Journal of Chromatography, 569 (1991) 421–447
Biomedical Applications
Elsevier Science Publishers B.V., Amsterdam

CHROMBIO. 5946

Review

DNA typing in hereditary disease

S. WOOD* and S. LANGLOIS

Department of Medical Genetics, University of British Columbia, 6174 University Boulevard, Vancouver, British Columbia V6T 1W5 (Canada)

(First received December 3rd, 1990; revised manuscript received March 8th, 1991)

ABSTRACT

An increasing number of hereditary diseases are becoming amenable to diagnosis by analysis of DNA as the responsible genes are located and identified. Gel electrophoresis of DNA fragments plays a central role in the diagnosis of hereditary disease. Electrophoretic separation of differently sized fragments enables the characterization or typing of normal variants which are known to be genetically linked to disease genes. For some diseases it is possible to directly detect mutations by DNA electrophoresis. Deletion mutants may be detected by a restriction fragment of altered size or by a failure to amplify a coding region with the polymerase chain reaction. Carriers of small deletions, involving a few base pairs, may be identified by DNA amplification which produces heteroduplexes that show characteristic, anomalous electrophoretic migration. Mutations that alter restriction sites also alter the sizes of restriction fragments. Common disease mutations that alter a single base pair may be detected using a pair of reactions with normal and mutant oligonucleotides under conditions where a perfect match is necessary for hybridization, amplification or ligation. Alternatively a mismatched oligonucleotide primer may be designed to generate a restriction site with either the normal or mutant allele, following DNA amplification. Finally a number of techniques are available that are useful as screening tools for novel mutations.

CONTENTS

0378-4347/91/$03.50

LIST OF ABBREVIATIONS

ACD	Acid citrate dextrose
ASO	Allele specific oligonucleotides
bp	Base pair
CCM	Chemical cleavage of mismatches
CF	Cystic fibrosis
CVS	Chorionic villus samples
DGGE	Denaturing gradient gel electrophoresis
DMD	Duchenne muscular dystrophy
EtBr	Ethidium bromide

HVR	Hypervariable region
kb	kilobase (*i.e.* 1000 base pairs)
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PIC	Polymorphism information content
RE	Restriction enzyme
RFLP	Restriction fragment length polymorphism
SDS	Sodium dodecyl sulphate
SSCP	Single-strand conformational polymorphism
STS	Steroid sulphatase
TEMED	N,N,N′,N′-tetramethylethylenediamine
TSD	Tay–Sachs disease
VNTR	Variable number of tandem repeats

1. INTRODUCTION

1.1. Genetic variation in humans

Genetic variation is a feature we all recognize implicitly in the resemblance of offspring to one parent or the other. Benign differences account for the vast majority of genetic variation. This is true regardless of whether the variation involves facial features, serum proteins or differences in the molecular sequence of DNA.

Some genetic variation is maladaptive, in our current environment, and leads to disease which may vary from a minor discomfort to being incompatible with life. Such hereditary diseases generally conform to the well recognized patterns of inheritance so that statistical figures for recurrence can be given to families at risk.

Both pathological and benign variation are determined by sequence differences in the molecules of DNA that constitute the chromosomes. These chromosomes are transmitted from parent to offspring in a predictable manner. Each parent contributes only one member of a chromosome pair, but that single transmitted parental chromosome is a composite of two grandparental chromosomes. The unique character of the transmitted chromosome is determined by recombinational exchange during parental meiosis. Consequently if a child inherits a disease from an affected parent and grandparent then the transmitted parental chromosome will carry the grandparental disease gene together with surrounding material that carries benign variants present in the grandparent. The closer together the disease gene and a variant are on the chromosome, the greater the probability that they well be transmitted together. A variant which is close to a disease gene may be used as a marker to indirectly infer the presence of the disease gene. This well established principle of genetic linkage [1] has in recent years been widely applied to humans, first to locate genes, then to provide linked diagnostic

markers and finally to clone the gene in order to define the disease mutations.

Any variant, *i.e.* inherited feature that exists in at least two forms, may be used as a marker. The two forms of the variant can serve to distinguish between a pair of homologous chromosomes in an individual who is carrying a disease gene. It is important to bear in mind that neither form of the variant is diagnostic of the disease and the presence of a particular variant in any individual does not necessarily mean the presence of a disease gene. Linked genetic markers are the basis of a unique form of laboratory test that requires family studies and whose interpretation depends upon known principles of inheritance.

1.2. Genetic markers in DNA

Genetic linkage was first observed in humans between sex–linked haemophilia and colour blindness [2]. The application of genetic linkage principles to diagnosis was, however, limited until the wealth of genetic markers present in DNA could be readily detected using the DNA molecule itself [3].

Two major types of DNA variation among different individuals or between a pair of homologous chromosomes in the same individual can be detected. Single nucleotide substitutions may replace one base pair with a different base pair. Alternatively a difference may exist in the number of base pairs. This difference may involve the deletion (or insertion) of only a single base pair or may be more extensive, involving hundreds or thousands of base pairs. The size of the human genome (about three billion base pairs) and the frequency of sequence variants/deletions lead to the expectation that the total number of common variants should be in the range of tens of millions. In theory, markers linked to any disease gene may be analysed by DNA typing. However, as disease genes become located by linked markers, this linkage approach to genetic diagnosis is being rapidly superceded by direct DNA typing of the disease mutations themselves. Direct diagnosis of mutation will play an ever increasing role in genetic diagnosis as more disease genes are isolated and their mutations discovered.

1.2.1. Restriction fragment length polymorphisms

Restriction fragment length polymorphisms (RFLPs) were the first markers to be described where variation in the structure of a DNA molecule is detected [3]. The length of a restriction fragment, generated by cutting DNA with a particular restriction enzyme (RE), will vary if the enzyme's recognition sequence has a nucleotide substitution preventing cleavage (Fig. 1A) or if the number of bases between two RE sites varies due to insertion/deletion (Fig. 1B). Traditionally RFLPs have been detected by hybridizing DNA probes, radiolabelled with ^{32}P, to Southern blots of restriction fragments separated by agarose gel electrophoresis. One of the first applications of marker RFLPs in hereditary disease was for the prenatal diagnosis of sickle cell anaemia [4].

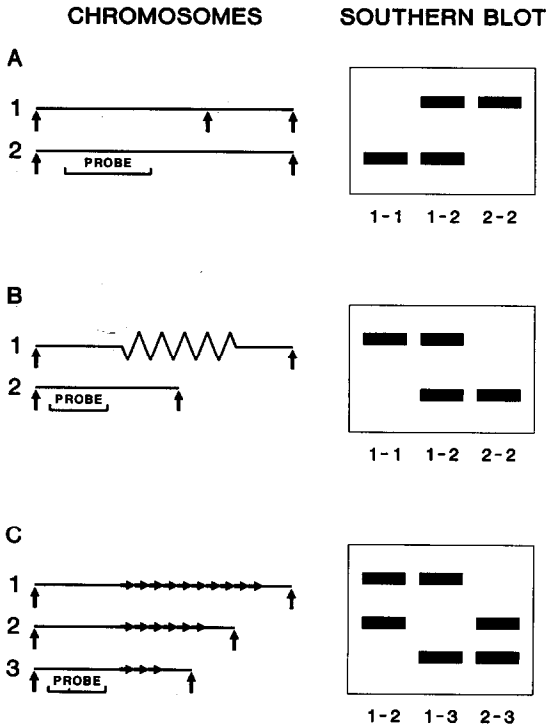

Fig. 1. Restriction fragment length polymorphisms. (A) Loss or gain of a restriction site. (B) Additional sequence inserted between two sites within identical regions. (C) The two restriction sites flank an HVR, hypervariable region. These markers are detected by Southern blotting. Following agarose gel electrophoresis of restriction digests and transfer to a solid support, the hybridizing ^{32}P-labelled probe is detected by autoradiography.

1.2.2. Minisatellite markers

Some RFLPs are due to the presence of minisatellite sequences and represent an insertion/deletion type of variation. These minisatellites consist of a hypervariable "core" sequence [5] of 17–35 bp which is present as a variable number of tandem repeats, (VNTR) [6], giving rise to a highly polymorphic or hypervariable region, (HVR) [7]. The core sequence, which is related to the prokaryotic recombination site, is present in multiple copies throughout the genome and therefore gives a fingerprint pattern when used as a hybridization probe. Unique probes that are adjacent to the HVR give a simple pattern on Southern blot analysis. Two bands representing each member of a chromosome pair (Fig. 1C) are frequently observed. An HVR near the α-globin gene has been useful for diagnosis of adult polycystic kidney disease [8].

