Journal of Chromatography, 618 (1993) 167–179
Biomedical Applications
Elsevier Science Publishers B.V., Amsterdam

CHROMBIO. 6824

Review

Checking of individuality by DNA profiling

R. Brdička*

Institute of Haematology and Blood Transfusion, Prague (Czech Republic)

P. Nürnberg

Institute of Medical Genetics, Humboldt University, Charité, Berlin (Germany)

(First received March 15th, 1993; revised manuscript received April 16th, 1993)

ABSTRACT

A review of methods of DNA analysis used in forensic medicine for identification, paternity testing, etc. is provided. Among other techniques, DNA fingerprinting using different probes and polymerase chain reaction-based techniques such as amplified sequence polymorphisms and minisatellite variant repeat mapping are thoroughly described and both theoretical and practical aspects are discussed.

CONTENTS

List of abbreviations . . . 168
1. Introduction . . . 168
2. Phenotype determination . . . 168
   2.1. Anthropological methods . . . 168
   2.2. Biochemical methods . . . 168
3. Direct genotype analysis . . . 169
4. What are the principles of identification? . . . 170
5. How can the DNA technology solve the problem? . . . 172
6. Technical guide . . . 174
   6.1. Isolation of DNA . . . 174
   6.2. Targeted fragmentation of DNA . . . 175
   6.3. Separation of DNA fragments . . . 175
   6.4. Blotting . . . 175
   6.5. Hybridization . . . 176
   6.6. Probes . . . 176
   6.7. Detection . . . 176
   6.8. Multiplication of DNA segments . . . 176
   6.9. Special techniques . . . 177
   6.10. Evaluation . . . 178
7. Experience and views . . . 178
8. Secondary information sources . . . 179
References . . . 179

* Corresponding author.

0378-4347/93/$24.00

LIST OF ABBREVIATIONS

cDNA	Complementary DNA
cM	Centimorgan, unit of genetic distance between loci
HLA	Human leucocyte antigens
PCR	Polymerase chain reaction
RFLP	Restriction fragment length polymorphism
TAE	Electrophoretic buffer (Tris–acetic acid–EDTA)
TBE	Electrophoretic buffer (Tris–boric acid–EDTA)
VMR	Variant repeat mapping
VNDR	Variable number of dinucleotide repeats
VNTR	Variable number of tandem repeats

1. INTRODUCTION

Characteristic features of the human body have been used since time immemorial to establish the identity of an individual. The story of Jacob and Esau (Genesis 27:15–27) demonstrates how easy it was for a determined fraudster to change his phenotype when the investigator had to rely on the fallible human senses unaided by any technical tools. Since then many techniques exploiting different individual characteristics have been analysed for their effectiveness in distinguishing one person from another.

The main aim of all these attempts has been to have on hand the most objective, easy-to-use and foolproof tool for the following tasks:

(1) To distinguish an individual from all other existing people.

(2) To define relationships between two persons.

(3) To prove that traces, remains or severed body parts belong to a particular individual.

2. PHENOTYPE DETERMINATION

2.1. Anthropological methods

For decades, comparison of skin patterns called fingerprints (dermatoglyphics) has been used with good results, and this method is still widely used today.

Two aspects qualify *classical fingerprints* as examined by dactyloscopy for the purposes of identification of individuals: the objectivity of the method — different observers obtain the same findings; and the stability of the patterns — there is no dramatic change within a lifetime. Also, the variability of this characteristic in the population is so high that it is nearly impossible to find two persons with the same pattern. From the starting point of one-finger pattern analyses, which were used for sample identification documents, patterns from all fingers and toes, as well as palm and plantar configurations, were later added. Taken together, these patterns reduced the risk of false identification to practically zero.

Unfortunately, the area of the body surface suitable for pattern analysis is not very large, and it can be destroyed or damaged wilfully or by accident. Thus, it is sometimes difficult or even impossible to use conventional fingerprints for identification.

Furthermore, at present many sources other than fingerprints, such as stains of blood, semen, saliva, etc., are even more important than fingerprints in the identification of victims and offenders.

Finally, the inheritance of classical fingerprints is not the simplest one. Because of the polygenic control of dermatoglyphic patterns, methods of quantitative genetics have to be applied and rather sophisticated mathematical calculations are necessary when classical fingerprints are used to solve paternity cases [1].

2.2. Biochemical methods

At present, there is a long list of characteristics than can be studied and used for identification purposes. However, their efficiency strongly depends on the polymorphic character of the marker analysed. For blood group antigens, enzymes and the majority of other proteins, the relatively low level of polymorphism does not allow them to be used as single markers. It is necessary to combine the results of individual examinations of

several marker systems [2]. Nearly any source of material, *e.g.* red blood cells or serum, and in some cases semen or hair, etc., can be utilized, although sometimes only fresh tissues are suitable. In particular, when enzyme polymorphisms are to be studied this requirement has to be taken into account. Antigenic substances extractable from hair or bones seem to be much more stable. As mentioned by King [3], Adam Dalgliesh, a fictional representative of modern detectives, already knows how to judge the validity of the biochemical evidence.

For a long time, the best results have been obtained by analysing the human leucocyte antigenic (HLA) system [4]. A large number of alleles with different antigenic properties make HLA typing a very effective tool in forensic medicine. Normally, white blood cells are used for HLA typing, but histocompatibility antigens are also detectable on the surface of other cells. An advantage of the biochemical and immunological markers mentioned above is their simple inheritance, which is of the classical Mendelian type. They are referred to as *polymorphisms*. According to the criteria generally used in population genetics to define a polymorphism, the frequency of the alleles or morphs has to exceed 1%. Thus a polymorphic system (locus) is always represented by alleles found at sufficiently high frequency, in contrast to a rare mutation.

