and its implementation as a means of achieving standardisation, without necessarily following identical methodology (i.e. the same software).

- 5.2.7 Forensic service providers (FSPs) in the UK have now implemented probabilistic genotyping software into their procedures that is a major step forward, enabling the interpretation of many complex DNA profiles. The need for stochastic thresholds no longer exists since the probabilistic software as implemented circumvents the requirement. Analytical thresholds remain in order to filter background noise (this is standard practice that is unlikely to change).
- 5.2.8 Since 2014 the implementation and use of more discriminatory and sensitive profiling kits has increased the incidence of obtaining complex and LT profiles.

6. Standardisation

- 6.1.1 Forensic science in the UK is provided by a mix of police, government and commercial laboratories. In consequence, a diversity of validated methodologies co-exist. At the same time, it is desirable from the court perspective to demonstrate that the interpretation methods utilised within the UK produce broadly similar statistical results. This means that the value of the likelihood ratio derived from a DNA profile should be within an acceptable range (between laboratories) for a given set of propositions.
- 6.1.2 Interpretation methodology should be based on validated probabilistic method(s), whether produced using an expert system, software or through manually based methods.
- 6.1.3 FSR-G-223 ‘Software validation for DNA mixture interpretation’ issued by the Regulator describes the requirements for software validation in much greater detail.

6.2 Assessment Framework

6.2.1 Any framework that monitors and assesses techniques should allow the development of novel methods that can directly compare results across the many different methodologies utilised by forensic service providers.

6.3 Setting Technical Standards

6.3.1 This is achieved through a number of routes, these include:

- a. The Forensic Science Regulator (FSR);
- b. The European Network of Forensic Science Institutes (ENFSI);
- c. International Society for Forensic Genetics (ISFG) recommendations;
- d. The Scientific Working Group on DNA Analysis Method (SWGDM);
and
- e. Individual national DNA databases.

6.4 Monitoring Technical Standards And Competence

6.4.1 These are achieved through a national accreditation body. This is achieved by the requirements for:

- a. Accreditation to ISO/IEC 17025:2017 by the FSR and legally through the implementation of The Council Framework Decision 2009/905/JHA enacted in March 2019, as The Accreditation of Forensic Service Providers Regulations 2018 (SI1276/2018); and
- b. The inclusion of expressions of opinions and interpretations in ISO/IEC 17025:2017 accreditation scopes, as set out in the United Kingdom Accreditation Service (UKAS, 2019) publication LAB 13.

6.5 External Performance Testing

6.5.1 In the USA the National Institute of Standards and Technology (NIST) has undertaken several studies to investigate the variability between laboratories. These studies showed that random match probability estimates (from the same electropherogram) varied by multiple orders of magnitude between different suppliers.

- 6.5.2 The FSR commissioned a UK DNA mixtures collaborative study of forensic service providers and defence experts in order to cross-compare the results (Barber et al., 2014). NIST was the organisation that prepared the study samples, which included DNA extracts for processing through the whole process and electronic DNA data for those participants who only carried out genotyping and/or statistical evaluation.
- 6.5.3 The study showed that there was variation in the terminology used, the methodology used, the statistical evaluation and the reporting of the results. The lessons learned from the study were used as the basis for generating guidance on mixture interpretation, which was published as FSR-G-222.
- 6.5.4 The proficiency testing scheme operated by the Home Office Forensic Information Databases Service (FINDS) for suppliers to the National DNA Database™(NDNAD) includes simple mixture cases in the undeclared trials annual program.
- 6.5.5 As a result of conducting the DNA mixtures collaborative study (Barber et al., 2014), the requirements for producing proficiency tests and collaborative exercises covering the DNA analysis process and use of software was set out in FSR-G-224 Proficiency Testing Guidance for DNA Mixture Analysis and Interpretation.
- 6.5.6 There is still a requirement for proficiency testing in order to ensure that there is consistency between laboratories.
- 6.5.7 Since the DNA mixtures collaborative study was conducted forensic service providers have implemented changes and accessed probabilistic software. The data from the 2014 FSR study have therefore been re-analysed to measure progress since the original study. It is anticipated that the outcome will be published in the near future.

7. Stochastic Effects

- 7.1.1 In comparing questioned profiles against the profile of a known individual, the Forensic Units (FUs) encounter two phenomena: alleles may be missing,

because of ‘allele drop-out’ and/or additional alleles may be present. There are several causes of additional alleles.

- a. Allele drop-in – a term to describe one or two ‘foreign’ alleles per DNA profile (Gill et al., 2000).
- b. Gross-contamination – where a partial or complete DNA profile is obtained as a result of a laboratory contamination event.
- c. Stutters – a small artefact peak in an allelic position.
- d. A mixture of two or more individuals – often one or more contributors may be unknown.

7.1.2 FUs deal with these phenomena using their organisational interpretation guidelines and reporting policy, which may differ between organisations and DNA experts.

7.1.3 When the starting DNA template is at a low level, the efficiency of the entire process is reduced. There is an increased variability in the process and this leads to more variable peak heights/areas in the resulting profile. The increased variance is due to stochastic or random effects. A more marked heterozygote imbalance and allele drop-out (where an allele is missing from the profile) are examples associated with stochastic effects.

