



# DNA profiling

UNIT 2

*European Initiative for Biotechnology Education*

---

## **Contributors to this Unit**

Dorte Hammelev (Unit Co-ordinator)

Dean Madden, Søren Nørby, Jill Turner



**The European Initiative for Biotechnology Education (EIBE) seeks to promote skills, enhance understanding and facilitate informed public debate through improved biotechnology education in schools and colleges throughout the European Union (EU).**

## EIBE Contacts



### AUSTRIA

| Rainhart Berner, Höhere Bundeslehr- und Versuchsanstalt für Chemische Industrie Wien, Abt. für Biochemie, Biotechnologie und Gentechnik, Rosensteingasse 79, A-1170 WIEN.



### BELGIUM

| Vic Damen / Marleen Van Strydonck, R&D Groep VEO, Afdeling Didactiek en Kritiek, Universiteit Antwerpen, Universiteitsplein 1, B-2610 WILRIJK.



### DENMARK

| Dorte Hammelev, Biotechnology Education Group, Foreningen af Danske Biologer, Sønderengen 20, DK-2860 SØBORG.  
| Lisbet Marcussen, Biotechnology Education Group, Foreningen af Danske Biologer, Lindevej 21, DK-5800 NYBORG.



### EIRE

| Catherine Adley / Cecily Leonard, University of Limerick, LIMERICK.



### FRANCE

| Gérard Coutouly, LEGTP Jean Rostand, 18 Boulevard de la Victoire, F-67084 STRASBOURG Cedex.  
| Laurence Simonneaux / Jean-Baptiste Puel, Ecole Nationale de Formation Agronomique, Toulouse-Auzeville, Boîte Postale 87, F-31326 CASTANET TOLOSAN Cedex.



### GERMANY

| Horst Bayrhuber / Eckhard R. Lucius / Regina Rojek / Ute Harms / Angela Kroß, Institut für die Pädagogik der Naturwissenschaften an der Universität Kiel, Olshausenstraße 62, D-24098 KIEL.  
| Ognian Serafimov, UNESCO-INCS, c/o Jörg-Zürn-Gewerbeschule, Rauensteinstraße 17, D-88662 ÜBERLINGEN.  
| Eberhard Todt, Fachbereich Psychologie, Universität Gießen, Otto-Behaghel-Straße 10, D-35394 GIEBEN.



### ITALY

| Antonio Bargellesi-Severi / Alessandra Corda Mannino / Stefania Uccelli, Centro di Biotecnologie Avanzate, Largo Rosanna Benzi 10, I-16132 GENOVA.



### LUXEMBOURG

| John Watson, Ecole Européenne de Luxembourg, Département de Biologie, 23 Boulevard Konrad Adenauer, L-1115 LUXEMBOURG.



### THE NETHERLANDS

| David Bennett, Cambridge Biomedical Consultants, Schuytstraat 12, NL-2517 XE DEN HAAG.  
| Fred Brinkman, Hogeschool Holland, Academy for Communication, Postbus 261, NL-1110 AG DIEMEN.  
| Liesbeth van de Grint / Jan Frings, Hogeschool van Utrecht, Educatie Centrum voor Biotechnologie, FEO, Afdeling Exacte Vakken, Biologie, Postbus 14007, NL-3508 SB UTRECHT.



### SPAIN

| María Sáez Brezmes / Angela Gómez-Niño / Rosa M. Villamañán, Facultad de Educación, Universidad de Valladolid, Geologo Hernández Pacheco 1, ES-47014 VALLADOLID.



### SWEDEN

| Margareta Johansson, Föreningen Gensyn, PO Box 37, S-26881 SVALÖV.  
| Elisabeth Strömberg, Östrabo Gymnasiet, S-45181 UDDEVALLA.



### THE UNITED KINGDOM

| Wilbert Garvin, Northern Ireland Centre for School Biosciences, NIESU, School of Education, The Queen's University of Belfast, BELFAST, BT7 1NN.  
| John Grainger / John Schollar / Caroline Shearer, National Centre for Biotechnology Education, The University of Reading, PO Box 228, Whiteknights, READING, RG6 6AJ.  
| Jill Turner, Department of Science and Technology Studies, University College London, Gower Street, LONDON, WC1 6BT.  
| Paul Wymer, Society for General Microbiology, Marlborough House, Basingstoke Road, READING RG7 1AE.

## EIBE Co-ordinator

Horst Bayrhuber, Institut für die Pädagogik der Naturwissenschaften an der Universität Kiel, Olshausenstraße 62, D-24098 KIEL, Germany. Telephone: + 49 (0) 431 880 3166 (EIBE Secretary: Regina Rojek). Facsimile: + 49 (0) 431 880 3132.



# DNA profiling

# UNIT 2

*European Initiative for Biotechnology Education*

MATERIALS

## Contents

★ ★ ★ ★ ★ ★ ★ ★ ★ ★ ★ ★ ★ ★ ★ ★

<b>I Development team, Copyright, Safety</b>	4
<b>I DNA profiling</b>	
The background	6
The technology	10
<b>I Classical profiling</b>	12
Examples	14
<b>I Modern profiling</b>	16
<b>I Reliability</b>	20
<b>I Discussion</b>	24
<b>I References</b>	25
<b>I Practical activity</b>	
Fingerprint simulation	26

Edition 2/2 completed January 1998

## World Wide Web

★ ★ ★ ★ ★ ★ ★ ★ ★ ★ ★ ★ ★ ★ ★ ★

Few areas are developing as rapidly as biotechnology. So that they can be revised and kept up-to-date then distributed at minimum cost, the EIBE Units are published electronically.

These pages (and the other EIBE Units) are available throughout Europe and the rest of the world on the World Wide Web. They can be found at:

<http://www.eibe.reading.ac.uk:8001/>

All of the EIBE Units on the World Wide Web are Portable Document Format (PDF) files. This means that the high-quality illustrations, colour, typefaces and layout of these documents will be maintained, whatever computer you have (Macintosh - including Power PC, Windows, DOS or Unix platforms).

PDF files are also smaller than the files from which they were created, so that it will take less time to download documents. However, to view the EIBE Units you will need a suitable copy of the Adobe Acrobat® Reader programme.

The Acrobat® Reader 3 programme is available free-of-charge. It can be downloaded from:

<http://www.adobe.com/>

With this software, you can view or print the EIBE Units. In addition, you will be able to 'navigate' around and search the documents with ease.

PLEASE NOTE: Adobe and Acrobat are trademarks of Adobe Systems Incorporated, which may be registered in certain jurisdictions. Macintosh is a registered trademark of Apple Computer Incorporated.

## EIBE contributors

- Dorte Hammelev (Unit Co-ordinator)  
Frederiksberg HF kursus  
Copenhagen, Denmark
- Dean Madden  
NCBE, The University of Reading  
U.K.
- Søren Nørby  
Lab of Biological Anthropology  
Institute of Forensic Medicine  
Copenhagen University, Denmark
- Jill Turner  
Dept. of Science and Technology  
Studies, The University of London  
U.K.

Design, illustration and typesetting: Dean Madden and Caroline Shearer, NCBE, The University of Reading, RG6 6AJ

## Acknowledgements

We are very grateful to Niels Morling, Hanna Hansen and Birthe Eriksen, all of the Department of Forensic Genetics, University of Copenhagen, especially for the provision of photographs and research data.

The practical activity in this unit was developed by John Schollar at the NCBE.

Dorte Hammelev (Frederiksberg HF Kursus, København, Denmark), Wilbert Garvin (Northern Ireland Centre for School Biosciences, The Queen's University of Belfast, The United Kingdom), John Schollar (National Centre for Biotechnology Education, The University of Reading, The United Kingdom) and Jill Turner (University College, London, The United Kingdom) arranged and ran a multinational workshop in which the materials in this Unit were tested. EIBE

would like to thank them and the teachers from Denmark, Eire and Germany who took part and gave many helpful comments on the draft materials.

The workshop participants were:

From Denmark: Lisbet Leonard; Lene Tidemann; Mario Bro Hassenfeldt; Greta Grønqvist; Jytte Jørgensen; Tine Bing; Per Vollmond; Anker Steffensen.

From Eire: John Lucey; Michael O'Leary; Bruno Mulcahy; Tim O'Meara; Tom Moloney; Brendan Worsfold; Frank Killelea.

From Germany: Ulrike Schnack; Werner Bährs; Jürgen Samland; Cristel Ahlf-Christiani; Erhard Lipkow; Hubert Thoma.

From the EIBE team: Eckhard R. Lucius; Catherine Adley; Jan Frings; Wilbert Garvin; Jill Turner; Dean Madden; John Schollar; Dorte Hammelev.

## © Copyright

This EIBE Unit is copyright. The contributors to this Unit have asserted their moral right to be identified as copyright holders under Section 77 of the Copyright, Designs and Patents Act, UK (1988).

Educational use. Electronic or paper copies of this EIBE Unit, or individual pages from it may be made for classroom use, provided that the copies are distributed free-of-charge or at the cost of reproduction, and the contributors to the unit are credited and identified as the copyright holders.

Other uses. The Unit may be distributed by individuals to individuals for *non-commercial* purposes, but not by means of electronic distribution lists, mailing (listserv) lists, newsgroups, bulletin board or unauthorised World Wide Web postings, or other bulk distribution, access or reproduction mechanisms that substitute for a subscription or authorised individual

access, or in any manner that is not an attempt in good faith to comply with these restrictions.

