

Chapter 2

The interpretation
of bioinformation



The interpretation of bioinformation

Introduction

2.1 This chapter examines fingerprinting and the science underlying DNA profiling, and considers how they are developing. Early problems with interpretation and reliability appear to have been overcome, but new procedures that push techniques to their limits raise questions about the accuracy, reliability and interpretation of these technologies. Widespread automation using computers and robots has lessened reliance upon humans, but with both fingerprints and DNA, human fallibility may still result in flaws even if the most robust scientific techniques are used.

Fingerprint comparison

2.2 The genesis of the 'science' of dactyloscopy – commonly known as 'fingerprinting' – has been well documented, if remaining disputed.¹ Although DNA is hailed as the 'gold standard' of identification techniques, the more humble fingerprint retains its status as the most commonly used method of identification and is a cornerstone of forensic crime scene investigation. Despite controversy over the standard required before a 'match' can be declared, fingerprint evidence generally remains unassailable in the criminal courts as a unique identifier. However, disputes can and do arise about methods of collection or negligent interpretation, and there have been several criminal cases where fingerprint evidence has proven flawed (see Box 2.1).

2.3 Finger skin is made of friction ridges, with pores (sweat glands). Friction ridges are created

Box 2.1: Fingerprint flaws?

- Mr Chiori was arrested and charged with burglary after a fingerprint expert claimed 'no doubt' about a match between Chiori's fingerprints with those at a crime scene, this being the only evidence. Two independent experts found the match to be wrong. At trial, the prosecution explained that their expert had made an error of judgment and the case against Chiori would not be pursued. The defence expert claimed that there was evidence of gross negligence and incompetence on the part of the Metropolitan Police Fingerprint Bureau (although without malicious intent). Chiori sought compensation at the European Court of Human Rights.² This was unsuccessful as it was ruled that police and fingerprint officers were immune from prosecution.
- In 1998, Scottish police officer Shirley McKie was tried and convicted of perjury for testifying that a fingerprint on a doorframe at a murder scene was not hers. Four fingerprint experts testified to a match with McKie, while two American experts gave contrary testimony. McKie's conviction was overturned. Controversy continues over how the erroneous 'match' got through the Scottish Criminal Records Office procedures. Shirley McKie recently received £750,000 compensation.³
- In the United States of America, an attorney from Oregon, Brandon Mayfield, was arrested for involvement in the Madrid train bombings after a fingerprint 'match' by the Federal Bureau of Investigation (FBI). The FBI reported a match against one of 20 fingerprint candidates returned by their automated fingerprint database. The FBI called the match '100 per cent positive' and 'absolutely incontrovertible'. The Spanish National Police examiners concluded the prints did not match Mayfield. The FBI subsequently acknowledged the error, and Mayfield was released after two weeks in custody. In January 2006, the US Justice Department found the misidentification was due to misapplication of methodology by examiners and ordered the FBI pay Mayfield US\$2 million in compensation.⁴
- Research by Dr Itiel Dror of the University of Southampton demonstrates that the context in which an examiner undertakes a fingerprint comparison can influence the expert's conclusions.⁵ He concludes that fingerprint experts can be influenced by the context in which they are asked to evaluate the evidence (particularly the emotional context) and that the same expert can come to different conclusions about the same fingerprint when asked to evaluate the match on successive occasions without the expert knowing of his or her earlier conclusion.

1. For example, see Beavan C (2003) *Fingerprints* (London: Fourth Estate) and Cole S (2001) *Suspect Identities* (Massachusetts: Harvard University Press).

2. *Chiori v UK* (1999) ECHR App. No. 44926/98.

3. See www.shirleymckie.com for details.

4. Eggen D (2006) US settles suit filed by Ore. Lawyer: \$2 million will be paid for wrongful arrest after Madrid attack *Washington Post*, 30 November, pA03.

