David I. Hoar

Molecular Diagnosis: New Horizons in Medicine

SUMMARY

Recombinant DNA technology, one of the major controversial areas of biological research in the late 1970s, is now rapidly providing new avenues for diagnosis and treatment. With the early recognition that extensive DNA variation exists in human populations, molecular genetic diagnosis of a variety of common hereditary diseases has become a reality. Recent identification of the location of the gene (or genes) for cystic fibrosis and adult polycystic kidney disease, and characterization of the region of the Duchenne muscular dystrophy gene will lead us towards a better understanding of the basic defects in these diseases. The identification of large multi-generation families with genetic diseases that are useful for identifying gene locations will require the co-operative participation of clinicians, medical geneticists and molecular biologists. (*Can Fam Physician* 1987; 33:401–404.)

SOMMAIRE

La technologie de recombinaison de l'ADN, l'un des domaines de la recherche biologique les plus controversés vers la fin des années 1970, offre actuellement, à un rythme accéléré, de nouvelles avenues diagnostiques et thérapeutiques. Avec la reconnaissance précoce de l'existence d'une grande variation des taux d'ADN chez les êtres humains, il est maintenant devenu possible, grâce à la génétique moléculaire, de diagnostiquer une variété de maladies héréditaires courantes. L'identification récente de la localisation du gène (ou des gènes) de la fibrose kystique et de la polykystose rénale chez l'adulte, et la caractérisation du site du gène de la dystrophie musculaire de Duchenne nous amèneront à mieux comprendre les anomalies fondamentales de ces maladies. L'identification des familles nombreuses dont de multiples générations ont présenté des maladies génétiques, lesquelles seront utiles pour identifier les localisations de gènes, nécessitera la collaboration des cliniciens, des généticiens médicaux et des biologistes moléculaires.

Key words: molecular genetics, DNA banking

Dr. Hoar is an associate professor in the Departments of Medical Biochemistry and Pediatrics at the University of Calgary. He is also the Director of the Molecular Diagnostic Laboratory, a component of the Alberta Hereditary Diseases Program (Southern Region), located at the Alberta Children's Hospital. Reprint requests to: Dr. David I. Hoar, Departments of Medical Biochemistry and Pediatrics, University of Calgary, 3330 Hospital Drive N.W., Calgary, Alta. T2N 4N1

DURING THE PAST 10 YEARS there has been a remarkable expansion in our knowledge about the molecular basis of a wide variety of human hereditary diseases. These developments have taken place primarily because of the powerful investigative advantages imparted by the techniques of modern molecular biology. The techniques of recombinant DNA technology have permitted the laboratory research scientist to dissect the human genome down to the level of its basic coding unit, the single-base pair. These new technologies have permitted us to discern simple alterations that lead to genetic diseases.

Although scientists are today discussing the task of obtaining a total DNA sequence of the human genome, our most powerful approach to the predictive diagnosis of genetic diseases, at the present time, is through the exploitation of our natural variation in DNA sequence. It has been estimated that the human genome contains between three and four billion base pairs of genetic information. Various studies suggest that as many as one in every few hundred base pairs may be different between individuals and hence, there must be significant inter-individual variation in DNA sequence.

It has become possible to discern this variation through the combination of three discoveries that form the cornerstones of modern molecular genetics. The first and most important involved the identification of the molecular scalpel or the restriction enzyme.[1] These enzymes, usually of bacterial origin, recognize specific DNA sequences and, hence, can be used to generate reproducible patterns of fragments from human DNA. There

now exist several hundred different restriction enzymes, each with a characteristic specificity. These restriction enzymes provided the cornerstone for the development of molecular cloning.[2] This second important discovery was perhaps the most controversial, as there was considerable discussion of the possibilities for abuse of this technology. The techniques for cloning specific DNA segments into bacterial hosts and amplifying them billions of times have provided the reagents for molecular hybridization that permit the identification of unique fragments or subsets of fragments from among the millions generated by a single restriction digestion of human DNA. The third important step towards the characterization of the human genome was the separation and immobilization of these million or more DNA fragments according to size, using a technique developed by Ed Southern.[3] This technique, referred to as the 'Southern Transfer', has become the standard technique employed in the molecular diagnosis of human diseases.