Commercial use. Should you wish to use this material in whole or in part for commercial purposes, or to republish it in any form, you should contact:

EIBE Secretariat  
c/o Institut für die Pädagogik  
der Naturwissenschaften  
Universität Kiel  
Olshausenstraße 62  
D-24098 Kiel  
Germany

Telephone: + 49 431 880 3166  
Facsimile: + 49 431 880 3132  
E-Mail: rojek@ipn.uni-kiel.de

## About this Unit

These materials have been devised by practising teachers and educationalists from several European countries, brought together with financial support and encouragement from DGXII of the European Commission, under the auspices of EIBE, the European Initiative for Biotechnology Education.

The EIBE materials have been extensively tested in workshops involving teachers from across Europe.

The views expressed in this Unit and the activities suggested herein are those of the authors and not of the European Commission.

Particular attention should be paid to the general safety guidelines given in the introduction to this Unit, and to the specific safety guidelines throughout the text.

## Safety

In all of the EIBE Units, we have tried to check that all recognised hazards have been identified and that suitable precautions are suggested.

Where possible, the proposed procedures are in accordance with commonly-adopted general risk assessments. If a special risk assessment may be necessary, this has been indicated.

However, users should be aware that errors and omissions can be made, and that different employers and educational authorities adopt different standards. Therefore, before doing any activity, users should always carry out their own risk assessment. In particular, any local rules issued by employers or educational authorities **MUST** be obeyed, whatever is suggested in the EIBE Unit.

Unless the context dictates otherwise, it is assumed that:

- practical work is carried out in a properly equipped and maintained science laboratory;
- any mains-operated equipment is properly maintained;
- care is taken with normal laboratory operations such as heating substances;
- good laboratory practice is observed when chemicals or living organisms are used;
- eye protection is worn whenever there is any recognised risk to the eyes;
- pupils and/or students are taught safe techniques for activities such as handling chemicals and microorganisms.

# DNA profiling - the background



## Introduction

The analysis of human DNA has applications, apart from in pure research, in two main areas.

1. Health care: including diagnosis of hereditary disease, chromosome aberrations and cancer.
2. The judicial system: identification of suspects in criminal cases (especially murder, rape and other violent assaults) and analysis of family relationships in disputed paternity and immigration cases.

The teaching material in this unit deals mainly with the applications of DNA analysis within the judicial system.

A DNA analysis carried out to clarify a legal problem, be it a criminal case or a question of disputed family relations, only deals with a tiny part of the human genome. In essence, the routine analysis consists of the determination of the lengths of four or five selected chromosomal DNA segments. The result is an individual combination of letters and numbers called a DNA profile. Some have compared the DNA profile to a bar code. In fact, the original method, developed for legal purposes, gave its result as a series of dark bands on an X-ray film. This was called 'genetic fingerprinting'. The method which is used today, however, is carried out differently. It is called DNA profiling.

## Application of DNA profiling

Every day newspapers feature stories of murder and rape, and DNA profiling is now routinely used in solving these cases. DNA profiles of blood or semen stains from the scene of the crime are compared with those of possible suspects.

In cases of disputed paternity, an initial comparison between the DNA profiles of a mother and child will show which parts of

the child's profile must have been contributed by the father. DNA analysis of the putative father makes it possible to decide if the man in question may be the child's father or could not be, i.e. if his DNA profile is compatible with the part of the child's profile not contributed by the mother. This can also be applied to immigration cases when the family connection between an established citizen and a new immigrant requesting citizenship on the basis of family relationships, may be established or disproved.

DNA profiling may also be used to aid the identification of bodies, by comparison with possible relatives. It has proved valuable in sea and air accidents when victims are beyond recognition (1). It has also been used for victims of terrorism in Israel and the former republic of Yugoslavia, and of war, e.g. in the Gulf war in 1990.

Identification of long dead human remains is also now possible due to the development of methods allowing the extraction of DNA from bones, including teeth. This has allowed identification of the remains of past victims of crime and has been used in more notorious cases such as the Nazi doctor Mengele and the last Russian tsar and his family (2,3).

In the case of the Russian tsar it was possible, about 75 years after the execution of the family, to clarify which of the nine skeletons that were found in the mass grave in question came from the body of the tsar, his wife and three of their children. Similarly it was possible by DNA analysis to prove that a woman who throughout her adult life had claimed to be Anastasia, the youngest of the daughters of the tsar and his wife, had been an impostor (*see page 23*)(4).

Archaeologists and anthropologists have also found uses for DNA analysis. A stunningly successful example is the recent analysis of DNA from the prehistoric human skeleton found in Neanderthal near the German town of Düsseldorf in 1856. This skeleton is estimated to be between 30,000 and 100,000 years old and represents an extinct group of human beings. The results of the DNA analysis show that the Neanderthals and modern man, who during several thousand years coexisted in Europe and elsewhere, are two separate species whose most recent common ancestors lived about 500,000-700,000 years ago (5).

DNA analyses are also being used in the study of viruses, bacteria, plants and animals. They have proved valuable both in research (i.e. to elucidate relationships among species) and for more practical purposes such as diagnosis of infectious diseases, tracking stolen or illegally transported animals, and surveillance of genetically manipulated plants.

### **The human genome**

Most of the human genome is located in cell nuclei (nuclear DNA). A small, but important and interesting part, is found in the mitochondria (mitochondrial DNA). This will be dealt with separately.

### **Nuclear DNA**

In man the nuclear DNA comprises 46 long molecules, each of which forms 'the genetic backbone' of a chromosome. The 46 chromosomes of each cell consist of 23 pairs, one chromosome from each pair having been inherited from the individual's mother the other from the father. This is the basis of Mendel's universal first law of inheritance and means that it is possible by DNA analysis to trace the biological relations within a family, including the segregation of disease genes.

One pair of the 23 pairs of chromosomes, the sex chromosomes, differ between the two sexes: two X chromosomes (XX) in the

female, an X and a Y (XY) in the male. Analysis of blood stains and tissue samples for the presence of Y-chromosomal DNA can therefore be used to identify the sex of the donor. The remaining 22 pairs of chromosomes are called autosomes.

It is estimated that the human nuclear genome has between 50,000 and 100,000 genes which code for the many different proteins of the organism. However, only a minor part of the nuclear DNA contains such coding sequences. The non-coding sequences are located within individual genes, forming intervening sequences or introns, as well as between genes. The DNA sequences used in DNA profiling belong to the non-coding sequences. The complexity and variation of genetic material means that each human being, other than monozygous individuals (identical twins, triplets etc.), has a unique genome and thus a personal DNA profile.

### **Mutations**

DNA is stable but not static. Occasionally mutations happen in DNA, changing the sequence of base pairs. A mutation can have serious consequences if it changes the synthesis or function of an important protein. However, most mutations have no immediate consequences for the organism, as they take place in non-coding sequences, either within or outside a gene. Such harmless, neutral, mutations form the basis for most of the differences between individual genomes.

### **Highly variable regions**

In routine forensic DNA profiling, regions of the genome are analysed in which a specific sequence of base pairs is repeated a variable number of times in a continuous sequence (Fig. 1). These regions are called VNTR regions, VNTR being an abbreviation of *variable number of tandem repeats*.

The length of a given VNTR region of a certain chromosome will thus depend on

**Figure 1. A VNTR region in a chromosome pair:** *the two chromosomes have 10 and 18 repetitions respectively.*

*paternal chromosome:*



*maternal chromosome:*



the number of base pairs in the repeat sequence and the number of times it is repeated. In DNA profiling the lengths of a number of different VNTR regions are determined and the DNA profile is thus a simple combination of these lengths expressed in a format agreed upon by the international forensic genetics community.

### The discovery of VNTR regions

These especially variable regions of the genome were discovered by professor Alec Jeffreys and his co-workers at the University of Leicester, UK. They were studying the gene coding for myoglobin, the red oxygen-binding protein in muscle cells. During this work they discovered that one of the introns in this gene contains a sequence of base pairs tandemly repeated a number of times and that the number of repeats can vary among myoglobin genes and thus among individuals. Jeffreys and his group wanted to use this variable region as a marker in their effort to find the chromosomal location of the myoglobin gene. They therefore isolated a piece of DNA containing this region and used it as a probe in DNA analyses.

A probe is a DNA molecule that can base-pair, or hybridise, with the DNA strands in the gene of interest. The probe is labelled, either chemically or radioactively, so that it can be detected in the analysis in which it is being used, e.g. by means of a sensitive photographic film (see Fig.5).

When Jeffreys and his co-workers used the probe they discovered to their surprise that

every analysed sample resulted in numerous 'bands' on the photographic film, due to the fact that the probe had bound to a large number of DNA regions from each person. The band pattern turned out to be very personal and thus 'the genetic fingerprint' was discovered. Later the various VNTR regions causing the many bands were identified and characterised and it became possible to analyse them separately which is the principle behind today's DNA profiling.

Alec Jeffreys has received much recognition, including a knighthood, for this discovery and its subsequent development, laying the grounds for forensic DNA analyses. His work is one of many instances of an important discovery being made during an investigation with quite a different goal. It also shows how basic research can make discoveries which can quickly have valuable practical applications.