1.2.3. Microsatellite markers

Other simple sequences form tandemly repeated arrays. One example consists

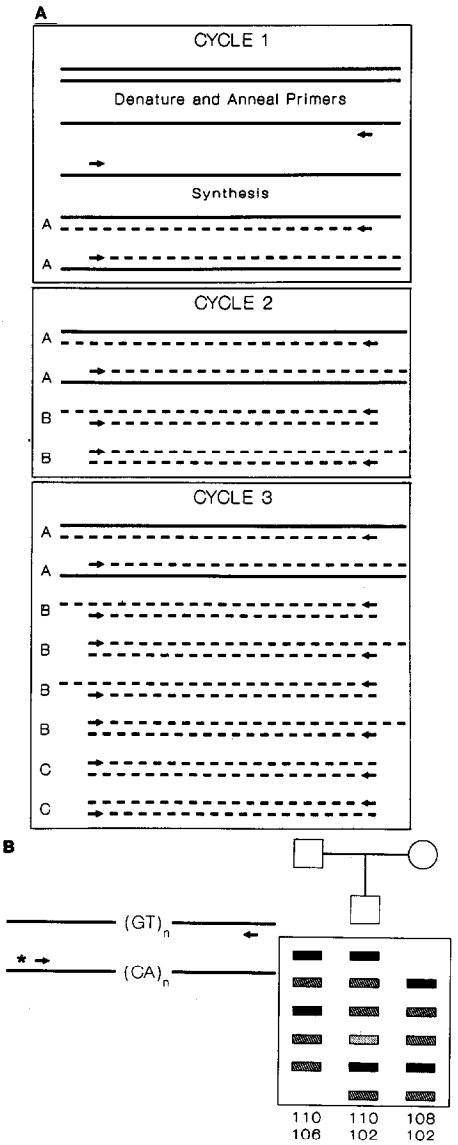

Fig. 2. (A) DNA amplification by the PCR, polymerase chain reaction. Each cycle consists of denaturation, annealing of primers and primer extension synthesis by *Taq* polymerase. The A product from the original target remains at a constant amount. The B product, first produced in cycle 2, increases linearly. The C product, first produced in cycle 3, increases exponentially and represents the final amplification product after the usual 25 to 35 cycles. (B) Typing a CA repeat by the PCR. One primer is end-labelled with ^{32}P, indicated with an asterisk. The products are electrophoresed on a denaturing, sequencing gel and detected by autoradiography. Minor bands form ladders whose members vary by 2 bp due to replication slippage. Allele sizes, indicated below the gel, are determined relative to a convenient reference marker such as sequencing reactions with known size bands. The figure is a representation of an autoradiograph of a dried gel.

of a GT repeat on one strand and a CA repeat on the complementary strand. These short, microsatellite, stretches of about 10–20 repeat units vary by a multiple of 2 bp between members on homologous chromosomes. These markers can be readily detected by using unique sequence, flanking, oligonucleotide, primers in a polymerase chain reaction (PCR) for DNA amplification (Fig. 2). One member of this recently described marker type [9,10] has been found within the dystrophin gene [11,12] that is mutated in both Duchenne and Becker muscular dystrophy.

1.2.4. Other markers

Length variation may be a common feature of simple sequence families such as repeating mono-, di-, tri- and tetranucleotides. Length variation of the polyA tract of Alu elements [13,14] has been reported. Sequence polymorphisms involving regularly occurring elements such as Alu may be expected to provide useful diagnostic markers. Sequence differences may be detected by denaturing gradient gel electrophoresis (DGGE) or as single-strand conformational polymorphisms (SSCP) or by chemical cleavage of mismatched (CCM) bases. These methods also provide powerful techniques for detecting disease mutations (see Section 4.2.2).

1.3. Detection of genetic markers

The detection of genetic markers for DNA typing involves two problems. Firstly, it is necessary to be able to visualize the region of particular interest from among the three billion base pairs of human DNA. Secondly it is necessary to be able to discriminate between slightly different variants of essentially the same piece of DNA.

1.3.1. Visualization

Visualization has been traditionally based upon the exceptional affinity of a single-stranded DNA probe molecule for its complementary strand. DNA probes, radiolabelled with ^{32}P, have been hybridized to single-stranded DNA immobilized on nitrocellulose or nylon membranes, the bands corresponding to regions of hybridization being detected by autoradiography.

An alternative approach uses the PCR to amplify specific fragments that lie between a pair of primers located on opposite, complementary strands (Fig. 2A). The product may be directly visualized followed by electrophoresis on an agarose gel containing ethidium bromide (EtBr). Under UV illumination the EtBr dye that is bound to DNA will produce yellow fluorescence.

1.3.2. Discrimination by size

DNA markers have been distinguished most frequently by size differences. Agarose gel electrophoresis resolves fragments throughout the range of several

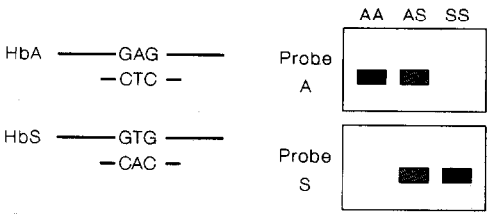

Fig. 3. Detection of sequence variation using allele specific oligonucleotides (ASOs). Southern blots of restriction digests (or PCR products) are probed with ASOs under hybridization conditions that have been determined to give allele specific hybridization. Generally oligomers that are 19 bp in length are used under conditions where a single base mismatch eliminates hybridization. The probe specific for the haemoglobin A allele (CTCCTGAGGAGAAGTCTGC) hybridizes to DNA from both normal and carrier individuals while the haemoglobin S specific probe (CTCCTG*T*GGAGAAGTCTGC) hybridizes with DNA from carriers and individuals affected with sickle cell anaemia.

thousand to a few hundred base pairs and discriminates between fragments where size differences are greater than about 5%. Polyacrylamide gel electrophoresis extends this range to fragments smaller than a few hundred base pairs and allows the resolution of single bp differences.

1.3.3. Discrimination by sequence variation

Sequence differences can be detected by changes in hybridization properties, gel mobility or chemical reactivity. Short allele specific oligonucleotide (ASO) hybridization probes that detect the sequence difference generated by a disease mutation [15] are commonly used, especially in conjunction with PCR amplification [16]. DNA from unaffected individuals will hybridize the normal ASO only, DNA from affected individuals will hybridize the mutant ASO only, while DNA from carriers will hybridize both ASOs (Fig. 3). An alternative procedure uses an ASO as a PCR primer. The mismatch is placed at the 3′ end of the ASO so that only perfectly matched samples will hybridize the ASO, act as a substrate for DNA polymerase and be amplified [17]. This allele specific amplification procedure eliminates the need for radioisotopic detection.

Methods of detecting sequence variation by changes in gel mobility have been developed. Polyacrylamide sequencing gels are denaturing and separate single-stranded DNA on the basis of size alone. Non-denaturing gels will separate both on the basis of size and conformation. Sequence differences in single-stranded DNA can often be resolved as SSCPs on non-denaturing gels [18]. An alternative gel method is DGGE. PCR products amplified from DNA containing a disease mutation can be mixed with normal PCR products. The heteroduplex that is produced will have a region of mismatch consisting of at least a single base pair mismatch. This mismatch will destabilize the molecule and lead to strand separation. If this mismatch is in an early-melting domain then the molecule will assume an altered conformation under denaturing conditions sooner than a nor-

mal molecule. Thus the heteroduplex will melt sooner as it migrates into the gradient, will assume an altered conformation and its mobility will decrease, allowing separation from the normal molecule [19].

Single base pair mismatches are also chemically more reactive and will react specifically under appropriate conditions [20]. This CCM procedure results in strand cleavage at the site of the mismatch. The resultant smaller fragments are detected on a standard sequencing gel.

2. SAMPLE PREPARATION FOR DNA ISOLATION

High-molecular-mass DNA can be successfully obtained from many tissues. Leukocytes, amniocytes and chorionic villi are commonly used for DNA typing. However, protocols have been developed to isolate DNA from paraffin-embedded tissues and tissues fixed in ethanol or formaldehyde, which can be extremely valuable when a crucial family member is no longer available for study. Dried blood spots, *i.e.* Guthrie spots used for newborn screening of inborn errors of metabolism, can provide the small amount of DNA required for PCR amplification.

2.1. Blood samples

Venous blood should be drawn in the presence of an anticoagulant. Acid citrate dextrose (ACD) is slightly superior to EDTA or heparin for preserving the yield of DNA when blood is kept at room temperature for more than three days [21]. EDTA, an inexpensive and convenient anticoagulant, is generally used. The difference in yield and quality of DNA between ACD and EDTA in the first two days is minimal. Clotted blood is much less suitable for DNA extraction, often giving a low yield, and should where possible be avoided.

Few studies have addressed the effect of transport, temperature and storage on DNA quality and yield. The yield of high-molecular-mass DNA was reduced by 20% when blood was stored for five days at room temperature or 4°C [22]. Storage at −70°C is preferable, although high-quality DNA has been successfully extracted from blood stored at −20°C for up to two years. Repeated freezing and thawing should be avoided as it significantly reduces DNA yield and integrity. Once DNA is extracted it can be stored frozen for an indefinite period either in solution, or as dried sample, or at 4°C in a low salt buffer containing a small amount of chloroform.

