Jumping Genes Help Trace Inherited Diseases

Molecular biology comes to the aid of human genetics with the discovery of markers that can be used to construct a genetic linkage map

A new approach to tracking—and therefore to potential prevention through counseling—the inheritance of genetic diseases is currently being developed in a collaborative effort between researchers at Massachusetts Institute of Technology and the University of Utah. David Comings, editor of *American Journal of Human Genetics*, heralded the still embryonic technique in a recent issue with the following words: "Since the degree of departure from our previous approaches and the potential of this procedure are so great, one will not be guilty of hyperbole in calling it the 'New Genetics.' "

"That's perhaps putting it a little strongly," demurs MIT biologist David Botstein, one of the initiators of the work. "Yes, the ultimate aim is to produce a genetic linkage map which will be the basis of tracing inheritance of genetic diseases through families," he admits. "But," he cautions, "there is still a very long way to go. In order to construct the linkage map we need to find 150, maybe more, good markers on the human genome. And so far we've found just one."

The markers to which Botstein refers are DNA polymorphisms, genetic loci occupied by one of a group of readily identifiable variants, or alleles. Virtually all polymorphisms so far detected in humans are enzyme variants, or isozymes, and cell surface markers, or antigens. A key point about the MIT-Utah technique is that the *function* of the polymorphic DNA is of no account. What is important is that the locus can readily be excised from the genome and its overall structure analyzed so as to determine which variant it represents.

"In principle, there should be thousands of such loci in the human genome," claims Raymond White of the University of Utah. "We need just a couple of hundred for a fully saturated map for the disease linkage." The markers would be scattered throughout the genome, each marker being separated from its neighbor by about 20 million base pairs. Once this basic complement of markers is established, the genetic linkage between them is to be determined, that is, which loci are associated with which others through the process of meiosis and crossing-over when gametes are formed. "We're very fortunate in having available the well-documented pedigrees of the large Mormon population here in Utah," comments White. "By studying the inheritance of the markers through several generations we will be able to see what linkage patterns exist," he says. "Mark Skolnick, here at Utah, will be overseeing this."

The last step is to discover with which markers the inheritance of each genetic disease of interest is associated. The gene or genes associated with the disease are not isolated or pinpointed, nor are their products isolated and characterized. It's a question of genetic guilt by association. A series of studies on family histories, involving tracing the path of inheritance of particular diseases with the inheritance pattern of the marker alleles, will be required for establishing this association.

There's a certain irony in the association between a visible but anonymous marker and an invisible but identified disease locus forming the basis of a powerful probe for genetic diseases. Comings muses on this in his recent editorial. "One of the major goals when studying genetic diseases is to find the primary gene product," he says, "which in turn leads to a better understanding of the biochemical basis of the disorder. The bottom line often reads, 'This may lead to effective prenatal diagnosis and eventual eradication of the disease.' But now we have the ironic situation of being able to jump right to the bottom line without reading the rest of the page, that is, without needing to identify the primary gene product or the basic biochemical mechanism of the disease."

The benefits of going straight to the DNA are twofold, claims Botstein. "First, these loci are more polymorphic than conventional markers, and, ideally, each variant occurs at relatively high frequency in the population," he explains. "And second, the alleles are all codominant, that is, the two variants in a heterozygotic pair of chromosomes are equally detectable." In principle, therefore, in a mating between two individuals that are both heterozygous for a marker in question, there will be information about the fate of all four relevant chromosomes. These are the best possible circumstances and are achievable with the polymorphic DNA markers.

The seeds of the project were sown in April 1978. Botstein and fellow molecular biologist Ronald Davis, from Stanford, were at Skolnick's laboratory listening to a seminar on the linkage between the genetic disease hemochromatosis and the HLA, or histocompatibility, antigen locus. "It came out in the discussion that the only thing that prevents the really useful upgrading of genetic information on inherited diseases was the availability of adequate polymorphic markers," recalls Botstein.

> The potential is so great, we will not be guilty of hyperbole in calling it the "New Genetics."