### Mitochondrial DNA

Mitochondrial DNA (mtDNA) is a small, circular DNA molecule with about 16,600 base pairs. In comparison nuclear DNA consists of some 6 billion base pairs or

**Figure 2. Sir Alec Jeffreys**





## More about VNTR regions

Many of the VNTR regions that were first discovered had a rather large number of base pairs (typically 20-50) per repeat and a variation in the number of repeats from 50 to several hundred. Such a region can thus vary from 1,000 to 10,000 base pairs or more. The number of repeats, and thus the length of the VNTR region, is a characteristic which is inherited according to the fundamental Mendelian principle of segregation. In genetic terms a VNTR region is therefore a locus with a number of alleles, each characterised by a certain length. Such a locus, with a variety of alleles each showing a fairly high frequency, is said to constitute a genetic polymorphism (polymorphism = the occurrence of many forms; here: variants). In a highly variable VNTR locus more than 95 % of the population will have alleles of different lengths and thus be heterozygous in this locus, and the likelihood that two unrelated individuals have the same combination of alleles in a given VNTR locus is often much smaller than 1 %.

More recently, a vast number of VNTR regions have been identified which have repeated sequences of only 2-4 base pairs, showing a variation of, for instance, 5 to 15 repeats. A number of these so-called STR regions (short tandem repeats) form the basis of modern advanced forensic DNA profiling.

about two metres of DNA! As each mitochondrion contains 5-10 mtDNA molecules, and a cell can have hundreds or even thousands of mitochondria, the cells from an individual will, on average, contain thousands of mtDNA molecules with identical sequences of base pairs. This high copy number per cell explains why mtDNA is the part of the genome which is most likely to be recovered from ancient remains and from biological stains.

As mtDNA is located outside the cell nucleus, it is inherited exclusively through the female, the sperm only contributing nuclear DNA. This means that a woman, her children, mother, maternal grandmother, and other family members linked through an unbroken female line have identical mtDNA sequences. Mitochondrial DNA analysis is therefore useful in identifying biological relationships even across a separation of many generations, provided there is an unbroken female line.

Mutations can occur in mtDNA, as in nuclear DNA, and per base pair they do so at a higher rate. Mitochondrial DNA is less protected against damage and does not have access to the DNA repair mechanisms

present in the nucleus. Over millennia a significant variation in the mtDNA sequence of man has therefore arisen. About half of that variation is located in two minor, non-coding regions of the molecule and sequence analysis of these highly variable regions have become an important tool in both forensic and anthropological investigations. It was mtDNA analyses that made it possible to identify the skeletons of the last Russian tsar and his family (2,3) and reveal the identity of 'the false Anastasia' (4). It was also mtDNA analyses that gave us the first important genetic information about the Neanderthals revealing that they were a human species only remotely related to modern man (5).

# DNA profiling - the technology

★ ★ ★ ★ ★ ★ ★ ★ ★ ★ ★ ★ ★ ★ ★ ★

Two types of DNA profiling, classical and modern, will be described. Both use electrophoresis as an essential element of the technique so it is first necessary to have a clear understanding of the electrophoresis process.

## Electrophoresis

Electrophoresis means migration in an electric field. As DNA molecules contain numerous phosphate residues they are negatively charged when in an alkaline solution. In an electrical field they therefore migrate towards the positive electrode, the anode. If this migration takes place in a gel, the DNA molecules are at the same time separated according to size, as the small molecules move faster through the mesh of the gel than larger ones. The gel effectively acts as a molecular sieve. Gel electrophoresis is a simple and extremely useful technique for separation of DNA molecules according to size, and it is used in many forms of DNA analysis.

## Classical DNA profiling

Before reading any further, look at pages 12-13 for an outline of the procedures involved in classical DNA profiling.

Classical DNA profiling deals with larger VNTR regions and is carried out as an RFLP (*restriction fragment length polymorphism*) analysis. This term originated in the 1970's when it was discovered that human DNA, when treated with restriction enzymes, is cut into smaller molecules called restriction fragments.

After treatment with a restriction enzyme a particular autosomal VNTR region from a single individual will usually result in restriction fragments of two different sizes, one from the maternal and the other from the paternal chromosome (Fig. 3).

**Figure 3. Two VNTR alleles** (see Fig.1) **showing the sites of action of a restriction enzyme:** the length of the resulting DNA fragments of interest is indicated by the distance between the cuts (arrows).

paternal chromosome:



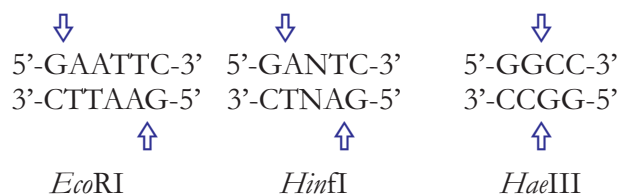
maternal chromosome:



It was mentioned earlier that the alleles of a VNTR locus are elements in a genetic polymorphism. As the polymorphic alleles used in classical DNA profiling are defined by the lengths of the corresponding restriction fragments, this type of polymorphism is called *restriction fragment length polymorphism* (RFLP).

## Restriction enzymes - fundamental tools in DNA analysis

Restriction enzymes are found in bacteria. They are characterised by their ability to recognise a specific sequence of usually 4-6 base pairs in a DNA molecule and 'cut' the two DNA strands either within or close to the recognition sequence. In precise terms restriction enzymes are sequence-specific DNA endonucleases. Consequently, by using a restriction enzyme, it is possible, reproducibly, to 'cut' long DNA molecules into well defined fragments. They have therefore become biochemical tools of fundamental importance in both gene technology and DNA analysis.



**Figure 4. Recognition sequences and 'cutting pattern' of three restriction enzymes:** most enzymes have symmetric (palindromic) recognition sequences.

## More about restriction enzymes

Bacteria defend themselves against attacks from bacterial viruses (bacteriophages or simply: phages) with restriction enzymes. The enzymes cut invading viral DNA and thus render it harmless. In this way they restrict the spectrum of viruses which can infect a particular bacterium, hence their name: restriction enzymes. (The bacterium protects its own DNA against the restriction enzyme by means of another enzyme which puts a methyl group on one of the bases in the recognition sequence, wherever it might occur in the bacterial genome. This modification prevents the sequence from being recognised by the restriction enzyme.)

Several hundred restriction enzymes, each with different recognition sequences are now known and can be bought from molecular biological suppliers. The enzymes are given italicised three-letter 'names' which are abbreviations of the name of the bacterial species in which the enzyme was originally found, e.g. *EcoRI* (pronounced 'ecco are one') was originally found in the intestinal bacterium *Escherichia coli*, better known as just *E. coli*. Other examples (Figure 4) are *HinfI* (pronounced 'hin ef one') from *Haemophilus influenzae* and *HaeIII* (pronounced 'hay three') from *Haemophilus aegypti*. In Europe *HinfI* is traditionally used in classical DNA profiling whereas *HaeIII* is used in the U.S..

## Multiple VNTR site location

If the quality of the band pattern on the autoradiogram, i.e. the exposed film, is good, the membrane is used for hybridisation with another probe recognising another VNTR region. The first probe is removed, usually by boiling the membrane. This procedure does not change the pattern of the bound single-stranded restriction fragments nor their capacity to base-pair with a new probe. Steps 5-7 can then be repeated using a probe which binds to another VNTR region. The membrane is usually reused in this way for analysis of the four or five VNTR loci included in DNA profile. It is, of course, a great advantage, both for simplicity and accuracy, that all loci can be analysed in a single gel electrophoresis run.

## Advantages and limitations

The greatest advantage of RFLP based DNA profiling is the high degree of variability in the classical VNTR regions. In other words there are a large number of alleles for each locus. Consequently it is extremely unlikely that the DNA profiles of two unrelated individuals will be identical. There is, however, an inherent weakness in this analysis due to the fact that the restriction fragments are rather large and therefore show a rather continuous size

distribution after electrophoresis (i.e. the bands may not be clearly separated). This makes it difficult to distinguish between fragments of similar, but different lengths, especially if the samples in question have been analysed on separate gels. Two gels given the same treatment may show slightly different results after electrophoresis on two different sets of equipment, or even when using the same equipment at different times.

At least 20 nanogram of purified and fairly intact DNA is needed for RFLP-based DNA profiling. This limits the usefulness of this method in criminal case work when there may be very little, rather degraded DNA from extracts of specimens such as blood or semen stains. Further problems may be caused by other compounds in the extracts, e.g. dyes from cloth, such as the blue colour of denim, which may stick to the DNA and influence the migration rate of the fragments, giving misleading results. This might cause the DNA profile of a stain to differ from that of a blood sample drawn from the same person, and the person to be wrongly excluded from the investigation. However, a consistent minor difference in migration of fragments from all analysed VNTR regions will alert the experienced investigator to this problem.

# Classical DNA profiling



## the steps of an RFLP analysis

1

Samples are received. DNA is extracted and purified. In blood only the white blood cells contain both nuclear and mitochondrial DNA. (In mammals, including man, the red blood cells do not contain DNA; platelets contain only mitochondrial DNA).

2

The DNA samples are treated with the restriction enzyme *HinfI* which cuts the DNA in several places including each side of the VNTR regions of interest.

3

The samples containing the restriction fragments produced in step 2 are pipetted into individual wells in an agarose gel immersed in an alkaline buffer (in an electrophoresis tank). Electrophoresis is started.

4

After electrophoresis the gel is removed. DNA fragments are transferred from the gel to a nylon membrane by placing the membrane on top of the gel and covering it with a layer of absorbent paper towels. This is called Southern blotting (after the biochemist, Ed Southern). The gel is treated with a strong alkaline solution (NaOH) to denature the DNA fragments (break the hydrogen bonds between the DNA strands) so that the fragments are bound to the membrane as single strands.