7.1.4 The drop-in phenomenon is typically associated with low-level DNA conditions. The environment is randomly ‘contaminated’ with fragmented DNA molecules. The drop-in phenomenon occurs when a fragmented DNA molecule contaminates a tube or other consumable that contains a sample extract. This typically results in the appearance of a single (or two) extra alleles that cannot be attributed to the known reference profile (Gill et al., 2000). Moore et al. (2020) have undertaken studies that increase understanding of the drop-in phenomenon and the difficulty in identifying drop-in versus very low-level DNA contamination due to extreme stochastic effects.

7.1.5 Drop-in is distinct from gross-contamination. The drop-in phenomenon is associated with random allelic events (the alleles are ‘independent’ of each other); whereas gross contamination refers to the transfer of a partial or full profile from a single person (these alleles are ‘dependent’). Consequently,

drop-in is routinely used to refer to the observation of just one or two extra alleles per profile.

- 7.1.6 The International Society for Forensic Genetics (ISFG) DNA commission published recommendations on the evaluation of short tandem repeat (STR) typing results that may include drop-out and/or drop-in using probabilistic methods (Gill et al., 2012).
- 7.1.7 Stutters are also considered to be in the class of ‘additional alleles’, especially if a major/minor mixture is present and the minor contributor is of evidential significance.
- 7.1.8 It is not possible to verify whether drop-in, drop-out or contamination have occurred in a given crime profile. The number of contributors may also be unknown. Statistical models can be used to calculate strength of evidence that takes these uncertainties into account.
- 7.1.9 To summarise, the effects typically observed with low-level DNA conditions are heterozygote imbalance, allele drop-out and allele drop-in.
- 7.1.10 Within the experienced DNA FUs, it is apparent that the effects typically observed with low-level DNA profiles are not restricted to a particular technique. They are also observed with standard analytical methods and, for example, 28 polymerase chain reaction (PCR) cycles, (SGM Plus®) typified by the partial DNA profile.
- 7.1.11 Partial profiles have always been observed with DNA profiles (since the historical beginning of the National DNA Database™ in 1995). But it was not until the year 2000 that the phenomenon was properly described and characterised (Gill et al., 2000).
- 7.1.12 The effects are manifest more often when low-level DNA is analysed, but they are not eliminated with high levels of DNA.
- 7.1.13 The introduction of capillary gel electrophoresis around 2000 resulted in increased sensitivity of the test. Other enhancement techniques were quickly developed, for example, increased injection time, or concentration of the sample. Hence it is probable that the effects associated with complex DNA profiles are more commonly observed today. The continued improvement in

instrumentation and profiling kits again increased the sensitivity leading to the generation of more complex DNA profiles.

- 7.1.14 It has proven difficult to provide a precise definition of the difference between a conventional and low-level target DNA profile. Although the methods used to generate profiles may have differences, no distinct threshold exists to define when such methods are applied; neither is it possible to distinguish between conventional and low-level profiles generated with any single method. It is more appropriate to consider a full profile obtained using standard methods, and a 'poor' mixed partial profile obtained using enhanced/low template (LT) DNA methods as opposite extremes of a continuous range of profile quality.
- 7.1.15 Because the effects increase progressively as the amount of DNA decreases, there is no natural delineator that can be used to differentiate between conventional and low-level DNA profiles (Gill et al., 2008a). In the opinion of a New York Frye hearing (The People v. Hemant Meganth, 2010) it was ruled that the LT-DNA method was a simple extension of existing methodology. If a delineator is chosen for pragmatic purposes, then the decision is based on an arbitrary criterion, usually the amount of template DNA added. Levels such as 100 picograms (pg) or 200pg per PCR reaction have previously been suggested as proxy delineators (Caddy et al., 2008; R. v. Reed and Reed, 2009).
- 7.1.16 The strength of the scientific evidence (to support a prosecution or defence hypothesis) is likely to be maximised with the full conventional DNA profile, and minimised with the poorest interpretable low-level DNA profile. Between the two extremes the strength of evidence is effectively represented on a 'sliding scale', typically as a likelihood ratio (LR).
- 7.1.17 Mixtures often comprise both major and minor contributors. The major contributor may provide a complete profile, whereas the minor contributor may be represented as a partial profile. Hence the sample may respectively exhibit characteristics of a conventional and a low-level target DNA profile.

- 7.1.18 A degraded (unmixed) sample may simultaneously exhibit low-level DNA characteristics in the high molecular weight region and conventional characteristics at the low molecular weight end.
- 7.1.19 Because there is no natural delineator, a quantification test cannot always be used to identify a low-level target DNA profile beforehand. If there is a mixture present, since the separate contributors combine to provide a result, the quantification test gives no information about the relative proportions of contributors. If the 'Y' chromosome is quantified then the relative proportions of male/female components can be determined, but the number of male versus female contributors is indeterminate.
- 7.1.20 The partial profile has less information and this is interpreted using statistical analysis such as a LR method, or match probability. In general, the mixed DNA profile is best interpreted using a LR method.
- 7.1.21 At LT-DNA levels, stutters are typically at the same level as the minor contributor alleles. Probabilistic models can be used to deal with stutters.

