Forensic DNA profiling

Introduction

DNA evidence made its courtroom debut in the case of Colin Pitchfork, who was convicted of murder in 1988. When it was first introduced, large volumes of material were required to obtain a result. However, as the realisation of the benefits of DNA testing gathered pace, so did the development of the methods used. By the time I started my career as a forensic reporting biologist in 2002 the SGM Plus DNA test was the routine test used in criminal casework and had already been around for 3 years. The capabilities of DNA testing were continuing to develop fast and the use of the National DNA Database as a source of assistance in identifying unknown individuals was routine. Throughout my career I have been trained in DNA processing methods and reporting results for presentation at court using a variety of DNA techniques. Looking back, it is impressive to see the range of developments; such as the increased sensitivity, the wider range of sample

By Dr Philip Avenell

types from which we can obtain a DNA profile and the wider investigative avenues available. It is also fair to say, with advances in Next Generation Sequencing (NGS) technology commonly seen in modern genetic research, we are about to go through another evolution in technology within the next 10 years or so.

In the same way that the DNA tests have advanced over the years, so has the interpretation and evaluation of the results obtained. With regards to whom the DNA detected could have originated, statistical evaluation has always been needed in order to assess how frequent, or common, a DNA profile is within a given population of individuals. However, the methods always became limited when the complexity of the result increased. That complexity could either be because there were multiple individuals who could have contributed to the sample, or there was very little DNA within the sample, commonly referred to as low template DNA. Advances in both statistical modelling and the necessary computer processing required to deal with the complex algorithms has led to several methods capable of dealing with these complexities.

However, despite these technological and statistical advances, the key area of interpreting the context of the DNA within the scenario in question remains of utmost importance. The principles behind these evaluations remain unchanged and will continue to do so. The increase in pressure for quicker and cheaper results means that this crucial area may be overlooked

or not addressed in initial reports (for example Short Format Reports). Vital questions include, '*How and when was the DNA deposited*?', '*Could there be an alternative explanation for why the DNA was detected here?', 'Just because their DNA is on that knife does that mean they handled or used it?*'. DNA detected using today's techniques is far more likely to identify the individual from which the DNA could have originated, however, does that mean they committed a crime? In this article we will address some questions of critical importance when considering the significance of a DNA result within the specific context of the circumstances and methods of testing which have been carried out.

How is DNA tested?

The aim of this article is not to make the reader a DNA profiling expert. It is, however, important that we first address what DNA profiling is and how it is undertaken. We cannot cover all eventualities in this article but will cover the most significant ones.

The DNA profiling process typically involves many steps (as shown in figure 1). The extent and specific technique applied in each of these steps can vary depending on the sample type or manufacturer's kit used. Some modern processes may even remove some of these steps, however the overall principles remain the same. Once a sample has been collected, the first step must be to extract the DNA or make it available from all the other material within the sample, for example unwanted biological material from the cells themselves, and any background material such as dirt or environmental chemicals, all of which may interfere with the DNA profiling technique.

Once the DNA is 'available' it is necessary to understand how much is present. This is to ensure that the right amount of DNA goes into the next stage of the process, too much or too little input DNA and the detection of DNA present in the sample becomes difficult or misleading.

The next stage is to target and copy the areas of DNA which are of interest in the DNA profiling technique. This process is often considered as similar to photocopying pages from a book, where we are concerned with only specific sections of the book and not copying and recreating the whole thing.

Finally, it is necessary to identify the DNA which is present, and this is effectively achieved by passing it through a sieve and determining the size of the copied sections of DNA, determined by how quickly they pass through.

