Methods of molecular analysis: mutation detection in solid tumours

I M Frayling

Most mutation detection techniques are unsuitable for routine use on solid tumours. Important parameters include sensitivity, specificity, efficiency, use of existing resources, and cost. In the UK, < 0.2% of service genetics laboratory activity involves mutation analysis in tumours (usually for family studies), mainly because it is time consuming/labour intensive (thus expensive) and DNA extracted from formalin fixed, paraffin wax embedded tissue is of low quality and yield. The small size of DNA fragments obtained from tissue blocks limits the polymerase chain reaction, the basis of most mutation detection methods. Other, biological, factors include: (1) heterogeneity of mutations within and between tumours, (2) variation in type and site of mutations in any one gene, (3) normal tissue harbouring mutations, (4) few genes are mutated in most of any one tumour type, and (5) few clinically useful correlations with genetic changes have been found. Present research is centred on correlating single gene mutations with various clinicopathological features, but the pattern of mutations in a combination of genes will probably prove more useful. Microsatellite instability, however, appears to be worth testing for in both familial and sporadic tumours, particularly of the colorectum.

The detection of genetic mutations in solid tumours is still largely a research tool, but nonetheless the indications are that it will become an important part of the service pathologist’s armamentarium. Although it might appear to be an alien technology, it is important not to consider its use in isolation, but to regard it as a complementary extension of existing techniques of pathological examination. This overview will consider the nature of mutations, current service practice, and practical aspects of detecting mutations, and then show how these can be combined in one particular application, the diagnosis of DNA microsatellite instability.

THE NATURE OF MUTATIONS

Human chromosomes are linear DNA molecules many tens of millions of nucleotides in length. At any given position may be an adenine (A), cytosine (C), guanine (G), or thymidine (T) residue. Because DNA is composed of two complementary strands it exists as a double stranded molecule, the length of which is measured in base pairs (bp). The genetic code, which ultimately determines the structure of the protein produced by a gene, is defined by the precise sequence of these four bases. Mutations are changes in this sequence, which may, or may not, be pathogenic by deleteriously affecting the function of the gene or its protein product. One complication of mutation detection is that genes are composed of several parts (fig 1). Upstream of the coding sequence, that part of the gene which defines the order of amino acids in the final protein, there are regulatory sequences, such as promoters and enhancers. The coding region of most eukaryotic genes is divided up into subsections called exons. In between exons lie stretches of DNA called introns. After RNA has been transcribed from the DNA template, the cell removes the introns so that the exons are spliced together to produce one continuous coding sequence in the mRNA. The sequence motifs that the cell uses to define the start and finish of each exon are not infrequent sites of mutation, and by attention to detail can be screened together with the protein coding exon sequence. Mutations affecting promoter regions can be similarly screened for, although their interpretation may be more complex. However, more problematical are mutations in introns that generate false exon start/stop signals. Introns can be up to tens of kilo base pairs (Kb) in size, and hence are not generally screened by most current methods of mutation detection, which are limited to concentrating on exons and short stretches of intron either side. The screening of enhancers for mutations is not undertaken, partly because most have not been defined, but also because they may be at a considerable distance from the promoters they regulate.

A second complication of mutation detection is that mutations themselves can vary by over eight orders of magnitude, and few methods cover more than two orders (table 1). Gross chromosomal rearrangements visible cytogenetically, and involving tens of millions of DNA bp (megabases; Mb) have been recognised to be associated with cancer since the start of the 20th century. The DNA in a chromosome may be around 1 cm long, but it is packaged into a length of less than 1 μm.

