INVITED REVIEW

DNA ANAYSIS FOR DIAGNOSTIC PURPOSES IN HUMAN GENETIC DISORDERS

(Received 3 June, 1996)

A.Tolun, Ph. D.*

** Professor, Department of Molecular Biology and Genetics, Boğaziçi University, Istanbul, Turkey.*

Molecular genetic techniques were developed greatly in the last decade, and currently are being applied to various fields because of their high sensitivity and the ease in application. DNA methods are used for an increasing number for genetic disorders for several purposes:

1. **Diagnosis of diseases.** Many of the genetic diseases are known to have genetic heterogeneity, and it is important to know the mode of inheritance and the locus of the gene in order to treat or cure the disease, if possible, and to perform carrier determination and prenatal diagnosis in those members of the family who are at risk. For example, two muscular dystrophies (Duchenne and Becker types) result from defects in the same gene on the X chromosome, while several other dystrophies are autosomally inherited, with gene loci on different autosomal chromosomes. It is not possible in most cases to differentiate the disorders only on the basis of clinical findings and pedigree information.

2. **Early diagnosis.** A patient may be too young for diagnosis on the basis of clinical findings or pedigree information. He can be tested using DNA analysis designed for the most likely disorder(s).

3. **Prenatal diagnosis.** Fetus of parents at risk can be tested during its first trimester.

4. **Carrier detection.** In families at risk, it is essential to identify carriers so that prenatal diagnosis can be performed for fetuses at risk.

Samples to be used in DNA analysis can be obtained easily from individuals. A few milliliters of peripheral blood from adults is sufficient for hundreds of assays. Embryonic DNA is obtained from cells in the amniotic fluid in the second trimester, or better yet in the first trimester, from chorionic villus samples.

Why is it necessary to perform DNA analyses for these purposes, and why is it preferred over other methods? Several different reasons may exist: First of all, there might not exist any other suitable method, simply because the biochemical defect may be unknown, as is the case with most of the genetic disorders for which the genes have not been identified yet. In most cases in which the genetic defect is known, no convenient biochemical assay exists. For example, there is no common method yet to detect the protein of the recently identified ataxia-telangiectasia gene. For some diseases, tissue samples other than the chorionic villus may be required as in the case of the muscular dystrophies. Muscle biopsy specimens are necessary to detect relevant proteins (1).

Different DNA assay strategies are employed depending on whether the gene responsible for the genetic disorder has been identified, or whether the gene mutation in a particular family concerned is known. In case of a genetic disease whose gene has not been identified, at least the gene locus should be known to perform prenatal diagnosis, and to identify carriers and affecteds. A common example to this situation was the Huntington's disease prior to the identification of its gene in 1993: DNA linkage analysis was performed in affected families for ten years after the localization of the gene to 4p16 in 1983 (2,3). A similar assay is used in families affected by diseases whose genes are known, but the mutation in the family has not been identified. DNA analysis is simplest in families in which the mutation is known: The members of the family who are at risk are analyzed for the presence of the specific mutation. To diagnose a sporadic patient for a suspected disease, mutation search is carried out in the disease gene. I would like to explain these points with examples from our experience in our laboratory on a few genetic diseases for which we have been carrying out the molecular genetics studies. Technical details will not be given since this paper is intended to be for non molecular geneticits, and the details of the analytical methods can be found easily in laboratory textbooks. However, it is impotant to emphasize that

the techniques employ mostly DNA fragments amplified by the polymerase chain reaction (PCR)(4). Hundreds of thousands of copies of specific gene regions can be obtained this way, rendering DNA assays much simpler than if the whole genomic DNA were to be used. There is intensive effort to improve the general methods constantly since the commercial benefit is formidable. Several private laboratories are offering DNA analysis services, along with laboratories in hospitals which can offer the service for a legal fee.

The various approaches for DNA analysis suitable for different situations are described below:

A. A familial case with a disorder for which either the gene has been localized but not identified, or the gene is known but the mutation in the family has not been identified. The only strategy is to use DNA linkage analysis, which requires a study of all family members to determine the inheritance of the disease allele by the individuals. Thus it is possible to identify the members who carry one or both parental disease alleles, and also those who are completely normal genotypically with regard to the disease. Linkage analysis employs DNA polymorphisms. A DNA polymorphism is found most commonly in the parts of our genome that do not code for proteins. Thus a DNA polymorphism does not affect a protein's structure. These polymorphisms are DNA variations that are thought to have arisen several thousands of years ago. It is estimated that one in every two hundred nucleotides (building blocks of the DNA) are different in any two unrelated individuals.