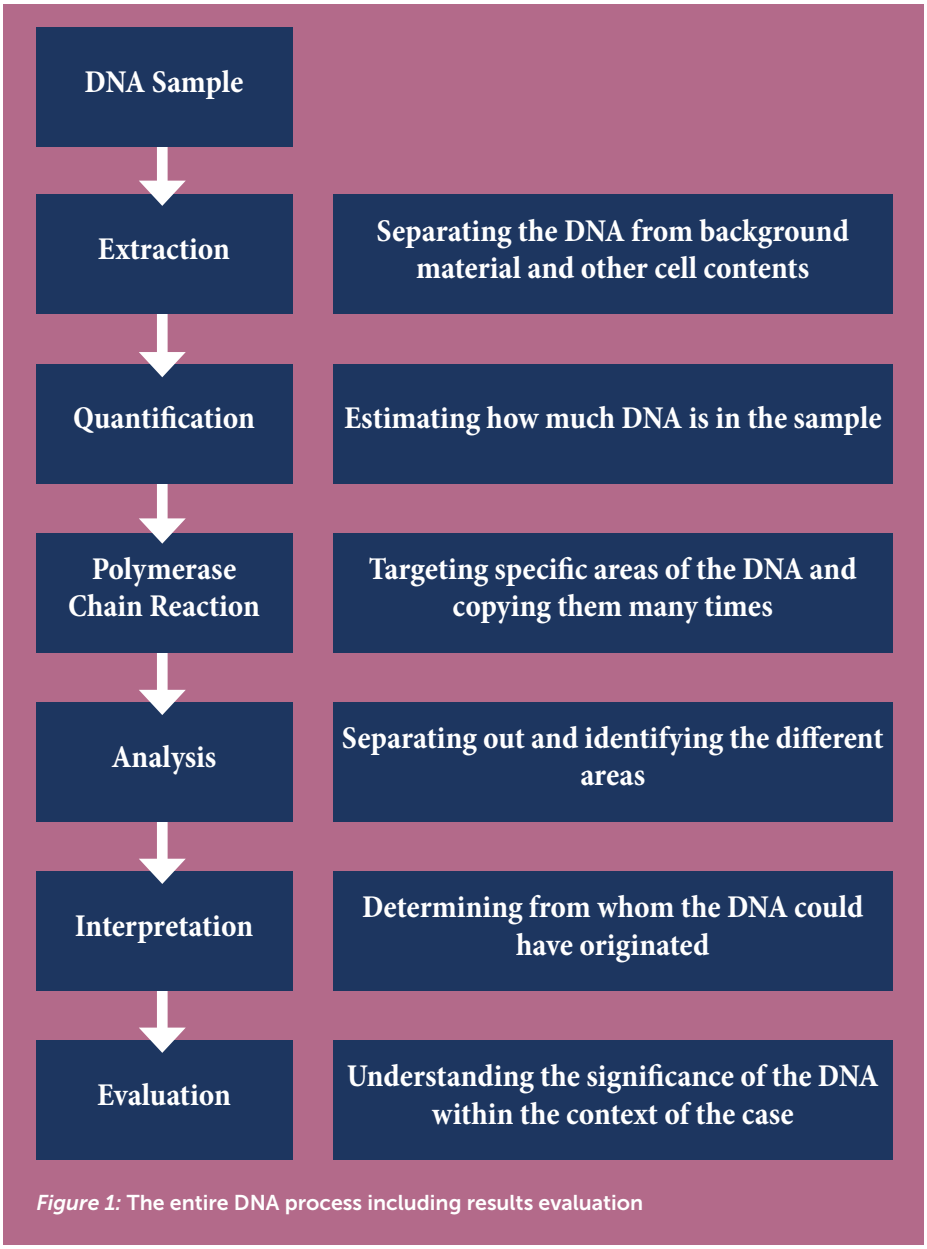

Figure 1: The entire DNA process including results evaluation

Considering the exact DNA process used and the results at each stage of the process can be critical in attributing detected DNA to a body fluid or evaluating the significance of how it was deposited. For example, the volumes of chemicals used in the various stages of the process may affect the actual amount of DNA detected in the original sample. This must be considered when comparing this to reference literature and assessing the likelihood of DNA transfer. Similarly knowing the amount of DNA can assist in determining whether a body fluid such as saliva was present in a sample rather than just large amounts of 'touch' DNA.