Abbreviations: ARMS, amplification refractory mutation system; DGGE, denaturing gradient gel electrophoresis; HNPCC, hereditary non-polyposis colorectal cancer; MMR, mismatch repair; MSI, microsatellite instability; PCR, polymerase chain reaction; SSCP, single stranded conformation polymorphism; TGGE, thermal gradient gel electrophoresis.
Methylation of C residues alters its molecular weight by
rences are now being defined, as will be seen later.
ciated with tumour development, and specific, testable, occur-
of that gene. Abnormal changes in gene methylation are asso-
gene promoters is generally associated with downregulation
expression of some genes is naturally regulated by the so
smallest end of the scale, there are mutations affecting the
analysing the PCR synthesised material. T echniques for this
requires different techniques, such as Southern blotting.
Loss or gain of discrete portions of chromosomes can be
detected by conventional cytogenetic techniques, such as G
banding and latterly comparative genomic hybridisation, with
a resolution approaching 10 Mb. Genes vary in size, but even a
1 Mb region would be expected to contain a few genes, thus
deletion or loss of a single gene, or substantial part thereof,
requires different techniques, such as Southern blotting.1
Mutations affecting one or a few bp can only easily be detected
by first amplifying that stretch of DNA in vitro, in the process
known as the polymerase chain reaction (PCR), and then
analysing the PCR synthesised material. Techniques for this
are covered later in this article in greater depth. Finally, at the
smallest end of the scale, there are mutations affecting the
pattern of cytosine methylation of a gene. In eukaryotes
the expression of some genes is naturally regulated by the so
called epigenetic phenomenon of cytosine methylation.
Methylation of C residues in the sequence context -CG- within
gene promoters is generally associated with downregulation
of that gene. Abnormal changes in gene methylation are asso-
ciated with tumour development, and specific, testable, occur-
rences are now being defined, as will be seen later. Methylation of a C residue alters its molecular weight by
< 5%, but, nonetheless, it can be readily detected.

Insertional mutagenesis covers an unusual class of muta-
tions caused by the integration of a viral genome into that of
the host. Insertion of a large foreign portion of DNA into a
host's chromosome may disrupt genes in that region, and/or
introduce novel oncogenes—classically, exemplified by the
Rous sarcoma virus in chickens, leading to identification of the
src family of oncogenes. Human examples include hepa-
tis B virus and hepatocellular carcinoma, and human papillo-
navirus and carcinoma of the cervix. Southern blotting or
PCR of a small, but characteristic, portion of the viral genome
The reasons for the small amount of activity directed at
solid tumours include the fact that the process is time
consuming and labour intensive, and therefore expensive, and
DNA extracted from formalin fixed, paraffin wax embedded
tissue is of low quality and yield. The small size of DNA frag-
ments obtained from tissue blocks (< 300 bp) limits the use of
PCR, the basis of most mutation detection methods. An addi-
tional and no less important factor is that analysis of solid
tumours requires the input of a histopathologist, who can
confirm that the block submitted for analysis is of the relevant
tumour type, and that sections cut from it for DNA extraction
contain a satisfactory proportion of representative tumour
cells.
Traditionally, clinical scientists in genetics laboratories have
had little or no need to interact with histopathologists, and
histopathologists have not required molecular genetics input.
Clearly, therefore, this is an area for change and links will need
to be forged between service genetics and histopathology
laboratories within pathology directorates.
Currently, most tissue samples available for analysis have been fixed in formaldehyde before being embedded in paraffin wax. Both under and over fixation are undesirable. It is universally found that the DNA in fixed material has been degraded by fragmentation and damage caused by the formation of covalent adducts. Paradoxically, incomplete fixation can result in greater loss of DNA as a result of autolytic or bacterial degradation, and may also spoil other adjunctive techniques, such as immunohistochemistry. Although the DNA in fixed and embedded tissue is not ideal, it is nonetheless possible to work around its limitations by—for example, not attempting to amplify DNA fragments greater than around 300 bp in PCR. In general recently fixed, but unembedded, tissue gives a better yield of DNA than embedded tissue, and it has even been possible to extract DNA from archived museum specimens. Lymph nodes are a good source of normal DNA because they contain little connective tissue and large numbers of small cells. Molecular detection of metastases is also possible, and is potentially more sensitive than standard histopathology.\(^4\)

Once a source of material has been identified, several analytes can be extracted. Intact nuclei can be obtained from fresh/frozen, or even fixed and embedded tissue, or they can be analysed in situ. DNA can be extracted from tissues by a variety of means. The increasing use of kits and shorter, simpler protocols is in contrast to some historical methods requiring prolonged digestion of tissues. The best quality and quantity of DNA is extracted from fresh or frozen tissue, followed by, in decreasing order of quality and quantity, tissue preserved but not fixed (for example, in ethanol), tissue fixed but not embedded, and finally fixed and embedded tissue. Buffered formalin is somewhat less damaging than unbuffered.

