MOLECULAR MEDICINE

MOLECULAR DIAGNOSIS

(Second of Two Parts)

BRUCE KORF, M.D., PH.D.

M ost of the mutations responsible for genetic disease are remarkably subtle, considering the dramatic effects they may have on the phenotype. A change in just one of many thousand bases in a gene can be sufficient to disrupt and thus alter or eliminate the expression of a protein product. On the other hand, genetic variation is common from person to person and is usually of no phenotypic consequence. Benign changes may occur in regions that do not encode protein or do alter the properties of the protein encoded by a gene. The challenge in the molecular diagnosis of gene mutations is therefore twofold: to find the rare change in sequence, and to distinguish important changes from benign polymorphisms.

In the research laboratory, the identification of gene mutations is a critical step in establishing that a cloned gene is actually responsible for a genetic disorder. Approaches have been developed to allow rapid screening of a segment of DNA for variations in the sequence of bases. These methods are sensitive to the presence of a difference in sequence between two samples, obviating the need to sequence the entire gene to find a mutation. The segment with the mutation is sequenced, and the sequence change is then analyzed to determine whether it is pathogenic or represents a polymorphism. A pathogenic mutation will segregate with the disease in a family and will occur only in affected persons or carriers. Moreover, it should be possible to predict the effects of a mutation on the expression of a protein. Mutations that abolish the expression of a protein, such as insertion or deletions that cause frame shifts or base substitutions that cause premature termination of translation, are likely to be pathogenic. The pathogenicity of mutations that substitute one amino acid for another in a protein may be more difficult to establish. Some may be predicted to have major effects on the chemical properties of the protein, but in other cases the effects may be more subtle. Ultimately, it may be necessary to reproduce a phenotype in an animal model or in cell culture to prove the pathogenic nature of a mutation.

These challenges explain why, despite all of the hundreds of genes that have been cloned, relatively few molecular diagnostic tests are in routine use. The process of scanning a large gene for mutations is generally confined to the research laboratory and is usually not practical for general diagnostic testing. Direct analysis of mutations is used in diagnostic testing for disorders in which a limited number of well-characterized mutations account for the majority of cases. Often, this is the result of a founder effect — the introduction of a mutation into a small population by one member, leading to a relatively high prevalence of that mutation in the population in future generations. Thus, in many cases, mutations may be specific to particular populations or ethnic groups.

A variety of approaches are commonly used to detect mutations (Fig. 3). One of the simplest takes advantage of the base-sequence specificity of restriction endonucleases (Fig. 3A). These enzymes recognize precise sequences of four to eight bases and cut double-stranded DNA only at these sites. A mutation at such a site will prevent the enzyme from cutting there; conversely, a mutation may result in the creation of a new enzyme-recognition site and lead to cutting where it normally should not occur. To detect the mutation, DNA surrounding the site of potential mutation is amplified by the polymerase chain reaction (PCR), the product is incubated with the restriction enzyme, and then the DNA is analyzed by electrophoresis. If the enzyme cuts, two fragments will result; otherwise there will be a single fragment. The presence or absence of the mutation can be inferred, depending on whether the mutation creates or destroys an enzyme-recognition site.

Although tests based on the digestion of DNA by restriction enzymes are relatively simple to perform and are reliable, they are limited by the need to find an enzyme that distinguishes mutant from wild-type sequences. Also, it is difficult to use this technique to detect multiple types of mutations simultaneously. Accordingly, two alternative approaches based on PCR amplification are now in wide use.

One involves amplification of a segment of DNA surrounding the site of a mutation and detection of mutant or wild-type sequences by hybridization with DNA probes specific to either sequence (Fig. 3B). The PCR product is separated into single strands and applied to two membranes. One is hybridized with a labeled oligonucleotide homologous with the wild-type sequence, the other with a mutant-sequence oligonucleotide. Under carefully controlled conditions, only the exact sequence match will be stable and can be detected when the unbound oligonucleotide is washed away. In this way, the presence of mutant or wild-type sequences in the DNA sample can be determined. This approach is reliable and lends itself to mass production. Multiple sequences can be amplified simultaneously in the same reaction, allowing different segments of a gene to be screened for mutation. Ingenious schemes have been devised to screen for several dozen mutations efficiently in a single sample, increasing the yield of detection of mutations even in the face of considerable genetic heterogeneity.

The other scheme for detecting mutations, which also permits simultaneous detection of multiple types of mutations (Fig. 3C), is based on the specificity of the PCR reaction itself. A PCR primer is designed that ends right at the site of a potential mutation. If the primer is homologous to the wild-type sequence, it will amplify only the wild-type sequence in conjunction with another primer some distance away in the gene. The wild-type primer will not amplify mutant DNA, however. Conversely, a primer that is homologous to the mutant sequence will amplify only mutant DNA. If
Figure 3. Detection of Mutations.

In Panel A, showing the detection of a point mutation by digestion of DNA with a restriction enzyme, the mutation creates a new recognition site. The region surrounding the mutation is amplified by PCR, and the resulting PCR product is incubated with the restriction enzyme and then analyzed by agarose-gel electrophoresis. Lane 1 shows DNA from a person without the mutation; there is only one band because the enzyme does not cut the DNA. Lane 2 shows DNA from a person homozygous for the mutation; the two bands represent the two fragments obtained after enzyme digestion. Lane 3 shows DNA from a heterozygote; there is one uncut fragment and two cut fragments.