5

The nylon membrane with the bound but invisible pattern of single-stranded DNA fragments is submerged in a solution containing a radioactive, single-stranded DNA probe. The nucleotide sequence of the probe is complementary to part of the sequence of the VNTR region to be characterised. The strands of the probe hybridise with the membrane bound, single-stranded fragments containing the corresponding VNTR region.

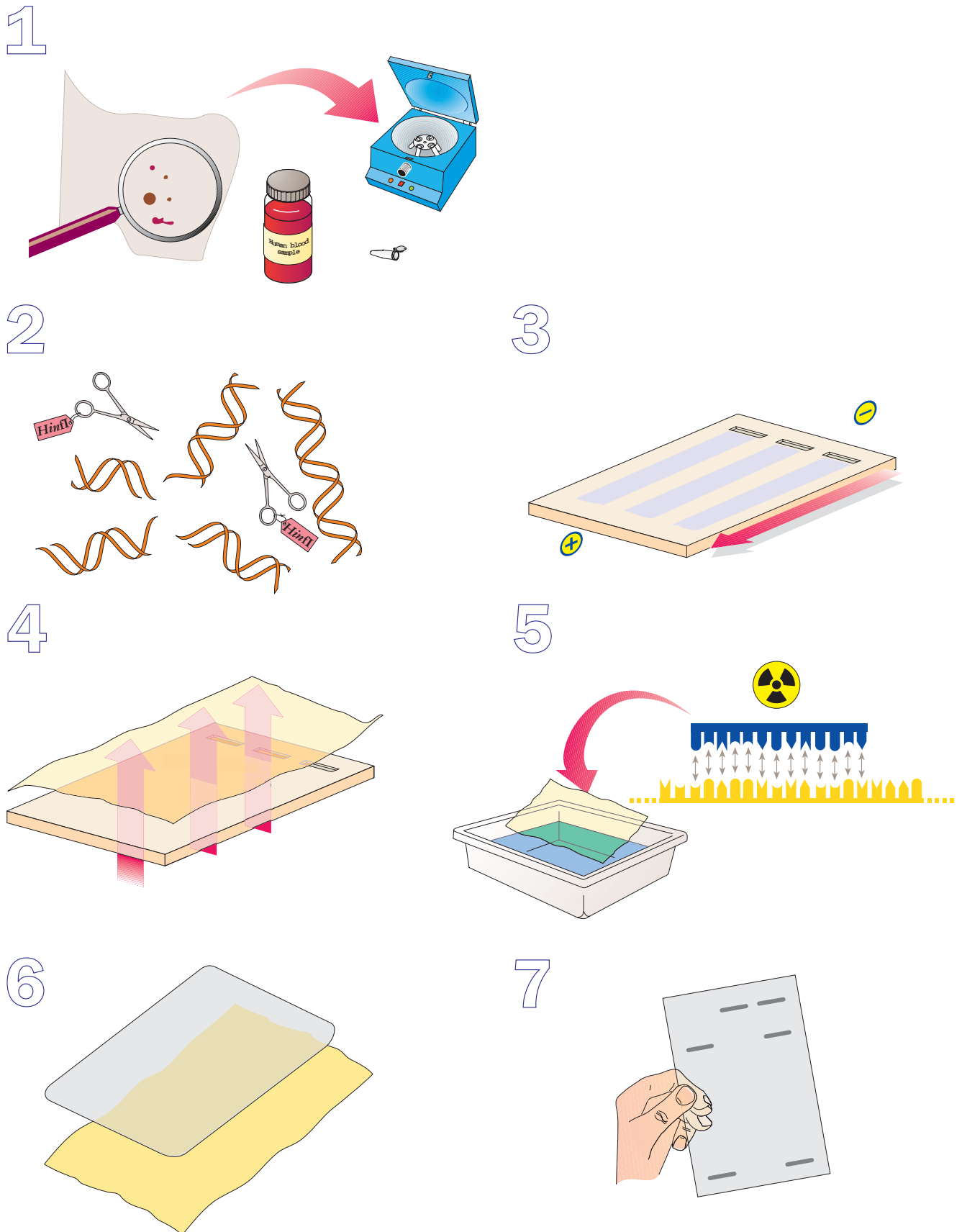
6

The nylon membrane is lifted out of the hybridisation solution, unbound probe is washed off and the membrane is dried. It is then placed in contact with a sensitive photographic film which will reveal where on the membrane the radioactive probe is bound and hence the positions of the fragments of interest.

7

The film is developed and 'bands' corresponding to the restriction fragments of interest appear. The approximate sizes of the fragments can then be determined from their distance of migration during electrophoresis, by comparison with DNA molecules of known size run in a separate lane on the gel.

**Figure 5. Classical DNA profiling** - the steps of an RFLP analysis



# Case-oriented examples



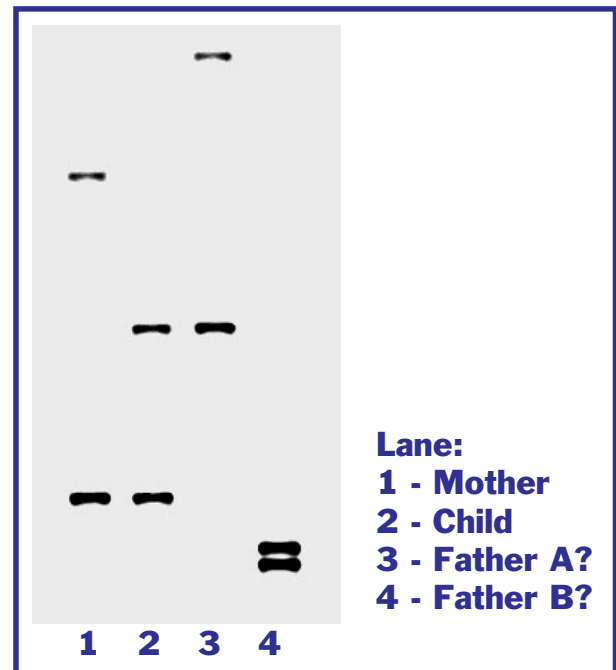
## EXAMPLES

### Problem no. 1: Disputed paternity

Figure 6 shows the result of an RFLP-based DNA profile in a case of disputed paternity. In interpretation of the result keep basic genetic principles in mind: a VNTR locus has been analysed in which every individual has two alleles, a paternal and a maternal one.

- In this case two possible fathers have been presented. Which of the following questions can be answered with the highest degree of certainty? (Give reasons for your answer).
  - Which man is the child's father?
  - Which man cannot be the child's father?
- Why are up to five VNTR regions routinely analysed in cases of disputed paternity?

Figure 6.

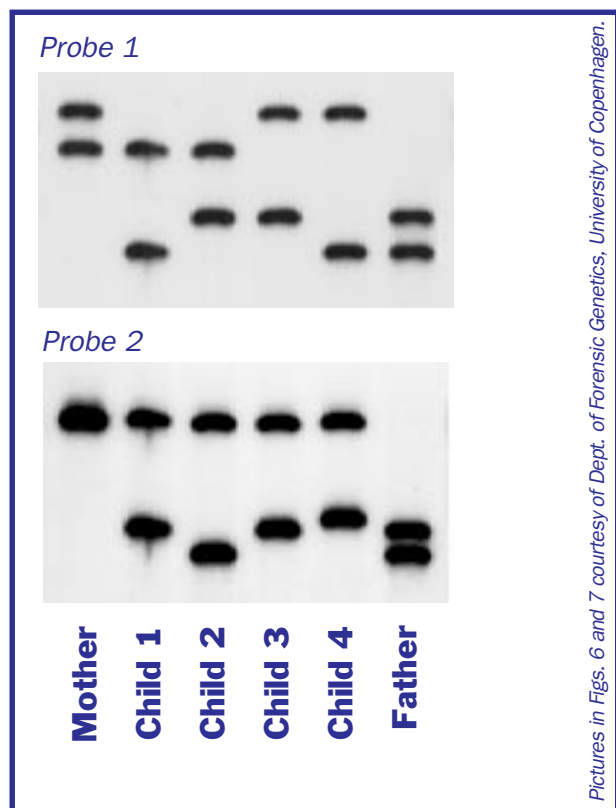


### Problem no. 2: Immigration

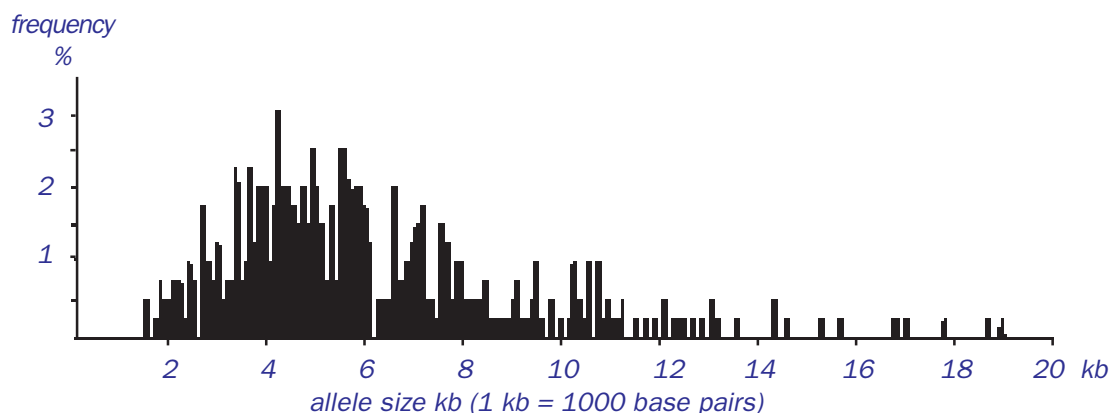
Figure 7 shows the DNA profile of a family where two VNTR regions have been analysed. The story is that a family of refugees with three children have been given asylum. Later, an adolescent boy arrives and asks for permission to be granted residence too. Both he and the family claim that he is the fourth child of the family. The authorities ask for DNA analysis in order to clarify the biological relationship.