"It is important that DNA should not be extracted from an unidentified or unconfirmed tissue source, and the input of a histopathologist is therefore essential."

It is possible to obtain DNA from sections stained with haematoxylin and eosin, but if microdissection is required then light staining with toluidine blue is preferred. In any event, microdissection should be carried out under the supervision of, if not by, a histopathologist. It may be possible to use laser based microdissection or related techniques.\(^5\) In this, the cells in tissue sections wanted for DNA extraction are protected from UV irradiation by the use of UV opaque ink. Paradoxically, the quality and quantity of DNA extracted from stained and mounted sections over about 20 years old can be superior to the original blocks. This may be because of the better protection of the DNA from atmospheric oxidation in mounted sections. The use of fixatives containing—for example, mercuric or picric acid—results in DNA that cannot be amplified by PCR. It is possible to extract RNA from fixed embedded tissue, but it has usually suffered more from the treatment of the specimen before fixation, than from the processing itself. For all practical purposes RNA should be extracted from snap frozen material stored at −70°C or below. Protein does not usually need to be extracted from tissues for analysis, given that analysis is usually by immunohistochemistry, but it may need to be made more available to antibodies (antigen retrieval) by—for example, microwaving in defined buffers.

**Techniques**

Most mutation detection methods rely on the technique of PCR to amplify the region of DNA of interest (fig 2). With DNA of the highest quality it is possible to amplify sections of DNA several kb long; however, as has already been mentioned, the quality of DNA obtainable from most tumours restricts the maximum length of PCR products to about 300 bp. This is not too much of a restriction, because most exons can be amplified in one or two PCRs of this length.\(^6\) However, there is the additional problem that the process of fixation and embedding can itself introduce DNA damage, which can be detectable as a mutation, especially if mutation at a particular site is being sought—for example, in K-ras.\(^7\)

Various PCR adjuncts are available. Where there is heavy contamination with foreign DNA, or mutant DNA is present in very low amounts compared with normal DNA—for example, in stool specimens—it is possible to enrich the mutant DNA species. This effectively improves the signal to noise ratio and lowers the limit of detection, and hence increases the chance of the process of PCR itself can introduce mutations, for a variety of reasons.\(^8\) This is not usually a problem, but when the amount of mutant DNA is low and enrichment is necessary before greater than usual PCR amplification it can become important. It is possible to amplify the whole genome when DNA from one or only a few cells is available, using the techniques of primer extension preamplification and degenerate oligonucleotide preamplification.\(^9\) It is also possible to obtain a digital signal by diluting DNA such that only a single molecule is amplified in each tube.\(^10\)

Having obtained a PCR product, there is then the task of determining whether or not there is a mutation in it and, if so, what the nature of that mutation is. There are many screening techniques available, which depend upon the mutant DNA having, or being capable of taking up, a different secondary structure (conformation) compared with that of the normal DNA.\(^11\) This difference in conformation is then exploited by the fact that it frequently, although by no means always, results in altered gel electrophoretic mobility of the mutant DNA species. It is then necessary to sequence the PCR product to determine the mutation. Alternatively, the PCR product can be directly sequenced, but this can be more expensive in reagents and data analysis time: most genes have several exons, all requiring screening, but there will usually only be one mutation, so most PCR products will be normal. Direct sequencing analysis is more efficient at detecting mutations, but it is not 100% efficient. The challenge is to screen out all the normal sequences, and minimise sequencing to a subset of suspect PCR products. Given the limitations of current techniques this is probably the better strategy. However, considerable time and effort is being devoted to improving direct sequencing.\(^12\)