Many protocols for DNA extraction from blood are minor variations and adaptations of a protocol developed for DNA diagnosis of haemoglobin disorders [23]. Some protocols have a first step which consists of red cell lysis with a solution of Tris–ammonium chloride [22]. Leukocytes are lysed with sodium dodecyl sulphate (SDS) and incubated in the presence of proteinase K, to digest proteins that would otherwise remain bound to the DNA. The standard method

of extracting DNA from these cell lysates uses phenol and chloroform. The organic and aqueous phases are separated by centrifugation and the upper aqueous layer containing the DNA is removed with a large-bore glass pipette. Recently the use of silica gel barrier tubes (commonly used to separate serum from cells) has been shown to increase the yield of high-molecular-mass DNA [24]. After incubation with the SDS–proteinase K solution, the cell lysate is added to the silica gel-containing tube and shaken for 5 min with phenol–chloroform (1:1). Centrifugation at 1000 *g* for 5 min results in trapping of the organic layer below the gel polymer. The aqueous phase containing the DNA remains above the gel and can easily be poured off. DNA yields were 40% greater than the conventional phenol–chloroform extraction [24]. DNA can also be extracted by salting out the cellular proteins [25]. After treatment with a detergent–proteinase K solution, saturated sodium chloride is added to precipitate the proteins. The DNA remains in the supernatant. The final step consists of precipitating DNA with an equal volume of isopropanol or two volumes of 95% ethanol.

2.2. Chorionic villus samples

DNA is isolated from chorionic villus samples (CVS) in a manner similar to that described above for blood. Critical to the preparation of foetal DNA derived from CVS is the careful removal of all maternal decidua. Chorionic villi are homogenized in a small volume (1–2 ml) of buffer and incubated with detergent–proteinase K at 37–50°C for a few hours. DNA is then isolated by phenol–chloroform extraction and ethanol or isopropanol precipitation [26].

2.3. Amniotic fluid samples

Amniotic fluid samples of 15–20 ml can easily be obtained as early as fifteen weeks of gestation. Amniocytes are pelleted by centrifugation and processed following the protocol described for CVS in a volume of approximately 5 ml. DNA can also be extracted from cultured amniotic fluid cells. The decision to extract directly from amniocytes or from a culture will depend on the amount of DNA required for analysis.

2.4. Other sources

Many tissues can be rich sources of DNA at necropsy. Liver, if fresh, yields large amounts of DNA but undergoes rapid autolysis even at 4°C. Spleen is a better source of DNA, for specimens not be stored at −70°C [27]. DNA has been successfully extracted from tissues fixed in ethanol [28]. The tissue is washed extensively in RPMI-1640 cell culture medium before lysis. DNA extracted from this source was of comparable purity and had an estimated size of at least 20 kb.

Formaldehyde-fixed and paraffin-embedded human tissue have also been

found to be a valuable source of DNA. Parameters that needed to be adjusted for successful isolation of DNA include use of ionic detergents, relatively high pH, high concentration of protease and long periods of proteolytic digestion and rehydration or dewaxing [29]. While the DNA obtained is somewhat degraded, it is double stranded, can be restriction digested and is amenable to Southern blot hybridization or PCR. The DNA has a size range of 100 bp to 10 kb and is therefore useful for DNA diagnosis of numerous genetic disorders.

Finally PCR technology allows the use of dried blood spots, routinely collected and stored for newborn screening programs [30]. The highest reproducible yields are obtained by rehydrating the white blood cells by shredding and mincing the blood-spotted filter paper and placing it in 3 ml of 0.85% sodium chloride. After 1 h of gentle shaking at room temperature the 3-ml fluid layer is pipetted off and the white cells are pelleted by centrifugation. DNA is then isolated using the routine protocol for blood.

3. SAMPLE PROCESSING FOR TYPING OF DNA MARKERS

In this section we first outline typical procedures for DNA typing using RFLPs and Southern blotting. We then describe PCR amplification of DNA, the direct detection of specific mutations and the use of different gel systems for marker characterization. More extensive protocols can be found in Sambrook *et al.* [31] as well as elsewhere in this volume.

3.1. Restriction digestion for RFLP detection

Isolated human DNA, irrespective of the source, is first cut with the restriction enzyme known to detect the polymorphism of interest. Generally 4–5 μg of DNA is cut with 5–10 U of enzyme during overnight incubation using the buffer supplied or recommended by the enzyme manufacturer. The reaction is stopped, by adding one fifth volume of 50 m*M* EDTA, 40% sucrose and 0.25% bromophenol blue, and loaded onto an agarose gel. This gel loading solution increases the density and adds colour to simplify the gel loading process. The dye migration approximates that of a 300-bp fragment.

3.2. Agarose gel electrophoresis and blotting

The fragments are resolved by horizontal gel electrophoresis. Agarose is added to the 0.045 *M* Tris–borate, 1 m*M* EDTA (pH 8.0) electrophoresis buffer and dissolved by bringing to a boil. A 0.8% agarose concentration provides good separation of fragments in the 0.5–10 kb range. The gel is cooled to 50°C, EtBr is added to a final concentration of 0.5 μg/ml using a 10 mg/ml stock solution, and the gel is poured into a mold with a comb being used to produce loading wells. EtBr stains DNA so that it can be visualized using UV light. Electrophoresis is

carried out at 1–3 V/cm until the dye marker has migrated an appropriate distance (about 16–24 h for a 20-cm-long gel). The migration of a size standard, such as restriction-digested bacteriophage lambda DNA, can be monitored by examining the gel with UV light. Following electrophoresis the gel is transferred on to a nylon or nitrocellulose membrane in order to transfer the DNA fragments to a solid support for hybridization. Nylon membranes are robust while nitrocellulose is brittle unless a sandwich nitrocellulose membrane that uses a polyester support is used. Protocols provided by membrane suppliers generally have depurination, denaturation, neutralization, transfer and bonding steps. Vacuum [32] and electrophoretic [33] transfer are fast although they lack the simplicity of the original Southern capillary blotting procedure [34]. Using an alkali transfer buffer [35] rather than salt reduces the time for capillary blotting.

3.3. DNA hybridization and autoradiography

The membrane is hybridized with a ^{32}P-labelled DNA probe overnight. The random priming labelling method [36,37] provides high specific activities (above 10^9 cpm/μg). Alternatively ^{32}P may be incorporated by nick-translation [38]. Hybridization in a solution of salts and filter-blocking reagents is carried out at 65°C. Alternatively, if formamide is added the temperature is reduced (to 42°C for 50% formamide). The membrane should be prehybridized for 1–2 h in the same solution before the addition of labelled probe. Dextran sulphate, which accelerates the reaction but increases background, may be added to 10% in the hybridization solution [39]. The membrane is washed and exposed to X-ray film.

3.4. Polymerase chain reaction

The PCR can be used to amplify DNA targets by factors of up to 10^8, thus simplifying subsequent analysis. A pair of primers, flanking the target sequence, are synthesized. Primers that are 20–24 bp in length are frequently used since they are likely to be unique (although a database search should be made routinely). The reaction containing target DNA, primers, deoxynucleotide triphosphates, buffer and Taq DNA polymerase (from a thermophyllic organism) is carried out in a thermal cycler. Following initial denaturation, 25–35 cycles of annealing, synthesis and denaturation are required to amplify a single-copy sequence from human DNA. The reaction should be optimized with regard to Mg^{2+} concentration. The annealing temperature is calculated from the primer composition.

The exponential chain reaction catalysed by DNA polymerase produces a fragment that can be visualized on an EtBr-stained gel. This fragment can be restriction-digested for RFLP typing. Primers can be designed for allele specific amplification or the product probed with ASOs for mutation detection. Microsatellite markers can be typed by end-labelling one primer followed by polyacrylamide gel electrophoresis (PAGE). Many compilations of protocols have been

published. We find PCR protocols [40] particularly useful since it presents numerous applications and variations of the basic procedure and useful information for setting up a PCR lab.

3.5. Polyacrylamide gel electrophoresis

A variety of detection procedures use vertical PAGE. These gels, generally 6–8% polyacrylamide with 1 part in 20 of bisacrylamide for cross-linking, are prepared in 0.090 *M* Tris–borate, 2 m*M* EDTA (pH 8.0). Polymerization is catalysed by N,N,N′,N′-tetramethylethylenediamine (TEMED), using 35 μl per 100 ml of acrylamide solution, and accelerated with ammonium persulphate which is present at a final concentration of 0.07%. The gels are poured between glass plates that are separated by 0.4-mm spacers, sealed with yellow electrical tape and clamped. A standard length is 50 cm but other sizes may be run.