We are rapidly moving into an era when molecular biology will be a part of the repertoir of clinical diagnosis. In addition to the uses of recombinant DNA technology in human genetics,[4, 5] the potential applications in the fields of virology,[6] microbiology[7] and oncology[8] are substantial. Of particular interest to the geneticist is the definition of genetic markers that would be useful in predictive diagnosis of disorders where the molecular basis remains unknown.[9-11]

The early developments of molecular diagnosis took place in the study of the hemoglobinopathies.[12] The lessons learned here have provided the basis for analysis of a wide variety of single-gene disorders, as it soon became evident that similar mechanisms of mutation have led to the diverse genetic diseases observed in humans.

One of the early observations that has proved most useful was that there are variant restriction fragments that are polymorphic. These Restriction Fragment Length polymorphisms (RFLPs) often vary with ethnic backgrounds, and specific variations may be found in association with mutations causing disease. Hence, in the case of sickle cell anemia, a particular beta globin RFLP was identified in association with the disease gene in blacks of African origin.[13] This RFLP could be used for predictive diagnosis of sickle cell disease in many such families. Subsequently, in the case of the sickle cell mutation, a particular restriction enzyme was identified that could differentiate between the normal DNA sequence and the sickle mutant. The direct mutation detection method has replaced RFLP linkage for prenatal diagnosis in this disease.[14] This progression from the use of a linked RFLP to a definitive marker for the specific mutations has led to more accurate predictive diagnosis in several diseases.

The Search for Linkage

It has been established that as few as 250–350 evenly spaced RFLP DNA markers would permit the mapping of any genetic disease in the human genome. At the present time, many more than 350 DNA markers exist; however, they are not evenly distributed. Using collections of RFLPs, scientists have mapped a variety of significant genetic diseases in the human genome, without any knowledge about their basic defects. Examples of such mapping discoveries are Huntington's Disease,[9] adult polycystic kidney disease,[11] cystic fibrosis[10] and Wilson's Disease.[15] A significant number of genetic diseases that are known to reside on particular chromosomes have been mapped more precisely and RFLPs or disease allele specific markers identified. Examples of the latter group would include the Duchenne and Becker forms of Muscular Dystrophy,[16] Chronic Granulomatous Disease,[17] X-Linked retinitis pigmentosum,[18] Norrie's Disease,[19] Hemophilia A[20] and Hemophilia B[21] on the X chromosome, and a variety of autosomal diseases such as Myotonic dystrophy,[22] phenylketonuria,[23] alpha-1-antitrypsin deficiency[24] and familial retinoblastoma.[25]

In some diseases the DNA markers used in diagnosis may reside within the region coding for the gene product (e.g., Factor VIII and Factor IX). In other instances the DNA marker may be close to the gene, and hence a possibility exists that recombination can occur between the marker and the disease locus. The accuracy of predictive diagnosis using RFLP markers is therefore a function of how closely the markers are linked to the mutation causing the disease. Recent data on the Duchenne Muscular Dystrophy locus[16] indicate that this is a very large region, and even with DNA markers within the gene, predictive errors of 5%–6% are to be expected.

The development of new genetic markers that are useful for analysis of linkage with diseases will require the co-operative efforts of the observant clinician, the molecular biologists and, in particular, the families that exhibit the genetic disease. In order to establish RFLP linkages to diseases, it is necessary to study the DNA from affected and unaffected individuals in a family unit over several (usually three) generations. When the RFLP is polymorphic within the family unit, the linkage can often be defined, and relative distance between the gene in question and the RFLP be calculated, to determine the usefulness of the RFLP for predictive diagnosis. The accumulation of data from many families is often necessary before such DNA markers can be used routinely in predictive diagnosis.

As many genetic diseases are extremely rare and large family units suitable for RFLP linkage analysis exceedingly infrequent, all clinicians should keep in mind the possibility that they may have an informative family in their practice. The techniques outlined below provide a means of banking DNA from such families for study.

