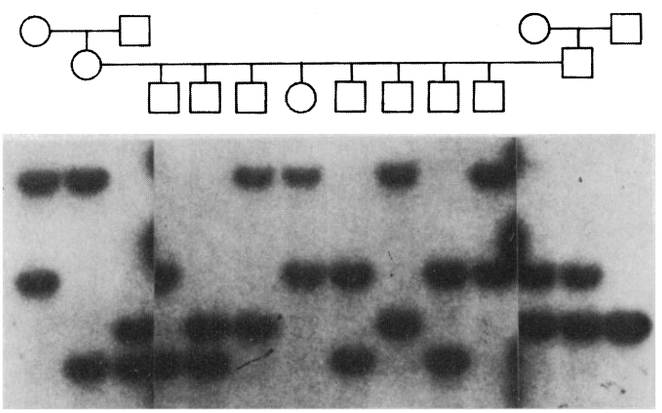
Putting the Human Genome on the Map

An unusual cooperative experiment for making a detailed linkage map of the human genome may produce a better understanding of genetic diseases

An unusual cooperative experiment, which is being performed under the aegis of Jean Dausset and the Human Polymorphism Study Center (CEPH in the French acronym) at the College de France in Paris, has as its goal the production of a detailed map of genetic markers encompassing all the chromosomes of the human genome. "The idea is to do for man what was done for other organisms, such as *Drosophila* and mice," explains David Botstein of the Massachusetts Institute of Technology.

Such a detailed map should enable investigators to zero in on the defective genes that cause hereditary illnesses, even in the absence of information about the biochemical effects of the mutations.



Polymorphism in a human ras gene

*The DNA from each member of the three-generation family was digested with a restriction enzyme and then probed with radioactive, cloned Harvey-*ras *DNA. The pattern of fragments detected by the probe is displayed for each individual under his or her position in the pedigree (females are represented by circles and males by squares). A close inspection of the patterns shows that the fragments are inherited in a Mendelian fashion.* [*Source: R. White* et al., Nature (London) **313**, *101 (1985)*].

"That is why mapping is so exciting an approach," notes Victor McKusick of Johns Hopkins University School of Medicine. According to a unique directory compiled by McKusick, some 1300 genetic diseases afflict human populations. For most of these diseases neither the cellular defects nor the affected genes have been identified. Examples include such serious conditions as cystic fibrosis and Huntington's disease.

But once the map is available, families afflicted with a particular genetic disease can be studied to see whether it is inherited in conjunction with any of the markers. If two traits are inherited together, then the corresponding genes are likely to be located near one another on the same chromosome. The marker would thus serve as a guidepost to the gene that causes the problem, possibly permitting its isolation. At the very least, the marker could be used for diagnosing the disease and for identifying carriers.

In addition, some conditions, including schizophrenia and diabetes, appear to be partially genetic in origin, although environmental factors may also play a role. A detailed marker map might help to identify the genes that contribute to the development of these diseases, too.

The human genome mapping project combines traditional methods of Mendelian genetic analysis with the newer techniques of modern molecular biology. Although genome mapping in *Drosophila* and mice began by following the inheritance of traits such as coat or eye color and determining the relative chromosomal locations of the genes encoding those traits, it is now possible to look directly for sequence variations in genomic DNA segments. These can also be mapped to their chromosomal locations and used as markers.

About 7 or 8 years ago, investigators learned that such variations can be detected with restriction enzymes that cut the DNA at specific sites. If the variation causes a particular segment of DNA to lose or gain a site that can be recognized and cut by one of the enzymes, then the sizes of the fragments produced by the enzyme will be altered. These size changes, which are called restriction fragment length polymorphisms (RFLP's) can be detected with a probe of cloned DNA that hybridizes to the fragments. Although RFLP's need not be located within a traditional protein-coding gene, they are inherited according to Mendelian principles.

Screening populations for the presence of RFLP markers is a relatively simple task. In fact, they have already been found linked to the genes for a small number of hereditary diseases, including Huntington's. Currently, however, suitable RFLP markers are few and far between. Finding a linkage between one and a particular genetic disease requires a certain amount of luck. The luck element would be much reduced if the human genome were more thoroughly mapped.

To facilitate the realization of that goal, Dausset has obtained permanent cultured cell lines from all the members of some 40 large multigenerational families that were identified by investigators from around the world. Many were supplied by Raymond White's group at the Howard Hughes Medical Institute of the University of Utah Medical School. Each family includes at least five children, although most have seven or more. About two-thirds of the families have all four grandparents. Daussent will provide samples of the family members' DNA, which can be screened for RFLP markers, to all investigators who agree to the conditions of the experiment.

The project has been under way for about a year and a half at a cost that Dausset estimates to run about $15,000 per month. So far, he has supplied DNA from the families to each of 15 laboratories in England, France, and the United States.

The first requirement for participating in the experiment is possession of a cloned DNA probe that can detect an RFLP. As a further condition, the investigator must agree to determine the inheritance patterns of the RFLP in all of the 40 families for which this is possible.