Routine denaturing gels that are used for microsatellite typing as well as sequencing and CCM contain 7 *M* urea. The sample is loaded in 95% formamide containing the tracking dyes 0.05% xylene cyanol FF and 0.05% bromophenol blue. These gels are run at a constant power, of about 50 W for a 50 cm × 20 cm gel (voltages around 2000V), for about 1 h at temperatures of 45–50°C produced by ohmic heating. The gels are then dried and fragments ranging from 40 to 300 bp detected by autoradiography. Denaturing gradient gels have a gradient varying from 0 to 100% denaturant (7 *M* urea and 40% formamide). These gels are maintained in a constant temperature bath at 60°C and run overnight at low voltages (16–18 h at about 70 V). Separated fragments can be visualized by staining with EtBr following electrophoresis.

Non-denaturing gels are used for typing SSCP markers and heteroduplex analysis. Samples are diluted into formamide dye solution to denature DNA into single strands for SSCP analysis. The gels are run at 4°C or room temperature overnight at low power, to prevent temperature fluctuations. Glycerol added to 10% sometimes allows the resolution of fragments that otherwise migrate together. For heteroduplex analysis the native, reannealed PCR product is loaded directly onto a non-denaturing 12% high-resolution 20-cm minigel and electrophoresis is carried out at 150 V for 20 min.

4. CLINICAL APPLICATIONS

A direct test for mutation should be carried out whenever possible. This simplifies both the laboratory testing and the counselling of the at risk family. Molecular characterization of the familial mutation is a necessary prelude to direct diagnosis of the disease mutation.

4.1. Diagnosis by family studies

DNA typing of linked markers can be used to determine the genetic risk of inheriting a disease-causing gene. These indirect tests can be very powerful and highly sensitive but lack the precision of a direct test for a known mutation. The counselling of families is more complex since probability of disease rather than an exact diagnosis is obtained. Maximum likelihood risks and range of risk based on 95% confidence intervals (for recombination fractions) both involve sophisticated statistical concepts not familiar to most families and hard to understand.

4.1.1. Determining who to test

Generally it is necessary to test at least one affected individual (proband) and additional relatives, although in some families it is possible to obtain the required information without testing a proband. Determining who to test depends primarily upon the mode of transmission of the disease.

4.1.1.1. Dominantly inherited disease. A typical situation is where one parent is affected and wants to know whether a child or foetus is carrying the disease gene. Examples include myotonic dystrophy, adult polycystic kidney disease and Huntington disease. DNA typing addresses a number of questions. Does the affected parent have different allelic forms of linked, ideally flanking markers? Different, segregating markers are required to distinguish between the two homologous chromosomes only one of which carries the disease gene. Does the genotype of the unaffected spouse at these marker loci allow an unambiguous determination of the markers transmitted from affected parent to offspring? This is the case if the spouse has only one form of the marker on both chromosomes or has different forms of the marker that are not present in the affected parent. If both parents are segregating the same marker pair then the mating will be only partically informative (*i.e.* informative in only half of the possible types of offspring). Finally which allele marks the chromosome carrying the disease gene in the affected parent? Determination of the coupling phase requires testing of the affected spouse's parents, one of whom will be affected and will transmit a particular marker allele together with the disease gene. Fig. 4 illustrates possible outcomes of such testing.

Frequently it may not be possible to test one parent of the affected individual (either I-1 or I-2 in Fig. 4), due perhaps to death, marriage breakdown, lack of consent, etc. It may, however, still be possible to obtain the necessary information. For example, if I-1 was not tested, marker system E still provides the necessary coupling information. If I-2 was not tested marker system D provides the necessary information.

If neither parent is available for testing it may still be possible to determine likely coupling phase by testing siblings of the affected parent or prospective parent. If an affected sibling is homozygous for a marker that is heterozygous in the affected prospective parent then the more likely coupling phase is with that

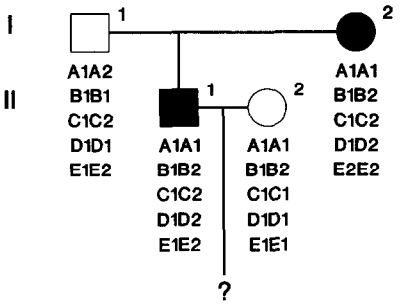

Fig. 4. Autosomal dominant inheritance: RFLP typing. Marker system A is not useful since the affected parent is not segregating at this locus (*i.e.* both alleles are identical). Marker system B is partially informative, the paternal contribution will be known only for children with genotypes B1B1 and B2B2. Marker system C is not useful since the linkage (coupling) phase in the affected parent cannot be determined. Marker system D is fully informative provided that I-1 is the father of II-1. Marker system E is fully informative and non-paternity does not affect the determination of coupling phase.

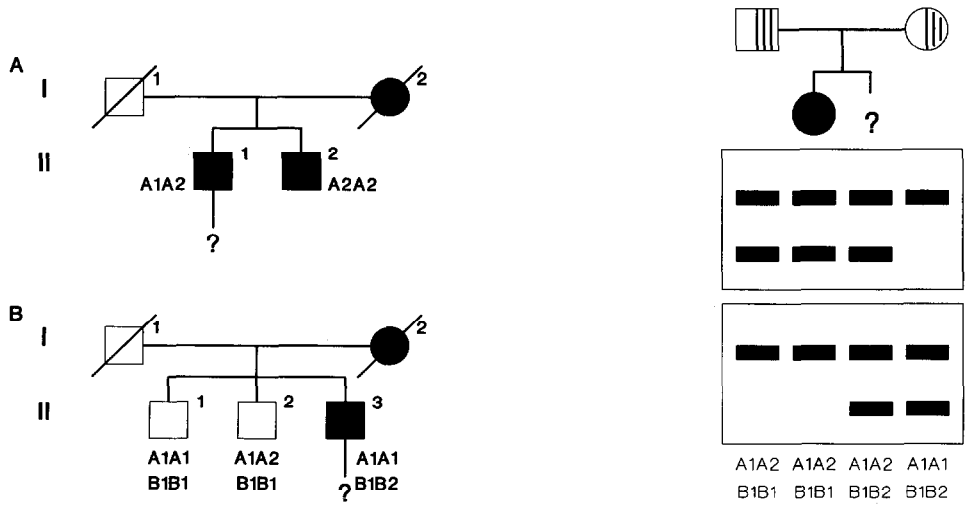

Fig. 5. (A) Autosomal dominant inheritance: RFLP typing. II-2 inherited the disease allele together with the A2 marker. Hence II-1 has also inherited the disease allele on the chromosome with the A2 marker provided that the disease and marker alleles have not recombined. (B) Autosomal dominant inheritance: RFLP typing. The genotype of II-1 indicates that each parent has a normal chromosome with the haplotype A1B1. The genotype of II-2 indicates that the unaffected parent also has a chromosome with the haplotype A2B1. The genotype of II-3 indicates that the fourth and last haplotype for the parental couple is A1B2 and this haplotype is inherited together with the disease allele (assuming no recombination between the markers within the haplotype and between the haplotype and the disease allele). Note the B marker system alone does not allow the assignment of linkage phase.

Fig. 6. Autosomal recessive inheritance: RFLP typing from Southern blots. The father is segregating for the A marker system while the mother is homozygous, therefore this mating is informative for the A system. Similarly the mother is segregating for the B marker and the mating is fully informative. The foetus has inherited the mutant allele from the father but the normal allele from the mother and is therefore a carrier. For markers within or close to the mutant gene the chance of misdiagnosis due to recombination is generally less that 1%.

marker (Fig. 5A). If there are no affected siblings it may still be possible to assign likely coupling phase by reconstructing haplotypes (Fig. 5B).

4.1.1.2. Autosomal recessive disease. A typical situation is the couple with one affected child who wish to determine whether a subsequent pregnancy will be affected. Examples include phenylketonuria and cystic fibrosis. If the parents do not carry any of the known mutants then the aim of DNA typing is to detect an informative marker for each parent (Fig. 6). As the markers lie within or close to the mutant gene the chance of recombination disrupting linkage phase is a fraction of 1%, and the markers transmitted together with the disease allele to the affected child have a probability close to 100% of being coupled to the disease allele in the parent (the overall average rate of recombination is 1% per million base pairs).

4.1.1.3. Sex-linked recessive disease. Typical situations include determining whether a woman with an affected child will have a subsequent affected pregnancy, and, determining whether female relatives are carriers and therefore themselves at risk for having affected children. Diagnosing a subsequent pregnancy in a woman with one affected child is most straightforward when there is a family history of the disease (Fig. 7). In this case the woman is an obligate carrier (because she has an affected child and at least one affected brother or maternal uncle). The linkage phase for an informative marker can be determined by testing her father who transmitted a non-disease allele to his daughter.