- Can the adult couple be the parents of children numbers 1, 2 and 3? (Give reasons for your answer.)
- Can the adult couple be the parents of child number 4? (Give reasons for your answer.)
- Give your opinion, based on your answers to the above questions, of whether child no. 4 should be granted residence in the country on the basis of his relationship to the family. (See also the questions for discussion on p. 24).

Figure 7.

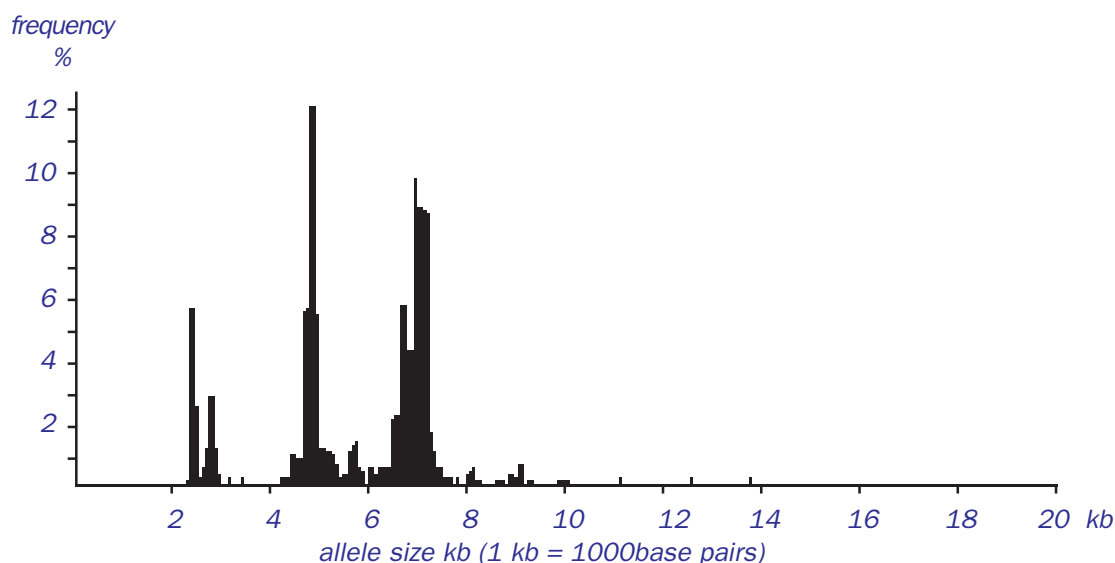


**Figure 8. VNTR locus D1S7 (probe: MS1).**



The abscissa shows allele size; the ordinate shows the frequency of the alleles as a % of those analysed. The repeat sequence consists of nine base pairs. The smallest alleles have a little more than 130 repeats, the largest ones over 2,000. The mutation rate of this locus is ca. 5 %, which is relatively high, even for VNTR regions. D1S7 is located on chromosome 1q, i.e. the long arm of chromosome 1. (Source: Cellmark sales catalogue)

**Figure 9. VNTR locus D5S43 (probe: MS8).**



The abscissa shows allele size; the ordinate shows the frequency of the alleles as a % of those analysed. This repeat sequence consists of 30 base pairs. The smallest alleles have about 80 repeats, the largest ones have several hundred. (Source: Cellmark sales catalogue)

### Problem no. 3:

Figures 8 and 9 show the distribution of alleles of the two VNTR loci D1S7 and D5S43 in a European population.

1. Describe in words what each graph shows.
2. Which allele is the most frequent in each of the two loci ?
3. Are these two VNTR regions equally useful for DNA analysis in
  - a) a murder case,
  - b) a case of disputed paternity and
  - c) an immigration case ?(Give reasons for your answer.)

# Modern DNA profiling

Modern forensic DNA profiling deals with the analysis of STR regions (short tandem repeats, see information box on page 9) where the repeated sequences are four base pairs long (see Fig. 11). Before any determination of the lengths of the alleles, the regions in question are amplified using the PCR (polymerase chain reaction) technique, a very efficient method for in vitro replication of DNA. Under suitable conditions, this technique makes it possible to produce several million copies of a specific DNA sequence within a few hours.

## Polymerase chain reaction (PCR)

In order to carry out DNA amplification by this method the following are necessary: DNA from the sample to be analysed, DNA polymerase (an enzyme which synthesizes copies of pre-existing DNA), and the four deoxyribonucleoside triphosphates, i.e. the 'building blocks' for the new DNA molecules. Two small, single-stranded DNA molecules, 20-30 nucleotides long, are also needed. These oligonucleotides are designed to hybridise, or

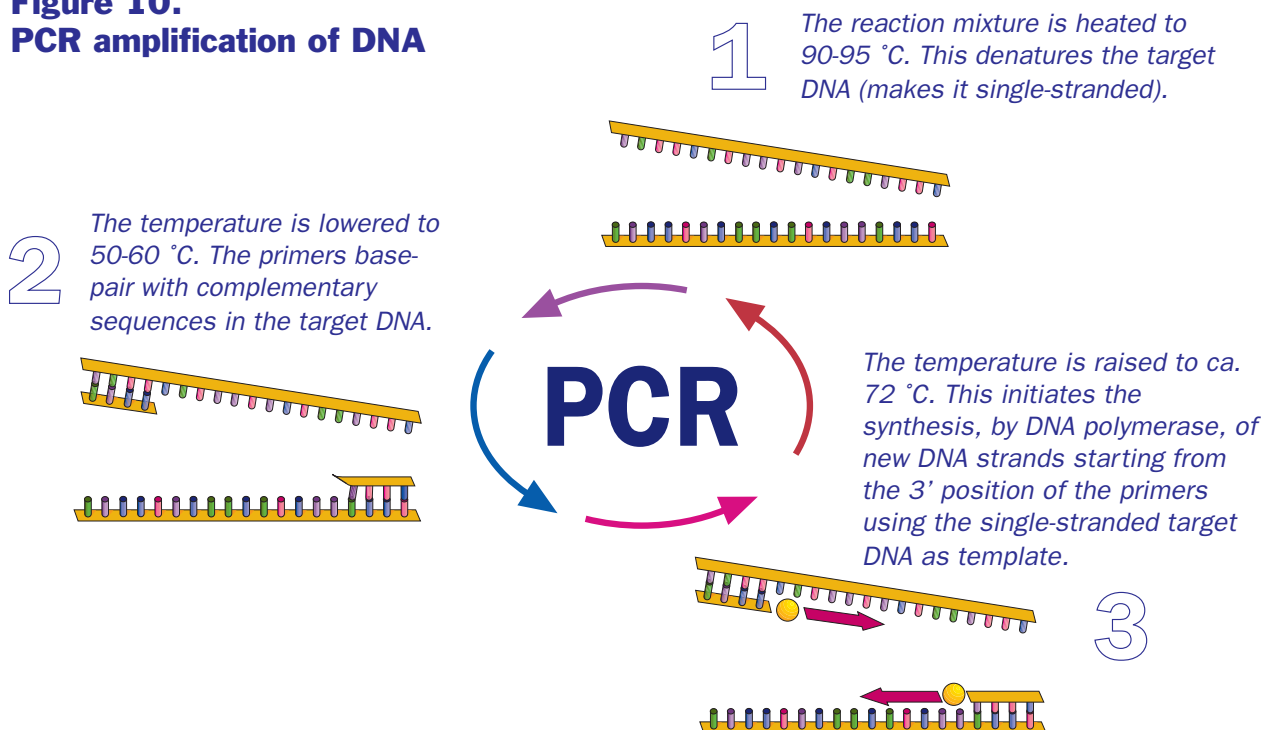
base-pair, to the DNA on each side of the region to be amplified (see Fig. 10). These oligonucleotides, called primers, form the starting points for synthesis of new DNA strands, i.e. they initiate polymerisation. PCR amplification of a DNA sample consists of the following steps which are repeated for 25-35 cycles (see also Fig. 10).

### A. Initial preparation

The two necessary primers, each complementary to the regions flanking the DNA segment to be amplified, are acquired, i.e. bought or synthesized.

DNA is purified from the sample to be analysed. The starting material may be a blood sample, cells from the buccal mucosa (obtained by a simple mouth wash or by light scraping with a spatula), blood or semen stains or hair. (Under suitable conditions it is possible to use the PCR technique on non-purified DNA, e.g. from a droplet of whole blood.)

**Figure 10.**  
**PCR amplification of DNA**





### B. Preparation of the reaction mixture.

To a suitable amount of the DNA sample the following is added: the two primers, the four deoxyribonucleoside triphosphates (dATP, dCTP, dGTP, and dTTP), reaction buffer, and DNA polymerase.

### C. Running the reaction - the PCR machine.

The PCR apparatus is programmed for a suitable number of cycles at the desired temperatures in order to accomplish denaturation of the DNA, binding of the primers, and synthesis of new DNA strands respectively.

The reaction is started by initiating the programme. A single cycle of the three temperature steps normally takes three to

four minutes. Each cycle will, theoretically, double the quantity of target DNA.

