REVIEW

Demystified . . . Molecular pathology in oncology

J Crocker

....

In the past 10 years, molecular biology has found major applications in pathology, particularly in oncology. This has been a field of enormous expansion, where pure science has found a place in clinical practice and is now of everyday use in any academic unit. This demystified review will discuss the techniques used in molecular pathology and then provide examples of how these can be used in oncology.

....

O ver the past decade or so, molecular biology has found major applications in pathology, notably as applied to oncology. This has been a field of enormous expansion, where pure science has found a place in clinical practice and is now of everyday use in any academic unit. Presumably, we are looking at a series of methods which are, like immunocytochemistry 20 years ago, novel in the diagnostic laboratory. The beginnings of molecular pathology lie partly with the labelling of ribosomal DNA by "hot" RNA to demonstrate ribosomal DNA on chromosomes. Subsequently, there has been interest in the application of molecular methodology to disease states and this is described in this introductory review in relation to oncology. The purpose is to give the novice a guide to the chief techniques and their uses in this field.

What then does "molecular pathology" comprise? Molecules include not only nucleic acids but also all other chemical compounds (apart, by definition, from uniatomic elements!), including water, the chemistry of which is far less well understood than that of DNA, and yet "molecular biology" has now become synonymous with the science of nucleic acid mechanisms and their applications. Therefore, we are constrained by the concept that molecular pathology is the study of disease at the fundamental level in relation to nucleic acid abnormalities or by means of techniques involving DNA or RNA analysis.

In this review in the *MP* "Demystified" series, I shall deal initially with the techniques of molecular pathology and then give a series of examples of how these are applied. It should be remembered by readers that molecular biology is shrouded in rather forbidding jargon but is essentially quite straightforward!

MOLECULAR TECHNIQUES
Nucleic acid extraction

Before the analysis of nucleic acids can be undertaken they must, of course, be extracted from the cells or tissues of interest! The ability to do this relates to whether the nucleic acid concerned is DNA or RNA because their lability is quite different. DNA is remarkably physicochemically stable in relation to RNA; this is perhaps not surprising (teleonomically speaking) in view of the required conservation of integrity of DNA. RNA has an almost "mercurial" function in the cellular environment and might be expected to be less stable.

The potential enemies to the extraction of the two types of nucleic acid lie in their catalytic enzymes. Thus, it is relatively easy to prepare DNA because the DNAses, inevitably present in tissues, can be removed by means of heat treatment or by contact with chemical agents, such as EDTA. Conversely, RNA is much more difficult to extract and prepare. This is because, unlike enzymes with DNAse activity, RNAses are very stable and difficult to remove from any extraction "soup". To look at RNA it is therefore necessary to overcome this problem. The first and, arguably, most obvious method is to remove RNA from any proteins and DNA present in a mixture of macromolecules. This can be done physicochemically. In addition, certain "lysis buffers", such as guadininium, inhibit RNases, although this is rather a violent treatment, which also damages many other molecules. This can later be overcome by removal of the guadininium molecule.

> "It is relatively easy to prepare DNA because the DNAses, inevitably present in tissues, can be removed by means of heat treatment or by contact with chemical agents"

Finally, for RNA extraction, contamination is a major problem because physical handling of the material involved may cause problems as a result of contamination by RNase from the fingers of the experimenter. Therefore, it is necessary to autoclave any equipment that is re-used and to be scrupulously clean in any RNA retrieval events. Arguably, this would also be worthwhile when using DNA techniques.

In practical terms, it is important to note that, although "message" is most often sought in analytical methods and rRNA is rarely of interest, rRNA has a practical use because it "binds" to mRNA and can be demonstrated indirectly on electrophoresis. This is because rRNA, unlike the

Correspondence to:
Professor J Crocker,
Department of
Histopathology,
Birmingham Heartlands
Hospital, Bordesley Green
East, Birmingham B9 5SS,
UK;
crockej@heartsol.wmids.
nhs.uk

Accepted 6 September
2002
....

....

Abbreviations: CGH, comparative genomic hybridisation; CIN, cervical intraepithelial neoplasia; EBV, Epstein-Barr virus; FACS, fluorescence activated cell sorter; FISH, fluorescent in situ hybridisation; HPV, human papillomavirus; ISH, in situ hybridisation; LOH, loss of heterozygosity; NHL, non-Hodgkin's lymphoma; PCR, polymerase chain reaction; PTLD, post-transplant lymphoproliferative disorders; RT, reverse transcriptase; SSCP, single strand conformation polymorphism analysis

bound mRNA, has a very close profile on electrophoretic methods of separation. In other words, on gels, very well defined bands are produced for rRNA, unlike mRNA.

Nucleic acids, once in the tube (or on the slide), can be used for analysis. But how available is such material? Well, it depends on the requirements of the case in hand. The uses of various separative and, especially, analytical methods for nucleic acid and chromosomal structures will be described below.

Nucleic acid analysis

Once the nucleic acids are available, how are they prepared for analysis?

The main methods of handling the molecules inevitably depend on traditional and novel methodologies. The classic means of assessing nucleotides is by measuring their optical density in ultraviolet light. They are assayed by a comparison of absorption at 260 nm, compared with 280 nm, which is the absorption wavelength of proteins. Other methods include column chromatography, selective precipitation, selective staining of RNA and DNA, and ultracentrifugation on caesium chloride layers. In particular, nucleic acids or their fragments are separated by gel electrophoresis.

Gel electrophoresis

Nucleic acids or their fragments are separated on a layer of gel (usually agarose). The gel is usually buffered at neutral pH, at whic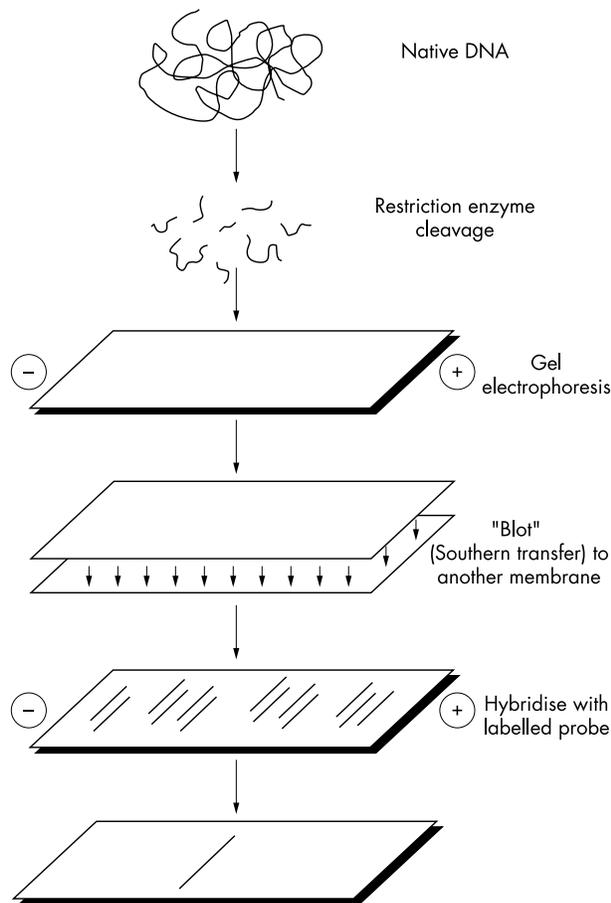h phosphate groups in nucleic acids are negatively charged and, of course, the molecules migrate towards the positive electrode. The speed that the molecules move depends on their sizes, and they separate into bands along the gel because the gel has "pores" that variably limit the rate of migration. The nucleic acid bands are then demonstrated by means of ultraviolet light or by reaction with certain chemical reagents (such as ethidium bromide).