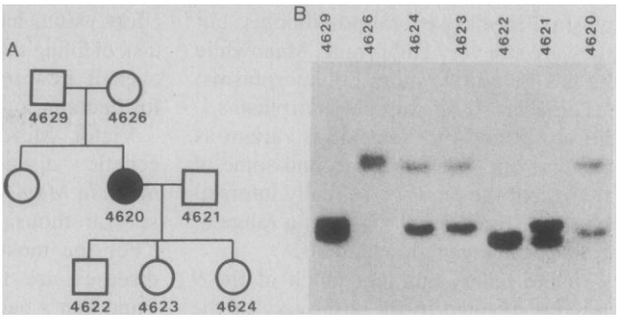
Both Botstein and Davis knew that such markers existed in yeast. And Davis had recently discovered the first of the jumping elements in nucleated, or eukaryotic, cells, elements that could be an important part of the origin of polymorphic loci. "We were therefore sensitive to the potential of polymorphisms," says Botstein, "and we decided something ought to be done about it."

At the time two other groups of workers—one in California, the other in Leicester, England—were investigating sequence variants in human globin genes with a view to their use in providing information for genetic counseling. By cutting cellular DNA with the restriction enzyme Hpa I, Yuet Wai Kan and Andrée Dozy, of the University of California, San Francisco, found variants in a stretch of DNA containing the β-globin

0036-8075/81/0213-0690$00.50/0

Inheritance of polymorphic DNA markers

Wyman and White examined the heritability of polymorphic markers by analyzing the digested, gel-smeared and hybrized DNA (B) of seven individuals over three generations (A). If one assumes that individuals 4629, 4626, and 4622 are incompletely resolved heterozygotes for fragment lengths of similar lengths, the fragment pattern of 4620 can be accounted for by the inheritance of the lower molecular size fragment from 4626 and 4629. If, for instance, a disease had been shown to be genetically linked with the lower molecular size fragment of 4626, then one could predict that 4620, 4623, and 4624 would have the disease. If associated with the higher size fragment, however, then no one else in the family would have it. [Proc. Natl. Acad. Sci. U.S.A. *77, 6754 (1980)*]



gene. It turned out that one of the variants was frequently associated with the sickle cell gene. In a paper on their work published in the November 1978 issue of *Proceedings of the National Academy of Sciences* they said that the polymorphism "could be considered as a new class of genetic marker" and that it may offer "a new approach to linkage analysis."

Alec Jeffreys, publishing a year later in *Cell*, reported three variants among globin genes using restriction enzymes to cut the DNA. He concluded that "such polymorphic variants should be detectable on all human chromosomes and could therefore ultimately be used for the indirect determination by linkage of any genetic disease suitably linked to a restriction enzyme cleavage site polymorphism."

Botstein and Davis were aware of this work, and they noted two crucial factors. One, the search for polymorphisms had focused on known genes. And two, the polymorphisms appeared to be the result of single base-pair changes that altered the position of the restriction enzyme cleavage site. "We thought it would be better not to start with the markers that Kan and Dozy and Jeffreys had described because of an intuitive belief that there was more polymorphism to be found elsewhere," recalls Botstein. If the yeast-type polymorphisms, which were based on jumping genetic elements, occurred in humans, then these in principle would give a richer array of variants, thus making the locus richer in information content. Random search for such elements was considered to be a potentially more profitable way to go. "Not everybody agreed," says Botstein.

In these early discussion stages, White, who had once worked with Davis and had maintained a long association with Botstein from a neighboring laboratory at MIT, became interested in the project. "White, who was then at the University of Massachusetts, jumped in with both feet and really made it work," says Botstein. He recruited postdoctoral fellow Arlene Wyman to the task.

The plan was to search for single-copy, as opposed to the more abundant repetitious, DNA that occurred as different alleles. To do this White and Wyman took DNA from many individuals, cleaved it with restriction enzyme Eco RI, distributed the fragments over a sheet of gel according to size, and then used a short section of single-copy DNA as a probe. The probe, which was radioactive, would then hybridize with its homologous sequence on the gel, showing up as a dark line on an autoradiograph. If an individual is homozygous for the DNA locus in the probe, then a single line appears in the radiograph. If, however, an individual has different length variants in his or her two chromosomes, then two closely associated lines stand out in the radiograph.