The most unpleasant problem for a geneticist arises from the fact that not all the alleles of a polymorphic system are accessible to identification by the same method. Variability at the phenotype (protein) level revealed by different methods, with the best results provided by electrophoresis, has led to the conclusion that the mean heterozygosity of the human population is generally around 5%. Not all alleles are, of course, electrophoretically detectable [5], and much of the naturally occurring variability remains undiscovered when this method is used exclusively. When polymorphisms are detected according to the enzymatic activity of the marker, a serious complication is caused by the existence of the so-called "null" alleles, *e.g.* enzyme variants that have lost their activity.

3. DIRECT GENOTYPE ANALYSIS

The adoption of methods offered by molecular genetics, *i.e.* the direct analysis of the human genome, has brought about a substantial leap in objectivity. At the genomic level, the number of differences that can be classified according to the conceptual framework of classical genetics as "alleles" is much higher. At this level we compare structural differences between molecules of deoxyribonucleic acid. It has been concluded that, on average, 1 in 100 nucleotides is different when two individuals are compared. This means that the whole human genome, consisting of 46 DNA molecules with a total of about $6 \cdot 10^9$ nucleotides, harbours at least $6 \cdot 10^7$ deviations (ignoring mtDNA). In reality, this number greatly underestimates the actual variability of human DNA. More and more regions are being found in the human genome that show much higher variability than previously expected.

Allelic differences at DNA loci can be visualized as restriction fragment length polymorphisms (RFLPs). RFLPs usually arise from point mutations or other processes (small deletions or insertions) that destroy an existing restriction site or create a new one. Typically, they are represented by only two alleles, reflecting the presence or absence of the affected cutting site for the specific restriction endonuclease. Another type of RFLP is characterized by a variant number of tandem repeats (VNTR) between two unaffected restriction sites in the flanking single-copy DNA. Such loci are represented by a multitude of different alleles (often more than ten) and form the structural basis of hypervariability.

The number of repeat units arranged in a head-to-tail fashion at these loci can vary from a few to several hundred. According to the number of base pairs (bp) in the repetition motif, we distinguish so-called "minisatellites", for which the term VNTR is often used as a synonym, from simple repeats or "microsatellites". The motif of the latter consists mostly of only 2–4 bp, while a sequence of more than 10 bp is repeated in the former (up to 64 bp have been reported [6]).

In recent years, microsatellites and especially

the very abundant CA dinucleotide repeats have become the most pursued sources of polymorphism [variable number of dinucleotide (VNDR)]. Several laboratories are trying to saturate the human linkage map with these highly polymorphic markers in order to achieve a density of less than 1 cM (centimorgan) distance, on average, by the end of 1993 [7]. The pattern of the repetition can be simple or complex. The goal of this paper is not to establish the systematics of these sequences, but to mention their existence and to stress their effect on variability. They are dispersed over the whole genome, but many of them not evenly. Some are preferentially found in the centromeric and subtelomeric regions of the chromosomes [8,9]. Others are localized within the introns of the genes or in the intervening sequences between the genes. Recently, several exonic trinucleotide repeats showing variability have been described within the coding region [10].

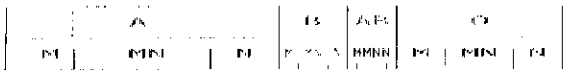
We can conclude this part of the introduction by the statement that the most pronounced differences between the genomes of two individuals are found in the so-called satellite DNA consisting of a variable number of repeat units of larger or smaller size. The diversity and polymorphism of these DNA sequences result from evolutionary dynamics. They arise by amplifications of a basic unit during evolutionary processes.

Even more important, the individual repeated units can lose their identity, change their internal structure and become more and more distinct from each other. To illustrate the process, let us take an example: at the beginning, there can exist a sequence from which different new alleles are produced by simply multiplying this copy. If the original allele is referred to as X, then the new alleles may be XX or XXXX for instance. Subsequently, one or the other of the X units will absorb a point mutation transforming X into X′. In this way, a new kind of allele is born, *e.g.* XXX′XX. Non-stop processes such as recombination and gene conversion will propagate X′ through the alleles, thus giving rise to an enormous increase in variability. Furthermore, a particular sequence of X and X′ motifs, such as XX′X, may act as a higher order repetition unit to form clusters like (XX′X) (XX′X) (XX′X). It is easy to envisage the enormous capacity for variability that could result from these sequences. Knowing all this, direct comparison of nucleotide sequences has to be accepted as the ultimate form in the description of individuality.

4. WHAT ARE THE PRINCIPLES OF IDENTIFICATION?

Identification attempts to prove that no other individual is likely to share the characteristic used for identification. Obviously, a high variability of the characteristic under study will facilitate the task enormously.

Imagine that a person has the blood group A. The probability that we will find another person with the same blood group depends on the prevalence of blood group A in the population. If we suppose that it is 50%, we must take into account the fact that every second person in the population will have the same blood group and, thus, cannot be excluded from consideration. We may also analyse a second blood group system, *e.g.* MN, and find that our person is A, M. In this case all individuals with the combination A, M are again indistinguishable, but the probability of finding individuals in the population bearing both A and M is much lower.



