NISTIR 8351-DRAFT
EXECUTIVE SUMMARY

DNA Mixture Interpretation:
A NIST Scientific Foundation Review

John M. Butler
Hari Iyer
Rich Press
Melissa K. Taylor
Peter M. Vallone
Sheila Willis*
*International Associate under contract; retired director of Forensic Science Ireland

The full report is available free of charge from:
https://doi.org/10.6028/NIST.IR.8351-draft

National Institute of
Standards and Technology
U.S. Department of Commerce

Executive Summary

All scientific methods have limits. One must understand those limits to use a method appropriately. This is especially important in forensic science as critical decisions impacting life and liberty are often based on the results of forensic analysis.

Forensic DNA technology brings immense benefits to society, and new tools and techniques can increase those benefits further. But as new technologies are implemented with increased detection capabilities, we believe it is important to periodically assess their impacts on the scientific discipline. We do so in this scientific foundation review by identifying scientific principles, reviewing the scientific literature, gathering other empirical evidence from publicly available sources, and receiving input from a group of forensic DNA practitioners and researchers. This scientific foundation review explores what is known about the limits of DNA mixture interpretation methods, including probabilistic genotyping software systems.

As with any field, the scientific process (research, results, publication, additional research, etc.) continues to lead to advancements and better understanding. Information contained in this report comes from the authors' technical and scientific perspectives and review of information available to us during the time of our study. Where our findings identify opportunities for additional research and improvements to practices, we encourage researchers and practitioners to take action toward strengthening methods used to move the field forward. The findings described in this report are meant solely to inform future work in the field.

Improvements in DNA testing methods have allowed forensic scientists to reduce the quantity of DNA required for profiling an individual. In the 1990s, an evidence sample needed to contain thousands of cells, such as from a visible blood or semen stain. Today, analysts can extract a DNA profile from the few skin cells that someone might leave behind when handling an object.

This increased sensitivity extended the usefulness of DNA analysis into new areas of criminal activity beyond homicides and sexual assaults. DNA on bullets or cartridge casings can reveal clues to crimes involving firearms. Swabbing objects that a perpetrator might have handled can yield evidence in property crimes. Cold case evidence previously analyzed with less discriminating methods can be re-opened and researched again to find new insights. However, because people constantly shed small amounts of DNA into the environment, and by touching objects, people can potentially transfer small amounts of DNA from one surface to another, including someone else's DNA. Analyzing small quantities of DNA can create challenges in interpreting the data.

Highly sensitive methods, now universally used across the forensic DNA community, often detect DNA from more than one individual in a sample. But distinguishing one person's DNA from another in these mixtures, estimating how many individuals contributed DNA, determining whether the DNA is even relevant or is from contamination, or whether there is a trace amount of suspect or victim DNA make DNA mixtures inherently more challenging to interpret than single-source samples. These issues, if not properly considered and communicated, can lead to misunderstandings regarding the strength and relevance of the DNA evidence in a case.

When laboratories analyze high-quality, single-source samples, decision-makers often have confidence in DNA test results in part because it has been demonstrated that different laboratories will arrive at the same result. This is true regardless of the specific instruments, kits, and software used. However, multiple interlaboratory studies conducted by different groups over the past two decades have demonstrated a wide range of variation in how specific *DNA mixtures* are interpreted.

This report is arranged into six chapters and two appendices. Chapter 1 introduces the topic of DNA mixtures (samples that contain DNA from more than one individual), the difficulties behind their interpretations, and the relevance of the issues explored in the other chapters of this scientific foundation review. Chapter 2 provides background information on DNA and describes principles and practices underlying mixture measurement and interpretation. The likelihood ratio (LR) framework and probabilistic genotyping software (PGS) are also discussed. Chapter 3 lists data sources used in this study and strategies to locate them. Chapters 4 and 5 cover the report's core concepts: reliability and relevance issues in DNA mixture interpretation. Chapter 6 explores the potential of new technologies to assist mixture interpretation and considerations for implementation. The two appendices provide context on how the field has progressed and strategies to strengthen it going forward. Appendix 1 presents the history of DNA mixture interpretation, while Appendix 2 considers various perspectives on training and continuing education.