Let us assume that there is a DNA polymorphism that is very close to a disease gene locus. In a family affected with the genetic disorder, it is possible to determine which of the alleles is inherited together with the disease allele, i.e. linked to it; thus the polymorphism is considered a marker. A specific allele of the marker has no general association with the disease allele; different alleles might be linked with the disease gene in different families. Suppose in a family affected by Duchenne muscular dystrophy (DMD), a fatal X-linked neuromuscular disorder, a patient carries allele 1 at the gene locus, and the mother carries allele 1, as expected, and also a different allele, say allele 2. In the mother's next pregnancy for a male fetus, the fetus will either carry allele 1, in which case he will be considered to be affected, or allele 2, indicating that he is normal. Similarly, her daughter who has inherited a maternal allele 1 is a carrier. However, to determine the maternal allele, the paternal allele needs to be known. If the father's DNA sample is not available, his allele may be deduced from those of several daughters'. The daughters' common allele represents the paternal allele. Let us assume that one of the daughters has allele 1 (A1) and allele 2 (A2), and the other A2 and A2. This wil show that the paternal allele is A2. The assay is not so simple in autosomally inherited diseases, since each parent carries one disease allele. Another example for linkage analysis we were obliged to perform in the lack of a key family member's DNA sample was when the patient was no longer alive. Two sisters applied during pregnancy, and the only other blood sample available was that of the mother's. Both the affected brother and the father had died several years ago. We found that the mother was homozygous for an allele at the gene locus (let us say A1), and both daughters had alleles A1 and A2, thus the paternal allele was deduced to be A2. Since the son had died of DMD (X-linked), mother was considered a carrier, and either of her alleles (both being A1) could have been linked to the disease. As a result, luckily, both of the daughters were informative at the locus, and the paternal allele was known. Since there was no way to identify the disease-linked allele, even if the mother were informative, maternal allele exclusion was the only way to realize prenatal diagnosis. Upon analysis of fetal DNA samples, it became obvious that the younger daughter had passed her maternal allele to her male fetus, thus he was at risk (50%), while the male fetus of the older daughter was more lucky, having inherited his mother's paternal allele with no risk.

An individual is said to be informative if he/she is a heterozygote at the locus of the polymorphism. This is essential in determining the inheritance (by the sibs) of the disease allele which is linked to the polymorphic marker. In the example above, the mother was not informative in the A locus. It may be necessary to test an individual for several polymorphic markers at or close to the gene locus to come across a marker for which the individual is informative. Thus polymorphisms with a wide distribution in the population are more useful. This property is observed more commonly at loci with more than two alleles.

DNA polymorphisms are of two types: restriction enzyme fragment length polymorphisms (RFLP) and microsatellites of variable repeat number. Most RFLP's have only two alleles, the less frequent being around 20-45 percent, resulting in a heterozygote frequency of 32-50 percent. Restriction enzymes are used to determine the alleles. Microsatellites, on the other hand, typically have a very large number of alleles (3-20 or more), raising the heterozygote frequency to over 80 percent. The polymorphic region of an individual, when amplified by PCR, contains fragments of two different sizes. These fragments can be resolved on suitable electrophoresis gels.

B. A familial case with a known mutation. When the mutation in a particular family is known, members can be screened for that specific mutation. With the available mutation detection methods, this can be done very efficiently in a short time. This approach has two advantages over the linkage analysis described above: A family study is not necessary, thus non availability of samples from all members would not hinder the analysis, and there is no risk of misdiagnosis due to DNA recombination (chromosomal crossing over) as posed by linkage analysis when the polymorphic marker is not very close to the gene locus.