The lower the frequency of an allele or a combination of alleles in the population, the higher the probability that no other person has the same genotype. Loci that have many alleles often offer the advantage that their alleles have low frequencies, whereas loci with two or a few alleles normally do not. What level of probability can we accept as sufficient for identification or exclusion? Usually we continue to combine the different polymorphisms tested until we reach the 99% threshold. By testing blood groups, enzymes or other proteins, sometimes more than ten different systems are necessary to approach this level, and yet not all cases can be solved in this way.

Introduction of the methods routinely applied

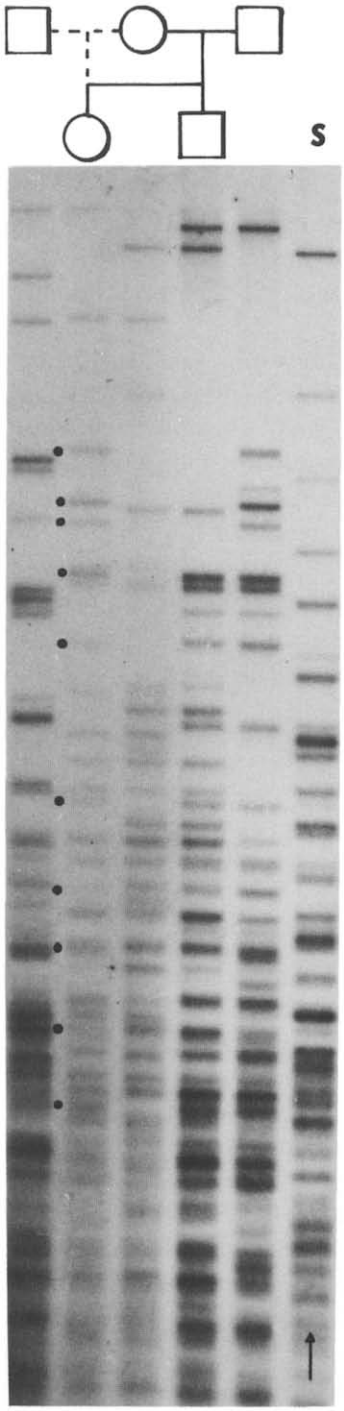


Fig. 1. DNA fingerprinting. Autoradiogram of a family of two children and two prospective fathers tested for paternity. Genetic symbols were used to describe the situation. S = standard pattern. DNA samples (5 μg) were digested with *HinfI* (5 units per μg) for 4 h at 37°C. Electrophoresis was carried out in 0.7% agarose (Pharmacia) and TAE buffer at 0.8 V/cm for 40 h. After electrophoretic separation the gel was dried (Bio-Rad gel dryer) and after denaturation (0.5 *M* NaOH, 0.15 *M* NaCl) and neutralization (0.5 *M* Tris–HCl, 0.15 *M* NaCl, pH 7.0) hybridized with radioactivity labelled (*5'* end labelling) oligoprobe $(GTG)_5$ according to the standard method (J. T. Epplen and H. Zischler, *DNA Fingerprinting with Oligonucleotides – Radioactive DNA Fingerprinting*, Martinsried, Max-Planck-Institute für Psychiatrie, Fresenius, 1990). Dots indicate bands shared by the true father and his daughter.

in modern genetics has dramatically changed the situation.

5. HOW CAN DNA TECHNOLOGY SOLVE THE PROBLEM?

What methods are used to pick an individual from the crowd? How can paternity be excluded or proved? What are the advantages of DNA analysis?

As we have mentioned before, DNA analysis studies polymorphisms (DNA polymorphisms), subsets of which represent the highest level of variability ever observed. There are two possibilities for analysing these polymorphisms. The first involves a stepwise approach with the use of single-locus probes, which is something like firing a rifle with one bullet at a time, and similar to the situation with protein polymorphism. The other approach is to use multilocus probes, which hit a target (the genome) at many sites simultaneously like a shotgun. The latter way is now known as DNA fingerprinting (Fig. 1). This method is based on the outstanding ability of nucleic acids to form duplexes, *i.e.* to hybridize. Of course, successful hybridization is dependent on certain external and internal conditions. Internal conditions, *i.e.* the degree of complementarity between the hybridizing strands, determine in the first place whether or not hybridization occurs. External conditions, *i.e.* ionic strength, pH, temperature, etc., are secondary and influence the dynamics of hybridization by favouring the single-stranded or double-stranded state. Several types of probes are used for DNA fingerprinting. As single strands, these sequences of DNA (or RNA) are capable of hybridizing to many different sites within genomic DNA. All the interspersed target sequences have something in common: their repetitive nature generates numerous multiallelic polymorphisms. Different sources of such probes have been described:

(a) Human genome [6,11,12].
(b) Invertebrate genome [13].
(c) Viral genome [14].
(d) DNA synthesizer [15].

Properly handled, all types of probes offer equally good results, but synthetic oligonucleotides seem to be preferable for technical reasons. The original probes of Jeffreys [6] were supposed to hybridize with at least 20–30 loci.

However, it is possible to dismantle the multilocus profiles into simpler patterns using appropriate single-locus probes. These probes can be derived from DNA fingerprint fragments by cloning the single-copy sequences that normally flank the repetitive elements at both sides [16–18].

The ability of multilocus probes to hybridize with many slightly different sequences distributed over the genome depends on the extended homology between these sequences. Short stretches of the so-called core sequences have been shown to form the structural basis of the homology among minisatellites [11,12].