DNA Banking for our Future

Although several approaches to storage of suitable samples have appeared in the literature, we believe that much of the information now available is insufficient, and that many of the published methods are cumbersome and have unacceptably low DNA yields requiring large volumes of whole blood. An average, healthy, adult blood sample has five to ten million white cells per ml. Each cell contains 6.6 picograms of DNA; therefore, each ml of blood should yield 30–60 μg of DNA, enough for up to 30 analyses at 2 μg/track. (Molecular diagnosis can be accomplished on 1 μg, under ideal conditions.) The technique reported here yields close to 100% of the theoretical DNA content of the sample, and the product is suitable for molecular diagnosis. Built into this technique is an early step at which the cell lysate can be stored under refrigeration for prolonged periods or shipped by mail without risk of degradation.

Cell lysate. Blood (7 mls) collected with ethylenediamine tetra-acetate (EDTA) (or, if unavailable, Heparin) as

an anticoagulant is mixed with exactly five volumes of warm (37°C) Tris/Ammonium chloride (17 mM Tris pH 7.65, 0.14 M NH_4Cl) (35 mls) and incubated for five minutes at 37°C. The tube is then centrifuged for 10 minutes at 2000 rpm in a swinging bucket table-top centrifuge, and the supernatant removed by suction to leave about 5 ml of liquid containing the soft pellet of nucleated cells. This pellet is immediately resuspended in 20 ml of normal saline (0.85% NaCl) and centrifuged (2000 rpm, 10 minutes) to wash the cells. The cells may be washed a second time with normal saline to reduce the amount of soluble protein, and the final washed pellet of nucleated cells from 7 mls of blood is thoroughly resuspended in 2 mls of TE 100, 40 (Tris 100 mM pH 8.0, EDTA 40 mM). This cell suspension is immediately lysed by the injection of 2 mls of TE 100, 40 containing 0.2% sodium dodecyl sulfate. Lysis is instantaneous and complete if the cells are resuspended thoroughly, and the lysis mix is injected using a 5 ml syringe fitted with a 16–18 gauge needle. The lysate can be extracted immediately, shipped to service laboratory centres or stored for extended periods in a sealed tube containing two or three drops of chloroform as a preservative, or the DNA extracted immediately. DNA isolated from lysates stored at refrigerator temperature for four years or longer is suitable for restriction analysis.

This preparation of lysates from whole blood works well on fresh blood from a wide variety of genetic diseases; however, where cell-membrane fagility is suspected or blood is more than four or five days old, there may be losses caused by increased osmotic fragility of white cells. In these situations, whole blood can be gently centrifuged, the pellet resuspended in an equal volume of TE 100/40, and the suspension lysed immediately as outlined above. Samples lysed in this fashion should have NaCl or LiBr added to a final concentration of 0.5 M before extraction. They may require several additional extractions with phenol in order to clean the DNA sufficiently.

Most techniques employed for the extraction of DNA from peripheral blood involve a proteolytic digestion step using pronase or proteinase K. We strongly advise that this not be done, as such treatment will produce aqueous soluble glycopeptides from the glycoproteins, and these will copurify with DNA, yielding a product that is sometimes refractory to restriction digestion.

DNA extraction. An equal volume of TE 100, 40 saturated, freshly distilled or stabilized phenol (0.1% 8 hydroxyquinoline) is added to the lysate, and extraction is carried out by mixing thoroughly for 10 minutes on a rotary mixer at 60 rpm. (A milky emulsion should be maintained throughout the extraction.) The phases are separated by centrifugation (2000 rpm for 5 minutes), and the upper, clear, nucleic acid-containing aqueous phase is saved. (All pipetting is done with large-bore pipets, and organic extractants should be handled with glass pipets.) The milky interface is mixed with four volumes of TE 100, 40, reextracted with phenol, centrifuged and the upper aqueous phase pooled with the original aqueous phase. The phenol extraction is repeated once more on the pooled aqueous phases. The upper phase is then removed to a new tube and extracted once with an equal volume of chloroform: isoamyl alcohol (24:1). Following centrifugation to separate the phases, the upper aqueous phase is mixed thoroughly with 1/10 volume of 4 M ammonium acetate, and then an equal volume of absolute isopropanol is added. As the mixture is swirled to mix, long strands of high molecular weight DNA will form and collect. These are "fished out" as a clump with a hooked glass pasteur pipet tip and washed by dipping into a clean tube containing 70% ethanol. The DNA is briefly air dried and dissolved overnight in 0.1 ml of Tris 10 mM pH 8.0, 1.0 mM EDTA, for each ml of whole blood used.