Panel B shows a mutation detected by oligonucleotide hybridization. The segment of DNA is amplified by PCR, divided into aliquots, and spotted onto separate filter membranes, which are hybridized with a labeled oligonucleotide corresponding to the wild-type or mutant sequence. The amplified segment of DNA from a person with the wild-type sequence (WT) hybridizes only with the wild-type oligonucleotide, whereas the DNA from a person homozygous for the mutant (MUT) sequence hybridizes only with the mutant oligonucleotide. DNA from a heterozygote (HET) hybridizes with both oligonucleotides.

In Panel C, showing the detection of a mutation by PCR, the mutant sequence differs from the wild-type sequence by the substitution of an A for a C. To search for the two kinds of sequences by PCR, two primers are necessary. A separate reaction is carried out with each, together with a common downstream primer. With a wild-type gene, the primer corresponding to the wild-type sequence yields a PCR product. Similarly, the mutant primer produces a product with the mutant sequence. However, with the wild-type primer and the mutant sequence, or the mutant primer and the wild-type sequence, there is no PCR product. The agarose-gel pattern shows that DNA from a person homozygous for the wild-type allele (WT) reacts only with the wild-type primer; DNA from a person homozygous for the mutant (MUT) sequence reacts only with the mutant primer; and DNA from a heterozygote (HET) yields PCR products with both primers.

Panel D shows an example of the detection of a triplet-repeat mutation by Southern blot analysis or PCR. In the fragile X syndrome, a CGG repeat occurs near the 5′ end of the gene. The number of repeats ranges from 5 to 50 in the general population and from approximately 50 to 200 in asymptomatic carriers and exceeds 200 in those with the fragile X syndrome. To detect the abnormality, DNA is treated with a restriction enzyme that cuts at recognition sites flanking the CGG repeat. Hybridization on a Southern blot with labeled DNA from the region of the gene reveals a single band in a normal male subject (wild type). An asymptomatic male carrier will have a band of higher molecular weight, and a subject with a full mutation will have a very large, diffuse band because of the instability of the full-mutation allele. The normal and asymptomatic-carrier alleles can also be detected by PCR (right). The full-mutation allele cannot be amplified by PCR because it is too large.

All these techniques are used to detect subtle point mutations or small deletions or insertions in DNA. Recently, a new class of mutation has been discovered that is responsible for a number of interesting genetic disorders. These involve triplet repeats — runs of three DNA bases that are tandemly repeated up to dozens of times within or adjacent to some genes (Fig. 3D). Although the number of tandem repeats may vary from person to person, expansion of the region beyond a threshold disrupts the expression of the gene and leads to a genetic disorder. Triplet repeats have been found to underlie myotonic dystrophy, the fragile X syndrome, Huntington’s disease, and a number of other disorders, most of which affect the nervous system.

Both Southern blot and PCR-based diagnostic systems have been used to measure the size of the triplet-repeat regions, providing accurate diagnostic tests for these disorders.

The advent of accurate means of testing directly for
mutations in many genetic disorders has raised the possibility of population-wide screening for carrier status. This possibility poses both technical and ethical challenges that are just beginning to be addressed. In some cases only a proportion of mutations can be detected, so a negative test does not entirely rule out carrier status. Even if all mutations can be detected, the ability to carry out screening in the laboratory may outstrip the resources available to provide counseling to those who are found to be carriers. Molecular diagnostic techniques are already being used in established screening programs, such as those for thalassemia and Tay–Sachs disease. Their possible application in other disorders, including cystic fibrosis and fragile X syndrome, is under study.

All the approaches described so far presume that we have the ability to identify a pathogenic mutation directly. What about families with a genetic disorder in which the mutations cannot be located? If the gene is cloned or its location in the genome is identified by genetic mapping, a limited form of genetic testing can be offered that is based on linkage analysis (Fig. 4). This involves the use of polymorphic DNA sequences that are closely linked to or found within the gene to track
the mutant sequence through the family. The technique can be used only if two or more generations of affected family members are available for study and requires extensive analysis of the family for informative markers. It cannot be used in families that have only a single affected member or in which relatives are unavailable for testing or are unwilling to be tested. This approach is also limited by potential errors due to genetic recombination or genetic heterogeneity (if more than one gene locus is potentially responsible for the same disorder in different families). Despite these limitations, linkage-based testing has been very useful for counseling families with Duchenne’s muscular dystrophy, hemophilia, spinal muscular atrophy, and many other disorders.

Molecular diagnostic testing will undoubtedly continue to grow in importance as more and more genes responsible for disease come to light. New forms of technology are being developed that will permit the simultaneous determination of a person’s genotype at multiple loci. These will include not only loci responsible for rare genetic traits, but also those that contribute to susceptibility to more common disorders, including cancer, atherosclerosis, and hypertension. The landscape of preventive medicine will be changed dramatically, and the principles of genetics will become integral to day-to-day medical practice.

**Recommended Reading**