DNA fingerprinting as a method of genome analysis offers many advantages as compared with other identification techniques. Of course, there are also several disadvantages. The most valuable advantage is the possibility of achieving the final results in a single step. Less time and money is consumed as compared with the separate examination of many loci. Among the disadvantages, problems of reading and evaluating the complex electrophoretic patterns have to be mentioned. It is, of course, impossible to identify the alleles of all loci involved in the multilocus profile characterizing an individual. For this reason we usually only try to distinguish between the bands that are shared (present in both patterns compared in the same position, *i.e.* at the same distance from the starting well) and those that are not shared by the two individuals. Thus, the patterns are evaluated as a whole and the proportion of shared bands among all bands checked is calculated. When two separate DNA samples from the same person are compared, complete band-sharing is expected. But some small differences among several fingerprints from the same person are possible. This can be caused by technical errors, *i.e.* different amounts of DNA applied for electrophoresis, or by improperly digested samples, but such artifacts can be reduced by accurate monitoring of DNA digestion and concentration.

Nevertheless, we also have to consider other than technical sources of pattern "abnormalities", especially when DNA from different tissues is analysed, or when samples for DNA isolation have been taken after long time intervals. For instance, tumour tissues are known to accumulate many deviations from the normal DNA fingerprint. Furthermore, chimaerism and mosaicism can result in tissue-specific patterns in samples from a single individual. Yet apart from these particular cases, somatic stability has been well documented for DNA fingerprints.

When considering the drawbacks, some technical requirements have to be taken into account:

(1) DNA fingerprint analysis consumes relatively large amounts of high-molecular-weight DNA ($\geqslant 5$ μg for each run).

(2) Exact quantification and optimal digestion of DNA samples are necessary.

(3) Good electrophoretic separation is a must.

Furthermore, it is difficult to compare samples from different gels. Therefore, several lanes on each gel have to be reserved for the standard DNA. For better resolution, low voltage ($\leqslant 1$ V/cm) and long gels ($\geqslant 25$ cm) are preferred.

There are several techniques for the evaluation of fingerprint patterns. In the first place visual inspection and comparison of the profiles has to be mentioned. Though this is laborious and troublesome, it is judged by many experts to give the best results. The problem of low objectivity can be overcome by having several experienced workers independently examining the same fingerprint.

But in the age of computers and robots, designed to do for us the tedious work, sometimes better and most often faster than we do, several manufacturers of laboratory equipment have developed reading and documentation systems composed of a television camera combined with a computer and a laser printer. They are not only able to produce the image but also to process it, *i.e.* to evaluate the density of spots and bands and to estimate their positions. The images can be stored and compared at any time with a new one (GDS 5000 from UVP or Bio Image, Model 60S-RFLP or 110S-RFLP, from Millipore).

As already mentioned above, single-locus probes hybridize to the alleles of one locus and individual alleles can be clearly identified. Very often identification according to the size of the corresponding restriction fragments is possible. Though normally loci with a high degree of polymorphism are chosen for the single-locus analysis, results obtained from one locus are seldom sufficient to solve an identification problem and combining the results obtained with several single-locus probes is the rule. Larger quantities of DNA are then required, as compared with the multilocus profiling.

Many single-locus probes exhibit a simple pattern, *i.e.* one allele corresponds to a single band, but others produce rather complex pictures, with constant bands in addition to the allelic ones, and with several bands corresponding to a single allele. Sometimes, the multiband patterns resemble fingerprints, but they do not have the same high information content (Fig. 2).

In principle, all methods of DNA analysis are superior to protein techniques because the substrate DNA is much more resistant to degradation and can be found practically everywhere in the human body. It is possible to use all tissues with nucleated cells for DNA extraction, including blood as a whole because it also contains white blood cells, which are nucleated. Several tissues can be used as the source of DNA for a long time after death. Hair roots and skin and bone tissues, especially when mummified by low humidity, retain enough DNA suitable for examination. These samples are often more or less degraded, but even partly fragmented DNA can be utilized by several techniques, mentioned below. DNA isolated from human mummies has been successfully analysed by Paabo *et al.* [19] and Hagelberg *et al.* [20].

The advantage of the DNA analyses is the high variability of marker systems and the ubiquity of DNA. As a consequence, many more traces are amenable to analysis.

The drawbacks are caused by the fact that relatively large amounts of DNA are needed and poor quality of the samples may become critical. However, the latest techniques of DNA analysis can overcome even these problems.

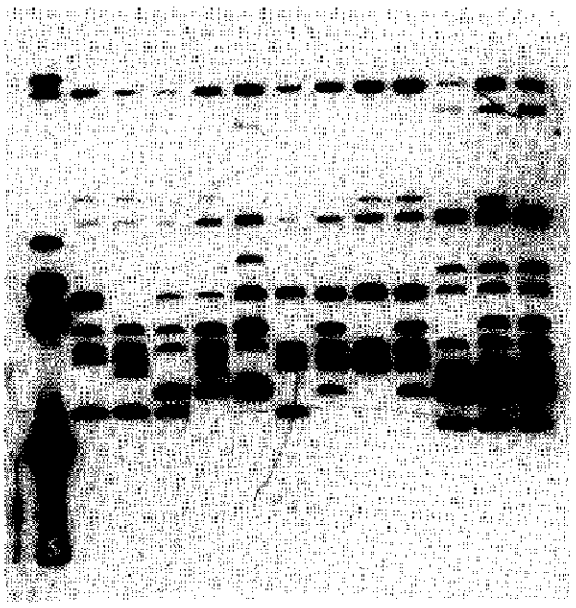


Fig. 2. Autoradiogram of genomic DNA (10 μg) digested with *Taq*I restrictase (MBI Fermentas) at 65°C for 4 h (2 units per μg of DNA), hybridized to a single-locus (DXYS20) probe, p362A, radioactively labelled (random priming method). Digested samples were electrophoresed without any further purification in 0.8% agarose gel (Pharmacia) in 0.5 × TAE buffer at 1.2 V/cm for 20 h (GNA 200, Pharmacia, Sweden). Gel was vacuum-blotted by alkali buffer onto the Nylon membrane (Hybond N +) as suggested by the producer (Amersham, UK). DNA samples from non-related blood donors. Individual alleles of this system were not yet established. 5'-lambda/*Bst*EII digest was used as a size marker. In all electropherograms fragments migrated towards the anode, which is situated at the top.