A DNA Mixture Resource Group (see Table 1.2), with extensive experience in public and private forensic DNA laboratories, reviewed an early draft of our report and provided valuable feedback, insights, and suggestions. However, they were not asked to sign off on our final report or endorse its conclusions. The NIST team is grateful for their dedication and contributions to our efforts.

Chapter 1: Introduction

New tools and techniques for analyzing and interpreting minor contributors to DNA mixtures are now routinely employed in everyday casework in the United States and around the world. These tools include DNA profiling kits, genetic analyzer instruments, and probabilistic genotyping software.

DNA mixtures can be partly understood by analogy to latent print examination. If multiple fingerprints are deposited on top of one another, it would be difficult to tease apart the individual fingerprints because it may not be clear which ridge lines belong to which print. In a DNA mixture it may not be clear which genetic components, called alleles, belong to which contributor. Interpreting the mixture requires an assessment of which alleles go together to form the DNA profiles of the individual contributors.

Forensic scientists interpret DNA mixtures with the assistance of statistical models and expert judgment. Interpretation becomes more complicated when contributors to the mixture share common alleles. Complications can also arise when random variations, also known as stochastic effects, make it more difficult to confidently interpret the resulting DNA profile.

Not all DNA mixtures present these types of challenges. We agree with the President's Council of Advisors on Science and Technology (PCAST) that "DNA analysis of single-source samples or simple mixtures of two individuals, such as from many rape kits, is an objective method that has been established to be foundationally valid" (PCAST 2016). Therefore, this scientific foundation review does not concentrate on interpretation of single-source DNA samples and two-person mixtures involving significant quantities of DNA from both contributors.

Instead, this review focuses on methods for interpreting data from complex DNA mixtures, which we define as samples that contain comingled DNA from two or more contributors in which stochastic effects or allele sharing cause uncertainty in determining contributor genotypes. The following factors contribute to increased complexity (see also Chapter 2):

- Number of contributors and the degree of overlapping alleles
- Low-quantity DNA from one or more minor contributors
- Degree of degradation or inhibition of the DNA sample.

It is important that users of forensic DNA test results understand that DNA evidence can vary greatly in complexity based on these factors, and that more complex samples involve greater uncertainty.

Chapter 2: DNA Mixture Interpretation: Principles and Practices

Successful analysis and interpretation of DNA results depends on crime scene evidence (the "Q" or questioned sample) being of suitable quality and quantity, and the availability of a reference sample (the "K" or known sample). When appropriate Q and K DNA profiles are available, forensic scientists can perform a Q-to-K comparison and report a likelihood ratio (LR) that is an evaluative interpretation of the strength of this association using specific assumptions and usually one of several statistical approaches. In testing forensic casework samples, a range of DNA profile qualities and quantities can exist. DNA mixtures are inherently more difficult to interpret than single-source DNA samples.

The process of DNA evidence analysis can be divided into two major steps: (1) *measurements* of relative abundances of polymerase chain reaction (PCR) products in a tested DNA sample that are displayed as an electropherogram (EPG), and (2) *interpretation* involving use of the EPG data to make a strength-of-evidence assessment when an evidentiary DNA profile is compared to a person of interest (POI). The outcome of interpretation includes an LR number that can range in value depending on the analyst's assumptions, protocols, algorithms, tools, and other variables. There remains a need to assess the fitness for purpose of an analyst's LR using empirical methods.

Forensic scientists interpret DNA mixtures with the assistance of statistical models and expert judgment. Interpretation becomes more complicated when contributors to the mixture share common alleles. Complications can also arise when reduced DNA template amounts are used in PCR, where random sampling, also known as stochastic effects, makes it more difficult to confidently interpret the resulting DNA profile.

This chapter describes 16 principles and includes 6 key takeaways.

KEY TAKEAWAY #2.1: DNA mixtures, where the DNA of more than one individual is present in a sample, are inherently more difficult to interpret than single-source DNA samples.

KEY TAKEAWAY #2.2: Generating a DNA profile involves measuring the inherent physical properties of the sample. Interpreting a DNA profile involves assigning values that are not inherent to the sample. To do this, the DNA analyst uses their judgment, training, tools (including computer software), and experience, and considers factors such as case context.