"For the probe we used a library of human DNA developed by Tom Maniatis and co-workers at Caltech," says Wyman. "They had sheared the DNA into fragments approximately 20 kilobases long, and attached them to two 'arms' of phage lambda so they could be easily cloned." The task then was to pick out the clones of single-copy DNA and use these one by one as probes until a polymorphism turned up.

"We prepared 30 or 40 clones of single-copy DNA initially," says Wyman. "We examined them systematically, seeing one nonpolymorphic region after another. And then one day we saw something different. This was the ninth probe we'd tested." What they saw from the DNA of a small number of individuals was a set of at least eight variants of the single-copy DNA contained in the probe; the variants differed in size between 14 and 29 kilobases. "This is a substantial degree of polymorphism," says White, "and it potentially has high information content for tracing inheritance." This was the fall of 1979, 9 months after the search had begun in earnest.

Wyman and White showed that the variants occur at high frequencies. And they confirmed that the locus follows Mendelian laws of inheritance. Both of these properties are crucial to useful application of the marker in genetic linkage analysis. These workers also determined what they had suspected: that the polymorphism was a result of a series of rearrangements of sections of DNA, rather than of simple base-pair changes. In their *Proceedings of the National Academy of Sciences* paper, published in November 1980, Wyman and White make the following comment: "The observation that the DNA structural basis for the polymorphism is due to a rearrangement series within the allelic restriction fragments suggests, by analogy with sites containing transposable elements found in *E. coli*, yeast, and *Drosophila*, that an active DNA element may reside at his locus."

One down, 149 to go. The challenge ahead could be steep. "When we first proposed the idea of a search for these polymorphisms, the human genetics community claimed that it would be impossibly laborious to find enough suitable markers on which to establish a linkage map," says Botstein. "With Arlene and Ray's paper, that objection is now lessened," he claims, "because it wasn't too laborious to find the first. One has to be aware, however, that this may have been a lucky break. None of us believes that, but it remains to be seen."

In November, White moved from Massachusetts to the University of Utah and Wyman joined Botstein at MIT. Workers at both laboratories are engaged in the steady search for further polymorphisms while also exploring new, more specific, ways of fishing. "Here at MIT we're trying to elucidate the molecular basis of the polymorphism that Arlene and Ray found," says Botstein, "not because it's important in the grand scheme of things, but because the information might help us find further polymorphisms more easily." If, for instance, a particular sequence or set of sequences is associated with the "active DNA element," then this could be used directly as a probe to pick up similar sequences elsewhere.

White is also developing new nonrandom methods of pinning down polymor-

phisms. He has several possibilities, but they are still very embryonic. Meanwhile he has found two more polymorphisms: "These are base-pair polymorphisms," he says, "and they are not as variant as the first one. We expect to find some of these, but the ones we're really interested in are those resulting from a range of transposable genetic elements."

White points out that much of the 9 months devoted to the discovery of the first polymorphism was concerned with the development of techniques. Nevertheless, it's clear that a considerable effort would have to be directed to the task of filling the required list of markers, even if new and faster techniques are forthcoming.

Victor McKusick's compendium of genetic diseases—*Mendelian Inheritance in Man*—lists all that is known for several thousand inherited conditions. "For the most part the data on these diseases are terrible," says Botstein, "and that's because they are so difficult to obtain." The system being nurtured at MIT and Utah offers the possibility of a real improvement here, mostly by circumventing the type of information classically required by geneticists.

Although Botstein acknowledges the potential of the work, he insists that "there is still a very long way to go, and there has to be a chance that it will turn out to be impractical." White's commitment to the work is demonstrated, he says, "by the fact that I've voted with my feet." He is fully optimistic about the final outcome, and comments that the project is "a delightful example of real fallout from basic research."