8. Negative Controls And Characterisation Of Drop-in

- 8.1.1 Laboratories already carry out some form of negative control monitoring. This provides confidence that the reported results are reliable. Furthermore, the negative controls log provides an indication of the kinds of contamination that are prevalent in the laboratory process, and can act as an early warning system to discover the presence of gross or continual contamination events. The use of staff elimination databases is especially important, since contamination by very definition generally refers to post-incident deposition of DNA material, for example, at collection, during item examination or DNA processing; therefore it is unsurprising that most contamination events are derived from staff members themselves.
- 8.1.2 The probability of drop-in can be estimated, for example, from the frequency of events observed in negative controls as drop-in will be observed in negative controls due to laboratory-based contamination; however drop-in can also result from environmental exposure, thus the probability of drop-in based on events observed in negative controls will be an under-estimate.

Examples of drop-in calculations are given by Gill et al. (2000) and Balding and Buckleton (2009). Other methods could be developed to perform the calculations.

- 8.1.3 The quality of consumables used in the recovery and processing of DNA material has been demonstrated as another route for the introduction of contamination. The implementation of compliance against ISO 18385:2016 ‘Minimizing the risk of human DNA contamination in products used to collect, store and analyse biological material for forensic purposes’ as of August 2020 largely addresses this issue.
- 8.1.4 Manufacturer staff elimination databases are now standard for major DNA kit and consumable manufacturers and kit assemblers. The International Commission on Missing Persons (ICMP) operates a secure international DNA exclusion database for staff manufacturer profiles and unsourced profiles for the forensic DNA community to maintain their own profiles and/or search.
- 8.1.5 Laboratories should incorporate anti-contamination control and monitoring of their process by the following.
- a. Maintaining a log of batch-testing reagents and negative control results to record drop-in and gross contamination events. The purpose will be to act as a monitoring tool and also to provide data that may be used in probabilistic models for reporting purposes.
 - b. Checking profiles against appropriate staff elimination databases, which should include all those who are associated with the collection/recovery of evidence, its analysis, and the processing environment. The requirement for the management of elimination databases is set out in FSR-P-302.
- 8.1.6 Drop-in can be taken into account by probabilistic genotyping software. Some probabilistic genotyping software allow contamination checks to be undertaken where contributors to mixtures can be compared with elimination databases and case-samples within and between sample batches.

9. Estimation Of The Probability Of Drop-out

- 9.1.1 The drop-out event occurs where an allele found in a reference profile is missing in the questioned profile (under the prosecution hypothesis [H_p]). If conventional statistical analysis is applied, this can be anti-conservative. The calculation can be better accommodated by a consideration of 'drop-out'. This parameter can be incorporated into probabilistic calculations that are used to assess the strength of evidence. Examples of methods that might be used are provided by Gill et al. (2000); Balding and Buckleton (2009); Tvedebrink et al. (2009); and Perlin and Sinelnikov (2009). Calculations can also be extended to include mixtures and replicates (Curran et al., 2005) and validation of interpretation strategies, including drop-in rates is described by Taylor et al. (2016).
- 9.1.2 Interpretation methodology should incorporate a probabilistic consideration of drop-out and additional alleles, such as drop-in, stutters, gross-contamination and additional contributors.
- 9.1.3 The quantitative (continuous) probabilistic genotyping software does not require a formal assessment of drop-out, since this is now accommodated by modelling heterozygote balance. Drop-in is taken account of by the software, along with backward and forward stutters. In addition, it is possible to take account of different marker dyes to specify different analytical thresholds.

10. The Purpose Of The Quantification Test

- 10.1.1 Quantification is applied in order to determine the best method to process a sample. The DNA profiling test works best when an optimal amount of DNA is utilised. The quantification test will indicate the volume of extract that contains this optimal amount. If sub-optimal DNA is recovered, then enhanced/low template DNA methods may be used to increase the sensitivity.
- 10.1.2 The quantification test is 'indicative' of the total amount of human DNA (and may also estimate the amount of male DNA) present in a sample. The test may indicate a large quantity of DNA to be present – but on processing, a

much smaller quantity may be recovered. Inhibition or degradation of the sample may account for this discrepancy. The quantification result forms part of the decision tree (how best to process and to interpret a given sample). Otherwise it has little impact on the actual interpretation of the DNA profile result.

- 10.1.3 The DNA profile electropherogram itself provides the best indication of the actual quantity of DNA per allele, per locus, per contributor. Each locus can be additionally assessed relative to the 'local' effects of degradation.
- 10.1.4 The routine use of quantification to determine the method to process a DNA sample is advisable.
- 10.1.5 This is obviated if the amount of available evidential material is deemed to be so low that there is a risk of there being insufficient remaining to provide a successful result.
- 10.1.6 If no profile is obtained, then the possibility of inhibition should be addressed.

11. The Consensus Interpretation Methodology

- 11.1.1 The consensus interpretation method was adopted for the early 'low copy number' casework as described by Gill et al. (2000). Two (or more) replicate amplifications are simultaneously processed per extract. Only those alleles that are replicated, or observed at least twice, are reported with an assignment of evidential strength. Historically, the consensus model was introduced in order to take account of the drop-in phenomenon. Early data suggested that drop-in events were essentially random and relatively rare (one tube phenomena) that were unlikely to be replicated in subsequent tests. Consequently the consensus method acted to filter rare drop-in events, whilst allowing the predominant profile to be reported.
- 11.1.2 The consensus method was validated against a statistical model to demonstrate that (in general) the method was conservative, provided that scientists were suitably trained, since the method relied heavily on expertise.
- 11.1.3 In order to underpin the consensus model, a statistical model was developed and described. This also enabled results to be combined into a single

likelihood ratio (LR) (Gill et al., 2000). This statistical model was the preferred method, at the time, but could not be implemented since the software had not been developed.