DNA samples isolated in this fashion are routinely stored at 4°C. Where prolonged storage is anticipated, one drop of chloroform should be added to the DNA sample and a tightly sealing tube used. It is not necessary to freeze such DNA samples, and occasional exposure to ambient temperatures will not be detrimental.

The advantage of this technique is that it yields clean, essentially RNA-free DNA from whole blood in a reasonably short time. In addition, at the lysate stage, the material can, in theory, be stored indefinitely in the cold, and it can be transported by mail without risk of degradation. If the lysate is to be split into several samples, these samples must be removed by "cutting" with clean scissors and forceps. It is imperative that all solutions, tubes and pipets be free of possible contaminating DNA, and we strongly recommend that only disposable pipets and tubes be used.

Other DNA Sources

A number of situations are encountered by family practitioners and obstetricians that could benefit from the storage of DNA samples. Of particular interest are stillbirths or early neonatal deaths where a biochemical basis might be expected. Tissue samples as small as 1 cubic centimetre, taken from the placenta or a short segment of the cord, can be stored for several days in sterile tissue-culture medium in the refrigerator or frozen at −80°C for extended periods of time. In the case of non-immune hydrops fetalis, it is possible to obtain a retrospective DNA diagnosis with such materials at any time in the future, thus permitting accurate counselling of the family.

At autopsy, it may be impossible to obtain a suitable blood sample for DNA preparation. In such situations we recommend obtaining tissue. For DNA, the tissue can be stored frozen; however, when skin or other tissue is obtained fresh and sterile, it can be preserved for tissue culture, chromosome analysis, or biochemical analysis by first mincing it finely in tissue-culture medium, then placing the pieces in medium containing 10% glycerol in a freezing vial, quick freezing, and storing in liquid nitrogen or at −80°C. Tissue preserved in this fashion can be used to establish tissue-culture cell strains at any time in the future simply by thawing the vial, immediately washing the tissue pieces in a large volume of pre-warmed media, using centrifugation, and explanting the pieces under coverslips.[26] Our laboratory routinely splits all skin biopsies into two portions, one of which is minced and frozen, while the other is set up immediately to establish the cell strain. We have established primary cell strains from frozen tissue pieces (skin, liver, kidney, lung, and tumour material), using this approach, up to seven years after the original biopsy was obtained.

In order to develop the "New Genetics"[27] to its full potential, clinicians should establish a mechanism for obtaining DNA samples from the potentially informative family units. With

very little practice, the techniques presented here or similar ones should be made available in the routine clinical laboratory. This will permit a collective co-operation in the banking of DNA for our future. ◉

References

1. Linn S, Arber W. Host specificity of DNA produced by *Escherichia coli* X. *in vivo* restriction of fd replicative form. *Proc Natl Acad Sci. U.S.A.* 1968; 59:1300–6.

2. Cohen S, Chang A, Boyer H, Helling R. Construction of biologically functional bacterial plasmids *in vitro*. *Proc Natl Acad Sci U.S.A.* 1973; 70:3240–4.

3. Southern EM. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* 1975; 98:503–17.

4. Antonarakis SE, Phillips III JA, Kazazian Jr. HH. Genetic diseases: diagnosis by restriction endonuclease analysis. *J Pediatrics* 1982; 100:845–56.

5. Motulsky AG. Impact of genetic manipulation on society and medicine. *Science* 1983; 219:135–40.

6. Arens MQ, Swierkosz EM. Simplified method for typing Herpes Simplex virus by endonuclease analysis. *J Clin Microbiol* 1983; 548–51.

7. Moseley SL, Hug I, Alim AR, So M, Samadpour-Motalebi M, Falkow S. Detection of enterotoxigenic Escherichia coli by DNA colony hybridization. *J Infect Dis* 1980; 145:892–8.

8. Birnie GD, Burns JH, Wiedemann LJ, Warnock AM, Tindle RW, Burnett AK, et al. A new approach to the classification of human leukemias: measurement of the relative abundance of a specific RNA sequence by means of molecular hybridization. *Lancet* 1983; 1:197-200.