Single strand conformation polymorphism analysis

Single strand conformation polymorphism analysis (SSCP) is a method that boasts a complex name but in reality is quite straightforward. It is used for the detection of neoplastic genes, especially when there is a very small mutation. It is a rapid technique and depends on the fact that small changes in single stranded DNA molecules can be detected by their changes in speed in gel preparations when compared with control DNA. The rationale of the method is that such alterations lead to a new molecular structure at the three dimensional level. This method can detect even single base changes. DNA samples are amplified by the polymerase chain reaction (PCR; see below) then denatured into single strands and separated on gels. This method allows screening for gene changes, followed by sequencing (also, see below).

A not dissimilar technique lies in "heteroduplex analysis". This depends upon the fact that if wild-type and mutated DNA are amplified together in the PCR, heteroduplices form, with one strand from each DNA type. As in SSCP analysis, these have a different speed of migration on gels, enabling separation and further analysis.

"Blot" methodology

In this group of techniques, nucleic acid molecules are moved from a separated gel to a transfer membrane and then analysed. Traditionally, the movement of molecules was that of "wick" absorption, although nowadays it is usually assisted by the application of an electrical field (figs 1, 2). The separated nucleic acid molecules can then be subjected to analytical methods, whose types are described below.

There are three main types of separation, namely:

(1) Southern blotting: in this method, DNA is denatured by restriction endonucleases and the resulting fragments are separated by gel electrophoresis. Their spatial distribution is preserved by "blotting" on to—for example, a nitrocellulose filter—and then demonstrated with either labelled RNA or DNA. Denaturing is important in these techniques because it uncoils and, where appropriate, splits nucleic acid strands one from another. This exposes regions of interest to chemical investigation. The denaturing agent(s) is(are) often present in the gel used for separation. A commonly used denaturing combination is polyacrylamide and urea. Historically, this method is named after its designer, Edwin Southern.
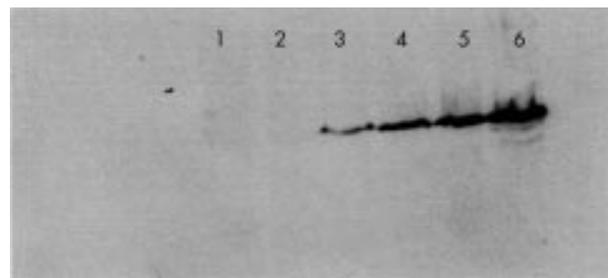

Figure 1 A schematic view of Southern blotting. The diagram shows how native DNA is cleaved, the restriction fragments are separated by electrophoresis, "blotted", and then labelled with probe by hybridisation. Artwork courtesy of Mrs S McGrory, Department of Medical Illustration, Birmingham Heartlands Hospital, UK.

Figure 2 An example of western blot analysis of expression of the Epstein-Barr virus encoded latent membrane protein 1 (LMP-1) in an inducible system. The Hodgkin's disease cell line, L428, was transfected with the LMP-1 gene and expression was induced by the addition of 6 mmol/litre cadmium chloride. Each lane represents increases in induction time, ranging from time zero (lane 1), through to 24 hours (lane 6). Figure courtesy of Dr K Baumforth, School of Health Sciences, University of Wolverhampton, UK.

(2) However, the scientific world cannot miss a catchy name and, accordingly, we now have the northern blot method, where mRNA is disaggregated on a denaturing gel, followed by blotting. The denaturing agents are characteristically formamide and formaldehyde and the RNA fragments are detected by hybridisation to nucleic acid probes.

(3) Western blotting is an analogous technique for the analysis of proteins.

In addition, we now also have "hybrid" terms such as "southwestern" and "northwestern" blotting, where DNA and RNA binding proteins are analysed, respectively.

In situ hybridisation

The ability to demonstrate nucleic acids in histological sections has certain advantages over liquid phase methods such as those described elsewhere in this review. Most importantly and rather obviously, the sequence of interest can be localised to a particular cell or tissue type. This applies to both DNA and RNA, although as discussed above, RNA is much more labile so that DNA is better suited to this series of chemical manipulations. Indeed, this raises yet again the concerns that must be expressed by all histopathologists, namely: there are so many potential artefacts introduced along the way during routine specimen processing.

The rationale behind in situ hybridisation (ISH) is that a probe sequence of nucleic acid complementary to that under investigation is applied to a tissue section or cell preparation. Conditions are manipulated such that the probe undergoes hydrogen bonding to its target. If the target nucleic acid is double stranded, then, of course, it must first be rendered into a single stranded configuration. This denaturation is usually performed thermally and the temperature needed for the process is called the "melting temperature".

> "There are so many potential artefacts introduced along the way during routine specimen processing"

Fortunately, both DNA and RNA can be detected in archival, formalin fixed, paraffin wax embedded tissues. However, delayed fixation can lead to loss of optimal localisation because autolytic enzymes tend to destroy nucleic acid integrity. This is especially the case for RNA. Another technical consideration is that formaldehyde is a so called "crosslinking" fixative and the tissue must be permeabilised to allow access of the probe. This is performed by incubation with—for example, endonucleases, HCl, or wetting agents.

In view of the comments above, it is perhaps no surprise that fresh frozen sections or cell preparations give stronger binding than fixed, wax embedded sections.

Consideration must be taken with regard to the nature of the probes used. These are usually cloned and either single or double stranded. Probes that are complementary to nucleic acid sequences in both directions along the molecule can be produced, so that both "sense" and "antisense" probes are available. Antisense probes will bind to the target sequence and the sense probes will act as negative controls because they will not bind. RNA probes are more delicate than DNA ones because of widespread cellular RNAse activity. Both single stranded and double stranded probes have their own distinct advantages over one another. Thus, single stranded molecules can be shown clearly because they hybridise more efficiently; however, double stranded probes form an arborising structure, leading to more signal amplification.

Probes can also be synthetic oligonucleotides, the production of which depends on knowledge of the target sequence, either from predetermined nucleotide information or from translation backwards from the amino acid sequence of the resultant protein. Hybridisation is especially efficient for oligoprobes that are rich in guanine and cytosine residues, which have especially strong hydrogen bonding.

Of course, to visualise the hybridisation site, we need a suitable labelling method. Traditionally, this was a radioisotope, such as ^{35}S, and was detected by means of autoradiography. However, the handling of such reagents is a problem, for obvious reasons and, accordingly, other, non-radioactive labelling methods have been sought. These include tagging with peroxidase, biotin, or digoxigenin. Digoxigenin is particularly useful because it is absent from human (and other mammalian) cells. These labels are visualised by means of chromogenic techniques.

Fluorescence in situ hybridisation

Fluorescence in situ hybridisation (FISH) is a specialised form of ISH and has become almost the province of the cytogeneticist. By using one or more probes, labelled with differently coloured fluorochromes, it is possible to demonstrate whole chromosomes or chromosomal loci in chromosome spreads or even in paraffin wax embedded archival tissue sections. This has revolutionised our analysis of tumour chromosomal abnormalities in tissue specimens at the molecular genetic level. The types of probes available for "chromosome painting" include:

- whole chromosome
- arm specific
- centromere specific
- gene sequence specific
- low copy sequences.

This technology allows us to look for chromosomal abnormalities in neoplastic cells and may be of academic or diagnostic value (fig 3). Figure 4 shows an example of the application of FISH. The assessment of the resulting preparations is usually performed today by means of digital image analysis, which enables objective analysis.

> "Hybridisation is especially efficient for oligoprobes that are rich in guanine and cytosine residues, which have especially strong hydrogen bonding"

Furthermore, chromosomes stained with two different fluorescent dyes can be identified by means of a fluorescence activated cell sorter (FACS) machine. The rationale here is that the FACS device separates the chromosomes by virtue of their degree of fluorescence, which in turn reflects their relative sizes.

The polymerase chain reaction

A drawback of blotting techniques is that a large number of cells is required to produce enough nucleic acid for hybridisation.

The PCR is based on a simple but ingenious concept that leads to the amplification of extremely small amounts of nucleic acid, so that they can be analysed. It relies on the presence of DNA polymerase to catalyse the reproduction of a specific DNA sequence; however, to enable DNA to undergo this process, it must be heated to a high temperature to split it into single strands. This temperature is between 94°C and 95°C, a temperature at which proteins such as the polymerases would be denatured. However, it is fortuitous that a bacterium (*Thermus aquaticus*), which lives in hot springs, has the enzyme in a form that still functions at high temperatures. This is now used in the PCR to great effect and is commonly known as Taq polymerase.