- 11.1.4 The consensus method is now redundant. Probabilistic genotyping software in current use is able to analyse replicate tests if carried out and drop-in is taken account of in the calculation of the LR.

12. Replication

- 12.1.1 The statistical method can be used to combine any number of replicates that are processed. Consequently, questions on the 'optimum number of replicates' have little meaning, since a suitable calculation will encapsulate the strength of evidence, irrespective of the number of replicate tests. Benschop *et al.* (2011) carried out a comparison of different methods (based on consensus and composite models) from two to six replicate tests, confirming the conservative nature of the commonly utilised replicate test where alleles must be observed twice before reporting; a decision to replicate a sample can be taken on an individual basis, particularly if the profile is complex.
- 12.1.2 However, careful consideration must be given to the compromised sample, where, despite all efforts, limited material is available. Splitting the sample into two parts may compromise the result, whereas a single analysis may provide the difference between a test result that can be reported and a test result that cannot. Maximising the sample size that is forwarded to polymerase chain reaction (PCR) will reduce the ambiguity inherent in the DNA profile, increasing the strength of the evidence. However, as previously indicated, there is no absolute rationale to support the compulsory replication of a test, provided that it can be supported by a suitable statistical analysis.
- 12.1.3 Replication (more than one PCR for a given DNA extract) of the complex profile is advisable wherever possible. If there is limited DNA then a single test result could be reported using a suitable statistical method.
- 12.1.4 Any replication methodology used should be able to combine replicate test results to produce a single likelihood ratio.

- 12.1.5 There is no specific recommendation made as to whether replicates should be analysed as a matter of routine as it will depend upon the quantity of material available.

13. Population Databases

- 13.1.1 The review led by Professor Brian Caddy (Caddy et al., 2008) noted that there were no standard population databases used within the UK at that time. This position was unsatisfactory. Recognising that forensic units in different areas of the UK may use databases that are derived from local populations, it is not necessary to recommend that ‘universal’ databases are used across the UK.
- 13.1.2 Forensic service providers should use population databases relevant within the UK. These databases should have undergone quality control (Bodner *et al.*, 2016), this can be done by using the short tandem repeats for Identity ENFSI Reference Database (STRidER). The use of other local and/or international population databases may be necessary; the quality and robustness of the data should be checked and any limitations be disclosed.
- 13.1.3 For the implementation of DNA-17 in 2014, UK population databases containing five populations groups were developed and published on GOV.UK. Further details can be found in FSR-G-213 ‘Allele frequency databases and reporting guidance for the DNA (Short Tandem Repeat) profiling’.

14. Summary Of The Basic Interpretation Method Principles

- 14.1.1 For the complex DNA profile, there is no single standard interpretation method. Some basic characteristics can be described to facilitate the development of interpretation methodology.
- 14.1.2 The basic characteristics are as follows.

- a. The interpretation method should consider the effect of additional alleles such as stutter, drop-in, mixtures, and artefactual peaks, ideally using probabilistic methods.
 - b. The interpretation method should consider the effect of ‘missing’ alleles, primarily those caused by the drop-out phenomenon (again ideally using probabilistic theory).
- 14.1.3 In a varied forensic unit environment, statistical methods may diverge. Even for a given (standard) electropherogram, different statistical results will be expected if the interpretation methods are different. The divergence is unknown unless monitored through proficiency testing (see Section 6.5).
- 14.1.4 The purpose of a statistical test is to evaluate the strength of the scientific evidence in relation to an alternative pair of propositions (if a likelihood ratio [LR] is used). It seems desirable, therefore, to be able to compare the relative effectiveness and robustness of statistical models (within scope of the specific claims made). Limitations of a statistical test should also be made clear and the reported strength should not be misleading.
- 14.1.5 In order to carry out the necessary evaluations, new methods need to be developed as a longer term objective, so that a comparative assessment of statistical tests may be carried out. For example, by simulation, rates of false inclusions ($LR > 1$ when the defence hypothesis [H_d] is true) versus false exclusions ($LR < 1$ when the prosecution hypothesis [H_p] is true) would provide an example of such a measure.
- 14.1.6 Methods to evaluate the robustness of statistical models continue to be required.
- 14.1.7 There is much standardisation between forensic service providers in the choice of software they use. There have been a number of publications (Manabe et al., 2017; Alladio et al., 2018; You and Balding, 2019) that show broad comparability between different software, but as of August 2020 these studies are limited and more are required. These comparative studies show that there is broad comparability between different models, but an order of magnitude difference is still expected when the same sample is analysed.