KEY TAKEAWAY #2.3: The process of generating a DNA profile can produce stochastic or random variation and artifacts that contribute to the challenge of DNA mixture interpretation.

KEY TAKEAWAY #2.4: DNA mixtures vary in complexity, and the more complex the sample, the greater the uncertainty surrounding interpretation. Factors that contribute to complexity include the number of contributors, the quantity of DNA from each contributor, contributor mixture ratios, sample quality, and the degree of allele sharing.

KEY TAKEAWAY #2.5: Continuous probabilistic genotyping software (PGS) methods utilize more information from a DNA profile than binary approaches.

KEY TAKEAWAY #2.6: Likelihood ratios are not measurements. There is no single, correct likelihood ratio (LR). Different individuals and/or PGS systems often assign different LR values when presented with the same evidence because they base their judgment on different kits, protocols, models, assumptions, or computational algorithms. Empirical data for assessing the fitness for purpose of an analyst's LR are therefore warranted.

Chapter 3: Data and Information Sources

This chapter contains sources of data and information used in conducting this review along with strategies to locate them. These sources include (1) peer-reviewed articles appearing in scientific journals, (2) published interlaboratory studies, (3) laboratory internal validation study summaries that are accessible online, and (4) proficiency test data available on test provider websites.

Chapter 4: Reliability of DNA Mixture Measurements and Interpretation

In this report, we divide the challenges presented by DNA mixtures into two main categories. The first involves the *reliability* of mixture interpretation methods when used with DNA evidence of varying complexity. (Chapter 5 deals with the second challenge: *relevance*.) In this report, we use the "plain English" definition of reliability as a measure of trustworthiness. A highly reliable method is one that consistently produces accurate results. Reliability is not a yes or no question, but a matter of degree. Understanding the degree of reliability of a method can help the user of that information decide whether they should trust the results of that method when making important decisions.

This chapter considers foundational issues related to reliability of DNA mixture interpretation. Reliability centers on trustworthiness established through empirical assessments of available data to evaluate the degree of reliability of a system or its components. We use the term “factor space” to describe the factors that influence complexity, measurement, and interpretation reliability – these factors include the number of contributors, the degree of allele sharing, the ratios of mixture components, and the amount and quality of the DNA tested.

We note that the degree of reliability of a DNA mixture interpretation system, such as a DNA analyst using a probabilistic genotyping software program, depends on sample complexity. Results cannot be simply categorized as “reliable” or “unreliable” without considering context. In addition, reliability cannot be established without validation tests using known samples of similar complexity. The results of such tests provide data that are considered accurate and reliable; only with such valid results can comparisons be made as to the reliability of unknown casework samples. We also emphasize that samples used in proficiency tests need to be representative of complex DNA mixtures seen in casework if these tests are intended to assess analysts’ ability to conduct dependable DNA mixture interpretation.

Finally, the theme of reliability is discussed throughout this report. Note that our original goal in this review was *external* and *independent* assessment of reliability based on publicly available data that met our selection criteria. These criteria evolved during this study as we became aware of the amount and type of data available to us. Laboratories and researchers may make claims or have their own understanding of reliability as it relates to their own work, but our findings are defined by the public information available at the time of this report.

This chapter includes eight key takeaways.

KEY TAKEAWAY #4.1: The degree of reliability of a component or a system can be assessed using empirical data (when available) obtained through validation studies, interlaboratory studies, and proficiency tests.

KEY TAKEAWAY #4.2: To enable effective use of any information, responsibilities exist with both providers and users of that information. While a provider explains the relevance and significance of the information and data, only the user can assess the degree of reliability, validity, and whether that information is fit-for-purpose.

KEY TAKEAWAY #4.3: Currently, there is not enough publicly available data to enable an external and independent assessment of the degree of reliability of DNA mixture interpretation practices, including the use of probabilistic genotyping software (PGS) systems. To allow for external and independent assessments of reliability going forward, we encourage forensic laboratories to make their underlying PGS validation data publicly available and to regularly participate in interlaboratory studies.

