A new strategy for mapping the human genome

DUNCAN J SHAW

From the Section of Medical Genetics, University of Wales College of Medicine, Heath Park, Cardiff CF4 4XN.

SUMMARY Recent advances in agarose gel electrophoresis of large DNA fragments raise the possibility of an entirely new approach to mapping mammalian genomes. In this article is discussed the potential of this technology for tackling problems such as construction of linkage maps, identifying chromosome translocation breakpoints, and moving from linked markers to genes causing diseases such as the muscular dystrophies and Huntington’s chorea.

Traditionally, human genetics has relied on family studies to construct linkage maps and on cytogenetics to identify chromosome structural abnormalities. More recently these approaches have been improved by the introduction of recombinant DNA to provide a source of new linkage markers, of improved chromosome banding methods to increase cytogenetic resolution, and of in situ hybridisation of cloned DNA sequences for accurate gene localisation. However, despite these advances, family studies and cytogenetics still have significant disadvantages, particularly in their inability to provide resolution at the level of a million base pairs (Mb) and smaller. Thus, there exists a gap in the range of resolving powers of the techniques currently available to the human geneticist, from family studies and cytogenetics at one level to DNA cloning and sequence analysis at the other.

In the last two years several strategies have been proposed to bridge this gap. In the opinion of the author, the most hopeful of these is the use of new forms of agarose gel electrophoresis capable of separating very large DNA fragments. By exploiting this technology, many problems could be tackled that are either difficult or impossible to solve by conventional means. These include the ordering of genetic markers and construction of complete linkage maps; the isolation of DNA sequences from chromosome translocation breakpoints; the detection of DNA deletions too small to be seen cytogenetically; and moving from closely linked markers to genes causing diseases whose biochemical basis is unknown.

Electrophoresis of large DNA fragments

In 1984, Schwartz and Cantor,1 and independently Carle and Olson,2 described a new form of agarose gel electrophoresis, which has since become known as orthogonal field alternation gel electrophoresis or OFAGE for short. Conventional agarose gel electrophoresis is one of the mainstays of the analysis of small DNA molecules such as cloned genes or viral genomes, but is limited by its inability to resolve DNA fragments longer than about 20 kilobases (kb). This limit arises because of the tendency of long DNA molecules to align themselves as rods in the agarose matrix and migrate end on, with a mobility that is independent of their length. In OFAGE, the DNA molecules are subjected not to one constant electrical field, but to two alternating fields, arranged at an angle of between 90° and 180° to each other (fig 1). Thus, the molecules are forced to turn a corner in the gel every time the field direction is changed. The speed with which a molecule can change its direction is a simple function of its length, and the size range of molecules that are resolved in an OFAGE gel is dependent on the time between each change of field direction. There is no theoretical upper limit to the size of DNA molecules that can be separated. In the original studies, OFAGE was used to separate whole yeast chromosomes whose lengths range from 30 to 2000 kb. It has since been applied to various other biological problems.

Mapping a human chromosome using OFAGE

A genetic map is a series of markers linked together in an orderly way and serves as a framework for investigating the structure and function of genes and
chromosomes. The markers used in human genetics include cytotypic bands, cloned genes and other DNA sequences localised by in situ hybridisation or somatic cell genetics, and polymorphic markers mapped by family studies and linkage analysis. As discussed earlier, these techniques do not provide resolution at the level of 1 Mb and below. The genomes of simpler organisms such as viruses can be mapped by a biochemical approach, namely restriction enzyme analysis. The markers in this type of map are specific DNA sequences recognised and cleaved by restriction endonucleases, and the distances between the markers are determined precisely by gel electrophoresis of the DNA fragments produced by restriction enzyme cleavage. If this approach could be applied to human chromosomes, it would provide a means of mapping at the 1 Mb to 10 kb level, the area not covered by current techniques.

Most restriction enzymes recognise sequences of 4, 5, or 6 base pairs and cut mammalian DNA into fragments a few hundred to a few thousand base pairs long. Therefore, a restriction digest of total human DNA produced by such an enzyme contains several million fragments. Specific fragments can, of course, be identified by Southern blotting and hybridisation with a radioactive probe sequence, but until recently there has been no way of using restriction enzymes to map large regions of DNA, such as whole human chromosomes.