Perhaps the most common conformation based mutation screening technique is single stranded conformation polymorphism (SSCP) analysis.\(^13\) Thermally denatured DNA is electrophoresed, and those single stranded DNAs that take up an altered conformation show up as aberrantly migrating bands on the electrophoretic gel. It is relatively simple and quick, and thus cheap, and does not require special apparatus. However, its main drawback is that it becomes increasingly inefficient at detecting mutations with increasing size of the PCR product tested, which limits it to PCR products less than 300 bp, and preferably less than 200 bp in size. The result of this is that to screen a gene composed of only a few exons requires rather more PCR products, and so what is gained in simplicity and speed is lost in the number of tests required to screen one gene.
Another technique, related to SSCP and often carried out at the same time, is DNA heteroduplex analysis. This exploits the difference in electrophoretic mobility between double-stranded DNAs composed of two normal sequences, compared with those composed of one normal and one abnormal sequence—that is, a heteroduplex. Like SSCP it is quick, cheap, and simple, but again suffers from inefficiency at detecting mutations. It is inherently better suited to finding small insertion/deletion mutations, rather than point mutations, but it is capable of screening large DNAs, even up to 1500 bp. Both SSCP and heteroduplex analysis can be carried out on fluorescent DNA analysers, although these may need the ability to be run at a controlled temperature.

DNA will denature—that is, separate into single strands—when heated. The exact temperature at which this occurs depends on the sequence of the DNA, and the concentration of salts and other small molecules in the solution. Mutant DNA has a different denaturation profile to its normal counterpart and this is exploited in the related techniques of denaturing gradient gel electrophoresis (DGGE) and thermal gradient gel electrophoresis (TGGE). In DGGE, a gel is set up with a gradient of increasing denaturant concentration from one end to the other. As the DNAs are electrophoresed through this gradient, at one point they will denature and their mobility alter. Mutant DNAs denature at a different point and thus are distinguished. In TGGE, the gel either has a spatial temperature gradient (from one end to the other), or the temperature of the whole gel is increased steadily as the electrophoresis is carried out. Either way, as with DGGE, at a certain point the DNAs distinguish themselves by their different denaturing behaviour. DGGE and TGGE are very efficient (> 90%), but they require specially designed PCR primers of increased length, which can be expensive. Neither technique can be performed on a fluorescent analyser.

Sequencing of the PCR product is almost universally carried out using dideoxy terminator chemistry. In the past, this was done with a radioactive label followed by autoradiography, but this is being increasingly replaced by sequencing on a fluorescent DNA analyser, the sequencing DNA products being fluorescently labelled. Either way, sequencing is time consuming and uses relatively expensive reagents, and the time spent comparing sequence data can be considerable. Nonetheless, fluorescent analyzers do make a very considerable saving in time, albeit at considerable cost. One drawback is that the software for sequence analysis is geared up to the absolute determination of sequence, rather than the comparison of a test with a reference (normal) sequence. This means that going through a set of sequences to look for a mutation has been labourious and not wholly efficient, but steps have been taken to improve this. It also partially explains the popularity of using a screening technique, like SSCP, first. A hybrid of sequencing and SSCP analysis called dideoxy fingerprinting has been used for the efficient screening of mutations. Another sequencing variant called minisequencing has been used to determine specific sequence changes of a recurrent nature, such as those in K-ras.

“Other practical limitation is that the development of molecular testing, especially in the service setting, is dependent on the availability of technology, equipment, and expertise, but above all on appropriate funding”

Modifications of the PCR process itself can be used to determine mutations. If the PCR oligonucleotide primers are designed to be specific for one mutant sequence variant or another, they can then be used to set up PCR reactions that will specifically detect one or the other variant. This is the basis of the amplification refractory mutation system (ARMS). As with minisequencing, ARMS may be useful if the mutation is a recurring one at a specific location, such as K-ras. Another technique uses thermostable DNA ligase to join oligonucleotide primers at a specific sequence site. This is the basis of the ligase chain reaction, and oligonucleotide ligation assay. But, again, it is restricted, although ideally suited to frequent mutations at a specific nucleotide.