5. Dror IE and Charlton D (2006) Why experts make errors *Journal of Forensic Identification* 56: 600–16.

during foetal development and only the general shape is genetically defined. Fingerprint patterns then vary from person to person (including identical twins) and no two people have been found with the same fingerprints. Fingerprints also remain unchanged throughout life. Identification by fingerprints relies on the matching of patterns and the detection of certain ridge characteristics, also known as Galton details (after Sir Francis Galton who originated the technique). These points of identity, or minutiae, and the comparison of the relative positions of these minutiae points with a reference print (either inked impression of a suspect's print or digital representation), are the basis of fingerprint comparison. Identification points consist of points such as bifurcations, ending ridges, and islands (see Figure 2.1). A single fingerprint may have as many as 100 or more identification points.

Figure 2.1: Image of a fingerprint with some ridge characteristics identified

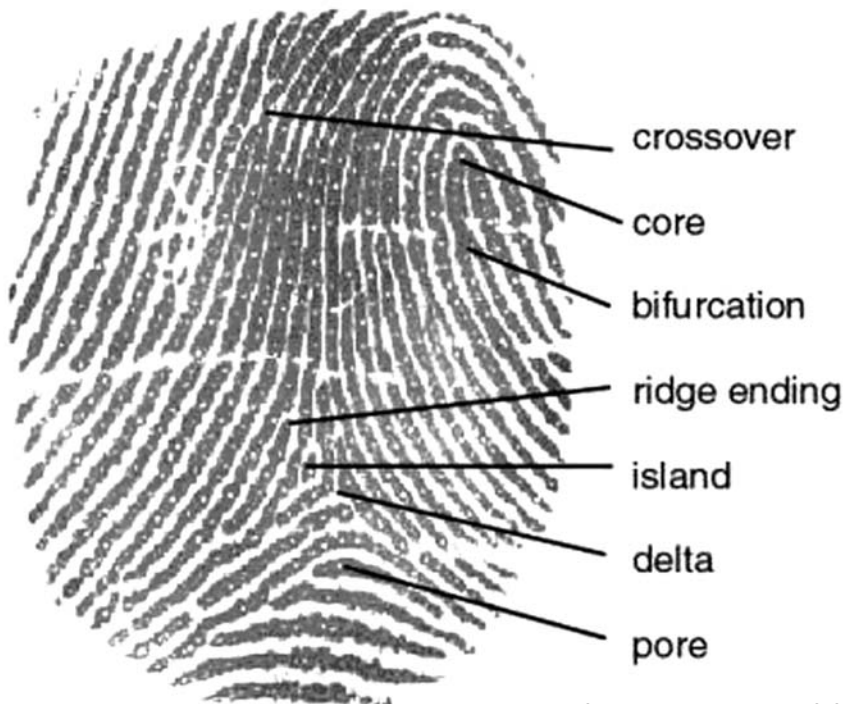


Image from <http://perso.orange.fr/fingerchip/biometrics/types/fingerprint.htm>

2.4 When fingerprints were first used, there were no set criteria to determine a 'match'. Practitioners originally argued that finding 12 similar ridge characteristics was the point at which to declare a match. However, in 1924, New Scotland Yard adopted a 16-point standard, and in 1953 it was agreed at a national policy level that this would become the common national approach. Other countries operate with 8, 10 or 12 points or have no numerical standard. In 1988, the Home Office and the Association of Chief Police Officers (ACPO) concluded that there was no scientific basis for the retention of the numerical standard. In 1995, a Home Office report stated that fingerprinting: "was not an exact science" but involved judgment.⁶ A target was set that by 2000, the numerical approach would be superseded. In 1999, this target was endorsed by *R v Buckley*,⁷ where the courts set down new guidance on the admission of fingerprint evidence (see Chapter 5). On 11 June 2001, national guidance was provided to all fingerprint bureaux by ACPO that the 16-point standard be replaced by a non-numerical standard which has 'objective criteria'.⁸ This guidance makes clear that whether or not a fingerprint 'match' has been found is a matter of opinion by a fingerprint expert, and cannot be presented as a fact.

6. Evett IW and Williams RL (1995) *A Review of the Sixteen Point Fingerprint Standard In England and Wales* (London: Home Office), p11.