Restriction enzymes whose recognition sequence consists entirely of G and C base pairs cut mammalian DNA rather infrequently, because stretches of DNA consisting only of G and C are under-represented in the genome. Enzymes such as NotI, whose eight base pair recognition sequence is GCGGCCGC, will cut DNA even less frequently. Although this has not been confirmed empirically, calculations based on dinucleotide frequencies indicate that NotI should cut human DNA, on average, once every million base pairs and enzymes with six base pair sites, such as XhoIII (CGGCGG), every 60 kb. Human DNA, digested with one of these enzymes, would be effectively fractionated by OFAGE yielding a few tens or hundreds of restriction fragments per chromosome. These fragments, after Southern blotting in the conventional way, could then be identified by hybridisation with chromosome specific probes. A complete restriction map could be generated by linking the fragments together as follows. A set of chromosome specific probes, each containing DNA sequences flanking a site for the infrequently cutting enzyme used to produce the genomic DNA fragments, is constructed. A procedure for doing this, currently being tested in our laboratory, is shown in fig 2. Each ‘junction...
probe' will hybridise to two different DNA fragments on the Southern blot of the OFAGE gel. These two fragments must originally have been contiguous in the genome (fig 3). Thus, by hybridising the blot with each junction probe in turn, a complete restriction enzyme site map of the chromosome in question may be generated.

It then becomes a relatively trivial matter to assign all of the existing chromosome specific markers to points on the map, by hybridising probes for these markers to the same set of OFAGE blots. The order of the markers and the distances between them (in physical terms) will be defined by their map positions.

A map constructed in this way is individual specific and, unlike genetic maps deduced from family studies, does not require one to assume that the order of markers is invariant. Indeed the technique is also potentially useful for studying long range variation in chromosome structure between persons.

Chromosome rearrangements

Deletions and translocations have been shown to have significant consequences, for example, in causing diseases such as Duchenne muscular dystrophy and retinoblastoma or in activation of oncogenes leading to neoplasia. Such chromosomal aberrations would show up as a disruption of the normal OFAGE restriction site map. Cloning of DNA from the breakpoint or deleted region would also be possible, as follows. The deletion or breakpoint will manifest itself as an OFAGE DNA fragment of abnormal length. This fragment, together with its normal homologue, can be isolated from the gel by standard procedures, and its structure analysed in greater detail by further restriction mapping and molecular cloning. In this way the site of the chromosomal rearrangement could be defined precisely and any gene whose normal function is being disrupted identified.

Linked markers and disease genes

Many genetic diseases whose biochemical basis is not understood have now been mapped using linked RFLPs. These include Duchenne, Becker, myotonic muscular dystrophies, Huntington's chorea, cystic fibrosis, and polycystic kidney disease. The problem now facing workers in this field is to 'walk', 'hop', or 'run' the few million base pairs from the linked marker to the disease gene itself. There are at least two ways in which OFAGE could be applied.

Firstly, it has been shown that in several Mendelian diseases, including Duchenne muscular dystrophy, and thalassaemias, and familial hypercholesterolaemia, DNA deletions at the disease locus are often involved. Such deletions would be readily identifiable by OFAGE, as described above. Deletions as small as a few kilobases should be detectable, since changes of fragment lengths of a few percent are easily resolved by electrophoresis. (Cytogenetic techniques can only detect deletions of several megabases in length). Thus, having identified closely linked markers for a genetic disease, it would be well worth screening a series of patients and controls for deletions by OFAGE.

Secondly, OFAGE can provide DNA from a small, precisely defined region of a chromosome, as a source of new probes or for complete molecular cloning or restriction enzyme mapping. This would be done merely by isolating the relevant DNA fragment from the OFAGE gel. Such an approach should be much simpler than other proposed methods for isolating small regions of chromosomes, such as use of somatic cell hybrids or chromosome walking.

Conclusions

OFAGE is a new technique that appears to have great potential in tackling a number of problems of interest to human geneticists. Many laboratories are now familiarising themselves with the procedures involved, and as a result it is likely that many significant discoveries will be made in the next few years.
References


Correspondence and requests for reprints to Dr Duncan Shaw, Section of Medical Genetics, University of Wales College of Medicine, Heath Park, Cardiff CF4 4XN.