Families where the affected child is the first occurrence of the disease are more difficult to assess, since the affected child may be the consequence of a new or recent mutation. The severity of the disease itself is related to the likelihood of the mother being a carrier. The carrier risk for the mild Becker muscular dystrophy is

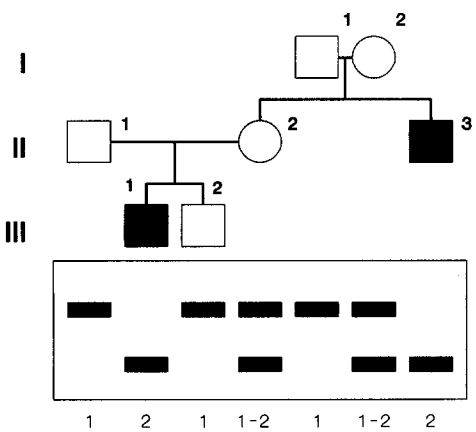

Fig. 7. X-linked recessive inheritance: RFLP typing from Southern blot. Individual II-2 has both an affected son and an affected brother indicating that both I-2 and II-2 are carriers. The disease is coupled to the 2 marker allele that II-2 inherits from her mother. Thus subsequent pregnancies of II-2 could be diagnosed prenatally using this marker.

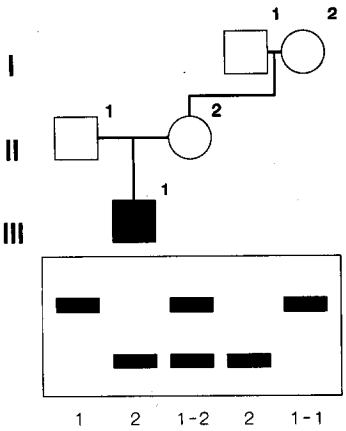

Fig. 8. X-linked recessive inheritance: RFLP typing from Southern blot. Individual II-2 is the mother of an isolated case of this X-linked disease. RFLP typing indicates that the marker allele 2 that was transmitted to the affected son was inherited from the normal maternal grandfather. In this situation II-2 may be carrier, or may be normal or may be mosaic. RFLP typing in subsequent pregnancies of II-2 is limited to making an exclusion diagnosis if the foetus inherits marker allele 1 and is unable to determine risk if the foetus inherits marker allele 2.

90% [41] while the carrier risk for the genetically lethal Duchenne muscular dystrophy (DMD) is 2/3 [42]. This risk is modified by both the extended family structure and by DNA typing. This information is taken into account, using Bayesian probability methods [43], to obtain the appropriate risks for family members. For example if the mother of the affected boy has many unaffected brothers then the chance that the maternal grandmother is a carrier is significantly reduced. This diminishes the carrier risk to the maternal aunts of the proband.

If the affected boy inherits a marker derived from the maternal grandfather (Fig. 8) then the mother either received a mutant chromosome from her father or the paternal chromosome underwent mutation in the mother before being transmitted to her affected son. Mutation in the mother may occur early in gestation leading to mosaicism [44], *i.e.* the mother would have cells with two types of paternal chromosome: normal and mutant. This means that peripheral leucocyte DNA from the mother may lack the mutation that is detected in her son while her risk of having an affected boy could be anywhere from 0 to 50%.

4.1.2. Detecting informative markers

There are for many disorders a large number of different markers that may be typed. A number of factors should be considered in determining which markers to type. A prime factor is the recombination fraction, between the marker and disease gene, which should be as small as possible. The aim of typing should be to identify flanking informative markers. Having markers on either side of the gene will restrict errors due to recombination to the more rarely occurring double

recombinant. The informativeness of the marker should be high. A marker that is heterozygous in 50% of the population is better than one that is heterozygous in 25%. The information content of a marker is measured by either the heterozygosity or the polymorphism information content (PIC). The PIC is the chance that any meiotic event will be informative. Heterozygosity and PIC value are identical for sex-linked markers since meiotic events in heterozygous females are always informative. These values should ideally be derived from the same racial group as the family being investigated. When multiple markers are being tested it is better to use markers that show no or limited linkage disequilibrium. An individual who is homozygous, non-informative, for a first marker will necessarily be homozygous for a second marker that is in complete linkage disequilibrium with the first. Conversely a marker known to be in linkage disequilibrium with a disease allele should be typed, since it is more likely to be informative. Expense should be minimized, where possible, by choice of restriction enzyme and by using enzymes that produce blots that can be probed either simultaneously or sequentially with multiple probes.

4.1.3. Calculation of genetic risk

DNA typing for hereditary disease is unlike any other laboratory test since it is not possible to provide a diagnosis based on the analysis of a sample from a single individual (in those cases where the mutation itself cannot be detected). Typing involves the extended family and the laboratory results are used for pedigree and segregation analysis. In many cases risk calculations are simple whereas in other cases they can become very complex requiring the use of Bayesian probability calculations, a treatment of which is beyond the scope of this article. We subscribe to the view that the calculation of the final risk is an integral part of the laboratory investigation [45] and that it is the responsibility of the laboratory to request specimens from those additional family members that may lead to the provision of a more precise risk. For these reasons the laboratory director should be trained in genetics and it is most efficient for the laboratory to liaise with the genetic consultant who is counselling the family.

4.1.4. Sources of error

A primary consideration for the laboratory is whether the genetic diagnosis is correct. Whereas many laboratory tests are designed to confirm disease, DNA typing of RFLPs is designed to give information about the segregation of regions of chromosomes in families. If the wrong chromosome region is being investigated in the laboratory the result will be meaningless. While the physician referring the sample bears the responsibility for the diagnosis, the laboratory should be sensitive to this issue and be aware of those conditions most likely to create problems.

Good laboratory and sample acquisition protocols should be established to minimize the chance of laboratory errors such as mislabelling. Non-paternity is a

frequent source of error and may require resampling to exclude possible laboratory error. Paternity testing is not routinely carried out by most diagnostic laboratories and such testing is a source of potential ethical problems since it is not part of the requested procedure. This again is an area where good liaison between laboratory and genetic counsellor is valuable. The final source of error, which is unavoidable, is due to recombination. The use of close and ideally flanking markers can frequently reduce this source of error to the region of 1%.

It may be argued that one should never give a zero risk. For example, if the laboratory has shown that a foetus does not carry a familial deletion causing DMD it has not excluded a possible new mutation which may have a 1 in 15 000 risk of occurrence.

4.2. Diagnosis by direct detection of mutation

The direct diagnosis of mutation has the advantages of simplicity and reliability. All new cases should be tested for those mutations known to be common for the disease in question and the ethnic background of the family.

4.2.1. Detection of known mutations

The known mutations are disease-specific but the procedures for their detection are general as outlined below.

4.2.1.1. Major structural rearrangements. Deletions, duplications or insertions alter the sizes of restriction fragments and Southern blots can therefore be used for direct detection of disorders known to be caused by major structural rearrangements.

The majority of homozygous α-thalassaemia, associated with hydrops foetalis, is due to a common deletion allele. Southern blot analysis using an α-globin probe gives no hybridization signal, while an abnormal-size fragment is seen using a ζ-globin probe [46]. Rapid PCR diagnosis has been developed [47] using a pair of primers that amplify a 136-bp region of the α-globin gene cluster between the ψα and α2 globin genes. This segment was chosen because it lies outside the repeat regions of the α-globin gene cluster and is deleted in α-thalassaemia. A 136-bp fragment is seen in the heterozygote or normal homozygote but no fragment is produced from homozygous α-thalassaemia DNA, so a control amplification is run simultaneously.

Deletions also account for a large proportion of the mutations in the steroid sulphatase (STS) gene which causes X-linked ichthyosis and the DMD gene [48,49]. Multiplex PCR methods have been developed to detect deletions in these large genes [50,51]. Two groups of 9 sets of amplification primers detect over 98% of deletions in the DMD gene [52]. This method will establish the status of a male foetus in families known to have a deletion. It cannot be used to determine the carrier status of females since their normal X will be amplified. Obligate deletion carriers have normal banding patterns.

Major structural rearrangements, however, account for a relatively low percentage of the mutations in most single-copy genes thus limiting Southern blot analysis as a method of direct mutation detection. The underlying molecular defects in β-thalassaemia, haemophilia A and B, and familial hypercholesterolaemia, for example, have been well characterized [53–56]. Although a deletion β-thalassaemia is common among Asian Indians [57], deletions generally represent [58] less than 10% of the overall mutations in these genes.

4.2.1.2. Altered restriction site. Disease point mutations that alter the recognition sequence of a restriction endonuclease can generally be detected by Southern blot analysis, especially if PCR amplification is used. This is especially useful for common diseases such as sickle cell anaemia which is caused by a single mutation that destroys an *MstII (OxaNI)* site. A PCR procedure has been developed where a 294-bp region of the β-globin gene around the *OxaNI* site is amplified [47]. After digestion with *OxaNI* and PAGE the amplified fragments can be visualized by EtBr staining. Only full-length, 294-bp DNA will be seen from a sickle cell homozygote whereas the normal β-globin gene will yield two fragments of 191 and 103 bp [47].