In practice this means that the DNA of the parents is screened first. Only those families in which at least one parent is heterozygous for the RFLP, that is, the individual has different variants on both members of the chromosome pair, need

be studied further. If both parents in a family are homozygous, it cannot provide useful information about the inheritance pattern of RFLP in the progeny.

Once the inheritance pattern of an RFLP has been traced in all the appropriate families, the data must be sent to Dausset at CEPH. The computer there can compare the results with the inheritance patterns of different RFLP markers that were determined in other laboratories to determine which markers are linked and how closely. "We hope to accelerate in this way the mapping of the human genome," Dausset says. "If many laboratories use the same families, it will be much easier."

The experiment has been arranged, Botstein points out, to obviate a consideration that might otherwise have been a problem. It requires the sharing of data, but not of the probes themselves, which investigators might have been much more reluctant to do. "What Dausset has done," Botstein explains, "is to make it possible for people to work together. He has reduced the competition problem."

How long it will take to complete the map is uncertain. This depends on the length of time required to cover the human genome with markers of suitable quality. The minimum required for future studies aimed at detecting linkages with disease loci is probably in the range of 100 to 150 evenly spaced markers.

Many additional RFLP's may actually have to be mapped to obtain markers of the desired quality, which will largely be determined by how many variants they have. Although in excess of 200 probes that detect RFLP's have been reported, perhaps only 10 to 20 percent have sufficient numbers of variants to be highly informative in gene linkage studies of human genetic diseases. In this context, more is definitely better.

Markers with many variants can provide more information because the variation helps to ensure that many members of the population will be heterozygous at the marker locus. This is a requirement for linkage studies. High polymorphism is especially important for showing linkages with genetic disease loci, a situation in which the families of interest are likely to be much smaller than those available for making the human genome map.

—JEAN L. MARX

Additional Reading

1. D. Botstein, R. L. White, M. Skolnick, R. Davis, *Am. J. Hum. Genet.* **32**, 314 (1980).
2. G. M. Lathrop, J. M. Lalouel, C. Julier, J. Ott, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3443 (1984).
3. R. White *et al.*, *Nature (London)* **313**, 101 (1985).

Pattern and Process in Life's History

Higher level selection processes are becoming acknowledged as important influences that shape the history of life

There are patterns in the history of life, of that there is no doubt. The question for evolutionary biologists is, how are these patterns to be interpreted? Specifically, if the perceived patterns are true deviations from chance clusterings, what is their configuration; and what processes might be responsible for shaping them?

The past several years has witnessed a good deal of public debate on some of these matters, which has ranged from proclamations of the supposed imminent demise of neo-Darwinism to suggestions that life on Earth is periodically assaulted by extraterrestrial bolides. Although the content of the public utterances has sometimes diverged startlingly from the exchanges between scholars of the subject, the razzmatazz accurately reflected the degree of intellectual foment.

Responding to the excitment of the times, the Dahlem Konferenzen, in West Berlin, last month held a workshop* that attempted to draw together some of the disparate approaches and viewpoints relating to these issues.

In common with Dahlem conference tradition, participants at this one were divided into four groups, each set the task of examining one major approach to the overall problem. The first group looked at some of the most obvious apparent patterns, such as the increase in complexity of life forms through time, the increase in body size within taxa through time, and so on, and discussed ways of testing their significance. The second concentrated on the causes and consequences of extinction, a very lively topic just now. The task of the third group, which included a mix of population geneticists and paleontologists, was to investigate the relationship between genomic and organismic evolution. And the fourth group concentrated on the evolution of communities, both living and fossilized. The overall goal was to see how far biological and paleobiological approaches could be integrated in a move "towards a new understanding of large-scale evolutionary change."

As ambitious a project as this was certain to lead to frustrations, not least because the languages and concepts employed by population geneticists, paleontologists, and ecologists are so disparate as to militate against ready communication. Moreover, the tone of disputation between population geneticists and paleontologists during the past decade has at times been harsh. In spite of these difficulties, however, the Dahlem meeting scored certain notable advances that mark an important turning point in the further development of evolutionary biology. There still are gaps between different intellectual approaches, to be sure, but there has been clear enhancement in mutual understanding, in addition to some very striking agreements.

By far the most notable of these agreements was the acceptance by the population geneticists that certain major evolutionary trends revealed in the fossil record might be the result of selection between species, a process that is analagous to selection between individual organisms within a species. This formulation, known as species selection, casts evolution as a hierarchical process and extends, but not replaces, conventional neo-Darwinism, which traditionally has focused on natural selection within species. The notion of species selection emerged from the development of the punctuated equilibrium hypothesis, initially advanced by Niles Eldredge of the American Museum of Natural History and Stephen Jay Gould of Harvard University, in 1972. Important developments of the theory's implications have been made by Steven Stanley of Johns Hopkins University and Elisabeth Vrba of the Transvaal Museum, Pretoria.

Until last month's Dahlem meeting most population geneticists strenuously

*The Dahlem Workshop on Phanerozoic, Life: Pattern and Processes, was held in Berlin from 16–21 June, 1985. The proceedings will be published by Springer-Verlag, Berlin/Heidelberg/New York/Tokyo.