In the PCR, DNA from patient tissue is processed in a "thermal cycler", where the polymerase is repeatedly allowed to act on the now single stranded nucleic acid at the reaction temperature, followed by cooling. This allows the sequence of interest to be doubled in each cycle, thus leading to very rapid amplification, such that after 30 cycles there has been an
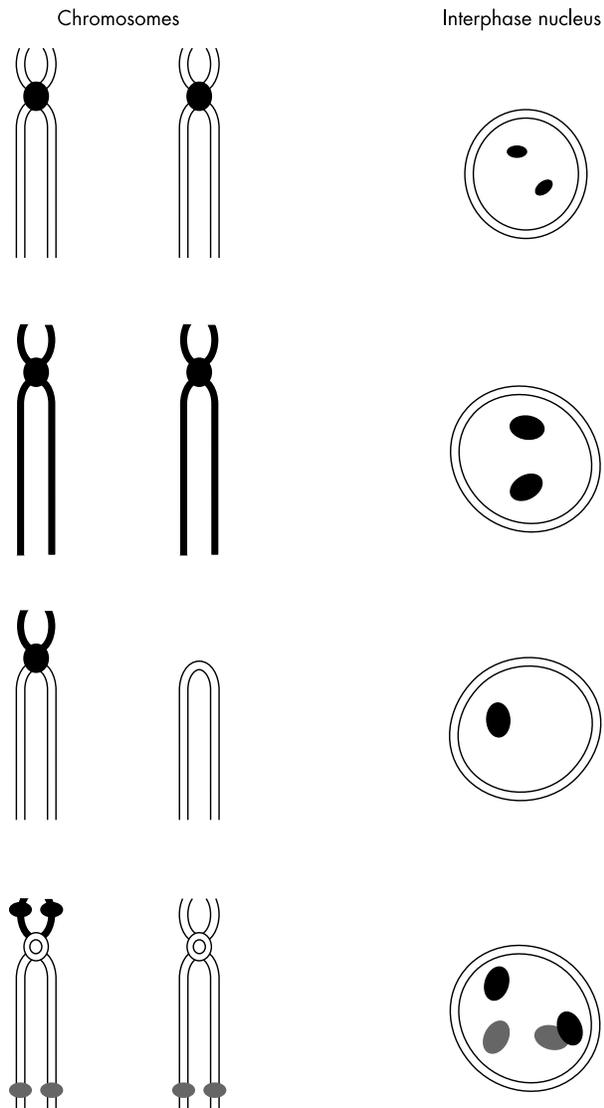
Chromosomes   Interphase nucleus

Figure 3 A schematic illustration of fluorescent in situ hybridisation, showing how metaphase and interphase chromosomal structures are localised. The top part of the diagram shows centromeric labelling in both types of preparation and below that whole chromosomes are labelled. Next, short arm deletion is depicted, and at the bottom there is an illustration of the use of two gene probes in combination. Artwork courtesy of Mrs S McGrory, Department of Medical Illustration, Birmingham Heartlands Hospital, UK.
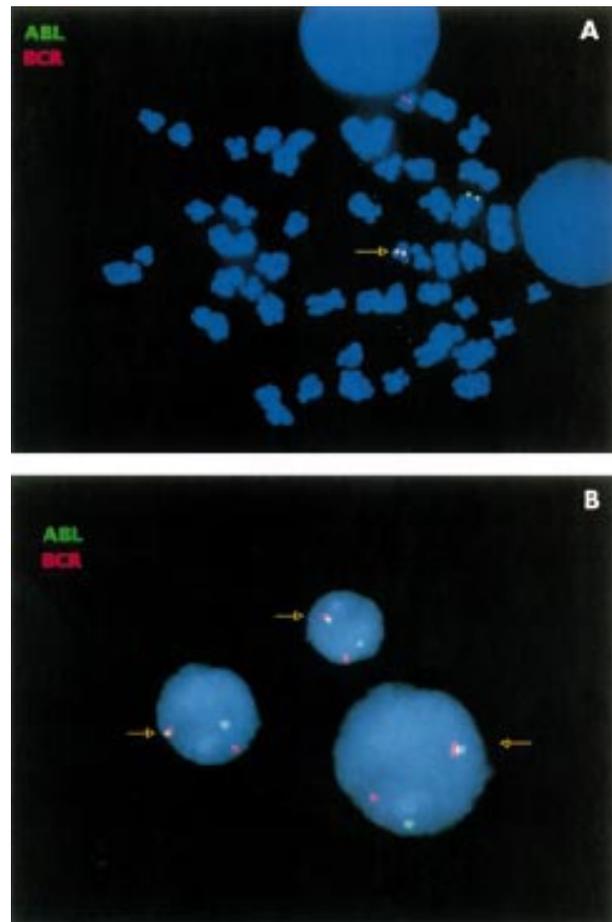

Figure 4 Fluorescent in situ hybridisation demonstrates BCR/ABL gene fusion in chronic myeloid leukaemia. (A) Metaphase: a BCR probe (red signal) hybridises to a complementary sequence on the normal chromosome 22 and on the der(22)t(9;22) translocation, known as the Philadelphia (Ph1) chromosome; arrowed). An ABL probe (green signal) hybridises to the normal chromosome 9 and the derived 22 (arrowed). (B) Interphase: for each informative interphase, a BCR probe (red signal) hybridises to a complementary sequence on normal chromosome 22 and on the der(22)t(9;22) Ph1 chromosome (arrowed). An ABL probe (green signal) hybridises to the normal chromosome 22 and to the derived 22 (arrowed). Note that with this probe set, a single fusion event is usually seen as a result of BCR/ABL fusion on the derivative Ph1 chromosome. This fusion event usually leads to a fluorescence colour change (red/green to yellow).

increase by about a billion fold! However, it should be noted that the effect eventually reaches a plateau, where further cycles are of little benefit (fig 5).

When the PCR is finished, the products of the reaction are analysed by (for example) gel electrophoresis, as above. However, how are the "target" areas of DNA found? The answer lies in the primers used. Primers are pieces of double stranded DNA that initiate or prime new DNA synthesis around areas of single stranded or exposed DNA and enable the synthesis of more nucleic acid. It is important that these primers are of the correct sizes to ensure that they are neither too small to be specific or of such a size that they can bind anywhere in the genome. Usually a size of 15 to 25 bp is ideal. The PCR can continue once these reagents are in place. Thus, the reaction mix contains DNA polymerase, deoxynucleotide triphosphates, buffer, primers, and template DNA.

> "A useful way of checking that amplifiable nucleic acid is present is to look for ubiquitous housekeeping genes"

As intimated above, once the PCR has completed, for however many cycles, the amplified DNA can be analysed by means of electrophoresis. The gels are compared with known controls, both positive and negative. Positive controls consists of a sample known to contain the target DNA and negative controls will omit the polymerase or primers, or even the target itself. A useful way of checking that amplifiable nucleic acid is present (and it may not be, as described below), is to look for ubiquitous "housekeeping" genes, such as those encoding β globin, glyceraldehyde-3-phosphate dehydrogenase, or β actin in the sample studied; if these are absent in the test, then it is likely that the specimen is not suited to PCR analysis. Reasons for inadequacy of specimens include:

(1) The material is aged and degenerate. In general, this is not a problem in current diagnostic specimens but the clinician must remember that biopsies may be small and lack sufficient "good quality" nucleic acid for analysis, even with the PCR. Furthermore, in the real world, the diagnosis of specimens is not always anticipated and, accordingly, tissue may lie in formalin fixative before the initiative is taken to perform the
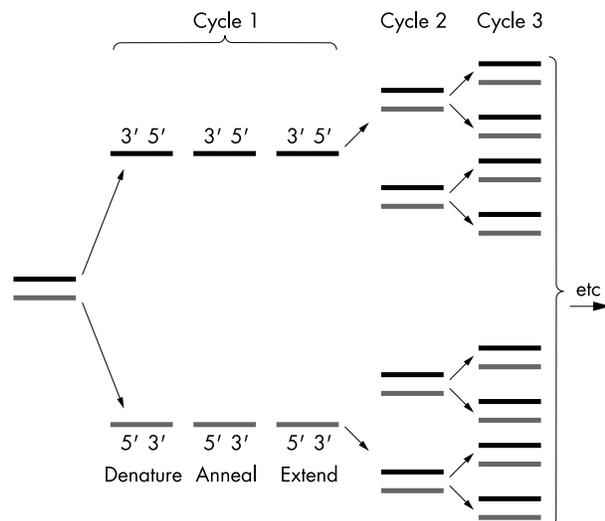
Figure 5 A depiction of the solution phase polymerase chain reaction, showing the amplification sequence induced by denaturation, annealing, and extension, followed by repeated cycles. Artwork by courtesy of Mrs S McGrory, Department of Medical Illustration, Birmingham Heartlands Hospital, UK.