C. A family with a single case. A definite diagnosis is very important, since this will determine both what should be done for treatment, as well as the strategies for carrier detection and prenatal diagnosis. For example, B-thalassemia in a child can be diagnosed easily by a protein assay. Although carrier detection can be performed with the same method, prenatal diagnosis preferably requires DNA analysis. Either linkage analysis for the family or mutation detection in the patient can be done. The gene is small, and reported sporadic cases are extremely rare. However, the situation is very different for most other diseases such as the muscular dystrophies. Protein assays are expensive, complicated and are offered in few centers. Diagnosis is not reliable if based only on clinical findings, because there are several types of the disease resulting from different genes. There are X-linked types, and there are autosomal types. Because the X-linked type (Duchenne/Becker muscular dystrophy: DMD/BMD) is the most common of the muscular dystrophies, mutation analysis designed specifically for this disorder is performed on all suspected cases. We find about 30 percent of the suspected cases to be X-linked. DNA deletions are common in this disease: About 60 percent of patients have deletions. Thus 60 percent of clinically uncertain DMD/BMD cases are diagnosed rapidly by deletion analysis. For the others, protein assay is necessary and is performed in the biochemistry laboratory of the neuromuscular disease clinics in Istanbul Medical School. A muscle biopsy section is immunostained for proteins of various genes for final diagnosis.

What can be done if clinical findings do not lead to a definite diagnosis, and if no chemical method is available in case of a patient with no family history? A search for mutation is launched in the most likely gene. This is the approach we take for cystic fibrosis patients whose clinical findings are not in total agreement with classical cystic fibrosis. However, the analysis takes a very long time, since it involves the analysis of 27 exons using the method of denaturing gradient gel electrophoresis for point mutations (mutations involving only 1-3 nucleotides).

Problems to be considered:

In some genetic disorders sporadic cases are very frequent. In DMD one third of all cases are estimated to be sporadic, corresponding to one in ten thousand boys. Thus in a family with a single affected boy, the sisters carrying the allele with the same marker may or may not be carriers, depending on when the mutation (the disease allele) arose. The mutation might have arisen in a germ cell of the mother or of any of her ancestors. In the former case only the affected has the mutation, while in the latter, a parent is a carrier. Also, the mutation might have arisen in a progenitor of a germ cell, thus there would be a large number of germ cells that harbor the mutation, a phenomenon known as germinal mosaicism. A germinally mosaic person possibly may have several progeny with identical mutations, even though she/he does not carry the mutation in the blood cells, which is the tissue commonly used for DNA analysis. In DMD, germinal mosaicism is estimated to be 10-15% in mothers of affected boys, and much less in the fathers. In these cases, a linkage analysis will not be useful for diagnosis. Unless the mutation in a seemingly sporadic case has been identified, usually after an extensive DNA analysis, and the members of the family at risk are screened for that mutation, it is impossible to say for sure who really carries the mutation in the family.

Another phenomenon that may render diagnosis by linkage analysis uncertain is a probable crossing over during meiosis in the mother's germ cells between a polymorphic DNA marker and the mutation site. As a result of the crossing over, the allele that is known to be linked to the disease allele is now linked to the normal allele. The longer the distance between the marker and the site of the mutation, the higher is the probability of crossing over. There are also other factors that increase the probability of crossing over. In the dystrophin gene which is the largest gene known and responsible for DMD/BMD, linkage analysis involves at least two intragenic markers in families without deletions.

Non paternity usually leads to misdiagnosis in cases in which paternal alleles need to be taken into account. Non paternity is estimated to be 10-15% in Western Europe and the North America. A paternity test using highly polymorphic markers resolves the issue.

Table I : DNA analyses available in Istanbul

	Genetic Disorder	Mutation Detection	Linkage
Bağaziçi University MBG Department			
	Cystic fibrosis	yes	yes
	B- thalassemia	yes	yes
	a- thalassemia	yes	yes
	Hemophilia A	yes	yes
	Hemophilia B	yes	yes
	Spinal muscular atrophy	yes	yes
	Charcot - Marie Tooth type I	yes	no
(together with DETAE)	DMD/BMD	yes	yes
Istanbul University DETAE			
	Spinal muscular atrophy	yes	yes
	Phenylketonuria	yes	yes
	Kennedy disease	yes	no

REFERENCES

1. *Matsumura K, Campbell KP. Dystrophin-glycoprotein complex: its role in the molecular pathogenesis of muscular dystrophies. Muscle Nerve 1994;17:2-15.*
2. *Gusella JF, Wexler NS, Corneally PM, et al. A polymorphic DNA marker genetically linked to Huntington's disease. Nature 1983;306:234-238.*
3. *MacDonald ME, Ambrose CM, Duyao MP, et al. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. Cell 1993;72:971-983.*
4. *Saiki RK, Gelfand DH, Stoffel S, et al. Primer directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 1988;239:487-491.*