15. Validation

- 15.1.1 Any given process using a particular multiplex is typically validated for set parameters of polymerase chain reaction (PCR) cycle number, extraction methodologies, any post-PCR treatment and capillary electrophoresis processes. A provider may use a standard process, but also have an enhanced or low template DNA method that is identical in every respect except, for example, that injection times for capillary electrophoresis or the number of PCR cycles are increased. These are two different processes that require separate validation and characterisation.
- 15.1.2 The components that are required to validate and characterise processes and that can be used to inform probabilistic models, should include an assessment of:
- a. Stochastic characteristics and associated thresholds (if used);
 - b. Heterozygote balance relative to peak height or DNA quantity, or other parameters; and
 - c. Stutter characteristics.
- 15.1.3 Recommendations on validation of probabilistic genotyping software have been published by the Forensic Science Regulator (FSR-G-223) and by scientific societies including the International Society for Forensic Genetics DNA commission (the Scientific Working Group on DNA Analysis Method (2015), Coble et al. 2016), and the European Network of Forensic Science Institutes (2017).

16. Determination Of The Homozygote Threshold And Its Impact On The National DNA Database®

- 16.1.1 The DNA profile signal is measured in relative fluorescent units (rfu). Historically, the homozygote threshold was set by the UK National DNA Database® (NDNAD) as a guideline to discern possible heterozygote peaks from the baseline (a single value used collectively for all loci and fluorescent dyes). The homozygote threshold is now determined by each forensic unit (FU) through validation of the profiling method and associated processes.

- 16.1.2 In practice, when a single allele 'a' appears at a locus and it is above the selected homozygote threshold then it is reported as a homozygote 'aa'. If it is below that homozygote threshold, then it is reported as 'aF', where the 'F' designation is used to signify potential allele drop-out. The effect of drop-out is to convert a heterozygote locus into a single-allele, which therefore appears to be an apparent homozygote. This cannot be distinguished visually from a true homozygote.
- 16.1.3 An example of a method to determine the homozygote threshold relative to the probability of drop-out (Pr[D]) of a surviving or present allele at a heterozygote is described in Gill *et al.* (2009). This method was used to standardise calculation of the homozygote threshold. Other methods may be preferred by providers (there is no intention to be prescriptive as to the method used).
- 16.1.4 In relation to the UK NDNAD, the 'F' designation is used to decide whether searches for potential matching loci are carried out using the 'F' as a 'wild card designation'.
- 16.1.5 A locus designated as a homozygote 'aa' will only match samples similarly designated, whereas a locus designated 'aF' will match any locus with at least one allele 'a'. The remaining allele can have any identity, including 'a'.
- 16.1.6 Therefore, if the contributor is 'ab' and a locus is wrongly designated as 'aa', then it will not match (although a near match [near miss] report [n-1 search] will capture the event provided that it occurs once only per profile [n-1 search] or twice [n-2 search]).

16.1.7 Previously, there has been no standard method to calculate an appropriate homozygote threshold i.e. the level employed is discretionary. Subject to further discussion and agreement, it would be possible to formalise determination of the threshold, for example, using logistical regression (Figure 1) as described by Gill et al. (2009). Many FUs have already incorporated this rationale as part of the validation of new and existing processes.

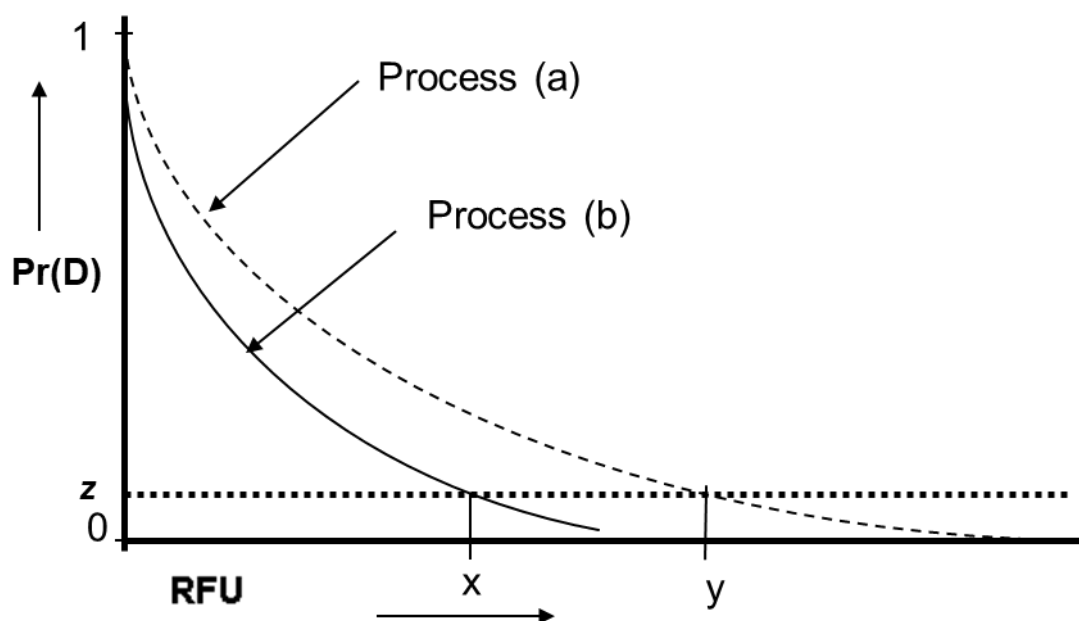


Figure 1: Determination of the homozygote threshold (T) by a method such as logistic regression

In Figure 1 the threshold (T) in rfu is determined with respect to a prescribed level of drop-out measured by $\text{Pr}(D)$. Process (a) is more sensitive than process (b); for example, 34 compared with 28 polymerase chain reaction amplification cycles. If a threshold based on a level of drop-out $\text{Pr}(D)=z$ is used, then for process (a) $T=y$ rfu and for process (b), $T=x$ rfu; the more sensitive the test, the greater the threshold.