### Analysis of the amplified DNA.

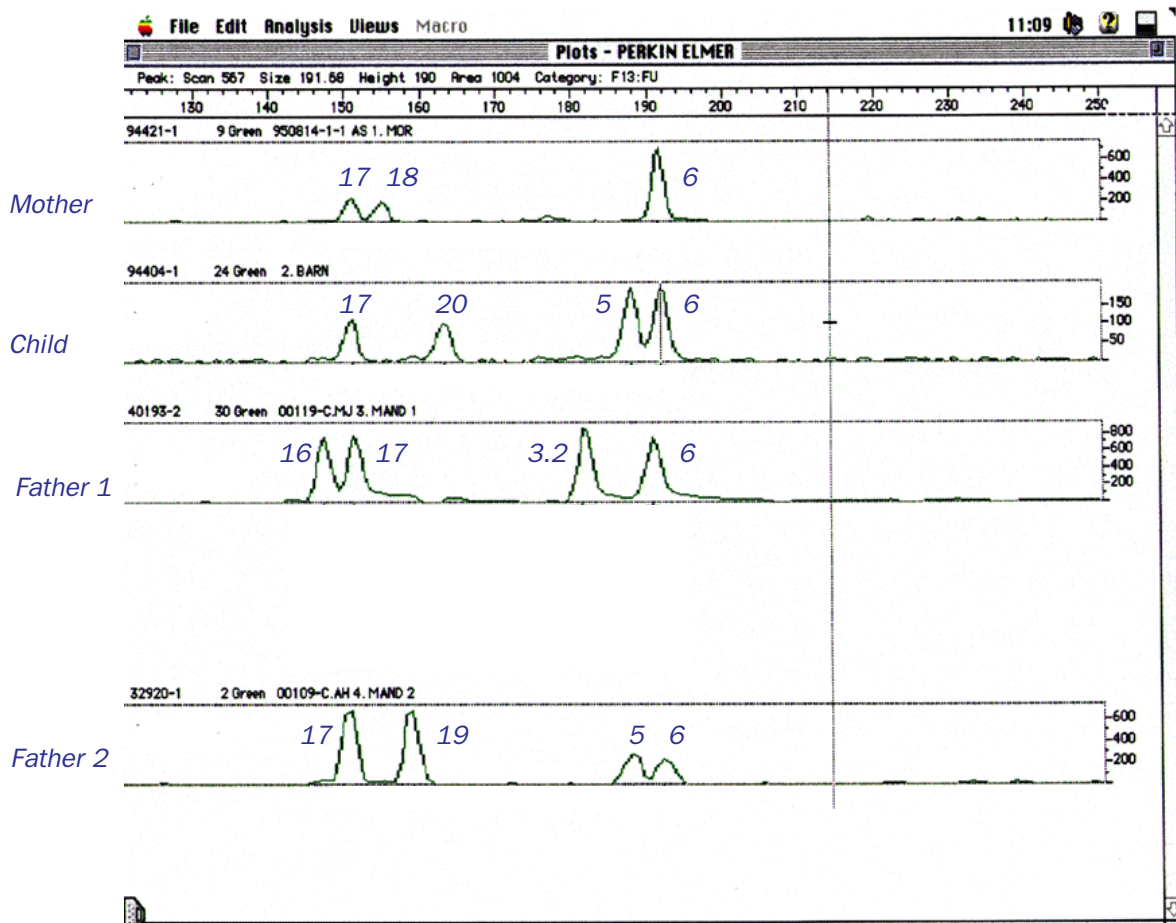
The sizes of the STR alleles are determined by electrophoresis of the PCR products. Sophisticated computer controlled laser equipment is now used to analyse the migration of the PCR products and compare them with DNA molecules of known size that co-migrate with the PCR products on the gel. Figure 11 shows a computer printout from such an analysis.

### Problem No. 4:

Discuss the results of the analysis shown in Figure 11 and give your opinion about the expected court ruling based on these results.

**Figure 11. Results from a PCR-based DNA profiling in a case of disputed paternity**

The two loci analysed: HUMvWA (peaks 16, 17, 18, 19, 20) and HUMF13 (peaks 3.2, 5, 6) both have a repeated sequence of four base pairs. Each peak represents an allele of the locus in question, the number indicates the number of repeats. Primers are chosen to create PCR products of distinct sizes so that several loci can be analysed in the same system without overlap. This explains why the smaller HUMF13 alleles (3 to 6 repeats) in the present analysis are represented by larger PCR products than the HUMvWA alleles (16 to 20 repeats). Peak 3.2 is a common allele (10 % frequency) that has three repeats and an extra two base pairs.



Picture courtesy of Dept. of Forensic Genetics, University of Copenhagen.

## Pros and cons of PCR

The PCR technique has revolutionized DNA analysis and PCR-based DNA profiling is about to replace RFLP analyses. Modern DNA profiling is both faster and less labour intensive as many of the steps can be automated. It is possible in a single reaction to amplify DNA from 4-6 STR regions.

The PCR technique also makes it possible to carry out analyses on samples containing very little DNA. Classical DNA profiling needs at least 20 nanograms of intact DNA, while one nanogram is sufficient for a routine PCR-based DNA analysis, although a single cell (less than 10 picograms) can be enough (9). This development has therefore been particularly important in criminal investigations where samples may be very small and so degraded that it is only possible to analyse STR loci and only after *in vitro* amplification. If only very little DNA is recovered, e.g. from very few cells to begin with, the amplification of some alleles may fail. This is called allele dropout (9).

Traces of unknown substances in extracts from stains can inhibit the DNA

polymerase thus hampering the PCR amplification. In criminal cases purification of the DNA is often a tedious challenge before a successful PCR amplification can be obtained.

The high sensitivity of PCR is also its weakness; contamination with DNA from other individuals may give rise to false results and wrong conclusions. Laboratories working with PCR amplification from small amounts of material have therefore had to develop demanding guidelines for the handling of samples. The risk of contamination of samples with foreign DNA, e.g. from the personnel handling them, must be minimal. Amplified DNA from other samples is another obvious source of contamination. Therefore, in forensic laboratories there is a clear-cut physical separation between areas where DNA is extracted and purified and areas in which the amplification and the subsequent analysis is carried out. Furthermore, much attention is paid to the destruction of amplified DNA after analysis and to limiting the possibility of amplified DNA escaping from the laboratory.

## The invention and development of the PCR technique

The first international paper describing the PCR technique was published in 1985 (6), but the method only became widely used in 1988 after the discovery and purification of a thermostable DNA polymerase and the automation of the process (7). PCR is now carried out on a computer-controlled thermoblock. Many 'PCR machines' are now commercially available. A number of different thermostable DNA polymerases, some genetically modified, are also available. The first of them, and still the most commonly used one, is the *Taq* polymerase named after the bacterium *Thermus aquaticus* in which it was found and which lives in hot springs. These enzymes tolerate repeated heating to over 90 °C thus making supplementary additions of enzyme to the reaction unnecessary. When PCR was first used in 1985 it was necessary to add DNA polymerase after each cycle because the only enzyme available was a thermosensitive enzyme from *E. coli*.

The PCR technique was invented by Kary Mullis in 1983. He had the idea while driving to his summer house in the mountains of California (8). Mullis got a bonus of 10,000 US\$ from the company he worked for. Later the company sold the patent right for 300 million US\$. Kary Mullis was awarded the 1993 Nobel prize in chemistry.

## Application in practice

For obvious reasons forensic laboratories try to agree on which VNTR regions they will analyse. Before deciding, an investigation is carried out in order to clarify which alleles are present, and at which frequencies, in the population in question. This information is essential for the evaluation of the usefulness of the loci for DNA profiling, either in criminal cases or for the elucidation of a disputed paternity. Figs. 8 and 9 show examples of the allele distribution in two VNTR regions among Europeans. The mutation rate of each locus of interest is also analysed and considered. This is essential knowledge when using DNA profiling in cases of disputed paternity or establishing other kinds of family relationships.

## Forensic genetics in Denmark

In Denmark all genetic analyses for the judicial system are carried out by the Department of Forensic Genetics, Institute of Forensic Medicine, at the University of Copenhagen. Until 1990 only results from traditional genetic markers such as blood groups (ABO, Rhesus, and MN), tissue antigens (HLA) and various blood enzymes and other proteins were accepted in the courts.

During the latter half of the 1980's RFLP-based DNA profiling was developed and used for the first time in court in 1990; initially as a supplement to traditional analyses. Today (late 1997), the traditional markers have been practically abandoned in criminal case work. In paternity and immigration cases DNA profiling has substituted tissue typing whenever extended genetic analyses are needed.

PCR-based DNA profiling became routine in paternity and immigration cases in early 1995. A single STR analysis replaced the RFLP-based analyses of the five larger VNTR regions used previously. The same STR analysis has also been used in criminal cases since the middle of 1996, as a

supplement to the RFLP-based analyses. More recently DNA profiling based on amplification and analysis, in one single procedure, of four more STR loci has been developed. This 'package' is now used routinely in paternity and immigration cases and is planned for routine use in criminal cases from January 1998.

There is no doubt that PCR-based DNA profiling will soon have substituted the RFLP-based analyses in Europe as in the USA. This is both for practical reasons (see *Pros and Cons of PCR*, p.18) and because the ability to discriminate between DNA profiles is as good or better using PCR-based analyses.