7. *R v Buckley* (1999) 143 SJ LB 159.

8. Typically, the finding of various points of similarity in ridge patterns. Lord Rooker, House of Lords, Hansard, 25 February 2002, column 2701.

Forensic DNA profiling

How is DNA bioinformation obtained?

- 2.5 The police can obtain a biological sample (containing DNA) either directly from a suspect, or from a consenting victim or volunteer (a 'subject sample'), or from a crime scene or exhibit (for example, clothing, footwear or a weapon). Bar-coded samples are received by the processing laboratory and put through a standard process to yield a few drops of water containing DNA in a bar-coded plastic tube. Any of the physical biological sample and extracted DNA that is not used up in the laboratory is stored indefinitely (the 'biological sample').
- 2.6 Subject samples are normally obtained by brushing the inside of a person's cheek with an instrument like a small toothbrush (a 'buccal swab'). This should yield an ample quantity of good quality DNA. Crime scene stains are much more variable. Bloodstains are the best source, although very small spots (1 mm or less in size) can be difficult. Other frequently used sources include cigarette ends, chewing gum, saliva obtained from drinking vessels, scarves, balaclavas, etc., hair and semen (Table 2.1 shows the proportions and success rates in obtaining a profile suitable for loading on to the NDNAD from various types of crime scene stains). Hard surfaces with which the offender is thought to have had skin contact can be swabbed and may yield minute amounts of DNA even when no stain is visible.

Table 2.1: The proportions and success rates in obtaining a profile suitable for loading on to the NDNAD from various types of crime scene stains, July–September 2005.

Relative recovery and profiling rates for DNA sources, July–September 2005

DNA Group	Percentage of Total Samples Processed	Percentage of Samples for Loading on NDNADB
Blood	26.1	86.1
Cigarette ends	24.0	73.5
Saliva	27.1	37.2
Chewing gum	0.9	71.2
Hair	1.1	18.1
Semen	5.2	92.1
Cellular	15.6	12.4

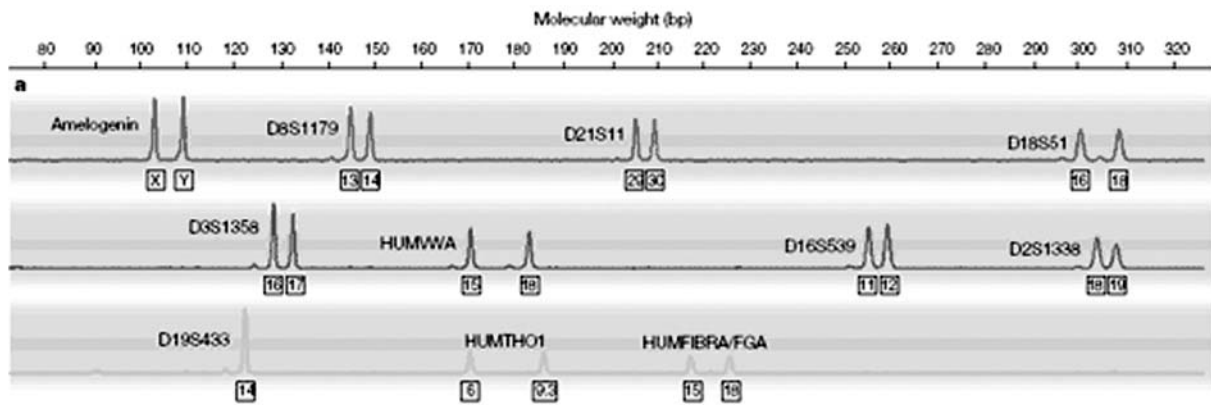
Source: Bond JW (2007) Value of DNA evidence in detecting crime *Journal of Forensic Sciences* 52: 128–36. Reprinted with permission from Blackwell Publishing.