16.1.8 In relation to case work reporting, the statistical calculation assigns 'F' to be neutral, since $\text{Pr}(F)=1$ but Buckleton and Triggs (2006) show that this assumption of neutrality is not necessarily conservative.

- 16.1.9 An example is described in detail by Gill et al. (2009) along with a method to carry out a concurrent risk assessment on any decision associated with a given homozygote threshold.
- 16.1.10 The homozygote threshold should be determined on a per locus basis. For simplification purposes an average value across loci can be used provided it can be demonstrated that there is not a significant variability between loci heterozygotes thresholds. This could impact on the level of ‘aF’ designations thus increasing adventitious matches or erroneously elevate the homozygote designations and thereby increase the number of near matches.
- 16.1.11 For data to be uploaded to the UK NDNAD the methods used to determine the homozygote threshold should be demonstrated by validation across the different processes used by the forensic service providers.

17. Acknowledgements

- 17.1.1 Special thanks to Professor Peter Gill as the principal author of the initial guidance and subsequent review and update with the Forensic Science Regulation Unit.

18. Review

- 18.1.1 It is the Forensic Science Regulator’s intention that, when possible, the parts of this document that are still relevant to current practice will be incorporated into other DNA guidance documents and that this document will be withdrawn.
- 18.1.2 The Forensic Science Regulator welcomes comments. Please send them to the address as set out at: www.gov.uk/government/organisations/forensic-science-regulator, or email: FSREnquiries@homeoffice.gov.uk

19. References

Alladio, E., Omedei, M., Cisana, S., D’Amico, G., Caneparo, D. and Vincenti, M. (2018) ‘DNA mixtures interpretation – A proof-of-concept multi-software comparison highlighting different probabilistic methods’

performances on challenging samples', Forensic Science International: Genetics 2018, vol. 37, pp 143–150.

Balding, D. J. and Buckleton, J. (2009) 'Interpreting low template DNA profiles', Forensic Science International: Genetics, vol. 4 (1), pp 1–10.

Barber, M., Evett, I., Pope, S., Sullivan, K., Tully, G., Whitaker, J. and Coble, M. (2014) DNA Mixtures collaborative study report (unpublished), Principle Forensic Services Ltd.

Benschop, C. C. G., van der Beek, C. P., Meiland, H. C., van Gorp, A. G. M., Westen, A. A. and Sijen, T. (2011) 'Low template STR typing: Effect of replicate number and consensus method on genotyping reliability and DNA database search results', Forensic Science International: Genetics, vol. 5, pp 316–328.

Bodner, M., Bastisch, I., Butler, J. M., Fimmers, R., Gill, P., Gusmão, L., Morling, N., Phillips, C., Prinz, M., Schneider, P. M. and Parson, W. (2016) 'Recommendations of the DNA Commission of the International Society for Forensic Genetics (ISFG) on quality control of autosomal Short Tandem Repeat allele frequency databasing (STRidER)', Forensic Science International: Genetics, vol. 24, pp 97–102.

British Standard BS EN ISO/IEC 17025:2017 General requirements for the competence of testing and calibration laboratories. Available at: shop.bsigroup.com [Accessed 25/08/2020].

Buckleton, J. and Triggs, C. (2006) 'Is the 2p rule always conservative?' Forensic Science International, vol. 159 (2–3), pp 206–209.

Caddy, B., Lineacre, A. and Taylor, G. (2008) A review of the science of low template DNA analysis. Available at: www.gov.uk/government/uploads/system/uploads/attachment_data/file/117556/Review_of_Low_Template_DNA_1.pdf [Accessed 25/08/2020].

Coble, M. D., Buckleton, J., Butler, J. M., Egeland, T., Fimmers, R., Gill, P., Gusmão, L., Guttman, B., Krawczak, M., Morling, N., Parson, W., Pinto, N., Schneider, P. M., Sherry, S. T., Willuweit, S. and Prinz, M. (2016) 'DNA Commission of the International Society for Forensic Genetics:

Recommendations on the validation of software programs performing biostatistical calculations for forensic genetics applications', *Forensic Science International: Genetics*, vol. 25, pp 191–197.

Council Framework Decision 2009/905/JHA of 30 November 2009 on accreditation of forensic service providers carrying out laboratory activities. Available at: <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32009F0905> [Accessed 25/08/2020].

Curran, J. M., Gill, P. and Bill, M. R. (2005) 'Interpretation of repeat measurement DNA evidence allowing for multiple contributors and population substructure', *Forensic Science International*, vol. 148 (1), pp 47–53.

ENFSI BPM DNA 01 (2017) Best Practice Manual for the internal validation of probabilistic software to undertake DNA mixture interpretation. European Network of Forensic Science Institutes. Available at: enfsi.eu/documents/best-practice-manuals/ [Accessed 25/08/2020].