The most important method to be mentioned here is a method of cyclic amplification of short pieces of DNA introduced by Saiki *et al.* [21]. It is known as the polymerase chain reaction (PCR). The sensitivity of this reaction is so high that it sometimes seems to be a curse rather than a blessing. It is possible to detect a single DNA molecule, and consequently contamination could become a serious problem. Nevertheless, genetic typing of a single cell is now feasible by this method [22].

Each PCR cycle doubles the amount of the DNA segment selected. This is achieved by the action of a thermostable DNA polymerase that is resistant to the heat-denaturing step at 91°C at the beginning of each cycle. After denaturation, which separates the two strands of the native DNA molecule, annealing of appropriate primers [short oligonucleotides of 17–25 base pairs (bp)] to the single-stranded DNA is achieved at a lower temperature (50–60°C). Now the polymerase is able to add further nucleotides to the 3' ends of the primers according to the sequence of the original DNA. This elongation step is continued at 72°C, the temperature optimum for the polymerase. As a consequence the DNA becomes double-stranded again in the region between the primer docking sites and another cycle can follow. Normally, there will be enough PCR product to be detected by ethidium bromide staining after 25–35 cycles. Which DNA segment will be amplified depends on the pair of primers chosen. Their sequences determine the sites of binding to the target DNA. Amplification will only occur when these binding sites are not too far apart (several hundred base pairs is the optimum).

Relatively short pieces of DNA molecule can be analysed by PCR-based methods. That means that partly degraded DNA can also be utilized as a template. We can operate one PCR for one locus, or in some cases several reactions can be combined in one tube. This is possible when the reactions have similar annealing temperature optima and the products can be clearly separated according to their size.

6. TECHNICAL GUIDE

6.1. Isolation of DNA

The first step in the analysis of any sample is the isolation of DNA. In some cases it can be a very simple procedure. The original method, still used in several laboratories and quite appropriate in some situations, is based on a phenol extraction of proteins from the DNA-containing lysate. DNA is then precipitated with ethanol or isopropanol. More recent methods no longer use phenol but high concentration of salts to precipitate the proteins while the DNA remains in solution. Treatment with proteases further improves the purity of the DNA.

For fingerprinting or other techniques dependent on the integrity of DNA to be analysed, careful handling is strongly recommended to prevent any mechanical fragmentation of DNA molecules. Only large pieces of DNA ($\geq$50 000 bp) should be subjected to the enzymic splitting into defined fragments.

Amplification by PCR may give good results also with DNA samples that are not very pure and/or highly degraded ($\geq$300 bp). Stripping DNA naked by boiling nucleated cells is sometimes sufficient to make the DNA substrate accessible to the polymerase. Even small amounts (old tissue specimens, blood smears) of DNA can serve as substrates for PCR.

6.2. Targeted fragmentation of DNA

The enzymic digestion is a very fundamental step in DNA analysis. Endonucleases of bacterial origin called restrictases are used to split double-stranded DNA in a sequence-dependent manner into fragments of defined length. This is achieved by the fact that the enzymes only cut within or nearby a unique sequence motif. Several hundreds of restrictases are known, and anyone can find a list of them in catalogues offered by suppliers of molecular biology products. A comprehensive list was also published in *Gene* 92 (1990) Nos. 1 and 2, but many new restrictases have been found and characterized since that time. Those most used for fingerprinting are *Hinf*I and *Hae*III.

6.3. Separation of fragments

The first step in identifying restriction or PCR fragments is to separate them according to their length. For this purpose, electrophoresis in agarose or acrylamide gels is commonly used. In contrast to proteins, mobility of longitudinal DNA molecules depends on their ability to conform their shape to the size of pores in the gel matrix, rather than differences in electrical charge; indeed, the longer the fragment, the more negative charges that it bears, but the more difficulty it has in being squeezed through the gel matrix.

It is, of course, possible to separate fragments by chromatographic methods, but electrophoresis offers many advantages:

(1) It can separate many samples of digested DNA simultaneously under nearly identical conditions.

(2) The fragments can be fixed at positions related to their mobility. This allows us to compare not only fragments present in one "line" (of digested DNA) but also to compare their relative mobility from lane to lane and even from gel to gel.

There are many types of electrophoretic procedures and equipment available, both horizontal and vertical. The most common technique is the submarine horizontal flat-bed agarose gel electrophoresis in Tris–acetic acid–EDTA (TAE) or Tris–basic acid–EDTA (TBE) buffers. In some cases (PCR), definitive results can be read out immediately after electrophoretic separation. Visualization of fragments at this stage can be done by staining the gel with ethidium bromide and UV light transillumination or by silver staining (polyacrylamide gels only).

If hybridization is necessary in order to detect specific DNA fragments, gels have to be dried or blotted. Both procedures result in an immobilization of DNA pieces, with the former within the collapsed gel matrix and with the latter at the surface of the blotting membrane.