9. Gusella JF, Wexler NS, Conneally PM, Naylor SL, Anderson MA, Tanzi RE, et al. A polymorphic DNA marker linked to Huntington's Disease. *Nature* 1983; 206:234–8.

10. White R, Woodward S, Leppert M, O'Connell P, Hoff M, Herbst J, et al. A closely linked genetic marker for cystic fibrosis. *Nature* 1985; 323–84. See also *Nature* 1985; 380–2 and 384–5.

11. Reeders ST, Breuning MH, Davies KE, Nichols RD, Jarman AP, Higgs DR, et al. A highly polymorphic DNA marker linked to adult polycystic kidney disease on chromosome 16. *Nature* 1985; 317: 542–4.

12. Orkin SH, Kazazian Jr. HH. The mutation and polymorphism of the human β-globin gene and its surrounding DNA. *Ann Rev Genet* 1984; 18:131–71.

13. Kan YW, Dozy AM. Polymorphisms of DNA sequence adjacent to the human β-globin structural gene: relationship to sickle mutation. *Proc Natl Acad Sci U.S.A.* 1978; 75:5631–5.

14. Geever RF, Wilson LB, Nallaseth FS, Bittner M, Wilson JT. Direct identification of sickle cell anemia by blot hybridization. *Proc Natl Acad* 1981; 78:5081–5. See also *N Eng J Med* 1982; 307:32–6.

15. Frydman M, Bonne-Tamir B, Farrer LA, Conneally PM, Magazanik A, Ashbel S, et al. Assignment of the gene for Wilson Disease to Chromosome 13: linkage to the esterase D locus. *Proc Natl Acad Sci U.S.A.* 1985; 82:1819–21.

16. Kunkel LM and co-authors. Analysis of deletions in DNA from patients with Becker and Duchenne Muscular Dystrophy. *Nature* 1986; 322:73–7.

17. Royer-Pokora B, Kunkel LM, Monaco AP, Goff SC, Newburger PE, Baehner RL, et al. Cloning the gene for an inherited human disorder—chronic granulomatous disease—on the basis of its chromosomal location. *Nature* 1986; 322:32–8.

18. Nussbaum RL, Lewis RA, Lesko JG, Ferrell R. Mapping X-Linked opthalmic diseases: II. Linkage relationship of X-Linked retinitis pigmentosa to X chromosomal short arm markers. *Hum Genet* 1985; 70:45–50.

19. Bleeker-Wagmakers LM, Friedrich U, Gal A, Wienker TF, Warburg M, Ropers HH. Close linkage between Norrie Disease, a cloned DNA sequence from the proximal short arm, and the centromere of the X chromosome. *Hum Genet* 1985; 71:211–4.

20. Gitschier J, Wood WI, Tuddenham EGD, Shuman MA, Gorlaka TM, Chen EY, et al. Detection and sequence of mutations in the factor VIII gene of hemophiliacs. *Nature* 1985; 315:427–30.

21. Choo KH, Gould KG, Rees DJG, Brownlee GG. Molecular cloning of the gene for human anti-hemophilic factor IX. *Nature* 1982; 299:178–80.

22. Davies KE, Jackson J, Williamson R, Harper PS, Ball S, Sarfarazi M, et al. Linkage analysis of myotonic dystrophy and sequences on chromosome 19 using a cloned complement 3 gene probe. *J Med Genet* 1983; 20:259–63.

23. Woo SLC, Lidsky AS, Guttler F, Chandra T, Robson KJH. Cloned human phenylalanine hydroxylase gene allows prenatal diagnosis and carrier detection of classical phenylketonuria. *Nature* 1983; 306:151–5.

24. Wilson-Cox D, Woo SLC, Mansfield T. DNA restriction fragments associated with α-antitrypsin indicate a single origin for deficiency allele PI Z. *Nature* 1985; 316:79–81.

25. Friend SH, Bernards R, Rogelj S, Weinberg RA, Rapaport JM, Albert DM, et al. A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. *Nature* 1986; 323:643–6.

26. Hoar DI. Familial mutagen sensitivities: are these the hallmarks of meiotic or mutator mutants in humans? *Can J Genet Cytol* 1979; 21:435–42.

27. Commings DE. Prenatal diagnosis and the "New Genetics". *Amer J Human Genet* 1980; 32:453–4.