PCR. A frequent but inevitable consequence of this is that analysis may be unsatisfactory because formalin can degrade the target molecule. (Note that it is better under these circumstances to process "extra tissue" to paraffin wax as soon as possible, where it is much better stabilised than in aqueous fixative.) Problems may also arise because the amount of neoplastic tissue is small in the sample available. That is to say, the quantity of normal or "reactive" cells greatly outnumbers that of their malignant counterpart. Under these circumstances, it is possible to dissect or scrape off the neoplastic area of a stained wax section and specifically use it for the PCR. Thus, the target DNA is not "diluted" by non-malignant DNA. A micromanipulator can be used for this purpose. Another approach is to coat the area of interest with protective dye and to ablate the DNA in the non-coated area with ultraviolet light. Alternatively, the new method of laser capture microdissection can be used. In this, the slide is placed under a special microscope, then transparent ethylene vinyl acetate is attached to the section, and a low power laser beam is projected on to the target cells. This briefly melts the material above this area and the cells are bound tightly, maintaining them in the now polymerised film, which is removed from the slide. The cells are then removed from the film and their nucleic acid can be amplified.

(2) We may be looking for RNA! This is mainly of importance when RNA viruses are being evaluated. Fortuitously, reverse transcriptase (RT) exists and can be used to convert RNA to DNA. This is a very powerful and sensitive method, where the RT reacts with the target RNA to yield cDNA, which can be submitted to conventional PCR for subsequent analysis. Again, a gift of nature enables this process in the form of another bacterial product, from the organism named *Thermus thermophilus*. This product, designated Tth polymerase, has the beauty of combining both thermostability and reverse transcription with DNA polymerase activity. Before this, RTs from avian or murine leukaemia viruses were used.

Contamination in the PCR
A further consideration when performing the PCR is that of contamination. Thus, precautions must be taken to avoid the ingress of exogenous sequences because the reaction is so sensitive. The risk of such contamination can be minimised by physically separating the PCR setup stage from the analysis phase, preferably in different rooms. In addition, all reagents and glassware should be autoclaved before use and rigorous hand washing and glove wearing adhered to.

Specialised forms of the PCR
Multiplex PCR
In this variant of the PCR, two or more sets of primers are used in the same reaction mix. Thus, more than one sequence can be detected simultaneously from one DNA specimen. A careful balance must be established to ensure that one amplification does not dominate another. In addition, of course, it is necessary that the reaction products are of different sizes, to enable their distinction on analytical gel electrophoresis.

Nested PCR
Nested PCR and nested RT-PCR can be used if the quality of the nucleic acid sample is suboptimal. The nested method improves the sensitivity and specificity of the technique. In nested PCR, there are two amplification sequences and in the first, a large reaction product is generated, which, in the second round is used itself as a template. This leads to a smaller second product. The nested method increases specificity because it relies on two sets of primers. Furthermore, because two runs of amplification, both of at least 25 cycles, are used the sensitivity is also greater than that of the conventional PCR.

Inverse PCR
In the inverse PCR, primers are annealed to either or both sides of the nucleic acid sequence of interest. The reaction is started by means of restriction enzymes that will attack only nucleic acid outside the target area. After this, the cut fragments are circularised and then another enzyme is used to straighten these fragments and to cut within them, generating nucleic acid with the two halves of the sequence of interest at either end. This then enables the "unknown" sequence in the centre to be amplified by means of conventional PCR (fig 6).

Quantitative PCR
Messenger RNA molecules can be quantified by means of a combination with "mimic" PCR. In this system, a mimic or competitive DNA fragment is used in serial dilutions with a fixed quantity of cDNA. These compete for primers in the PCR and the concentration of cDNA and therefore mRNA can be calculated because the amount of mimic is already known. Sophisticated optical technology is used to read and quantify the nucleic acid in "real time" in a device commonly known as a "light cycler".

The ability to quantify the results of the PCR is invaluable where—for example, viral load must be elucidated; however, inevitably with new technology, the equipment needed is costly.

In cell amplification
The purpose of this method is to combine the ability to localise target nucleic acid in cells or tissue sections and to combine it with the great sensitivity of the PCR. Thus, whereas solution phase techniques such as blotting or solution phase PCR require destruction of the tissue or cellular architecture, it is preserved in this group of methods. Unfortunately, many workers have reported poor success with in cell amplification, but this is probably for technical reasons, not the least being that of establishing the correct physicochemical conditions. However, the method is gradually being established in more centres and success is increasingly being reported.

DNA in cell amplification
Various modifications of the basic type of in cell DNA amplification exist. The original method, called "in situ PCR", uses an amplification mix, as in the solution phase PCR, which is placed on the suitably treated sections or cells. The probes used
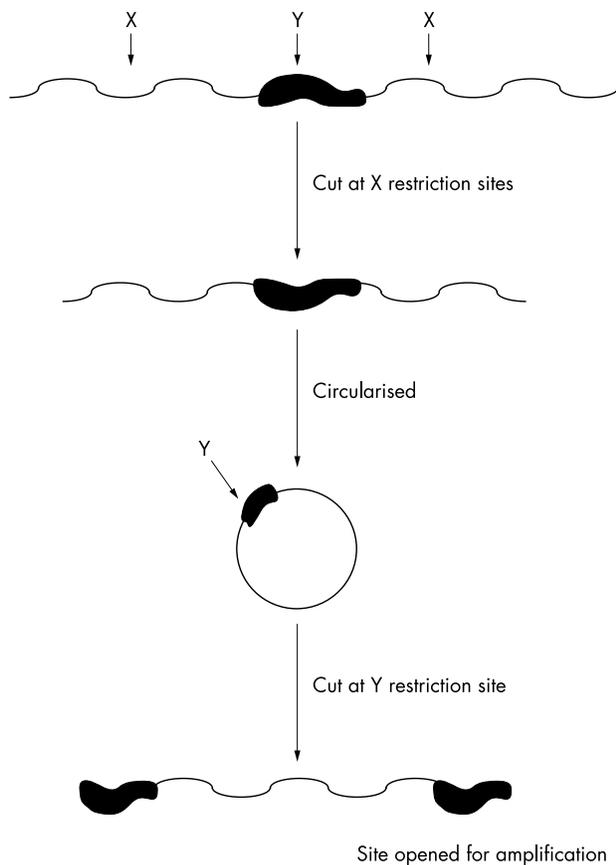
Figure 6 The inverse polymerase chain reaction, showing how unknown DNA sequences can be detected by selective cutting by enzymes, then circularised and "opened" again for analysis. Artwork courtesy of Mrs S McGrory, Department of Medical Illustration, Birmingham Heartlands Hospital, UK.

Figure 7 The sequence for in cell amplification. The cells or section undergo in situ amplification, eventually leading to indirect visualisation of the target nucleic acid. Artwork courtesy of Mrs S McGrory, Department of Medical Illustration, Birmingham Hospital, UK.

are labelled with a tag such as digoxigenin and access to the specimen is aided by pretreatment with a proteolytic enzyme. The label is then demonstrated, as usual, with a chromogenic immunocytochemical method (fig 7).