- 2.7 A small portion of the extracted DNA is used to obtain a DNA profile. A standard laboratory technique (the polymerase chain reaction, or PCR) is used to make millions of copies of specific parts of the original DNA, the 'markers'. These markers consist of repeated short sequences of DNA that vary in length between different people. The current standard profiling technique in the United Kingdom, SGM+, uses ten markers of a type called short tandem repeats (STRs) (see Box 1.1 and Appendix 4 for further details).
- 2.8 Each cell of a person contains two copies (alleles) of each marker, one from the father and one from the mother. These differ in various ways between individuals. The PCR product contains millions of copies of each marker in the SGM+ set. This fluid is put into a machine that plots the length of each marker sequence in the form of peaks on a graph (Figure 2.2). The graph is reviewed by a scientist to check its quality, completeness and any ambiguity. The profile is then entered on to the NDNAD. The profile stored on the NDNAD consists of a series of 20 numbers,

recording the size of each allele of each marker, plus a sex indicator, together with other information such as name and date of birth (see Box 1.3).

Figure 2.2: A DNA profile as obtained in the laboratory

The figure shows an electropherogram of the SGM+ profile from a male. The numbers beneath the peaks indicate the allele sizes in repeat units. The profile stored on the NDNAD consists of only these numbers. The diagram is included for illustrative purposes only.



Source: Jobling MA and Gill P (2004) Encoded evidence: DNA in forensic science *Nature Reviews Genetics* 5: 739–51. Reprinted by permission from Macmillan Publishers Ltd: copyright 2004.

How has the science of DNA profiling developed?

- 2.9 The 'DNA fingerprinting' method was developed by Sir Alec Jeffreys in 1984. The original DNA fingerprints consisted of a pattern of bands rather like a bar-code. For judging a match, bands were assigned to arbitrarily defined 'bins'. Each bin either contained a band or did not contain a band. The crime scene and subject samples were compared for the presence or absence of a band in each bin. The rarity, or match probability, was calculated from a database showing the probability that a random biological sample would have a band in any particular bin. This method, involving expert judgment of arbitrary patterns, is analogous to the way fingerprints are classified and compared.
- 2.10 This method has now been superseded by a quicker and more sensitive laboratory technique which is applicable to even very small and relatively degraded biological samples. Using 'STR profiles' allows a much more analytical and theoretically grounded approach to matching. Each marker can be assigned to a specific STR and characterised by its precise size. Sizes are not continuously variable, but consist of integer numbers of DNA units. Databases show the frequency of each size of allele of each STR in a variety of populations.
- 2.11 The match probability of a profile is the probability that an unrelated individual would have the same profile by random chance. It is calculated by multiplying together the chance of a match for each individual peak on the graph (see Appendix 4 for a discussion of this procedure). It depends on how common or rare each particular variant is in the population. For SGM+, match probability is normally much less than 1 in 1 billion (1,000,000,000). However, the chance of a match with a relative is much higher. Identical twins have identical profiles.
- 2.12 If requested, laboratories will 'type' a sample for a series of additional markers. This could be with the purpose of increasing the discriminating power of a profile or as an aid to familial searching (see paragraphs 2.15–2.16). Some of the most common additional analyses are detailed in Box 2.2.

Box 2.2: Additional DNA analysis techniques

Y-chromosome markers (Y-short tandem repeats (Y-STRs) or Y-single nucleotide polymorphisms (Y-SNPs)): The Y chromosome is present only in males. A man's specific combination of sequences on the Y chromosome is normally inherited unchanged from his father. The combination ('haplotype') can help with identification, and can provide probabilistic information about the geographical origin of somebody's male ancestors.

Variants in mitochondrial DNA (mtDNA): Unlike Y-chromosome variants, this small component of the DNA (which is not in the chromosomes but contained in a distinct component of each cell) is present in people of both sexes, but is inherited exclusively from the mother. Again, the particular combination of variants in mitochondrial DNA is normally inherited unchanged from an individual's mother, and may assist with identification or provide probabilistic information about ancestry, this time of the maternal ancestors. Additionally, because a single cell contains many copies of the mtDNA, this DNA can sometimes be analysed in samples that are too small or degraded to allow standard profiling. Y-chromosome and mtDNA data do not fit the template of standard NDNAD records, and we have not been able to establish whether, and if so where and how, this information is databased.