# How reliable is the conclusion from DNA profiling ?



## Criminal cases

In violent crimes (murder, rape) the challenge is to find out who is the source of a given stain, or other biological remains such as hair, found on the victim or at the scene of the crime. If the DNA profile of the stain is different from that of a suspect, the latter can normally be excluded from being the source. However, if the profile matches one cannot automatically conclude that the suspect is, in fact, the perpetrator. The possibility of more than one person fitting a given DNA profile cannot be completely excluded. As mentioned earlier identical twins must have identical DNA profiles, but sibs in general are also likely to have identical profiles if only a few loci are analysed. The probability of identical alleles among sibs is at least 25 % for each locus. If four loci are analysed the probability of identical profiles among two sibs is at least 0.4 % (the minimum values are for the most frequent situations where both parents are heterozygous and do not share alleles). The value of DNA profiling may therefore be somewhat limited if two closely related persons, e.g. two brothers, are both suspects in a case of murder or rape - or both are involved in a case of disputed paternity. However, by including more loci the likelihood of finding distinctive DNA profiles in non-monozygous sibs keeps increasing.

### Problem no. 5:

1. Explain why for a given VNTR locus the probability of two sibs having identical alleles is 25 % or more.
2. Explain why the probability of finding identical DNA profiles in two sibs is at least 0.4 % if four VNTR loci are analysed.

The critical question in a typical criminal case is the following: *If the DNA profile of the suspect matches that of the stain, what is the probability that the suspect is the source of the stain?*

The weight which can be attributed to a match between two DNA profiles depends on how likely it is for a person, picked at random from the population, to show the same profile. In order to estimate this it is necessary to know the frequencies of the alleles in the population. The expected frequency of a specific DNA profile can then be calculated. A large number of individuals have therefore been analysed in various countries in order to acquire information about allele frequencies. If the DNA profile is based on the four most frequently analysed RFLP-based VNTR loci, the probability that a person picked at random from the population would have the same DNA profile as a stain from a crime scene is somewhere between 1 in 100,000 and 1 in 100,000,000 (10).

In reporting a genetic analysis in a criminal case, where the DNA profiles from a stain and a suspect match, the routine statement would be: *The likelihood that a person picked at random from the population has the same DNA profile as the stain is less than 1 in 100,000 provided that this individual is not closely related to the perpetrator.*

STR loci have fewer alleles than the classical VNTR loci. It is therefore necessary to analyse more STR regions in order to achieve the same low probability of a random match of DNA profiles. At present five STR loci are analysed routinely, four of them in a single PCR amplification and subsequent gel electrophoresis. A technique for analysing six loci in one reaction is currently under development.

## Kinship analysis

For cases of disputed paternity, the task is to find a match for the alleles of the child that do not match those of the mother. It is obvious that, since each person has two alleles for every autosomal locus, many different DNA profiles may share combinations of alleles.

In short, calculations for classical RFLP-based DNA profiles show that the probability of a randomly picked male having a DNA profile with the necessary combination of alleles consistent with being a certain child's father is much below 1 in 10,000 in most cases. This means that the likelihood of picking, by chance, a man with a DNA profile compatible with fatherhood, and yet who is not the father, is less than 1 in 10,000 - again with the usual reservations for males closely related to the father.

Thus, in the case of a man with a DNA profile that does not exclude him from possible paternity, the conclusion from the analysis argues in favour of fatherhood with a weight higher than 10,000 to 1. Or, in terms of percentages, the likelihood that the man in question is the father is higher than 99.99 % (10).

Resolution of family relationships in immigration cases may be more complicated, depending on the problem to be addressed and on practical matters such as access to blood samples from key people.

## Mutations as a source of error

VNTR alleles, just as all other DNA, undergo mutations occasionally. This has to be considered in the evaluation of the results from a kinship analysis. Somatic mutations, i.e. mutations in ordinary body cells, will not influence the DNA profile as they only affect a negligible part of the analysed DNA (with the exception of mitochondrial DNA, see later).

Occasionally VNTR regions show a change

in the number of repeats in an allele from one generation to the next. This can take place during DNA replication in the precursors of the gametes and by so-called unequal crossing-over during meiosis. In such a case a child may inherit an allele not present in either parent (see Fig. 11).

VNTR loci used for kinship analyses are chosen from the more stable ones, but mutation rates of 0.1-0.5 % are common, i.e. 1 of every 1,000-200 gametes shows a mutation in the VNTR region in question. Consequently it is not possible to exclude a man from a paternity case with a 100 % certainty on the grounds of a mismatch in a single one of the regions analysed in the DNA profile.

In such cases the situation can almost always be clarified by including the analysis of three more VNTR loci. Calculations show that the probability of an unrelated man being incompatible in at least one more of the supplementary three loci is very high (99.9 %). The chance of two mutations occurring in two out of eight loci is less than 1 in 100,000 (10).

**Table 1. The distribution of HUMTH01 alleles in different populations.**

	Allele number (this is the number of repeats)						
	5	6	7	8	9	10	11
Danes		24.6	20.1	12.4	9.5	0.3	
Greenland Eskimos		10.9	68.7	5.1	3.7		
White Americans	0.5	22.6	15.9	11.0	14.3	0.5	
Black Americans		13.5	37.0	21.1	14.6		
Hispanic Americans		20.8	33.3	6.8	14.3		
Asians		10.4	26.0	5.2	44.2	4.6	0.7

### STR analyses and ethnic groups

HUMTH01 (human tyrosine hydrogenase) is one of the STR loci analysed in PCR-based DNA profiling. It is located on the short arm of chromosome 11 in the first intron (hence the 01) of the gene encoding the enzyme tyrosine hydrogenase. The repeated sequence contains four base pairs and so far six alleles have been found among unrelated Danes and five among a group of unrelated Greenland Eskimos. The frequencies of the different alleles are very different in the two populations.

Table 1 shows the different patterns of allele distribution in several populations. Allele frequencies are important to know because they are the basis for the calculations which, in a given criminal case or kinship analysis, try to answer the critical question: *What is the likelihood that a person with a matching DNA profile has nothing to do with the case in question, but just happens to have the same DNA profile by chance?*

### Problem no. 6:

1. Depict the information in Table 1 as histograms, cf. Figs. 8 and 9.
2. Describe the similarities and differences of allele distribution among these populations.
3. Evaluate the validity of concluding that a given suspect is the perpetrator (or father) if there is a complete HUMTH01 match in
  - a) a murder case
  - b) a case of disputed paternity. (Assume that analyses of other loci have not yet been made.)
4. If a justice system based its decisions on such a match (question 3), consider the significance for different populations.

### Analysis of mitochondrial DNA

The use of mitochondrial DNA (mtDNA) analyses is fast gaining ground; both by the judicial system, for identification, and in anthropological and archaeological investigations. Its value is partly due to the wide sequence variation found in mtDNA. This makes it highly probable that two individuals, not related through links of an unbroken female line, will have one or more differences in their mtDNA sequence. About half of the variation is located within two highly variable non-coding regions encompassing only about 750 base pairs. These are, therefore, the regions of interest in this kind of mtDNA analysis.

The mutation rate is higher in mtDNA than in nuclear DNA. Nevertheless it is fairly rare to find more than one mtDNA sequence in an individual. However it has now been shown, from the large amount of data collected, that certain base pairs in mtDNA mutate so often that finding, in one individual, mtDNA with an 'old' base pair coexisting with mutated mtDNA is not extremely rare. The mtDNA sequence in one tissue, such as hair, may thus be different from the mtDNA of blood from the same individual, in one of these highly mutable base pairs.

Consider the following situation. Hairs found at a scene of crime have been analysed for mtDNA. Comparison with the mtDNA of a suspect's blood sample shows a discrepancy in one base pair, known to be one that frequently shows mutations. This is not now considered sufficient evidence to exclude the suspect from the investigation.

## Tsar Nicholas II and his family - a mystery solved

The most famous example of the presence of two different mtDNA sequences in the same individual was found during the identification of the skeletal remains of the last Russian tsar and his family, a dramatic story in itself.

Together with his wife and their five children tsar Nicholas II was executed by the Bolsheviks in 1918. What happened to the bodies was kept secret

for many years but in 1979 some people succeeded in locating the grave in which the tsar, his family and members of the household had probably been buried. However, it was not until the dissolution of the Soviet Union that access was given for forensic and anthropological analyses which resulted in the identification of five of the nine skeletons in the grave as being those of the tsar, his wife and three of their daughters (2,3).

Initially sex determination and STR analyses of DNA extracted from the bones established that five of the skeletons, two from a middle-aged couple and three from young adult females, were likely to be from members of single family. Analyses of mtDNA showed that the four females had identical sequences, further evidence that these were the remains of a woman and her three daughters. The fact that the woman's teeth were very well kept with fillings of platinum added to the evidence that she had been the wife of the tsar.

The final identification of the middle-aged couple as being the tsar and his wife was made subsequently by comparing their mtDNA sequences with those of living persons who, from the available detailed pedigree information, were related to the tsar and his wife, respectively, through unbroken female lineages. HRH Prince Philip was the key person leading to the

identification of the tsar's wife. As Prince Philip is the son of a daughter of one of the tsarina's sisters he is related to the tsarina through an

unbroken female lineage. The mtDNA sequence of the four females in the grave should therefore be identical to his, if they were indeed the tsarina and three of her daughters.

It was.



As for the tsar, two living individuals were tracked down who were related to him through unbroken female lineages. Their mtDNA sequences were identical, as expected, and also matched the sequence found in the alleged tsar skeleton. However, a fraction of the mtDNA from the skeleton had a different base pair in one of the positions. To make a long story short: after finding two mtDNA sequences in the extract from the skeleton of the presumed tsar it was decided to seek permission to analyse mtDNA from the skeleton of the tsar's older brother Georgij Romanov who died in 1899 and was buried in St. Petersburg. His mtDNA showed the same two sequences which strongly supported the authenticity of the tsar skeleton (3).