PCR in situ hybridisation
In the PCR-ISH method, the tissue DNA is first amplified by means of the PCR, followed by detection by ISH.

Primed in situ synthesis
Here, a primer is annealed to nuclei or metaphase chromosomes, then amplified by the Taq polymerase method with incorporation of labelled oligonucleotide to give a reaction at the target site, which is single stranded. A more sensitive variant is described, where repeated cycles of amplification are used.

In cell methods
These include the "Taq Man" method, which is exquisitely sensitive. It runs in real time, with the release of fluorescent reporter molecules as the amplification progresses. The basis of this, in brief, is that one end of the probe has a green fluorescent tag and the other has a red fluorescent dye, which "quenches" the former. As the probe hybridises, it is arranged so that the quencher label is bound, leaving the reporter to be unopposed; as hybridisation continues, more and more of the reporter fluorescent dye accumulates.

RNA in cell amplification: in situ RT-PCR
Similar principles apply here to those in solution phase RT-PCR. Thus, as before, the cDNA template is created using a
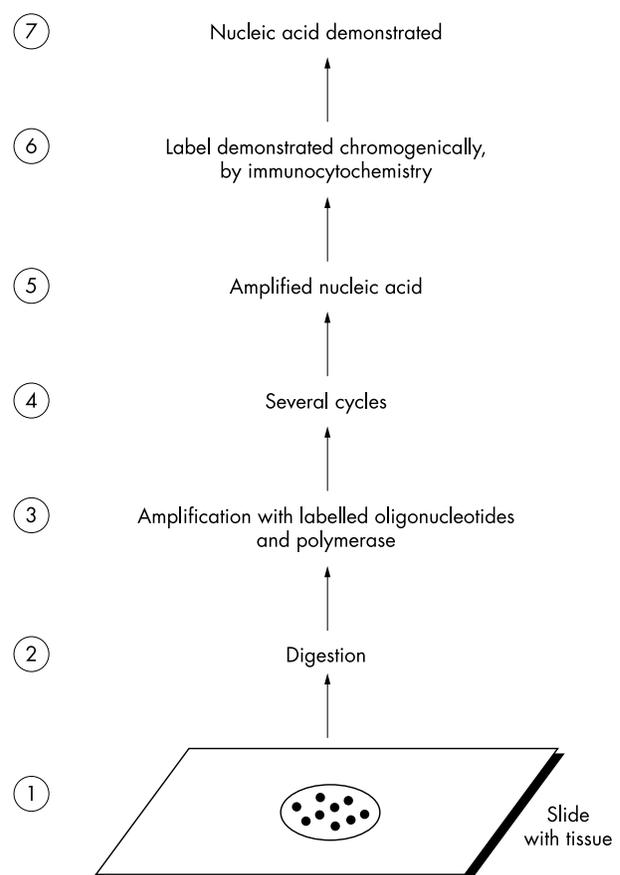
leukaemia virus RT enzyme. This is then amplified in the usual way by Taq polymerase and the product visualised in cell by a chromogenic technique. Other variants of in cell amplification exist but these need not trouble us here.

Comparative genomic hybridisation
Comparative genomic hybridisation (CGH) is a relatively new technique that enables the study of chromosomal aberrations in patients' tissues without the need to culture the constituent cells. It can be used for screening samples quite quickly, before more specific tests and analyses are performed. In short, the principle of CGH is that metaphase spreads are produced from normal human donor cells. Target mutant DNA is extracted from patient cells and tagged with a green fluorescent label; in addition, normal DNA is bound to a fluorochrome of contrasting colour, such as red, and the two nucleic acid aliquots are mixed 1/1. This mixture is applied to the normal metaphase chromosome preparations and the normal and mutant DNAs compete to hybridise with their complementary chromosomal loci. Digital image scanning is then used to quantify and compare the relative amounts of green and red signal. The ratio of green to red signal represents loss or gain of DNA, respectively. Thus, chromosomal abnormalities can be pinpointed.

Microarray analysis ("DNA chip technology")
This novel methodology is applicable to oncology and is about to become a major diagnostic and research tool; so much so that there is fierce commercial competition in this area, which, in the present climate, is likely to lead to patent lawsuits and so forth!

Microarray analysis is essentially a specialised form of CGH, where a metaphase spread is not used as the substrate, but an array of DNA fragments where the chromosomal locus is already identified and understood. Some 5000 DNA loci can be utilised and thus genomic analysis can be of high resolution. The array analysis is performed on "chips" and the data are computer analysed. So specific is the method that it can distinguish between a primary cancer in a specific patient and tell whether another tumour is an additional, synchronous primary or a secondary lesion. It can also be used to look for specific genomic features in relation to tumour typing and prognosis.

"Microarray analysis is applicable to oncology and is about to become a major diagnostic and research tool"

Another form of the method is microarray expression analysis, where gene expression rather than copy number is measured. The "chip" carries arrays of cDNAs of genes to which a mixture of labelled mRNA is hybridised. The ratio of signal intensities for each "spot" is measured and assays the intensity of expression for thousands of genes on each chip. Again, this can help to differentiate primary and secondary tumours from concurrent primary lesions and could be used to monitor resistance to chemotherapy.

Another variant lies in the detection of gene polymorphisms or mutations of individual nucleotides. Here, the probe DNA is hybridised to a chip array of target DNA sequences, which have minute variations in specific nucleotides. Thus, for example, mutations in oncogenes and anti-oncogenes can be detected. In addition, and of major medicolegal importance, the method could be used to detect risks for individual patients for the development of cancer. In addition, the method can be used for the typing of infectious organisms.

The chips themselves are generally produced photolithographically, with direct synthesis of oligonucleotides on to the chip surface. Alternatively, electrochemical or screen printing technology can be used. Recently, a novel approach has been the attachment of fully synthesised oligonucleotides to the chip.

DNA sequencing

Perhaps the ultimate goal of the molecular pathologist is to sequence the nucleotides in a gene of interest. This would have enormous implications in relation not only to the disease itself but would also impinge in a unique way on the particular patient involved. The main means of such an analysis led to a Nobel prize for its originators and has, almost inevitably, led on to automated methods.

Nucleotide sequencing is difficult, not least because the methods involved themselves can be fickle.

The conventional means of sequencing is that predetermined lengths of DNA are degraded in a particular way. Thus, the DNA concerned is taken out of the cell and translated to a bacterial plasmid. After isotopic labelling at the 3' or 5' end, reaction products can be detected, but this is after extremely complex chemical manoeuvres, whose details need not be described here. The difficulties with this method, described by Maxam and Gilbert, lie in the problems encountered in standardising and controlling the reactions involved.

Accordingly, a new approach was developed by Sanger and co-workers, where synthetic oligonucleotide primers anneal to single stranded DNA or a short segment of double stranded DNA. Isotopically labelled nucleotide analogues are incorporated as the polymerases in the mix catalyse the reaction and these, because of their non-conventional structure, terminate the elongation of nascent chains. If the natural nucleoside concentration exceeds that of the analogue, then chain extension will be quite long. If the converse is true, then the synthesised chain will be relatively short. However, some areas

of DNA may not behave on a gel in the same way as others in relation to their ease of denaturation, and this can lead to problems in comparison of mobilities on gel electrophoresis. The G, T, A, and C bases can be labelled with different fluorochromes and this can now enable automated gene sequencing. This depends on the ability of the machine involved to scan the gel products as they pass and to quantify the appropriate fluorescence colour peak. This system displays the DNA sequence for analysis by the researcher (fig 8). Unfortunately, these devices are still very costly.