Low copy number (LCN) analysis: The standard profiling technique is designed to be reliable for amounts of DNA equivalent to the content of 20 cells. The LCN method aims to produce SGM+ profiles from still smaller amounts of DNA.⁹ The standard SGM+ PCR reaction, which produces many copies of each marker allele, is run for longer with the aim of producing analysable numbers of copies from less starting material. LCN has allowed some high-profile 'cold' cases to be solved but the results are much more open to interpretation than profiles obtained on larger samples (see paragraphs 2.30–2.31). Usually several independent replicates of the LCN reaction are performed, and alleles are only reported if they appear in more than one replicate.

Pendulum: One of several computer programs for assisting the interpretation of mixed samples (from more than one person). Pendulum works by considering all possible solutions, identifying the most self-consistent ones and calculating the statistical support for each possible interpretation. It does not perform the interpretation, but helps the reporting officer form an expert opinion.

2.13 Crime scene stains are often much more difficult to analyse than subject samples. This is for one or more of three reasons:

- They may contain the DNA of several people (mixed samples).
- They may yield extremely small amounts of DNA. This challenges the technical limits of the PCR process, and increases the problems caused by possible contamination with extraneous DNA.
- The DNA may be fragmented into small pieces ('degraded') as a result of exposure to environmental agents. This can cause difficulties with the PCR process. Such difficulties may result in the profiling process producing a 'partial profile', that is a profile without the full set of ten markers.

2.14 Technical developments to assist with these problems include computer programs to help distinguish the individual profiles in a mixed sample, variants of the PCR process to analyse very small samples (e.g. the DNA low copy number (LCN) method used by the Forensic Science Service) and special markers (mini-STRs) that can be typed from highly degraded samples. As discussed below, interpretation of results obtained from such samples can be controversial.

What other uses may be made of DNA?

Familial searching

2.15 In difficult criminal investigations where a suspect cannot be identified, the police may request that 'near misses' – where there may 'nearly' be a match between DNA profiles – are further explored. If a crime scene profile does not completely match any subject profile on the NDNAD, it is possible to see if any subject profile could be from a relative of the person who deposited the crime scene sample. Familial searching is the process of comparing a DNA profile from a crime scene with subject profiles stored on the NDNAD, and prioritising them in terms of

9. The Forensic Science Service (2005) *DNA Low Copy Number*, available at: www.forensic.gov.uk/forensic_t/inside/news/fact_sheets.htm, accessed on: 2 July 2007.

'closeness' to a match. Parents and children always share one allele of each marker. For the ten SGM+ markers, there would be 1,024 possible combinations of shared alleles. For siblings (brothers and sisters) a search is simply made for profiles that share an unexpectedly large proportion of sequences, particularly very unusual sequences, with the crime scene sample. Unrelated individuals share on average six of the 20 SGM+ alleles in a profile, just by chance, whereas siblings share on average 13–14 alleles, but in each case there is a wide distribution around these averages.¹⁰ Currently a search of the NDNAD for profiles that share alleles with a crime scene sample might identify ten profiles that share 16 alleles, 20–30 profiles that share 15, several hundred that share 14, and several thousand that share 13 alleles.

- 2.16 The size of this pool must be limited geographically and/or by other means to produce a useful short-list. One way of reducing the list would be to 'type' both the crime scene stain and all retained subject samples on the shortlist for mitochondrial and/or Y-chromosome variants (depending on the type of relationship assumed). Eventually a person identified through familial searching would be asked to provide a biological sample, and would be identified as a suspect or eliminated from the enquiry depending whether this sample then matched the crime scene profile. The familial search results would not form part of the evidence presented in any prosecution. Nevertheless, there was considerable concern among respondents to our consultation about the potential of familial searching to be intrusive or improperly reveal family secrets. This is addressed in Chapter 6.