Only a small fraction of point mutations (5–10%) alter restriction sites. Reports of single base pair mutations within genes causing human genetic diseases suggest that CG dinucleotides are hypermutable sites. Restriction enzymes that contain CG dinucleotides in their recognition sequence (*e.g. TaqI, RsaI* and *MspI*) and therefore particularly useful for detecting point mutations [59]. Five different single-nucleotide substitution mutations that alter a *TaqI* site have been found in haemophilia A [54]. Point mutations that alter RE recognition sequences account for over 50% of the common β-thalassaemia alleles and involve over twelve different REs. These mutations can be detected by PCR amplification and restriction with appropriate endonucleases in separate reactions. After agarose electrophoresis and EtBr staining one can determine by the banding pattern if one is dealing with one of the common alleles that alters a restriction site [60].

The majority of mutations do not alter restriction sites. It is, however, possible to exploit the sequence difference between a normal and a mutant allele to introduce a restriction site into the PCR product of one of these alleles, by using a modified PCR primer with a base substitution adjacent to the point mutation [61]. This PCR-mediated site-directed mutagenesis approach has been used to detect the common cystic fibrosis (CF) mutation [62]. A novel *MnlI* site was introduced into the PCR product of the normal allele while the product of the CF allele lacked this engineered site, because of the sequence difference due to the CF mutation [62].

4.2.1.3. Allele specific oligonucleotides. ASOs are particularly useful for diagnosis of disorders that are well characterized at the molecular level. Base pair substitutions, deletions or insertions of 1–4 bp that do not alter an RE recognition sequence account for about half of the β-thalassaemia mutations. These mutations can be detected using oligonucleotide probes specific to these common

mutations [63]. The spectrum of β-thalassaemia mutations has been determined in Greek, Italian, Turkish, Spanish, Asian Indian and Chinese populations. The different β-thalassaemia mutant alleles have been found to be population-specific and for each population a small number (less than twelve) of mutations account for over 90% of the mutant alleles. Analysis of amplified β-globin DNA from carrier parents using ASO probes will allow determination of the mutant allele carried by each parent. Any unknown parental mutation can be determined by direct sequencing of the amplified β-globin gene [64]. Once the mutations are known for both parents, prenatal diagnosis can be carried out to determine the status of the foetus using the ASO probes specific for the family.

4.2.1.4. Allele specific amplification. Allele specific amplification has been developed for the analysis of mutations and DNA sequence polymorphisms [17]. It is particularly useful for diagnosis where one common mutation is responsible for a large proportion of mutant genes. A 3-bp deletion of codon 508 (Δ508) is responsible for 70% of CF mutant genes of Northern European ancestry [65,66]. Genomic DNA is amplified in one reaction with a set of primers that corresponds to the wild type allele and separately with a set of primers homologous to the Δ508 mutation. Another genomic sequence is concurrently amplified as an internal control. The amplified products are then electrophoresed on an agarose gel. The presence of a band with the wild type primer set only indicates that the individual is negative for the Δ508 mutation. The presence of the band with the Δ508 primer set only implies that the individual is homozygous for the Δ508 mutation. A positive result with both reactions means heterozygosity for the Δ508 mutation. This is a useful general method that may also be used to detect single base changes. It is amenable to automation through the use of fluorescent-labelled primers and colour complementation [67]. Oligonucleotides constructed with a 3'T mismatch make effective primers and should be avoided.

4.2.1.5. Allele specific ligation. Two oligonucleotides that are adjacent in the genome will ligate when annealed to template DNA. A mismatch at one base involved in the ligation will eliminate this reaction [68]. This alternative protocol for mutant allele detection has good potential for automation. A reporter group on one oligonucleotide can be joined to a ligand that facilitates isolation on the other oligonucleotide.

4.2.1.6. Heteroduplex analysis. Heteroduplexes are formed during the final cycles of PCR amplification from those samples where the amplified fragment differs in sequence between the two homologous chromosomes. Size differences between alleles may result in spurious migration on native polyacrylamide gels [69]. Heteroduplex analysis can be a useful tool for heterozygote detection when the mutant allele differs in length from the normal allele by a few base pairs (Fig. 9). Examples include Tay–Sachs disease (TSD), where a 4bp insertion in the HEXA gene is responsible for approximately 70% of TSD in Ashkenazi Jews [70] and CF [71], where the Δ508 deletion accounts for 70% of mutant alleles in Caucasians. Heterozygotes for these TSD and CF mutations have an extra, more

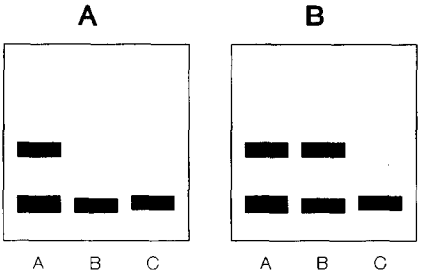

Fig. 9. (A) Heteroduplex analysis: detection of CF carriers of the Δ508 mutation. Lane A: heterozygous carrier of the Δ508 mutation. Lane B: affected homozygote for the Δ508 mutation. Lane C: homozygote for the normal allele. The heterozygous carrier DNA forms four types of double-stranded PCR products, mutant homoduplexes, normal homoduplexes and two heteroduplexes. Only the slower migrating heteroduplexes are well resolved from the faster migrating homoduplexes. The figure is a representation of an EtBr-stained polyacrylamide gel. (B) Heteroduplex analysis: detection of CF homozygotes for the Δ508 mutation. The samples are mixed with an equal amount of known normal homozygote PCR product. Lane A: heterozygous carrier of the Δ508 mutation. Lane B: affected homozygote for the Δ508 mutation. Lane C: homozygote for the normal allele. For prenatal diagnosis samples that produce only homoduplexes can be distinguished by mixing with an equal amount of known normal homozygote PCR product. A homozygous affected foetal sample will produce heteroduplexes while a homozygous normal sample will be unchanged. The products can be visualized by EtBr staining of the non-denaturing polyacrylamide gel which is represented by the figure.

slowly migrating, heteroduplex band formed by the annealing of a normal DNA strand with a mutant DNA strand (Fig. 9). Heteroduplexes and homoduplexes can often be resolved by electrophoresis on 4% NuSieve agarose minigels [72].

4.2.2. Screening for new mutations

Familial mutations of undefined sequence will continue to be detected indirectly by family studies of linked RFLPs. We anticipate, however, that the advantages inherent in the direct detection of mutation will result in mutation screening by some diagnostic laboratories. Presently we believe the following four techniques deserve serious consideration for laboratories contemplating screening to define new mutations (Fig. 10).

4.2.2.1. Denaturing gradient gel electrophoresis. DGGE, using a polyacrylamide gel containing a urea–formamide gradient, will frequently resolve mutant from normal homoduplexes. It is easier to detect heteroduplexes, where a single base pair mismatch has a greater, destabilizing effect [19]. β-Thalassaemia mutations have been detected by annealing mutant genomic DNA with single-strand, labelled, wild type probe prior to DGGE [73]. The DGGE method is simplified by gel analysis of PCR products that have a GC clamp [74] which provides a stable domain, thereby increasing the fraction of mutants that can be detected since the mutant will then lie in an early-melting domain. This procedure has been effective in detecting mutations in the haemophilia A gene [75].

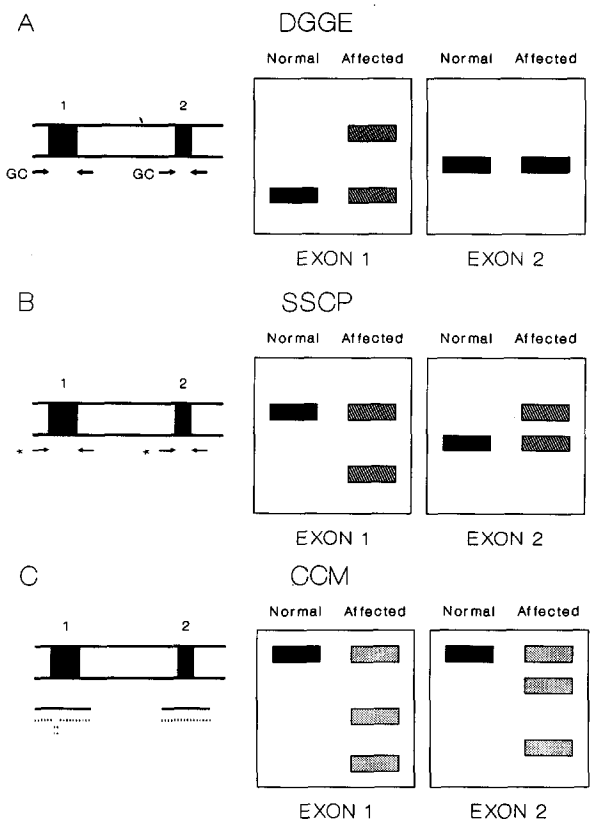

Fig. 10. Gel techniques for detecting mutant sequence variation. For illustrative purposes the sample is considered to be derived from an individual affected with a recessive disease. One allele has a 2-bp deletion in exon 1 while the other allele has an A → T substitution in exon 2. All figures represent autoradiographs of dried polyacrylamide gels. (A) Denaturing gradient gel electrophoresis: DGGE. PCR products are produced using a 40-bp GC clamp on one member of each primer pair. These products are electrophoresed as double-stranded molecules on a polyacrylamide gel into a gradient of denaturant. The exon 1 defect results in decreased stability and a novel band while the exon 2 defect remains undetected. (B) Single-strand conformational polymorphism: SSCP. PCR products are produced using an end label on one member of each primer pair. The double-stranded products are denatured with alkali and electrophoresed as single-strand products on a non-denaturing polyacrylamide gel. Conformers are observed for both the exon 1 and exon 2 mutations. (C) Chemical cleavage of mismatched bases: CCM. The PCR products are mixed with uniformly labelled normal sense strand and chemically treated to cleave at base mismatches. The products are electrophoresed on a standard, denaturing, sequencing polyacrylamide gel. The sizes of the novel products generated for both exon 1 and exon 2 indicate the position of the mutation.