KEY TAKEAWAY #4.4: Additional PGS validation studies have been published since the 2016 PCAST Report. However, publicly available information continues to lack sufficient details needed to independently assess reliability of specific LR values produced in PGS systems for complex DNA mixture interpretation. Even when a comparable reliability can be assessed (results for a two-person mixed sample are generally expected to be more reliable than those for a four-person mixed sample, for example), there is no threshold or criteria established to determine what is an acceptable level of reliability.

KEY TAKEAWAY #4.5: Current proficiency tests are focused on single-source samples and simple two-person mixtures with large quantities of DNA. To appropriately assess the ability of analysts to interpret complex DNA mixtures, proficiency tests should evolve to address mixtures with low-template components or more than two contributors – samples of the type often seen in modern casework.

KEY TAKEAWAY #4.6: Different analysts and different laboratories will have different approaches to interpreting the same DNA mixture. This introduces variability and uncertainty in DNA mixture interpretation. Improvements across the entire community are expected with an increased understanding of the causes of variability among laboratories and analysts.

KEY TAKEAWAY #4.7: The degree of reliability of a PGS system when interpreting a DNA mixture can be judged based on validation studies using known samples that are similar in complexity to the sample in the case. To enable users of results to assess the degree of reliability in the case of interest, it would be helpful to include these validation performance results in the case file and report.

KEY TAKEAWAY #4.8: We encourage a separate scientific foundation review on the topic of likelihood ratios in forensic science and how LRs are calculated, understood, and communicated.

Chapter 5: Context and Relevance Related to DNA Mixture Interpretation

The second major challenge posed by DNA mixtures involves the *relevance* of a DNA sample to the crime being investigated. The question of relevance arises because DNA can be transferred between surfaces, potentially more than once. This means that some of the DNA present at a crime scene may be irrelevant to the crime, and current DNA profiling methods increase the likelihood of detecting more DNA. Similarly, today's highly sensitive DNA methods increase the risk that very small amounts of contamination might affect DNA test results.

Chapter 5 focuses on questions of context and relevance: How and when was the DNA deposited, and is that DNA relevant to the crime being investigated?

The question of relevance arises because people readily shed DNA into the environment, and they can potentially transfer DNA between surfaces when touching objects or other people. Therefore, the DNA present at a crime scene or on a piece of evidence may be irrelevant to any crime. To assess relevance, in addition to knowing specific details of the case, one would need information on what factors make DNA more or less likely to transfer and to persist in the environment. This chapter reviews the scientific literature on DNA transfer and persistence and presents strategies for assessing DNA relevance.

The fact that DNA can be transferred between surfaces upon contact is a foundational principle of forensic DNA analysis. This is what makes the discipline useful for investigating crimes in the first place. This has several implications for DNA found at a crime scene. First, that DNA might have been deposited before or after the crime was committed and therefore may not be relevant to the crime. Second, the DNA might have been deposited via secondary transfer, which occurs when DNA is picked up for one surface and deposited on another. For instance, a person might pick up DNA from a second person during a handshake, then deposit the second person's DNA onto an item or surface.

These possibilities mean that the presence of a person's DNA in an evidence sample does not necessarily mean that the DNA is relevant to the crime. Relevance should be assessed. If not, the evidence can be misleading.

By definition, highly sensitive methods are more likely to detect small quantities of DNA, including background DNA that may be present in the environment. In addition, highly sensitive methods are more likely to detect DNA mixtures, which by their nature usually include irrelevant DNA. Therefore, when assessing evidence that involves very small quantities of DNA, it is especially important to carefully consider relevance.

This report uses the word contamination to describe the transfer of irrelevant DNA during an investigation. For example, a fingerprint brush can potentially transfer minute amounts of DNA onto evidence at a crime scene. Such a small amount of DNA might have gone undetected in the past, but highly sensitive methods increase the likelihood that it might now be detected. This increases the likelihood that contamination might affect an investigation.

Forensic laboratories have been using procedures to avoid contamination since the advent of DNA methods. However, because the likelihood of detecting contaminating DNA has increased with the development of highly sensitive DNA methods, contamination avoidance in forensic laboratories is more important than ever. Furthermore, contamination avoidance procedures should be used during all stages of an investigation, including at the crime scene. Elimination databases that include DNA profiles of laboratory staff and police who go to crime scenes can help identify contamination and should be maintained.