Current DNA profiling Techniques

The current DNA profiling technique used for routine crime stain testing is commonly called 'DNA 17'. This technique was introduced by the Home Office and forensic service providers in August 2014 and was a significant change from the previous technique known as SGM Plus. Firstly DNA 17 looks at 16 different areas of DNA plus Amelogenin, a further area which indicates if the sample was from a male or female, whereas SGM Plus only looked at 10 areas plus Amelogenin (importantly these 10 areas are present within the DNA 17 test and therefore results can be compared between these tests).

Increasing the number of areas targeted allows for higher discrimination when

considering the source of the DNA, however, for simplicity, the figure quoted within UK Court rooms for full profile matches remains unchanged. Another improvement is the significant increase in sensitivity of the kits, allowing DNA profiles to be obtained from samples which previously contained too little DNA for detection. In addition, they are also much more resistant to inhibitors and therefore more likely to give a DNA result from difficult or dirty samples.

Whilst these advantages have led to an increase in the numbers of DNA profiles being obtained, the increase in sensitivity has also led to an increase in detection of DNA from more than one person and as a result greater difficulty for the scientist in interpretation. A further consideration is also a greater risk in the potential for contamination – the inadvertent or accidental transfer of DNA during the storage, transfer or examination of an item or sample. Improvements in protocols and cleaning regimes have had to be introduced to reduce the risk of contamination as much as is practically possible, but this risk can never be eliminated and so another key aspect of results evaluation in the context of a scenario is understanding exactly how the sample was collected and processed.

Defining small amounts of DNA

The new tests have been significantly improved such that they are much more sensitive than routine SGM Plus. This has implications for the reporting scientist when considering factors such as attribution (*from what body fluid did the DNA originate?*), transfer (*was the DNA deposited as a result of primary, secondary, tertiary transfer?*) and persistence (*how long has the DNA been on this item?*). All of these could be significant factors when considered in the context of the case.

The issue of DNA sensitivity was one that the Courts attempted to address in the cases of Reed and Reed and Garmson in 2009 (EWCA Crim 2698). In these cases, there was a significant issue with low template DNA. The court attempted to define low template DNA as any amount lower than 200pg (picogram is 10^{-12} of a gram). There are approximately 6.5pg of DNA within a single cell and so this equated to approximately 30 cells. Any result with less than this amount of DNA was to be considered low template and the interpretation to be treated with greater caution. However, this approach was not widely accepted by scientists because there are many factors which may affect the accuracy of this measurement. A sample may generate a quant value in excess of 200pg, but due to degradation / poor quality of the DNA present the result may exhibit 'low template' characteristics. Similarly, if the sample contains DNA from more than 1 individual this result could be misleading as it is a measure of total DNA in a sample and cannot determine what contribution comes from different individuals in a sample. Therefore, scientists must take a much more rounded view when considering if a sample contains low template DNA and consider all of the process from the extraction technique through to the precise volumes used at each stage of the DNA profiling process as well as the final result itself.

> ***The aim of this article is not to make the reader a DNA profiling expert. It is, however, important that we first address what DNA profiling is and how it is undertaken.***