6.4. Blotting

When blotting, we transfer DNA fragments from a gel onto a membrane by a streaming fluid. Originally, the membranes used were made of nitrocellulose, which is certainly less fragile than agarose gel, but it is flammable. Nowadays, nylon membranes of different types are the material of choice. The technique also underwent substantial improvements – from hand-made "devices" with filter paper for capillary blotting, to electroblotting, vacuum-blotting or positive-pressure blotting.

Before blotting, DNA is made single-stranded by immersing the gel in an alkaline solution. The alkali treatment is referred to as the denaturation

step. After blotting, DNA molecules are only loosely attached to the surface of the membrane and must be fixed by UV radiation (several minutes) or by dry-heat treatment (80°C for 2 h). Chemically modified nylon membranes now available bind DNA covalently and the fixation step has become unnecessary.

6.5. Hybridization

This step of DNA analysis allows the formation of hybrids between the fixed fragments and a probing molecule, a piece of nucleic acid of defined sequence, in solution. Provided both are single-stranded, the probe is disposed to bind to any complementary sequence present among the DNA fragments by forming hydrogen bonds under appropriate conditions (temperature, ions). Probe molecules are tagged with reporter groups or molecules. These can act as a source of signals due to radioactivity, enzymatic activity or antigenicity, to mention only some of the more common signalling systems. The method of labelling determines the way of detection. The signal informs us about the place where the probe has found a complementary sequence of nucleotides. Conditions for hybridization and washing away the excess of the free and non-specifically bound probe are very discriminating and allows us to tune the specificity of the reaction. The precise control of temperature and ionic strength of hybridization solution are most important conditions defining the stringency at which a hybridization occurs. High-stringency conditions, *i.e.* high temperature and low salt concentration, are used if cross-hybridization with only partly complementary sequences has to be suppressed.

6.6. Probes

An extremely large number of probes is now available (it is growing every day), both genomic, that is fragments of genomic DNA cloned in bacterial plasmids, or cDNA probes, reverse transcribed from mRNA and cloned. Furthermore, chemically synthesized oligonucleotides specific for simple repeat motifs such as $(GATA)_4$, $(CAC)_5$, $(CA)_8$, etc. have been shown to be extremely useful in DNA fingerprinting [23]. These oligonucleotides are small enough to penetrate into the agarose gel matrix and so are perfectly suited for the in-gel hybridization procedure.

6.7. Detection

The technique of signal detection depends on the kind of the signal used. Radioactively labelling probes with ^{32}P is the most widely used technique, and the signal is then detected by autoradiography on X-ray film or by the Betascope 603 Blot Analyzer from Betagen or by phosphoimaging. Non-radioactive labels are detected by specific techniques corresponding to the label used. Enzyme activity is detected by a histochemical colour reaction or by chemiluminescence, antigenic reporter groups such as digoxigenin by antibody-bound enzyme reaction. Biotin-labelled probes can be visualized by avidin-conjugated enzyme.

6.8. Multiplication of DNA segments

In spite of the relatively recent introduction of the polymerase chain reaction into the collection of techniques used in molecular genetics [21], today hardly any laboratory working in DNA diagnostics is not equipped with a thermal cycler — an apparatus with programmable controls for incubation temperature. The bibliography on PCR, *e.g.* as edited by Perkin-Elmer-Cetus, also demonstrates impressively how common this method has become. The use of PCR is closely bound to the synthesis of short oligonucleotide sequences serving as primers in the reaction. The sequences of these primers are chosen according to the sequences flanking the DNA segment of interest. Only the segment enclosed by the primers is multiplied. Customary PCR provides results similar to those obtained by hybridization with single-locus probes in hybridization experiments. However, a multitude of special applications have been developed, some of which resemble multilocus profiling.

Many different VNTR polymorphisms suita-

ble for the identification of individuals in criminal or paternity cases by PCR techniques have been described. Well-characterized sets of primers are listed in ATCC/NIH Repository Catalogue (1991), pp. 70–71. The primers most commonly used are those directed to the APOB, HRAS or HVT-Ig VNTR [24]. Fig. 3 shows that PCR multiplication of several targeted sequences can also be accomplished simultaneously in one tube. It is obvious that the set of primers to be used in a multiplex PCR has to be carfully chosen. Otherwise, individual reactions could interfere with each other due to the complementarity between some of the primers and the discrepancies in the optimal annealing temperatures. Furthermore, the products of the different reactions should be of significantly different size to facilitate their subsequent electrophoretic separation.