Restriction fragment length polymorphism

Another, much simpler technique that can be used for gene mapping is restriction fragment length polymorphism. In this method, restriction enzymes are used to cut target DNA into fragments of varying lengths. Alterations in sequences within the restriction fragments alter their lengths and hence mobilities on gel electrophoresis. This affords far less information then formal gene sequencing but can be useful—for example, for tracing a genetic abnormality that might predispose to cancer through a family line. The polymorphisms occur by a change in a base pair, which can create or remove a restriction site. Alternatively, a segment of DNA may be inserted or lost. Thus, the genetic analysis of markers linked to genes that predispose to a particular type of cancer, such as retinoblastoma or Wilm's tumour, is possible.

APPLICATIONS OF MOLECULAR TECHNIQUES IN ONCOLOGY

In this section of the review, the uses of molecular methodology are described. This is largely as a series of examples because the uses of these techniques are legion, but the areas illustrated should give a good idea of how we can apply these methods to the analysis of tumours. Thus, for example, the PCR can be used for the detection of gene rearrangements in lymphomas, DNA sequencing can be used in mutational analysis, and FISH in the chromosomal analysis of leukaemias.

Oncogene analysis

It has been known for some years that retroviruses can cause cancer in certain animals, transcribing RNA to DNA and inserting gene sequences into the host genome, the transcription of which can lead to tumour development. The viral genes leading to this change are called "v-onc" or viral oncogenes. However, homologous sequences are found universally in vertebrate genomes and these are given the names of "c-onc" or cellular oncogenes. These have a vital role in transcribing to proteins that are essential factors in cell proliferation and the cell cycle. The pathobiology of human oncogenes is now well understood and more is discovered every year. This field is now enormous but examples of cellular oncogenes include:

- c-erb transcribes to an epidermal growth factor and is homologous to v-erb, which causes an erythroid-type leukaemia in fowl
- ras human oncogenes are related to cell cycle signalling and are homologous to ras viral oncogenes, which can induce sarcomas in rats
- fas oncogenes regulate cell death and gene expression in humans and are homologous to v-ras oncogenes, which cause murine osteogenic sarcoma.

Human oncogenes can be detected and studied by many of the methods detailed above, such as the solution phase PCR, ISH, ISH-PCR, etc. Furthermore, their protein products may often be demonstrated conveniently by means of conventional immunocytochemistry. Oncogene detection can be extremely valuable—for example, the analysis of c-erb B2 expression in breast cancer may provide useful information with regard to prognosis and metastasis. Another most useful example lies in
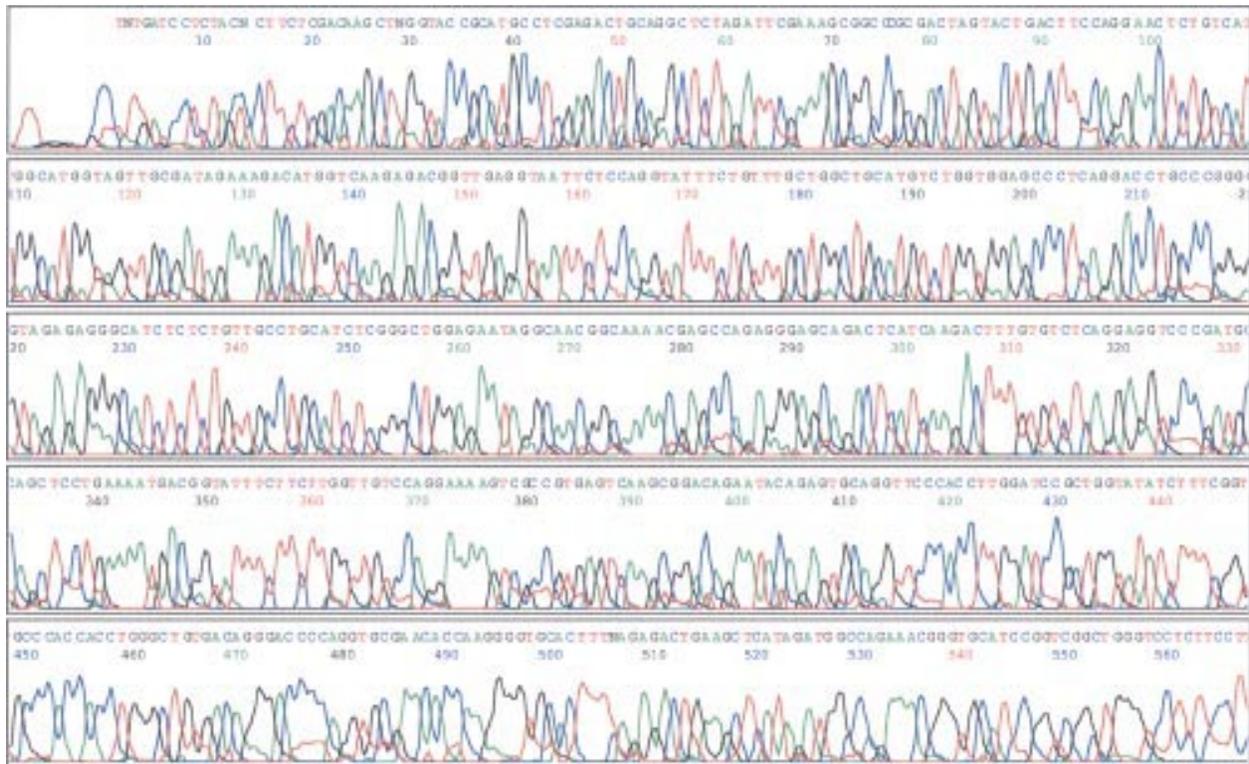
Figure 8 Gene sequence corresponding to phospholipase D2 (hPLD2). The different coloured peaks represent the different nucleotides. This enzyme is one of those responsible for the hydrolysis of phosphatidylcholine to give the signalling molecule phosphatidate. It has been shown to be upregulated in cancer cells. Plate courtesy of Dr D Power, Cancer Research Institute, Birmingham University, UK.

the demonstration of bcl-2, which is a gene that suppresses programmed cell death, or apoptosis. If transfected into lymphoid cells in vitro, this gene immortalises them, and the presence of bcl-2 in lymphoid tissue can be a most useful discriminator between benign follicular hyperplasia and follicular follicle centre cell lymphoma. Furthermore, the presence of bcl-2 can be of prognostic value in high grade B cell non-Hodgkin's lymphomas. At a more "academic" level, the study of fas and its ligand is providing important information with regard to tumour cell apoptosis.

> "The presence of bcl-2 in lymphoid tissue can be a most useful discriminator between benign follicular hyperplasia and follicular follicle centre cell lymphoma"

Related to cellular oncogenes are tumour suppressor genes. These are present in normal cells and inhibit the occurrence of malignant change. When they mutate, this effect is lost and cancer can occur more readily. Perhaps the most cited molecule in the current medical literature is the p53 protein. The p53 gene, which lies on the short arm of chromosome 17, encodes the protein that has rather glibly been described as "the policeman of the genome" because it suppresses tumour cell growth by preventing the progression of genetically damaged cells. The investigation of mutations of the p53 gene can help to assess the prognosis, metastasis, etc of malignant tumours. Some further applications of the study of oncogenes in human cancers follow and are listed in table 1.

Lung cancer
Amplification of myc, leading to overproduction of the p62 protein, has been observed in small cell and non-small cell carcinomas. This is associated with tumour progression and a poor response to chemotherapy.

Cervix
Overexpression of myc has been associated with poor prognosis and increased relapse.

Prostate
The formation of the ras associated protein, p21, is associated with tumour grade.

Testicular teratoma
Overproduction of the myc associated protein p62 has been related to better differentiation of the tumours.

Breast
Increased expression of ras is seen in over 60% of breast cancers and has been linked to poor prognosis and rapid progression. ErbB2 amplification is seen in up to 60% of breast cancers and, again, has been associated with poor prognosis.

Pancreas
Up to 85% of pancreatic cancers have mutations of ras. Such mutations can be detected in stool specimens and could provide a screening method.