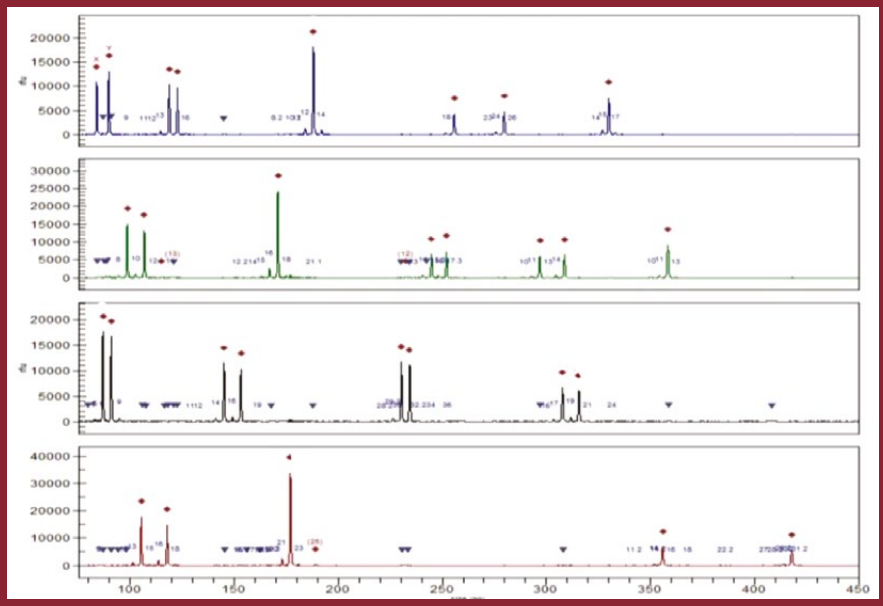

Figure 2: Example of a DNA 17 profile

Contamination

The risks of contamination are significantly increased as the sensitivity of the DNA profiling process increases and therefore careful consideration of all aspects of the sample and its route through DNA profiling are critical. Contamination could occur at the crime scene, in transit, in the examination process or in the DNA testing laboratory itself. The scientist must consider each of these areas as well as the possibility of person to item, item to item and transfer via an intermediate surface.

Laboratories go a long way to identify and eliminate contamination by utilising techniques such as elimination databases, which compare the results of all samples processed against staff DNA profiles and other samples processed within the same batch, and strict cleaning and exhibit handling protocols. Whilst these approaches are highly effective at minimising the risk of an undetected contamination event, they cannot eliminate it. It is worth noting that events which can cause contamination are rarely obvious and identification requires a wider knowledge of a case to identify them.

To illustrate this let us consider an example involving a sexual assault. An individual was charged with the assault which he strongly denied, but for which he was unable to provide an explanation for his innocence. The only information he could provide was that he had been arrested for an unrelated alleged offence around the same time as the reported sexual assault. On review of his case it was identified that a sample of carpet submitted to a forensic laboratory for examination in an unrelated case contained his semen (although this had never been tested). At the same time this item had been examined, the intimate samples from a victim of alleged rape were also being processed in the same laboratory. Due to the nature of the examinations in these cases it was identified that fibres from the carpet could have transferred via laboratory equipment such that they were detected on the intimate samples. Given that the semen on the carpet was not DNA tested this was not picked up by laboratory contamination checks and so went undetected until we reviewed the case notes.

Another issue we have encountered is that it is often believed by operative staff that sufficient anti-contamination measures are being taken, without an appreciation of the sensitivity of modern DNA testing methods. This might be, for example, when items are recovered from a scene and an individual is wearing protective gloves. The individual may be wearing suitable clean gloves and consider themselves to be taking anti-contamination measures, but they are only protecting themselves from depositing their DNA onto an item and not from transferring DNA between items, unless they regularly change or clean their gloves. A lack of regular glove changing, or cleaning means material picked up on the gloves whilst handling one item may then be transferred to another item. Careful consideration of continuity records and examination/

recovery notes may be required to ensure the risk of contamination between items has been minimised.

Statistical Evaluation

The evaluation of DNA results can be achieved using one of two methods, the match probability or likelihood ratio. The match probability evaluation is used when the components of interest can be unambiguously said to have originated from a single individual, whether this be a single source result or a 'major/minor' result. This method assesses the probability of a random individual in the given population having the same DNA profile as the one detected by chance. The fewer the number of DNA components detected, the more likely another individual in a defined population could also have that DNA profile.