Many interpretation methods, including probabilistic genotyping, address questions about who might have contributed DNA to a crime scene profile and express the strength of evidence in the form of a likelihood ratio. This statistic does not provide any information about how much DNA was present, or how or when the DNA was deposited. For instance, a large blood stain might produce a very similar likelihood ratio to a swab from a light switch, yet the two types of evidence might vary greatly in terms of their evidential value. Therefore, the likelihood ratio should not be used in isolation. It is imperative that the likelihood ratio be considered in the context of other evidence in the case.

The fact that DNA can transfer does not mean that DNA is useless as evidence. To the contrary, this is what makes DNA useful to criminal investigations in the first place. However, the possibility of DNA transfer may raise questions of relevance that need to be addressed, especially in cases that involve very small amounts of DNA. These questions can be addressed by considering DNA evidence in the context of case circumstances, including other evidence in the case.

More research is needed on DNA transfer and persistence. In addition, to make use of the studies that are available, individual laboratories would need to know how the sensitivity of methods used in their laboratory compares to the sensitivity of methods employed in the studies being considered.

This chapter includes six key takeaways.

KEY TAKEAWAY #5.1: DNA can be transferred from one surface or person to another, and this can potentially happen multiple times. Therefore, the DNA present on an evidence item may be unrelated (irrelevant) to the crime being investigated.

KEY TAKEAWAY #5.2: Highly sensitive DNA methods increase the likelihood of detecting irrelevant DNA. When assessing evidence that involves very small quantities of DNA, it is especially important to consider relevance.

KEY TAKEAWAY #5.3: Highly sensitive methods increase the likelihood of detecting contaminating DNA that might affect an investigation. Contamination avoidance procedures should be robust both at the crime scene and in the laboratory.

KEY TAKEAWAY #5.4: DNA statistical results such as a sub-source likelihood ratio do not provide information about how or when DNA was transferred, or whether it is relevant to a case. Therefore, using the likelihood ratio as a standalone number without context can be misleading.

KEY TAKEAWAY #5.5: The fact that DNA transfers easily between objects does not negate the value of DNA evidence. However, the value of DNA evidence depends on the circumstances of the case.

KEY TAKEAWAY #5.6: There is a growing body of knowledge about DNA transfer and persistence, but significant knowledge gaps remain.

Chapter 6: New Technologies: Potential and Limitations

New technologies are often investigated to assess whether they can provide solutions to existing problems in the forensic community. The adoption and implementation of these technologies depends upon a cost/benefit analysis within forensic laboratories. Appreciating fundamental challenges with DNA mixture interpretation can assist in considering whether new approaches can bring desired improvements to mixture interpretation.

The ability to analyze short tandem repeat alleles by sequence in addition to length promises to bring some new capabilities to forensic DNA laboratories, including the potential for improvements in DNA mixture interpretation. Next-generation sequencing platforms also enable additional genetic markers to be examined, some of which, such as microhaplotypes, have been pursued with the potential to improve DNA mixture interpretation. Additionally, cell separation techniques offer the potential to separate contributors prior to DNA extraction.

The ultimate decision to implement new technologies in forensic laboratories should be driven by a real-use case and by those responsible for producing and reporting the information. A vendor or members of the general public may encourage forensic DNA laboratories to adopt a new approach or technology without appreciating investments required to make a change. Consideration should be given to whether supporting factors and resources will be available upon implementation (e.g., allele frequencies, analysis software, interpretation methods, training, and support for potential admissibility hearings). An overall assessment of 1) how a new technology works, 2) what its limitations are, and 3) how it might specifically address the problem to be solved (e.g., DNA mixture interpretation) is important and a key component of evaluating whether implementation will be worthwhile.

This chapter includes two key takeaways.

KEY TAKEAWAY #6.1: Fundamental measurement and interpretation issues surrounding DNA mixtures, as described in Chapter 2, should be understood before attempting to apply a new technology.

KEY TAKEAWAY #6.2: Implementation requires a thorough understanding of the benefits and limitations of the new technology as well as the practical investment of time and effort put forth for its adoption by the laboratory.

To read the entire report, which is available free of charge, see https://doi.org/10.6028/NIST.IR.8351-draft