Table 1 Some commonly recognised oncogenes

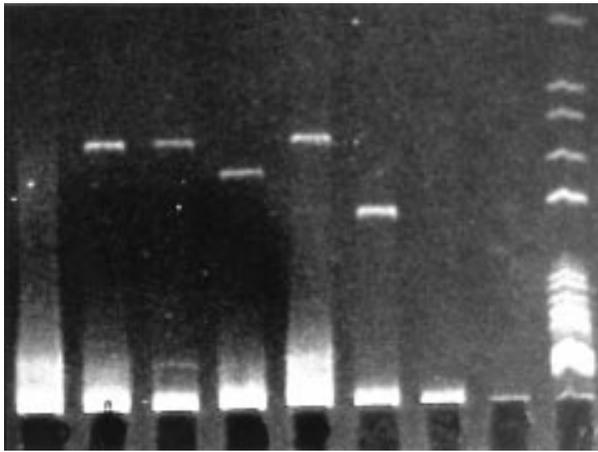| Tumour type	Oncogene
Lung cancer	myc
Cervical cancer	myc
Prostatic cancer	ras
Teratomas of testis	myc
Breast cancer	ras, erbB2
Pancreatic cancer	ras
Neuroblastomas	myc
Colorectal cancer	Ras

Figure 9 A typical polymerase chain reaction gel, showing strict DNA bands for the immunoglobulin heavy chain gene (IgH) rearrangement in a series of chronic lymphocytic leukaemias (CLL). The band on the far right is a molecular weight marker, the next two are negative controls, and the six to the left are from CLL cases with distinct bands, showing the IgH rearrangement. Gel courtesy of Mrs J Starczynski, Birmingham Heartlands Hospital, UK.

Neuroblastoma

Myc mutations have been shown to be associated with lesions with a poor prognosis.

Colorectal cancers

Up to 60% of colorectal carcinomas have ras mutations, but these are rarely seen in small adenomas. However, these mutations have not been related to survival. As with pancreatic cancer, the mutated ras oncogene can be found in stool samples.

Gene rearrangements and chromosomal translocations

Many of the methods described above can be used to investigate gene and chromosomal abnormalities in malignancy.

One of the most useful applications of molecular biology in oncology to date lies in the ability to demonstrate—usually by the solution phase PCR—immunoglobulin heavy chain (fig 9) or T cell receptor gene rearrangements in suspected B or T cell non-Hodgkin's lymphomas, respectively. If sufficient material is available to the histopathologist, then the application of these methods can render otherwise impossible diagnoses possible. These methods are often applicable to archival, paraffin wax embedded material and can also be applied to cytopathology specimens, such as pleural fluids.

Chromosomal translocations are also very important in the understanding and diagnosis of malignant disease. These have traditionally been examined in metaphase spreads from cultured tumour cells but molecular techniques are supplanting these today. The classic example lies, of course, in the Philadelphia chromosome of Down's syndrome and its possible sequel, acute myeloid leukaemia. Perhaps the most useful current examples of the recognition of translocations lie in the detection of the 2;5 translocation of anaplastic, large cell lymphoma and the 14;18 translocation of follicular, follicle centre cell lymphoma. These can be shown by means of the PCR (fig 10), FISH, or in cell hybridisation. The 14;18 translocation is related to bcl-2 expression and can be used to distinguish between the tumour and "reactive" lymph nodal hyperplasia, where morphology is difficult. The detection of the 2;5 translocation can be most useful in the diagnosis of anaplastic large cell lymphoma, which can be mistaken for Hodgkin's disease because both express CD30 in their neoplastic cells.

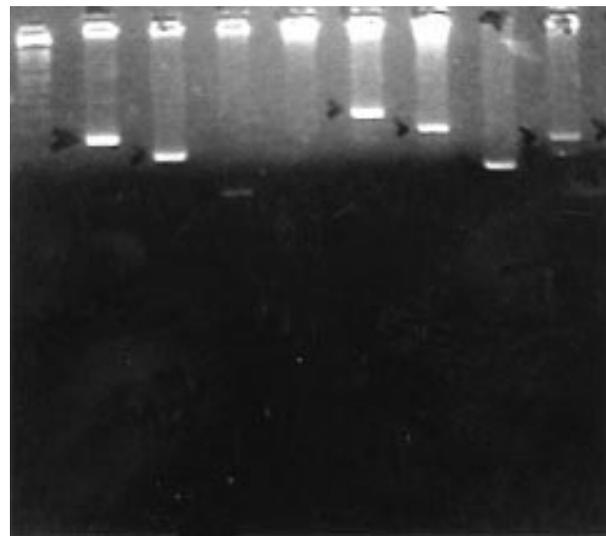Many chromosomal abnormalities have now been related to different malignancies and examples are given in table 2.

Figure 10 A polymerase chain reaction gel, showing, on the right, a molecular weight marker and, in the lane to the left of it, a band representing a cell line known to posses the t(14;18) translocation. Further to the left, there is a series of bands from several cases of follicular, follicle centre lymphoma, all with the 14;18 band. Gel courtesy of Mrs J Starczynski, Birmingham Hospital, UK.

Chromosomal analysis is important in detecting patients who are potentially susceptible to "malignancy syndromes". These include the ataxia telangiectasia mutation (11q 22.3), xeroderma pigmentosum (various translocations), and Gorlin's syndrome of "naevoid" basal cell carcinomas (9q22.1–q22.3).

Common indications for genetic testing of patients with established malignancy include: breast and ovarian cancers (BRCA1 and BRCA2), breast carcinoma and childhood sarcomas (p53), colonic carcinoma and polyps (MSH2 and MLH1), and retinoblastoma (Rb1).

Loss of heterozygosity

A technique that can be applied to the investigation of clonality in tumours is that of loss of heterozygosity (LOH). Certain gene products or chromosomal sites are generally expressed in a heterozygous way in somatic cells. If it is possible to show that this heterozygosity is lacking, then it is implied, for obvious reasons, that clonality has occurred. Thus—for example, microdissection has shown that LOH occurs in lobular carcinoma in situ of the breast in relation to chromosome 17. The polymorphic sites are amplified by the PCR and the products are analysed in the usual way.

Table 2 Examples of chromosomal abnormalities in malignant disease

Tumour type	Chromosomal abnormality
Large cell anaplastic NHL	t(2;5)
Follicle centre cell NHL	t(14;18)
Burkitt's lymphoma	t(8;14)
Mantle cell NHL	t(11;14)
Acute myeloid leukaemia	t(8;21)
Chronic myeloid leukaemia	t(9;22)
Breast carcinoma	t(9;22)
Synovial sarcoma	t(X;18)
Ewing's sarcoma	t(11;22)
Alveolar rhabdomyosarcoma	t(2;13)

NHL, non-Hodgkin's lymphoma.

Infectious agents

Molecular methods have been applied extensively to the detection of infectious agents in oncological pathology. This field encompasses the determination of aetiological organisms and the diagnosis of opportunistic microbes in immunosuppressed patients.

Aetiological agents

Several instances can be given where molecular techniques, ranging from ISH to the PCR and blotting, have been used to identify microorganisms, notably viruses, as the causative agents in neoplastic diseases. Examples are given below.

Human papillomaviruses

Human papillomaviruses (HPV) have been very highly implicated in and associated with the aetiology of cervical cancer and cervical intraepithelial neoplasia (CIN). Certain subtypes (HPV-16 and HPV-18) may be more important in CIN and invasive cervical carcinoma than others and, indeed, some authorities have stated that HPV typing may be a way forward in screening and could even replace cervical cytology. HPV subtypes have also been implicated in other forms of (mainly squamous epithelial) neoplasia, including that at sites such as the oropharynx and oesophagus.

> "Several instances can be given where molecular techniques, ranging from in situ hybridisation to the polymerase chain reaction and blotting, have been used to identify microorganisms, notably viruses, as the causative agents in neoplastic diseases"

Epstein-Barr virus

The Epstein-Barr virus (EBV) is an almost ubiquitous virus, affecting over 90% of the world's population. Commonly, it causes infectious mononucleosis (glandular fever), but has been highly implicated in recent years in some cases of Hodgkin's disease and other lymphomas, including Burkitt's lymphoma, post-transplant lymphoproliferative disorders (PTLD), and some T cell non-Hodgkin's lymphomas (NHLs). Why some people with EBV should develop these neoplasms is uncertain, but it is of interest that in these patients the virus can often be shown to be clonal by molecular methodology. Nasopharyngeal carcinoma of poorly differentiated type, with a high lymphocyte content, is also almost certainly caused by EBV. In addition, it has recently been suggested that multiple myeloma may be caused by EBV.