> *...forensic DNA profiling has developed significantly in its ability to assist in solving criminal investigations. The capabilities of the technology have and continue to improve year on year.*

Recent guidance from the Forensic Science Regulator has led to the match probability approach only being applied to results where the scientist is confident that a complete profile from one person has been obtained.

When results fall outside this category the biggest question is '*why can't we apply a routine match probability statistic?*' This can be due to a number of reasons, including that the result is too weak to accurately determine whether a DNA component is truly present; there are too many contributors of DNA to the result; or, there is uncertainty as to the number of contributors. In these cases, the most appropriate methods to assess these results are those of the probabilistic models using a likelihood ratio. This approach compares the likelihood of obtaining the result given each of two competing hypotheses, one considered as the prosecution hypothesis and the other the defence. The result is presented as being X times more likely if one hypothesis is correct rather than the other. There are numerous probabilistic models available, such as LiRa, Resolve, STRmix™ and EuroForMix.

Finally, it is worth noting that despite the advances in software, not all results are suitable for statistical evaluation, meaning that there are occasions where a DNA finding with respect to an individual cannot be evaluated further. In these instances, the DNA findings are evidentially neutral, meaning they provide support for neither the prosecution, nor defence.

Understanding the significance of the results

The evaluation of the significance of a DNA finding within every case should consider a number of key questions. These address each stage of the DNA process, from the recovery of an item and collection of the DNA sample through to the evaluation of the result within the context of the case, to ensure appropriate consideration has been given to the accuracy of the results presented.

The key questions to address will depend on sample type, nature of the result and specific scenario being presented. For example, where there is only a small amount of DNA in a sample then it is key to understand factors such as whether any assumptions as to the number of contributors are accurate, or if it is even possible to accurately determine how many contributors there are. If there is any doubt, then this must be addressed appropriately as it can have a significant effect on the strength of the statistical evaluation which should be presented.

Understanding the possible significance of contamination may be a key factor. Whilst all laboratories put in place strict protocols and procedures to minimise any such risk, it is never possible to completely eliminate the possibility of contamination. Specific examples such as that described above and others have demonstrated that in certain specific situations it is right to question and challenge the possibility that contamination could be an explanation for the findings.

The final key area for consideration is that of transfer and persistence. Given the right circumstances it is entirely possible for DNA to be transferred from one surface to another. It is also possible, depending on circumstances, for this transferred DNA to be subsequently transferred and detected on a further surface, giving the impression that an individual has handled or come into contact with this surface even when in reality this has never been the case.

One final example surrounds an individual accused of having oral sex with a teenager in a bed. 'Saliva' was detected in the crotch of a pair of boxer shorts and a mixed DNA profile was obtained. This finding was claimed as positive support for the prosecution case. However, it is known that other substances can also give positive test results to the test for saliva. Because of the way the boxer short material was sampled it was not possible to tell if the 'saliva' was on the inside or outside of the shorts. The defendant had a chest infection and was constantly coughing and spluttering. Both admitted sharing a bed and so it was not possible to determine how the saliva had been deposited on the shorts.

Conclusion

In summary, forensic DNA profiling has developed significantly in its ability to assist in solving criminal investigations. The capabilities of the technology have and continue to improve year on year. The developments in statistical evaluation have also made the ability to robustly assess whether DNA could have come from an individual routine for most cases despite the strength or complexity of the result.

However, whilst the ability to detect and link DNA recovered from a sample to a specific individual has improved dramatically over the years it has, and always will, remain critical to step back from the actual DNA result itself and consider it within the overall context of the case. A statistical DNA match in the order of 1 in a billion can provide extremely strong support in one case, however, depending on case circumstances, it could be evidentially worthless in another. Issues regarding contamination, attribution, transfer and persistence will always be critical in assessing whether or not the DNA evidence can help address the actual events which took place.

FORENSIC ACCESS
Dr Philip Avenell is Managing Director at Forensic Access