Ethnic inferences

- 2.17 All alleles of the SGM+ markers can be found in all ethnic groups, but particular alleles are more or less frequent in different groups. The particular combination of alleles in an SGM+ profile can be used to assess the most likely ethnic origin of the donor. These calculations are based upon research that was completed by the Forensic Science Service, using a sample of DNA profiles taken from the NDNAD. Such an 'ethnic inference' is occasionally made from a crime scene stain in the hope of narrowing the pool of suspects (i.e. to indicate that the suspect is likely to have a particular ethnic appearance). A more detailed inference could be made by analysing the actual DNA rather than the profile, using markers chosen because the allele frequency varies considerably between groups ('ancestry informative markers'). Any ethnic inference may be more or less specific, but it is unlikely ever to be unambiguous. Global mobility and ethnic mixing limit the value of such inferences. In the light of the social factors and policing practices that lead to a disproportionate number of people from black and ethnic minority groups being stopped, searched and arrested by the police, and hence having their DNA profiles recorded on the NDNAD, there are concerns that inferring ethnic identity from biological samples risks reinforcing racist views of propensity to criminality. These issues are discussed in Chapter 6.

A DNA photofit?

- 2.18 In principle, other physical characteristics might be predictable by more extensive analysis of the biological sample (but not of the DNA profile). The striking physical similarity of identical twins demonstrates that a person's DNA sequence has a very significant influence upon their general physical appearance. Given sufficient scientific knowledge, and a sufficiently complete DNA analysis, a photo-fit picture of the suspect could in theory be generated from a crime scene sample.¹¹ This is not currently anywhere near possible. Most human characteristics, even if they are partly determined by our DNA sequence, depend on the combined effects of many variables, each of which has only a small effect. It has proved extremely difficult to predict

10. Information provided by Richard Pinchin, Forensic Intelligence Bureau, 16 February 2007.

11. Although this would also rely on assumptions being made about the environment in which a person has been brought up (nutrition during childhood, childhood diseases, etc.).

observable human characteristics from analysis of DNA, except for certain genetic diseases that affect only a small proportion of the population. Research is continuing, including into the possibilities of deriving useful descriptive information from crime scene samples. Past experience suggests that progress is likely to be slow and unspectacular. The following deductions have been widely discussed:

- **Hair colour:** The FSS claims to be able to identify 84 per cent of redheads by analysis of the sequence of the *MC1R* gene. Other hair colours are not readily predictable. Of course, any such test simply provides information on genetic hair colour - it could not tell if somebody is bald or has dyed hair.
- **Eye colour:** Several genes have been identified that contribute to determining eye colour, which may allow limited probabilistic predictions.¹²
- **Skin colour:** This is known to be determined by a series of different genes, some of which have been specifically identified. The ethnic inferences described above would give some suggestions about possible skin colour.

Surnames

2.19 Since a man's Y chromosome and, often in the United Kingdom, his surname are both inherited down the male line, there is some correlation between Y-chromosome marker haplotypes and surnames of individuals born in the UK. For uncommon surnames the correlation might be sufficiently tight to help narrow a pool of suspects. The accuracy of surname inferences is untested in practice, but has been supported in a pilot study, at least for unusual surnames.¹³ Such inferences would of course fail if a person's surname did not reflect his true biological paternity as in the case of children adopted, not given their father's surname or not the biological offspring of their 'social' father.¹⁴

Health-related information

2.20 It is very unlikely that present forensic analysis of crime scene stains could yield health-related information about the donor, nor is it obvious what purpose would be served by attempting to obtain such information. The DNA profiles stored on the NDNAD do not predict any characteristics of a person, apart from their sex. We note, however, that the sex test that is part of the standard SGM+ profile will reveal rare individuals whose chromosomal sex does not match their physical sex or social gender, while testing Y-chromosome markers has the potential to inadvertently reveal certain deletions that can cause infertility in rare individuals. Either test might reveal sex chromosome abnormalities. Such inadvertent discoveries are perhaps analogous to the occasional inadvertent discovery of wrong paternity in the course of clinical genetic testing. General practice among clinical geneticists in the United Kingdom is not to inform the subject.