Abnormal cytosome methylation has already been mentioned as a form of mutation in tumours, and there are two main ways of testing for it. The DNA to be tested can be pretreated with a methylation sensitive restriction endonuclease. If the sequence is methylated it will not be digested, and will therefore allow PCR amplification across the methylation site. If, however, the sequence is unmethylated it will be cut and so not allow PCR. Alternatively, the test DNA can be pretreated with bisulphite and hydroquinone, which converts 5-methylcytosine into uracil. Uracil in DNA base pairs with A, and thus 5-methylcytosine residues are effectively converted into T residues. An ARMS PCR specific for either a C or T at the given position can then be used to diagnose if the C was methylated.

LIMITATIONS

Some of the issues of testing for mutations in solid tumours have already been touched upon, such as the necessity to work...
Mutation detection in solid tumours

from paraffin wax embedded tissues, and the need for the histopathologist and molecular geneticist to work together. However, a considerable limitation is the fact that very few clinically useful molecular correlates have been found so far. Most studies have naturally concentrated on one or a few genes, and it would be surprising if any one such variable gave as much information as the host of variables that are assessed in standard pathological examination.

The process of carcinogenesis is multistep, and a large number of mutations are possible in each of a large number of different genes. Therefore, it is to be expected that studies that bring together a large number of molecular variables, in addition to those determined by conventional pathological assessment, will be more successful.

“It is likely that the determination of patterns of gene mutation, allied with patterns of gene expression, will prove to be the way forward”

Finally, mention must also be made of biological limitations. The great variety of mutations in terms of site and type has already been discussed. However, there is also mutation heterogeneity; both within and between tumours: a tumour may have overall abnormal p53 expression, but different parts of the same tumour may harbour different mutations in p53 underlying this. Mutations have also been detected in apparently normal tissue. Although it is debatable whether this is a true or artefactual finding, nonetheless, such potentially misleading findings are possible. It is also the case that few genes are commonly mutated in any one tumour type. In colorectal tumours, APC is only mutated in perhaps 85% at most, and although p53 is mutant in most cancers, a substantial fraction do not have p53 mutations, although they might have equivalent changes in other genes. More typical perhaps is K-ras, which is only mutated in perhaps 25–30% of colorectal cancers. For this reason, and those mentioned above.

ONE PRACTICAL APPLICATION

Over the past few years, molecular genetic studies of colorectal cancer prone families have shown that such cancers can arise along one of two main pathways. The classic pathway is that alluded to in studies of cancer from nearly 100 years ago—that is, gross chromosomal derangement, now termed chromosomal instability. This is observed in about 80% of sporadic colon cancers, and about 98% of sporadic rectal cancers. The remainder arises along an alternative pathway, characterised by not having large scale chromosomal disruption, but rather a surfeit of small point mutations in their DNA.

It is now known that this results from loss of DNA mismatch repair (MMR) in such tumours, but other factors may also be involved. Normal individuals have two copies of each MMR gene and it takes on average 70–80 years for mutations to accrue in both copies of the same MMR gene.
in one cell, and then only in a small minority of individuals. However, individuals in families with underlying germline mutations in one of the MMR genes are, if they inherit the mutant gene, one step closer to a cancer. Therefore, they tend to develop colorectal cancer at an average age of about 40 years. This, together with a propensity to other tumour types, especially endometrial cancer, characterises hereditary non-polyposis colorectal cancer (HNPCC). In HNPCC associated tumours, and the 20% of sporadic colorectal cancers that have also lost MMR, frequent insertion/deletion mutations can be found in the stretches of repetitive DNA known as microsatellites. Hence, these tumours are characterised by the phenomenon of microsatellite instability (MSI). It is important to realise, however, that only a small proportion (1–2%) of colorectal cancers is the result of HNPCC. Most cancers with MSI arise sporadically.