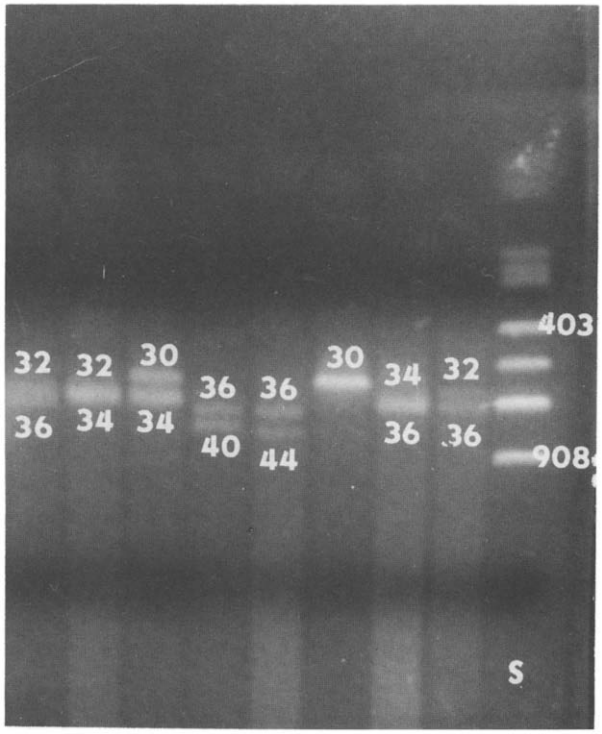


Fig. 3. Single-locus (ApoB) polymorphism detected with PCR. Primers (50 p*M*) [mentioned in E. Boerwinkle, W. Xiong, E. Fourest and L. Chan, *Proc. Natl. Acad. Sci. USA*, 86 (1989) 212] and cycle programming were as follows: 1 – denaturation at 94°C for 1 min; 2 – annealing + polymerization at 60°C for 6 min. The reaction mixture consisted of: sample, 1 μg of genomic DNA; buffer, 67 m*M* Tris–HCl, 2 m*M* $MgCl_2$, 10 m*M* 2-mercaptoethanol, 16.6 m*M* $(NH_4)_2SO_4$, 0.17 mg/ml bovine serum albumin, pH 8.4; 200 μ*M* of each deoxynucleoside triphosphate (dATP, dCTP, dGTP, dTTP) and 2.5 U of *Taq* DNA polymerase (MBI Fermentas). Thirty-five cycles of the two-step programme were performed in PREM III (LEP Scientific, UK) cycler. The products were separated by electrophoresis in agarose gel (1.5%) and visualized by ethidium bromide staining and transillumination with UV light (305 nm) and photographed with a Polaroid camera. Blood donors were used as non-related individuals for population surveys. Most of the samples analysed were heterozygous; individual alleles are shown. S – pBR322/*Alu*I digest with two fragments, size indicated (908 bp and 403 bp).

Complex patterns of products, very similar to DNA fingerprints, are also obtained by the so-called arbitrarily primed PCR (AP-PCR) [25]. In AP-PCR only one primer is used for extension in both directions. It starts with a low-stringency temperature allowing the primer to anneal also at sites that are not perfectly matching. At the next step the annealing temperature is raised to a higher stringency level. The products can be also instantaneously labelled by the addition of [α^{32}P]dCTP of dATP to the PCR mixture and the gels can be then used directly for autoradiography without the need for blotting and hybridization.

6.9. *Special techniques*

In this section several ingenious techniques will be mentioned that are useful for special applications. Some of them use procedures and equipment very different from those used in classical DNA fingerprinting. Nevertheless, they have been designed for DNA typing and hence given identical names. Two methods based on the amplification and separation of products on automated sequencing apparatus should be briefly mentioned. These are ASPs (amplified sequence polymorphisms) described by Skolnick and Wallace in 1988 [26] and the semiautomated DNA fingerprinting developed by Carrano *et al.* [27].

The next step in the strategy to maximize the resolution of individual variability is the direct genomic or sampled sequencing technique [28]. Finally, we want to mention a method of Jeffreys *et al.* [29,30], which brought new insight into the variability of repeat units within their clusters. As

described above, there may be slight differences in the sequences of the basic elements in such a cluster. The technique allowing detection of this variability among the repeat units in a cluster is based on the combination of PCR amplification of individual alleles in a VNTR locus and the partial digestion of the products by two restrictases. Subsequent separation of the fragments on an agarose gel results in a ladder-like pattern from which the sequence of different repeat units can be determined. This technique is called minisatellite variant repeat mapping (MVR). Recently, it has been further developed into an exclusively PCR-based method, MVR–PCR, which avoids the need for troublesome partial digest.

Most often, two types of repeat units are distinguished differing only at 1 bp. But this variability of 1 bp is decisive for the presence (1) or absence (0) of an additional restriction site. For each allele the sequence of (1) and (0) units can be determined now, *e.g.* 11010 for one allele and 01100 for the other. This binary code of both alleles can be combined to form a ternary code that is highly specific for the genotype of the diploid organism: 11010 + 01100 gives the sum of 12110.

The extraordinarity high capacity to distinguish individual alleles (more than 10^{70} for D1S8) at one VNTR locus makes MVR a very promising technique.

6.10. Evaluation

The naked eye, a ruler and a pencil together with the human brain remain the most effective evaluating tools in many cases. What we do is to transform the complexity of personal individuality into a few bands seen on the screen of a transilluminator. If we compare two pictures and it seems to us that the banding patterns match, *i.e.* no differences are present, we deduce that the pictures are identical, and conclude that both samples analysed come from the same body. Solving a problem of relationship we believe that a child gets a half of its genotype from each parent and, the other way round, that what we find in a child must have come from its parents only. This is of course true, but not absolutely, as mutations can occur in the genome during gametogenesis and ontogenesis [31–33]. Although a lot of the evaluation work is still done by hand and by eye, some of the methods mentioned, and especially those separating the products on sequencing gels, are becoming increasingly dependent on automation and on the evaluation with the aid of computers. Within the next few years, this will generally become the most effective way of getting quick and reliable results adequately documented in easily accessible databases [28].

7. EXPERIENCE AND VIEWS

The authors appreciate the recommendation made by one of the referees to include an evaluation of the reviewed methods based on personal experience. We both consider the introduction of DNA profiling for the identification of individual bodies (cells) the most significant achievement of forensic medicine in the last two decades. However, we might differ in assessment of the usefulness and effectivity of the particular approaches. One of us (P. Nürnberg) favours fingerprinting, while the other (R. Brdička) prefers single-locus techniques. These views are based on external conditions as well as on our personalities. Everybody prefers to play games in which he is more experienced, and this is also true in research and science. Good-quality DNA samples and considerable technical skills are necessary in order to get reliable DNA fingerprints. In addition, evaluation of the results is often complicated and requires profound experience. On the other hand, single-locus techniques are less skill-dependent (PCR also gives reliable results with somewhat deteriorated DNA). Because they identify individual alleles, their strength can be enhanced when the interpreter has a knowledge of population genetic data.