Forensic Science Regulator, Codes of Practice and Conduct for Forensic Science Providers and Practitioners in the Criminal Justice System. Available at: www.gov.uk/government/collections/forensic-science-providers-codes-of-practice-and-conduct [Accessed 25/08/2020].

Forensic Science Regulator, Allele frequency databases and reporting guidance for the DNA (Short Tandem Repeat) profiling, FSR-G-213. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/915100/FSR-G-213_DNA_profile_interpretation_Issue_2_Final.pdf [Accessed 07/09/2020]

Forensic Science Regulator DNA contamination detection: The management and use of staff elimination DNA databases, FSR-P-302. Available at: www.gov.uk/government/publications/dna-contamination-detection [Accessed 25/08/2020].

Forensic Science Regulator, DNA mixture interpretation, FSR-G-222.

Available at: www.gov.uk/government/publications/dna-mixture-interpretation-fsr-g-222 [Accessed 25/08/2020].

Forensic Science Regulator, Proficiency Testing Guidance for DNA Mixture Analysis and Interpretation, FSR-G-224. Available at:

www.gov.uk/government/publications/proficiency-testing-guidance-dna-mixture-analysis-and-interpretation [Accessed 25/08/2020].

Forensic Science Regulator, Software validation for DNA mixture interpretation, FSR-G-223. Available at:

www.gov.uk/government/publications/software-validation-for-dna-mixture-interpretation-fsr-g-223 [Accessed 25/08/2020].

Forensic Science Regulator, Validation, FSR-G-201.. Available at:

www.gov.uk/government/publications/forensic-science-providers-validation [Accessed 25/08/2020].

Gill, P., Whitaker, J. P., Flaxman, C., Brown, N. and Buckleton, J. (2000) 'An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA', *Forensic Science International*, vol. 112, pp 17–40.

Gill, P., Brown, R., Fairley, M., Lee, L., Smyth, M., Simpson, M., Irwin, B., Dunlop, J., Greenhalgh, M., Way, K., Westacott, E. J., Ferguson, S. J., Ford, L. V., Clayton, T. and Guinness, J. (2008a) 'National recommendations of the technical UK DNA working group on mixture interpretation for the NDNADB and for court going purposes', *Forensic Science International: Genetics*, vol. 2, pp 76–82.

Gill, P., Curran, J., Neumann, C., Kirkham, A., Clayton, T., Whitaker, J. and Lambert, J. (2008b) 'Interpretation of complex DNA profiles using empirical models and a method to measure their robustness', *Forensic Science International: Genetics*, vol. 2 (2), pp 91–103.

Gill, P., Puch-Solis, R. and Curran, J. (2009) 'The low-template DNA (stochastic) threshold – its determination relative to risk analysis for national DNA databases', *Forensic Science International: Genetics*, vol. 3 (2), pp 104–111.

Gill P., Gusmão, L., Haned, H., Mayr, W. R., Morling, N., Parson, W., Prieto, L., Prinz, M., Schneider, H., Schneider, P. M., Weir, B. S. (2012) 'DNA commission of the International Society of Forensic Genetics: Recommendations on the evaluation of STR typing results that may include drop-out and/or drop-in using probabilistic methods', *Forensic Science International: Genetics*, vol. 6 (6), pp 679–688.

DNA Exclusion Database, International Commission on Missing Persons. Available at: edb.icmp.int/index.php?w=log&ret_to=&l=en [Accessed 25/08/2020].

ISO 18385:2016 Minimizing the risk of human DNA contamination in products used to collect, store and analyse biological material for forensic purposes. Geneva, Switzerland: International Organization for Standardization. Available at: www.iso.org/standard/62341.html [Accessed 25/08/2020].

Manabe, S., Morimoto, C., Hamano, Y., Fujimoto, S. and Tamaki, K. (2017) 'Development and validation of open source software for DNA mixture interpretation based on a quantitative continuous model', *PLoS One*, vol. 12 (11), p e0188183.

Moore, D., Clayton, T. and Thomson, J. (2020) 'A comprehensive study of allele drop-in over an extended period of time', *Forensic Science International: Genetics*, vol. 48, 102332. doi.org/10.1016/j.fsigen.2020.102332.

Perlin, M. W. and Sinelnikov, A. (2009) 'An information gap in DNA evidence interpretation', *PLoS One*, vol. 4 (12), p e8327.

R. v. Reed and Reed and R. v. Garmson (2009) Neutral citation number [2009] EWCA Crim 2698. England and Wales: Court of Appeal.

SWGDM (2015) Guidelines for the Validation of Probabilistic Genotyping Systems. Scientific Working Group on DNA Analysis Method. Available at: www.swgdam.org/publications [Accessed 25/08/2020].

Taylor, D., Bright, J. A., McGoven, C., Hefford, C., Kalafut, T. and Buckleton, J. (2016) 'Validating multiplexes for use in conjunction with

modern interpretation strategies’, *Forensic Science international: Genetics*, vol. 20, pp 6–19.

The People v. Hemant Meganth (2010) Ind. No. 917/2007, Frye Hearing 2010. USA: New York.

The Queen v. Michael Scott Wallace (2010) CA590/2007 [2010] NZCA 46. New Zealand: Court of Appeal.