Finally, in the wake of the elucidation of the tsar story another 'mystery' found its solution. For decades a woman had claimed to be Anastasia, the youngest of the tsar's daughters, who had allegedly escaped execution in 1918. She died in the USA in the 1980's. A preserved sample from her small bowel was located at a hospital in which she had undergone surgery in 1979 and DNA was extracted and analysed. The mtDNA sequence turned out to be quite different from that of the tsarina so the claimed relationship could be ruled out. The woman's identity was finally established when it was shown that her mtDNA sequence was identical to that of a living Polish male whose maternal grandmother had been, by historical evidence, the impostor's sister (4).

# Discussion points



## The quality of samples

DNA from corpses, blood stains and other biological remains can be more or less degraded and sparse. What kind of problems does that bring for the analysis and interpretation of results?

## Errors in taking and handling samples

Human beings collect, register and store the samples and whenever humans are involved there is a possibility of error. There are even examples of deliberate swapping of samples. How can these problems be addressed?

## DNA analysis as the only evidence

What would your verdict be if the result of DNA profiling was the only evidence of guilt?

## Privacy

DNA analyses of families occasionally reveal that the social parents are not the biological ones. It has been estimated that a few percent of children in western society have a biological father who is not the person they think is their father. Some cultures have a broader understanding of family ties.

How can one ensure that sensitive data like this does not become accessible to the wrong people.

## Databases of DNA profiles

Should it be legal for the police to establish a database containing the DNA profiles of every individual analysed (convicted or not) in certain criminal cases? If so, who should be authorised to have access to and use this information?

## Independence of the DNA laboratories

DNA profiling can be carried out with relatively inexpensive equipment and materials. Should there, in your opinion, be special compulsory guidelines for laboratories that undertake such analyses, or should the market be open so that the prosecution and defence teams can hire competing laboratories to carry out the analysis?

## Ethnic problems

The distribution of allele frequencies in various VNTR loci can vary among different ethnic groups. How should one deal with a situation in which the police decides to investigate a certain ethnic group because the stain found on a victim has a DNA profile which is more common in that ethnic group?

## Surveillance

It has been suggested that DNA profiles should be printed, or encoded, in passports or social security cards. Discuss both the negative and the positive aspects of this. Could it be compared to making a complete photographic survey of the population in order to be able to publish photographs of wanted persons?

## Scientific evidence

Results of analyses, such as DNA profiling, are often presented in court as scientific evidence. How can it be ensured that the doubts which are always present concerning the validity of such results will also be adequately discussed in court?



## References

1. Olaisen B., Stenersen M. and Mevåg B. 1997. *Identification by DNA analysis of the victims of the August 1996 Spitsbergen civil aircraft disaster*. Nature Genetics **15**: 402-405.
2. Gill P., Ivanov P.L., Kimpton C. et al. 1994. *Identification of the remains of the Romanov family by DNA analysis*. Nature Genetics **6**: 130-135.
3. Ivanov P.L., Wadhams M.J., Roby R.K. et al. 1996. *Mitochondrial DNA sequence heteroplasmy in the Grand Duke of Russia Georgij Romanov establishes the authenticity of the remains of Tsar Nicholas II*. Nature Genetics **12**: 417-420.
4. Stoneking M., Melton T., Nott J. et al. 1995. *Establishing the identity of Anna Anderson Manaban*. Nature Genetics **9**: 9-10.
5. Krings M., Stone A., Schmitz R.W. et al. 1997. *Neandertal DNA sequences and the origin of modern humans*. Cell **90**: 19-30.
6. Saiki R.K., Scharf S., Faloona F. et al. 1985. *Enzymatic amplification of  $\beta$ -globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia*. Science **230**: 1350-1354.
7. Saiki R.K., Gelfand D., Stoffel S. et al. 1988. *Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase*. Science **239**: 487-491.
8. Mullis K. 1990. *The unusual origin of the polymerase chain reaction*. Scientific American **262** (April): 56-61.
9. Findlay I., Taylor A., Quircke P., Frazier R. and Urquhart A. 1997. *DNA fingerprinting from single cells*. Nature **389**: 555-556.
10. Eriksen B. and Hansen H.E. 1996. *Genetikken i retssalene*. Dansk Kemi **5**: 8-11.
11. Agesen H. et al. (Ed.) 1991. *Experimental Gene Technology*. Novo Nordisk.

# DNA profile simulation



There are various methods by which DNA profiling can be simulated in the school laboratory. Here is a method for teachers who are fairly familiar with the practice of DNA gel electrophoresis.

## Multiprobe profile simulation

This method uses four DNA samples, one of which has supposedly come from the scene of the crime, the other three belonging to suspects. A suspect's DNA will match that recovered from the scene of the crime; hence three different types of DNA are needed, each giving distinctive patterns (once restricted and run on a gel).

To minimise the expense, this exercise uses readily available DNA from bacteriophage lambda, and a single bacterial plasmid. Some deception is necessary to provide the three distinctive restriction patterns. This is achieved by restricting one of the lambda DNA samples with an additional enzyme before it is given to students for further restriction with *Hind*III and analysis. Students should not be made aware of this pre-treatment!

Teachers will be able to devise a suitable scenario (e.g. a description of a suitable crime) to enliven the classroom activity.

Any suitable DNA gel electrophoresis equipment may be used for this work. The National Centre for Biotechnology Education supplies low cost kits containing room temperature stable lambda DNA and restriction enzymes, together with all the necessary equipment, chemicals and instructions for DNA gel electrophoresis. Detailed guidelines are given for treatment of DNA with restriction enzymes and the separation of the resulting DNA fragments using gel electrophoresis.

Further details can be obtained from the NCBE website at:

<http://www.rdg.ac.uk/NCBE>

or,

NCBE

The University of Reading,  
Whiteknights

Reading, RG6 6AJ, UK.

Phone: (44) 118 987 3743

Fax: (44) 118 975 0140.

or,

Email: C.Shearer@rdg.ac.uk.

For this DNA profile simulation a suitable plasmid is needed in addition (e.g. pUC18, available from molecular biological suppliers - but expensive).

## Aim

- To simulate DNA profiling and show some of the essential procedures involved in DNA restriction analysis.

## Advance preparation

DNA samples must be rehydrated and mixed with *Eco*RI or a suitable plasmid, as appropriate, before the practical activity starts.

## Organisation

This activity takes about 50 minutes to complete, including a period of 30 minutes during which the DNA samples are incubated with the enzyme. If using low voltage electrophoresis equipment such as that supplied by the NCBE, running the gel takes between 6 and 12 hours.

## Equipment and materials

Required by each student or group of students (quantities and concentrations are as used in the NCBE protocols).

Three 'suspect' DNA samples:

- tube A, containing  $\lambda$  DNA solution: 2  $\mu\text{g}$  in 20  $\mu\text{l}$
- tube B, containing  $\lambda$  DNA solution: 2  $\mu\text{g}$  in 20  $\mu\text{l}$ , mixed with *EcoRI* (10 units, dried)
- tube C containing  $\lambda$  DNA solution: 2  $\mu\text{g}$  in 20  $\mu\text{l}$  plus a suitable plasmid\* (e.g. pUC18 or pBR322): 0.2  $\mu\text{g}$  in 2  $\mu\text{l}$

A 'scene of crime' DNA sample, i.e. one of the three listed above - the cheapest and simplest would be:

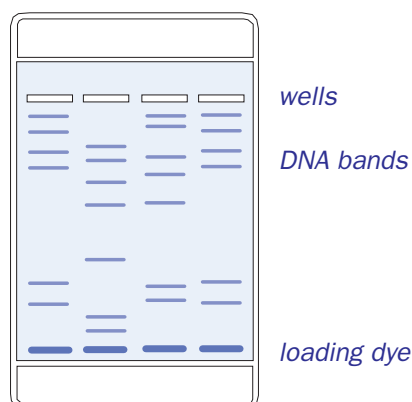
- tube D, containing  $\lambda$  DNA solution: 2  $\mu\text{g}$  in 20  $\mu\text{l}$

Also needed:

- 4 tubes, each containing 10 units of *HindIII* restriction enzyme
- an electrophoresis tank
- a 4-well gel comb
- 15 micropipette tips (this number includes a couple of spares)
- micropipette or microsyringe
- bromophenol blue loading dye with sucrose (about 10  $\mu\text{l}$ )
- TBE buffer solution
- 0.8 % agarose gel, made up using TBE buffer solution, molten
- about 10 ml 0.04 % Azure A solution, made up in 20 % ethanol (a safe alternative to ethidium bromide)
- marker pen
- power supply or batteries (12–25 volts)
- water bath or incubator, set at 37 °C

\* any plasmid of 4-6 thousand base pairs is suitable.

**Figure 12.**  
A DNA gel after electrophoresis and staining



## Procedure

Follow an established procedure for incubation of the samples with *HindIII*, separation of the fragments by gel electrophoresis and staining of the resulting gel to visualise the bands of DNA (see the instructions with the NCBE DNA kits).

## Results

Figure 12 shows the sort of band pattern to be expected, the 'scene of crime' sample can be matched with one of the 'suspect' samples. The pattern can be likened to the 'fingerprint' method of Alec Jeffreys' original work although it should be stressed that it is only a simulation. Single probe techniques would produce one or two bands per sample.