4.2.2.2. Chemical cleavage of mismatched bases. The bases cytosine and thymine are particularly sensitive to cleavage by hydroxylamine and osmium tetroxide when present as single base mismatches [20]. This CCM method has been used to detect point mutations in haemophilia B [76]. It should always detect single

base pair insertions or deletions or other small insertions [77]. When both wild type and mutant sense and anti-sense strands are used as probes [78] all mutations should be detectable.

4.2.2.3. Single-strand conformational polymorphisms. Small single-strand fragments of DNA that differ by a single base may assume different conformations that give rise to differing mobilities on native, non-denaturing polyacrylamide gels [18]. Such differences can be readily screened using the PCR [79].

4.2.2.4. Genomic sequencing. Mutation detection by any of the above procedures can be followed up to define the exact nature of the mutation by direct genomic sequencing [58].

4.2.3. Sources of error

Most diagnoses by direct mutation detection rely on amplification of DNA by the PCR. One advantage of this technology is its sensitivity. However, this can also be a major disadvantage as the PCR will detect trace contamination. The latter can be in any of the reagents. Minimal precautions require a negative control consisting of all PCR reagents without the DNA template. Segregating a set of pipettors for the PCR and the use of positive displacement pipettors or tips with an aqueous impermeable barrier is recommended.[a]

Of greatest concern in DNA typing for prenatal diagnosis is the possibility of maternal contamination of chorionic villi. If the CVS is not carefully cleaned of all maternal decidua, maternal sequences can be amplified, resulting in possible erroneous diagnosis of the foetal status. It has been estimated that if maternal contamination is kept to less than 5%, amplification of foetal sequences will predominate, allowing correct assignment of foetal status [50].

5. CONCLUSION

DNA typing in hereditary disease is a recent innovation. Points to consider when establishing a DNA typing laboratory [40,80] and for DNA banking [45,81] have been published. The field is one of rapid change. For any single hereditary disease there is a progression from gene location to gene identification to characterization of mutant alleles. Correspondingly DNA analysis progresses from typing closely linked markers to typing markers within the disease gene to direct detection of disease mutations, often within a few years. At each stage the laboratory is confronted with a variety of choices. There may be numerous closely linked markers available for typing as well as additional new markers under development. Choice may be determined by linkage distance, informativeness

[a] Continental Laboratory Products (San Diego, CA, USA) sells a pipette tip that contains a barrier of high-density polyethylene mixed with a cellulose polymer. This creates an aqueous impermeable barrier and prevents deposition of aerosols within the pipettor barrel.

and cost. Laboratories may develop a routine set of test markers and secondary set for those families that are uninformative with the routine set. With increasing knowledge this marker set will change. The genome data base [82] is a valuable source of genetic mapping and disease data for the human genome.

When a disease gene, together with a set of common mutations, is identified there is still a choice of detection methods. While any allele specific detection method is applicable, simplicity, cost, experience and avoidance of radioisotopes may play a role in determining the detection method of choice. The choice is sometimes largely determined by the nature of the disease gene. Diseases that are due to sex-linked genes are caused by many different mutations. Therefore each affected family is likely to have a novel mutation. In the case of the dystrophin gene the majority of mutations are deletions so that multiplex PCR amplification of exons is the method of choice. Also decisions have to be made concerning whether to set up procedures for screening for novel mutations and which procedures to select.

Genetic technology has seen a rapid growth in its biomedical applications. We except this rapid growth to be sustained, and even accelerate, as more is learned of the role of genetic predisposition to common diseases, that do not exhibit simple Mendelian inheritance.

6. ACKNOWLEDGEMENTS

We wish to thank Drs. P. J. Goodfellow and R. T. A. MacGillivray for their critical review of the manuscript and helpful comments.

REFERENCES

1 A. H. Sturtevant, *J. Exp. Zool.*, 14 (1913) 43–59.
2 J. Bell and J. B. S. Haldane, *Proc. R. Soc. B*, 123 (1937) 119–150.
3 D. Botstein, R. L. White, M. Skolnick and R. W. Davis, *Am. J. Hum. Genet.*, 32 (1980) 314–331.
4 Y. W. Kan and A. M. Dozy, *Lancet*, ii (1978) 910-912.
5 A. J. Jeffreys, V. Wilson and S. L. Thein, *Nature*, 314 (1985) 67–73.
6 Y. Nakamura, M. Leppert, P. O'Connell, R. Wolf. T. Holm, M. Culver, C. Martin, E. Fujimoto, M. Hoff, E. Kumlin and R. White, *Science*, 235 (1987) 1616–1622.
7 D. R. Higgs, S. E. Y. Goodburn, J. S. Wainscoat, J. B. Clegg and D. J. Weatherall, *Nucleic Acids Res.*, 9 (1981) 4213–4224.
8 S. T. Reeders, M. H. Breuning, K. E. Davies, R. D. Nicholls, A. P. Jarman, D. R. Higgs, P. L. Pearson and D. J. Weatherall, *Nature*, 317 (1985) 542–544.
9 J. L. Weber and P. E. May, *Am. J. Hum. Genet.*, 44 (1989) 388–396.
10 M. Litt and J. A. Luty, *Am. J. Hum. Genet.*, 44 (1989) 397–401.
11 C. Oudet, R. Helig and J. L. Mandel, *Hum. Genet.*, 84 (1990) 283–285.
12 A. H. Beggs and L. M. Kunkel, *Nucleic Acids Res.*, 18 (1990) 1931.
13 E. P. Economou, A. W. Bergen, A. C. Warren and S. E. Antonarakis, *Proc. Natl. Acad. Sci. U.S.A.*, 87 (1990) 2951–2954.
14 G. Zuliani and H. H. Hobbs, *Am. J. Hum. Genet.*, 46 (1990) 963–969.
15 B. J. Connor, A. A. Reyes, C. Morin, K. Itakura, R. L. Teplitz and R. B. Wallace, *Proc. Natl. Acad. Sci. U.S.A.*, 80 (1983) 278–282.