—ROGER LEWIN

There Is More to "Acid Rain" Than Rain

Significant amounts of acid and other pollutants reach the ground without the aid of precipitation, but most of it goes unrecorded

Scientists are not happy with the term "acid rain." It serves well in newspaper headlines and as a catchword for a growing international pollution problem. But "acid rain" neglects a part of the acid that has left mountain lakes barren of fish and corrodes car finishes. Some of that acid and numerous toxic chemicals, spewed from smokestacks and tail pipes, are being deposited on lakes, streams, trees, and fields without being washed out of the air by rain or snow.

The proportions of dry and wet deposition depend on the distance from the source of pollution and the wetness of the climate, but some scientists estimate that 10 to 30 percent of the acid problem may result from dry deposition. In some places, the figure seems to be at times greater than 50 percent. Researchers also agree that the dry deposition of pollutants, especially acid, is monitored crudely at best. While research on practical monitoring methods slowly progresses, a dispute continues about the usefulness of present methods, which were developed in the 1950's to deal with radioactive fallout.

One problem with measuring dry deposition is that acid, lead, arsenic, cadmium, and other pollutants do not always simply fall out of the sky, like raindrops or motes of dust. If they did, researchers could simply put out a bucket to catch them. Instead, some pollutants form particles so small that they fall slowly or, for all practical purposes, do not fall at all. A 0.1-micrometer particle, which can be formed by the chemical reactions of combustion gases, will be wafted around a flat surface rather than fall on it. Gases that can produce acid when dissolved in water, such as sulfur dioxide, do not fall at all.

Such finely divided particles and gases may not "fall out" onto the landscape, but they still reach the ground without being washed out by precipitation, apparently because natural surfaces can "catch" pollutants that a bucket cannot. In southern Ontario, near the ore smelting industry in Sudbury, conventional collectors miss as much as 60 percent of the sulfate—and presumably much of the sulfuric acid—entering lakes there, according to Peter Dillon of the Ontario Ministry of the Environment in Rexdale. He concludes that not all of the inputs of sulfur have been accounted for because more sulfate flows out of lakes than appears to enter them. Also, too little measured acid enters a watershed to account for the observed chemical compositions of lakes and streams. Much of the missing sulfur must add acid to a lake's watershed, Dillon says, because more calcium, magnesium, sodium, and potassium, which are leached from rocks and soil by acids, leave the lake than could be accounted for by the measured input of acid. The missing sources of sulfate amount to more than 60 percent of the total in watersheds 5 kilometers from the source at Sudbury, 40 percent at 10 kilometers, and 30 percent at 40 kilometers.

Most of the sulfur not collected by samplers as it enters the watersheds seems to be in the form of sulfur dioxide. That gas would not be collected efficiently by the wet deposition sampler—a pot that opens only during precipitation—or by the bulk sampler, a pot that remains open at all times to collect precipitation plus any particles that settle out. Sulfur dioxide would dissolve in the lake and in moisture on vegetation or in the soil, and it can be taken up directly by vegetation. Enough was in the air over the area to account roughly for much of the missing sulfur, Dillon says.

Farther from sulfur sources, the apparent contribution of sulfur dioxide decreases. Dry deposition remains significant, however, because the farther polluted air travels, the more sulfur dioxide gas is chemically transformed into sulfuric acid that forms submicrometer droplets or coats other particles. If not washed out of the atmosphere by precipitation or neutralized by basic substances, these particles can also carry acid to the surface. In the Haliburton-Muskoka area of southern Ontario, 200 kilometers from any large urban or industrial centers, about 20 percent of the acid was collected in bulk samplers, but not in wet-only samplers, according to a report by Wolfgang Scheider, Warren Snyder, and Bev Clark of the Ontario Ministry of the Environment.

Researchers have found that other watersheds having different climates and different sources of pollution also receive significant amounts of pollutants through dry deposition. The Hubbard Brook watershed in northern New Hampshire receives about one-third of its sulfur in a dry form, says Gene Likens of Cornell University. Most of it arrives

0036-8075/81/0213-0692$00.50/0