Finally, we would like to add a note about dealings with the legal profession and giving the evidence at courts. Law courts expect unequivocal results, *i.e.* 100% certainty. This is an unrealistic demand in biological sciences as all scientists absorb a certain level of agnosticism and scepticism from their daily experience. As a result

of these different attitudes, some disappointment on the part of the lawyers is almost inevitable.

8. SECONDARY INFORMATION SOURCES

B. Amos and J. Pemberton (Editors), *Fingerprint News*, First issue in 1989, University of Cambridge, Cambridge, UK.

Perkin Elmer Cetus (Editor), *PCR Bibliography*, First issue in 1990, Perkin Elmer, Norwalk, CT, USA.

REFERENCES

1 L. S. Penrose, *Sci. Am.*, 221 (1969) 72.
2 E. Giblet, *Genetic Markers in Human Blood*, Blackwell, Oxford, 1969.
3 M.-C. King, *Am. J. Hum. Genet.*, 44 (1989) 179.
4 M. J. Giphart and J. J. VanDerPoel, *Baillière's Clin. Haematol.*, 4 (1991) 975.
5 E. Nevo, *Theor. Popul. Biol.*, 13 (1987) 121.
6 A. J. Jeffreys, V. Wilson and S. L. Thein, *Nature*, 316 (1985) 76.
7 P. N. Goodfellow, *Nature*, 359 (1992) 777.
8 N. J. Royle, R. E. Clarkson, Z. Wong and A. J. Jeffreys, *Genomics*, 3 (1988) 352.
9 J. Wu and K. K. Kidd, *Hum. Genet.*, 84 (1990) 279.
10 G. J. Riggins, L. K. Lokey, J. L. Chastain, H. A. Leiner, S. L. Sherman, K. D. Wilkinson and S. T. Warren, *Nature Genet.*, 2 (1992) 186.
11 Y. Nakamura, M. Leppert, P. O. Connell, R. Wolff, T. Holm, M. Culver, C. Martin, E. Fujimoto, M. Hoff, E. Kumlin and R. White, *Science*, 235 (1987) 1616.
12 Y. Nakamura, M. Carlson, K. Krapcho, M. Kanamori and R. White, *Am. J. Hum. Genet.*, 43 (1988) 854.
13 M. Georges, P. Cochaux, A.-S. Lequarre, M. W. Young and G. Vassart, *Nucleic Acids Res.*, 15 (1987) 7193.
14 G. Vassart, M. Georges, R. Monsieur, H. Brocas, A. S. Lequarre and D. Christophe, *Science*, 235 (1987) 683.
15 J. T. Epplen, *J. Hered.*, 79 (1988) 409.
16 G. Chimini, M.-G. Mattei, E. Passage, C. Nguyen, J. Boretto, J.-F. Mattei and B. R. Jordan, *Genomics*, 5 (1989) 316.
17 A. J. Jeffreys, A. MacLeod, R. Neumann, S. Povey and R. N. Joyle, *Genomics*, 7 (1990) 449.
18 H. Zischler, C. Kammerbauer, R. Studer, K. H. Greschik and J. T. Epplen, *Genomics*, 13 (1992) 983.
19 S. Paabo, J. A. Gifford and A. C. Wilson, *Nucleic Acids Res.*, 16 (1988) 9775.
20 E. Hagelberg, I. C. Gray and A. J. Jeffreys, *Nature*, 352 (1991) 427.
21 R. K. Saiki, S. Scharf, F. Faloona, K. B. Mullis, G. Horn, H. A. Erlich and N. Arnheim, *Science*, 230 (1985) 1350.
22 A. J. Jeffreys, V. Wilson, R. Neumann and J. Keyte, *Nucleic Acids Res.*, 16 (1988) 10953.
23 J. T. Epplen, in A. Chrambach, M. J. Dunn and B. J. Radola (Editors), *Advances in Electrophoresis*, VCH, Weinheim, p. 59.
24 R. Decorte, H. Cuppens, P. Marynen and J.-J. Cassiman, *DNA Cell Biol.*, 9 (1990) 461.
25 J. Welsh and M. McClealand, *Nucleic Acids Res.*, 18 (1990) 7213.
26 M. H. Skolnick and R. B. Wallace, *Genomics*, 2 (1988) 273.
27 A. V. Carrano, J. Lamerdin, L. K. Ashworth, B. Watkins, E. Branscomb, T. Slezak, M. Raff, P. J. de Jong, D. Keith, L. McBride, S. Meister and M. Kronick, *Genomics*, 4 (1989) 129.
28 S. Brenner and K. J. Livak, *Proc. Natl. Acad. Sci. U.S.A.*, 86 (1989) 8902.
29 A. J. Jeffreys, R. Neumann and V. Wilson, *Cell*, 60 (1990) 473.
30 A. J. Jeffreys, A. MacLeod, K. Tamaki, D. L. Neil and D. G. Monckton, *Nature*, 354 (1991) 204.
31 P. Nürnberg, L. Roewer, H. Neitzel, K. Sperling, A. Popperl, J. Hundrieser, H. Proche and J. T. Epplen, *Hum. Genet.*, 84 (1989) 75.
32 H. Ritter, *Hum. Genet.*, 89 (1992) 363.
33 M. Krawczak, *Hum. Genet.*, 89 (1992) 363.