It is useful to study tumours from families affected by HNPCC because this can allow full and proper diagnosis and greatly assist with germline mutation detection. Several techniques can be applied, and their overall interpretation gives more powerful results than any one technique alone. The three different, but complementary, techniques that can be applied to suspected HNPCC tumours are MSI analysis, immunohistochemistry of MMR proteins, and MLH1 promoter methylation status. It has been found that most sporadic MSI positive colon cancers arise because of methylation of the promoter of the major MMR gene MLH1. Conversely, the study of HNPCC colon cancers has shown that this is not a mechanism of mutation that occurs in HNPCC. Thus, an MSI positive colorectal cancer that has MLH1 methylation can be diagnosed as sporadic.

Because only a minority of even MSI positive colorectal cancers are caused by HNPCC, testing for MSI alone does not give a high predictive value of HNPCC. In addition, it is known that about 5% of cancers in families affected by HNPCC do not exhibit MSI, for reasons that are presently unclear. However, it is possible to exploit information on site and type of tumour. Rectal cancers, for example, have a very low rate of MSI, about 2%, and most (5/6) of those that do exhibit MSI are the result of HNPCC. Colorectal adenomas similarly have a very low rate of MSI (2%), and again most of these arise in individuals with HNPCC. Thus, in these two instances (rectal cancers and colorectal adenomas) MSI positivity carries a substantially higher positive predictive value for HNPCC.

“A variety of molecular techniques can be used as an adjunct to standard pathological examination to produce a powerful diagnostic combination”

Immunohistochemistry of the major MMR proteins MSH2 and MLH1 can give important extra information. When studying a tumour from an individual strongly suspected of HNPCC, loss of either MSH2 or MLH1 indicates that their propensity to such tumours is the result of HNPCC and, moreover, is caused by a germline mutation in that particular gene (fig 3). Consistent loss across more than one tumour, either from the same individual and/or across a family, gives even more reliable results. This enables substantially more economic mutation detection to be carried out. Conversely, if loss of MMR proteins is not seen then it is unlikely that HNPCC is the cause.

Flowcharts illustrating how MSI, immunohistochemistry, and MLH1 methylation analysis might be combined for maximum advantage in cascade testing are shown in figs 4 and 5. These highlight the usefulness of using a variety of molecular techniques as an adjunct to standard pathological examination to produce a powerful diagnostic combination. Although this is currently directed at the diagnosis of a relatively rare cancer genetic syndrome, this may not always be the case. Early studies suggesting improved survival of those with colorectal cancers as a result of HNPCC have recently been extended to looking at the importance of MSI in sporadic colorectal cancers. This is increasing the weight of evidence suggesting that colon cancers with MSI may be much more sensitive to treatment with chemotherapy than those without MSI. MSI analysis, therefore, may become part of the standard pathological examination of all colorectal cancers.

**SUMMARY**

The detection of mutations in solid tumours is certainly possible, but is mostly restricted to research at present. The small amount of service activity that currently takes place is largely within regional genetics laboratories, and is directed at the diagnosis of familial cancer predisposition syndromes. This is because the methods for finding mutations are generally very time consuming and expensive, and apart from their usefulness in diagnosing familial tumours, definite correlations with clinicopathological variables are lacking. In addition, the analysis of samples is to some extent limited by current methods of sample handling—that is, formalin fixation and wax embedding of tissues. However, it is likely that mutation detection of so called sporadic (non-familial) solid tumours will become more important as data accrues on its usefulness, and methods of mutation detection become easier and cheaper. Such testing will include greater use of pattern analysis and cascade testing, and may necessitate a change in how samples are handled. It will certainly depend on and encourage the formation of stronger links between departments of genetics and cellular pathology.

**REFERENCES**

Mutation detection in solid tumours