16 S. C. Kogan, M. Doherty and J. Gitschier, *N. Engl. J. Med.*, 317 (1987) 985–990.
17 C. R. Newton, A. Graham, L. E. Heptinstall, S. J. Powell, C. Summers, N. Kalsheker, J. C. Smith and A.F. Markham, *Nucleic Acids Res.*, 17 (1989) 2503–2516.
18 M. Orita, H. Iwahana, H. Kanazawa, K. Hayashi and T. Sekiya, *Proc. Natl. Acad. Sci. U.S.A.*, 86 (1989) 2766–2770.
19 L. S. Lerman and K. Silverstein, *Methods Enzymol.*, 155 (1987) 482–501.
20 R. G. Cotton, N. R. Rodrigues and R. D. Campbell, *Proc. Natl. Acad. Sci. U.S.A.*, 85 (1988) 4397–4401.
21 S. Gustafson, J. A. Proper, E. J. W. Bowie and S. S. Sommer, *Anal. Biochem.*, 165 (1987) 294–299.
22 K. Madisen, D. I. Hoar, C. D. Holroyd, M. Crisp and M. E. Hodes, *Am. J. Med. Genet.*, 27 (1987) 379–390.
23 M. Goossens and Y. W. Kan, *Methods Enzymol.*, 76 (1981) 805–817.
24 S. M. Thomas, R. F. Morena and L. L. Tilzer, *Nucleic Acids Res.*, 17 (1989) 5411.
25 S. A. Miller, D. D. Dykes and H. F. Polesky, *Nucleic Acids Res.*, 16 (1988) 1215.
26 J. M. Old, R. H. T. Ward, F. Karaglozu, M. Petrou, B. Modell and D. J. Weatherall, *Lancet*, ii (1982) 1413–1416.
27 J. R. W. Yates, S. Malcolm and A. P. Read, *J. Med. Genet.*, 26 (1989) 245–250.
28 L. J. Smith, R. C. Braylon, J. E. Nutkis, K. B. Edmundson, J. R. Downing and E. K. Wakeland, *Anal. Biochem.*, 160 (1987) 135–138.
29 S. E. Goelz, S. R. Hamilton and B. Vogelstein, *Biochem. Biophys. Res. Commun.*, 130 (1985) 118–126.
30 E. R. B. McCabe, S.-Z. Huang, W. K. Seltzer, and M. L. Law, *Hum. Genet.*, 75 (1987) 213–216.
31 J. Sambrook, E. F. Fritsch and T. Maniatis, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 2nd ed., 1989.
32 E. Olszewska and K. Jones, *Trends Genet.*, 4 (1988) 92–94.
33 N. Arnheim and E. M. Southern, *Cell*, 11 (1977) 363–370.
34 E. M. Southern, *J. Mol. Biol.*, 98 (1975) 503–517.
35 K. C. Reed and D. A. Mann, *Nucleic Acids Res.*, 13 (1985) 7207–7221.
36 A. P. Feinberg and B. Vogelstein, *Anal. Biochem.*, 132 (1983) 6–13.
37 A. P. Feinberg and B. Vogelstein, *Anal. Biochem.*, 137 (1984) 266–267.
38 P. W. J. Rigby, M. Dieckman, C. Rhodes and P. Berg, *J. Mol. Biol.*, 113 (1977) 237–251.
39 G. M. Wahl, M. Stern and G. R. Stark, *Proc. Natl. Acad. Sci. U.S.A.*, 76 (1979) 3683–3687.
40 M. A. Innis, D. H. Gelfand, J. J. Sninsky and T. J. White, *PCR Protocols*, Academic Press, San Diego, CA, 1990.
41 S. Wood, R. J. Shukin, S. L. Yong, D. Wilson, D. Kalousek and A. Chudley, *Clin. Genet.*, 31 (1987) 45–57.
42 J. B. S. Haldane, *J. Genet.*, 31 (1935) 317–326.
43 E. A. Murphy and G. A. Chase, *Principles of Genetic Counselling*, Year Book, Medical Publishers, Chicago, IL, 1975.
44 S. Wood and B. C. McGillivray, *Hum. Genet.*, 78 (1988) 282–284.
45 J. G. Davis, P. T. Rowley, C. D. Boehm, P. M. Conneally, M. E. Hodes, N. A. Holtzman, H. H. Kazazian, Jr., K. W. Klinger, P. R. Reilly, S. D. Rose and M. R. Seashore, *Am. J. Hum. Genet.*, 42 (1988) 781–783.
46 S. E. Antonarakis, H. H. Kazazian Jr. and S. H. Orkin, *Hum. Genet.*, 69 (1985) 1–14.
47 F. F. Chehab, M. Doherty, S. Cai, Y. W. Kan, S. Cooper and E. M. Rubin, *Nature*, 329 (1987) 293–294.
48 A. Ballabio, G. Sebastio, R. Carrozzo, G. Parenti, A. Piccirillo, M. G. Persico and G. Andria, *Hum. Genet.*, 77 (1987) 338–341.
49 M. Koenig, E. P. Hoffman, C. J. Bertelson, A. P. Monaco, C. Feener and L. M. Kunkel, *Cell*, 50 (1987) 509–517.
50 J. S. Chamberlain, R. A. Gibbs, J. E. Ranier, P. N. Nguyen and C. T. Caskey, *Nucleic Acids Res.*, 16 (1988) 11 141–11 156.
51 A. Ballabio, J. E. Ranier, J. S. Chamberlain, M. Zollo and C. T. Caskey, *Hum. Genet.*, 84 (1990) 571-573.

52 A. H. Beggs and L. M. Kunkel, *J. Clin. Invest.*, 85 (1990) 613–619.
53 S. H. Orkin and H. H. Kazazian, Jr., *Ann. Rev. Genet.*, 18 (1984) 131–171.
54 S. E. Antonarakis and H. H. Kazazian, Jr., *Trends Genet.*, 4 (1988) 233–237.
55 F. Giannelli, P. M. Green, K. A. High, J. N. Lozier, D. P. Lillicrap, M. Ludwig, K. Olek, P. H. Reitsma, M. Goossens, A. Yoshioka, S. Sommer and G. G. Brownlee, *Nucleic Acids Res.*, 18 (1990) 4053–4059.
56 D. W. Russel, V. Esser and H. H. Hobbs, *Arteriosclerosis*, 9 (Suppl. I) (1989) 8–13.
57 S. Wood, M. Daya, J. E. Allanson, L. Kirby, R. Coupland and G. R. Gray, *Can. J. Genet. Cytol.*, 26 (1984) 296–301.
58 S. Langlois, J. J. P. Kastelian and M. R. Hayden, *Am. J. Hum. Genet.*, 43 (1988) 60–68.
59 D. N. Cooper and H. Youssoufian, *Hum. Genet.*, 78 (1988) 151–155.
60 H. H. Kazazian, Jr. and C. D. Boehm, *Blood*, 72 (1988) 1107–1116.
61 A. Haliassos, J. C. Chomel, L. Tesson, M. Baudis, J. Kruh, J. C. Kaplan and A. Kitzis, *Nucleic Acids Res.*, 17 (1989) 3606.
62 K. J. Friedman, W. E. Highsmith Jr., T. W. Prior, T. R. Perry and L. M. Silverman, *Clin. Chem.* 36 (1990) 695–696.
63 M. S. Ristaldi, M. Pirastu, C. Rosatelli, G. Monni, H. Ehrlich, R. Saiki and A. Cao, *Prenatal Diagn.*, 9 (1989) 629–638.
64 C. Wong, C. E. Dowling, R. K. Saiki, R. G. Higuchi, H. A. Erlich and H. H. Kazazian Jr., *Nature*, 330 (1987) 384–386.
65 J. R. Riordan, J. M. Rommens, B-S. Kerem, N. Alon, R. Rozmahel, Z. Grzelczak, J. Zielenski, S. Lok, N. Plavsic, J.-L. Chou, M. L. Drumm, M. C. Iannuzzi, F. S. Collins and L.-C. Tsui, *Science*, 245 (1989) 1066–1072.
66 C. R. Newton, M. Schwarz, C. Summers, L. E. Heptinstall, A. Graham, J. C. Smith, M. Super and A. F. Markham, *Lancet*, 335 (1990) 1217–1219.
67 F. F. Chehab and Y. W. Kan, *Proc. Natl. Acad. Sci. U.S.A.*, 86 (1989) 9178–9182.
68 D. Y. Wu and R. B. Wallace, *Genomics*, 4 (1989) 560–569.
69 C. M. Nagamine, K. Chan and Y.-F. C. Lau, *Am. J. Hum. Genet.*, 45 (1989) 337–339.
70 B. L. Triggs-Raine and R. A. Gravel, *Am. J. Hum. Genet.*, 46 (1990) 183–184.
71 J. Rommens, B.-S. Kerem, W. Greer, P. Chang, L.-C. Tsui and P. Ray, *Am. J. Hum. Genet.*, 46 (1990) 395–396.
72 S. Shore and R. Myerowitz, *Am. J. Hum. Genet.*, 47 (1990) 169.
73 R. M. Myers, T. Maniatis and L. S. Lerman, *Methods Enzymol.*, 155 (1987) 501–527.
74 V. C. Sheffield, D. R. Cox, L. S. Lerman and R. M. Myers, *Proc. Natl. Acad. Sci. U.S.A.*, 86 (1989) 232–236.
75 M. D. Traystman, M. Higuchi, C. K. Casper, S. E. Antonarakis and H. H. Kazazian Jr., *Genomics*, 6 (1990) 293–301.
76 A. J. Montandon, P. M. Green, F. Gianelli and D. R. Bentley, *Nucleic Acids Res.*, 17 (1989) 3347–3358.
77 D. W. Howells, S. M. Forrest, H.-H. M. Dahl and R. G. H. Cotton, *Am. J. Hum. Genet.*, 47 (1990) 279–285.
78 R. B. Pilz, C. Steglich and I. E. Scheffler, *J. Biol. Chem.*, 265 (1990) 8880–8886.
79 M. Orita, Y. Suzuki, T. Sekiya and K. Hayashi, *Genomics*, 5 (1989) 874–879.
80 R. Harris, R. Elles, D. Craufurd, A. Dodge, A. Ivinson, K. Hodgkinson, R. Mountford, M. Schwartz, T. Strachan and A. Read, *J. Med. Genet.*, 26 (1989) 219–225.
81 J. R. W. Yates, S. Malcolm and A. P. Read, *J. Med. Genet.*, 26 (1989) 245–250.
82 *Genome Data Base*, Johns Hopkins University, Baltimore, MD.