2.21 Any deliberate attempt to uncover health-related data would be contrary to the Council of Europe's Recommendation on the Protection of Medical Data, Principle 4.8 of which states for forensic DNA analysis that: "The data should only be used to establish whether there is a genetic link in the framework of adducing evidence, to prevent a real danger or to suppress a specific criminal offence. In no case should they be used to determine other characteristics which may be linked genetically."¹⁵ (See Appendix 3.)

12. Duffy DL, Montgomery GW, Chen W et al. (2007) A three-single-nucleotide polymorphism haplotype in intron 1 of OCA2 explains most human eye-color variation *American Journal of Human Genetics* **80**: 241–52.

13. King TE, Ballereau SJ, Schürer K and Jobling MA (2006) Genetic signatures of coancestry within surnames *Current Biology* **16**: 384–8.

14. Jobling M (2001) In the name of the father: surnames and genetics *Trends in Genetics* **17**: 353–7.

15. Council of Europe (1997) Recommendation R (97) 5 on the Protection of Medical Data.

Future scientific developments

- 2.22 Whether there is any benefit in adding more markers to the standard SGM+ set is debatable. Further markers can always be analysed in particular cases where more precision is needed. The current panel of markers for routine profiling will not be readily abandoned because that would necessitate re-analysing all samples in the NDNAD. However, additions to the SGM+ marker set may be driven by harmonisation of marker sets across Europe and the introduction into service of 'mini-STRs' that give more reliable results with badly degraded DNA.
- 2.23 Eventually STR markers might be superseded by a different type of marker, single nucleotide polymorphisms (SNPs). Technology now exists for typing thousands of SNPs from a biological sample in a single automated operation and SNP analysis is widely predicted to have an important medical role in predicting disease susceptibility, aiding accurate diagnosis, and drug therapy. Such analyses may, in the future, enable more accurate prediction of hair colour, eye colour, ethnicity, etc. from crime scene samples. Furthermore, SNPs would be better than the current SGM+ markers for analysing degraded DNA. Currently such analyses require more DNA than would be yielded by many crime scene stains. In any event, a recent study suggests that SNPs will not replace STRs in national DNA databases in the foreseeable future.¹⁶
- 2.24 Experts generally argue that the most promising way of increasing the number of crimes solved by DNA profiling is to focus on obtaining DNA from more crime scenes, rather than modifying current laboratory methods. Trials of mobile DNA analysis laboratories (a Forensic Response Vehicle or 'FRV') have demonstrated the feasibility of obtaining profiles at a crime scene, which could speed up investigations. This might raise issues of the security of the NDNAD, because such a development might require the Database to be remotely accessible using wireless technology which may be easier for non-authorised persons to access.

The interpretation of DNA

- 2.25 Having described the technical basis of forensic DNA profiling, we note that there still remain issues that a court might raise when presented with DNA evidence. With the original DNA 'fingerprinting' method, there was room for debate over a DNA 'match'. Complete SGM+ profiles are much less ambiguous, although certain features may still require interpretation. Stutter bands – small extra peaks on the graph that are an artefact of the laboratory PCR process – are well recognised and should not normally pose a problem. However, real problems remain with the interpretation of evidence from mixed samples and partial profiles, as detailed below. In addition, a mobile object, such as a cigarette butt, may have originally been dropped somewhere other than where it was found. A forensically sophisticated criminal or a corrupt investigator might deliberately contaminate the crime scene with an innocent person's DNA.

Mixed samples

- 2.26 Mixed samples are samples that contain DNA from more than one person. Such samples always require interpretation, and a court will need to be satisfied of the validity of the analysis (see Box 2.2). Where two people's DNA is mixed in a 3:1 or greater ratio, the differing peak sizes allow the individual profiles to be disentangled fairly unambiguously (although identification of a very minor component may be uncertain). Interpretation of more equal mixtures depends heavily on the expert opinion of the reporting officer. Profiles provided by known innocent bystanders can be subtracted from the mixed profile to identify peaks of unknown origin. In rape cases, special techniques¹⁷ may identify a male-specific profile from a vaginal swab.