Others

Other examples of virally induced cancers, as demonstrable by molecular pathology, include: Kaposi's sarcoma and human herpesvirus 8, certain T cell NHLs and human T cell leukaemia virus type 1, and numerous others, of varying certainty of involvement, such as SV40 and pleural malignant mesothelioma.

Opportunistic infections

Obviously, the immunocompromised, immunosuppression treated oncology patient is at risk from many life threatening infections, some relatively commonplace and some more "exotic". The rapid diagnosis of these is essential in management, and molecular techniques have made this possible. Until recently, the detection of some of these organisms has depended upon lengthy and insensitive culture or assay methods, or even the examination of histopathological material. However, we can now subject fresh tissue, fluids, or even paraffin wax embedded histological material to—for example, the solution phase PCR—and obtain fairly rapid results. Some of the agents detected in this way include: hepatitis B virus,

---

Take home messages

- Molecular techniques are increasingly being used both in diagnostic and research pathology related to oncology
- The jargon is considerable but, once it is minimised, the concepts are relatively simple
- We are about to enter the time of "molecular cellular pathology", just as we entered the era of immunohistochemistry a quarter of a century ago
- The technology is increasingly important in the management of patients with cancer. This applies both to the study and assessment of their neoplasms and to any causative microorganisms, or those resulting from immunosuppression

---

cytomegalovirus, human immunodeficiency virus, pneumocystis and, of course, mycobacteria. To give an example of the finesse of the methodology, mycobacteria can even be typed by these methods because they contain heat shock proteins that can be amplified, then endonuclease digested to produce a series of fragments of nucleic acids, which are characteristic for the different types of mycobacterium.

In this area, the importance of "viral load" comes into play. This can be essential in the management of post-transplantation patients, such as those with PTLD, where—for example, the EBV load is crucial; in this context, the "light cycler" or "real time" PCR comes into its own. The amount of EBV present is said to predict the likelihood of PTLD developing so that using quantitative PCR can allow the amount of suppression used to be controlled. Other areas where quantitative PCR is important can also be found in the immunocompromised patient, where the assay of *Toxoplasma gondii*, cytomegalovirus, and certain adenoviruses may affect management. Furthermore, it should be remembered that this method is very rapid, even compared with the conventional PCR.

Minimal residual disease

All oncologists and clinical haematologists know that the management of their patients depends upon the awareness of the presence or absence of any residual disease in the circulating blood or—for example, the bone marrow. This may be a very small cellular load and impossible to detect by conventional techniques but yet can be crucial for the treatment of the patient. The ability to detect such small numbers of cells by molecular methods has enabled us to pick up residual disease. For example, we can look for the 14;18 translocation in bone marrow samples or the same translocation in circulating peripheral blood cells in follicular NHL. Gene rearrangement evaluation can be used in a similar context—for example, the PCR can be used to detect residual B-NHL. Quantitative PCR is likely to find widespread application in this area.

> "The ability to detect such small numbers of cells by molecular methods has enabled us to pick up residual disease"

Finally, a potentially most exciting methodology in the detection of minimal residual disease lies in the production of patient specific probes, synthesised to detect patient specific DNA sequences, which would enable highly specific tumour demonstration.

FURTHER READING

1 **Baumforth KRN**, Nelson PN, Digby JE, *et al.* Demystified . . . The polymerase chain reaction. *Mol Pathol* 1999;**52**:1–10.
2 **Bradley J**, Johnson D, Rubenstein D, eds. *Lecture notes on molecular medicine.* Oxford: Blackwell Science, 1995.
3 **Crocker J**, ed. *Molecular biology in histopathology.* Chichester: John Wiley and Sons, 1994.
4 **Latchman DS**, ed. *Basic molecular and cell biology.* London: BMJ Publishing Group, 1997.
5 **Lindblom A**, Liljgren A. Tumour markers in malignancies. *BMJ* 2000;**320**:424–7.
6 **Randhawa JS**, Easton AJ. Demystified . . . DNA nucleotide sequencing. *Mol Pathol* 1999;**52**:125–30.
7 **Sexton DM**, Bennett MA, Croke DT. PCR and molecular pathology. *Curr Diagn Pathol* 1994;**1**:226–33.
8 **Wallis L**, Macdonald F. Demystified . . . Oncogenes. *Mol Pathol* 1999;**52**:55–63.
9 **Waters JJ**, Barlow AL, Gould CP. Demystified . . . FISH. *Mol Pathol* 1998;**51**:62–70.
10 **Weiss MM**, Hermsen MAJA, Meijer GA, *et al.* Demystified . . . Comparative genomic hybridisation. *Mol Pathol* 1999;**52**:243–51.

ECHO

Polymorphs may play a part in sickle cell retinopathy
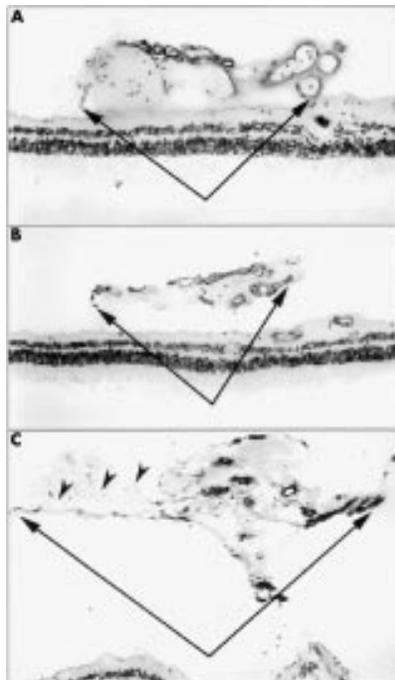
Various sea fans in retinal sections labelled with anti-perlecan antibody, specific for viable blood vessels. (A) Sea fan composed of mostly large caliper dilated vessels; (B) sea fan of small diameter capillary like vessels; (C) sea fan with viable and infarcted (arrowheads) vessels. Haematoxylin counterstain. (A), (B) 120x, (C) 80x.

New evidence indicates that leucocytes and leucocyte adhesion molecules feature in blocking retinal capillaries in sickle cell retinopathy. Retinas from patients with sickle cell disease showed significantly higher immunoreactivity of intercellular adhesion molecule 1 (ICAM-1), vascular adhesion molecule 1 (VCAM-1), and P-selectin than retinas from a control. ICAM-1 and P-selectin reactivities were highest in sections of retina whereas VCAM-1 reactivity was most significantly increased in "sea fans", especially compared with sections of retina from the control. Sea fans are histological features that result from neovascularisation in the pre-retinal area close to non-perfused peripheral retina.

Polymorphonuclear leucocytes (PMNs) were significantly more numerous in the retinas of the patients—about three times more than in the control—and in sea fans compared with retina in the patients.

The distribution of adhesion molecules was assessed by visually scoring histochemically stained thin sections of retinas in eyes obtained at necropsy and by densitometry. PMNs were counted in sections and counts expressed per mm^2 of retina. Three different areas of the retina were sectioned in each case. Nine samples were obtained from five patients with sickle cell disease; one was a vitrectomy sample. One subject without sickle cell disease, who had been a control in a previous study of diabetic retinopathy, provided the control samples.

Research in sickle cell disease has hinted at involvement of PMNs in adhesion to capillary endothelium in the retina, which is a precursor to retinopathy. Patients had higher amounts of cytokines, white cells, and soluble ICAM-1 and VCAM-1 in their blood.

▲ *British Journal of Ophthalmology* 2002;**86**:684–690.

Please visit the Molecular Pathology website [www.molpath.com] for link to this full article.