Tvedebrink, T., Eriksen, P. S., Mogensen, H. S. and Morling, N. (2009) ‘Estimating the probability of allelic drop-out of STR alleles in forensic genetics’, *Forensic Science International: Genetics*, vol. 3 (4), pp 222–226.

UKAS LAB 13: Guidance on the Application of ISO/IEC 17025:2017 Dealing with Expressions of Opinions and Interpretations, Edition 3, April 2019. Staines, England: United Kingdom Accreditation Service. Available at: www.ukas.com/download/publications/publications-relating-to-laboratory-accreditation/LAB-13-Edition-3-April-2019.pdf [Accessed 25/08/2020].

UK Statutory Instruments No. 1276 The Accreditation of Forensic Service Providers (Amendment) Regulations 2018 and 2019. Available at: www.legislation.gov.uk/uksi/2018/1276 [Accessed 25/08/2020].

You, Y. and Balding, D. (2019) ‘A comparison of software for the evaluation of complex DNA profiles’, *Forensic Science International: Genetics*, vol. 40, pp 114–119.

20. Abbreviations and Acronyms

Abbreviation	Meaning
BS	British Standard
DNA	Deoxyribonucleic acid
ENFSI	European Network of Forensic Science Institutes
FSR	Forensic Science Regulator
FU	Forensic unit

Abbreviation	Meaning
ICMP	International Commission on Missing Persons
IEC	International Electrotechnical Commission
ISO	International Organisation for Standardisation
LT	Low template
LR	Likelihood ratio
NDNAD	National DNA Database®
NIST	National Institute of Standards and Technology
PCR	Polymerase chain reaction
pg	Picograms
PT	Proficiency test
rfu	Relative fluorescent units
STR	Short tandem repeat
SWGDM	Scientific Working Group on DNA Analysis Method

21. Glossary

Allele:

A genetic variant at a particular location within an individual's DNA. DNA profiling tests examine a range of alleles that are known to vary widely between individuals. Alleles are represented by peaks in a DNA profile.

Allelic Drop-in:

Additional alleles present in a profile originating from random fragmented sources and are regarded as independent events (no more than two events per profile allowed).

Allele Drop-out:

Alleles may be missing from a DNA profile, so that it is partially represented.

Complex DNA Profile:

A crime-related stain profile that may exhibit drop-out/drop-in phenomena and may be a mixture. The complexity may only become apparent when the DNA profile does not exactly match the reference profile from a known individual under the prosecution hypothesis (H_p).

Contamination:

A spurious DNA profile(s) in a crime-related stain comprising three or more alleles from one or more individual(s). The contributors are considered to be of no relevance to the case (for example, may be introduced into plastic ware during the manufacturing process or may have originated from a scientist processing the samples in the laboratory). It is distinct from drop-in.

Conventional DNA Profile:

A simple, good quality profile.

DNA-17:

Short tandem repeat (STR) multiplex system (kit) with 17 STR loci (including the gender marker amelogenin).

Electropherogram (epg):

The graphical representation of the automated sequencer DNA profile data in a peak format, including information on allele peak molecular weight, peak height/area, and allelic designation relative to an allelic ladder.

Enhancement:

Where the technique sensitivity is increased. Examples include increasing the polymerase chain reaction (PCR) cycle number, increasing capillary electrophoresis injection time, or concentrating the sample for analysis.

Forensic Unit

A forensic unit (FU) is a legal entity or a defined part of a legal entity that performs any part of the forensic science process. [SOURCE: ILAC-G19:08/2014 'Modules in a Forensic Science Process'].

It refers to all providers of forensic science, whether commercial, public sector or internal to a police service. FUs can be small teams in large organisations, sole practitioners or large providers; they can be instructed by the prosecution or the defence.

Homozygote Threshold:

A threshold used to delineate the decision making process in relation to assignment of the 'F' designation to signify drop-out at a heterozygote locus.

Logistic Regression:

An example of a statistical method to determine the probability of an event (drop-out in the example described in figure 1) as a function of another quantity (relative fluorescent units [rfu] of a surviving allele).

Low Copy Number (LCN):

A (commercial) term originally used to describe the application of 34 polymerase chain reaction (PCR) cycles to analysis.

Low-level Target DNA:

A term describing very low amounts of DNA of interest for amplification (PCR).

Low Template DNA (LT-DNA):

A generalised term, also used in the Caddy review (Caddy *et al.*, 2008) to describe the various enhanced methods for analysing low-level DNA (including additional polymerase chain reaction [PCR] cycles, concentration of PCR products and capillary electrophoresis modifications).

Near Match:

Also called a 'near miss' or 'n-1', a near match describes a pair of DNA profiles that differ by one allele.

Polymerase Chain Reaction (PCR):

PCR or amplification of specific short DNA sequences.

Relative Fluorescent Units (RFU):

Florescent markers incorporated into the PCR product are detected during electrophoresis and displayed graphically in rfu.

SGM Plus[®]:

A multiplex system comprising ten short tandem repeat (STR) loci, previously in use in the UK for DNA profiling and loading to the national DNA database[™].

Published by:

The Forensic Science Regulator

5 St Philip's Place

Colmore Row

Birmingham

B3 2PW

www.gov.uk/government/organisations/forensic-science-regulator