16. Gill P, Werrett DJ, Budowle B and Guerrieri R (2004) An assessment of whether SNPs will replace STRs in national DNA databases *Science and Justice* 44: 51–5.

17. For example 'laser capture microdissection' or 'Y-specific amplification'.

Partial profiles

2.27 The interpretation of partial profiles or degraded biological samples are of concern. These are profiles that show fewer than the standard 20 peaks of a full SGM+ profile. Sometimes this is because the two marker alleles inherited by an individual from each parent happen to be of identical length.¹⁸ This should be readily apparent from the doubled height of the peak in the profile. Often, however, the cause is a failure of the PCR reaction in the laboratory to work with every allele of every STR ('allele drop-out'). Because there are fewer alleles on which to base a match, the match probabilities are less decisive than for a full profile. Partial profiles require much more skill in interpretation and should always be examined in detail by both DNA experts and ultimately a court.

Contamination

2.28 Contamination, whether deliberate or accidental is another major issue, especially with very small samples. Samples can easily be contaminated with DNA from one of the police or laboratory team if strict preventive measures are not taken. Elimination databases are maintained for the police, for those technicians that visit crime scenes, and those who visit forensic laboratories or who may be involved in the manufacture of items used in the DNA analysis process. These hold the profiles of potential 'innocent donors' of DNA, and hence enable their DNA profiles to be excluded from the investigation. The Police Elimination Database (PED) covers the police and Scenes of Crimes Officers (SOCOs), and the Manufacturers Elimination Database (MED) has the profiles of those who produce the equipment used in DNA analysis, as it was found that DNA profiles were being produced of the people who had been involved in the manufacture of the swabbing kits. Each private company also has databases of its staff.

2.29 These elimination databases ideally contain the profiles of all the people who might contaminate a sample. In theory this should allow such contamination to be identified and the contaminating profile to be 'subtracted' from the crime scene profile. However, many individuals are involved in the transfer of DNA from a crime scene through the process of collection, storage, transport and laboratory analysis. There is also the possibility of transfer of DNA from one crime scene sample to another, or indeed from the biological samples from the crime scene to clothing or other material taken from a suspect. Potential contamination can only be avoided by meticulous control of every aspect of sample collection and processing. It also requires adherence to strict custody controls of exhibits, so that crime scene exhibits cannot come into contact with a suspect at any point.

Very small samples

2.30 While DNA technology has advanced to be able to analyse smaller and smaller samples (see Box 2.2), very small samples present most of the problems mentioned above in acute form. Both allele drop-out and allele drop-in (extra artefactual peaks) can occur, and the risk of contamination is high. Computer programs for identifying individual profiles in mixed samples may be ineffective with very small samples. Selective reporting is also an issue, as laboratories may repeat the PCR reaction several times, and only report those peaks that are seen in at least two replicate runs. There are issues about how such analyses should be reported in court (see Chapter 5).

2.31 However far the science may develop in the future, its current uses are largely restricted to indicating the presence of a person's biological traces at one or several locations at a crime scene. Any determination of other relevant factors, such as how the biological material came to be left there, the time and duration of the person's presence, or the sequence of actions that took place, will need to draw on other kinds of evidence to be taken into account by investigators, advocates, judges and triers of fact (see paragraphs 5.23–5.26).

18. The person is 'homozygous' for that allele.

Conclusion

2.32 Generally speaking, the science and technology of DNA profiling is increasingly robust and reliable. However, there are several continuing risks, especially when the science and the technologies are pushed to their limits. In particular there are dangers of deliberate or accidental contamination, misinterpretation of mixed samples, and mistaken interpretation of partial profiles (and the potential for the misuse of statistics, which we consider in Chapter 5) in establishing the probability of a match. Our recommendations in the following chapters are designed to reduce the risks of mistaken identification and wrongful conviction resulting from (relatively rare) cases of flawed interpretations of the information provided by DNA profiles.