Genes *by Mail*

Participants discuss the Gene Library Project and its benefits to research on the genetic endowment of the human species.

An interview with L. Scott Cram, Larry L. Deaven, Carl E. Hildebrand, Robert K. Moyzis, and Marvin Van Dilla

SCIENCE: *Could you start by explaining exactly what the National Laboratory Gene Library Project provides to scientists studying human chromosomes, germ, and DNA?*

VAN DILLA: We always have to start with an explanation because the name "gene library" is misleading. It conjures up a vision of a storehouse stocked with many little vials, one for each of the hundred thousand or so human genes and all neatly arranged and ready for checking out. This vision will probably be realized in the future—the distant future—but for the time being researchers must be content with something less.

What we do provide is a selection from only twenty-four vials—twenty-four libraries. Each library contains DNA specific to one of the twenty-four types of human chromosomes. the two scx chromosomes and the twenty-two pairs of autosomes. The DNA is in the form of a set of fragments cut by a restriction enzyme. So the contents of a vial are a library only in the sense of being a collection—of DNA fragments—and arc in noway labeled or organized for retrieval.

DEAVEN: The fragments are not truly a "gene" library either, since each may incl udc one gene, more than one gene, a portinn of' a gene. or no gene at all.

MOYZIS: The vials we mail out to users don't contain naked DNA fragments though. There's even more to making a chromosome-specific Ii brary than culturing cells, sorting the chromosomes from the cells. and letting a restriction enzyme chop up the DNA molecules from the chromosomes. A library must contain enough copies of the set of restriction fragments that the user has a reasonable amount of material to work with. One could obtain that many copies by chopping up the DNA from lots and lots of sorted chromosomes from lots and lots of cultured cells. but that's a hard and expensive way to go. 1 nstead we put a biological amplifier to work and clone the fragments.


LIVERMORE, CA 94550
MAR
29
1985
USPS
Folk Art USA 22 Folk Art USA 22
Dr. J. L. Hsueh
Institute of Genetics
Fudan University
Shanghai
People's Republic of China
AIR MAIL

First the human DNA fragments are each spliced into a DNA molecule of' a bacterial v irus—a bacteriophage. or phage for short. Then these recombinant DNA molecules are each wrapped up in the bodily trappings of the phage, and the phages arc allowed to infect some unfortunate bacteria. The bacteria then proceed to work themselves to death making marry duplicate viruses, each containing a human DNA fragment in a viral DNA molecule. We gather up the plaques. or colonies of viruses, suspend them i n some liquid, and ship them off. So a library actually consists of' human DNA fragments encased in bacteriophages.

HILDEBRAND: The phage bodies make convenient, protective packages for the fragments. And more important. if a user wants to work one at a time with members of the set of restriction fragments i n a 1 ibrary —and that's usually the case—all he needs to do is infect some more bacteria with the phages in his library and pick off one plaque at a time.

SCIENCE: *How is such a library used?*

HILDEBRAND: It depends on what the user bas in mind. In some cases the fragments may be invest igated directly--their base sequences determined, for example—and the information pieced together to gi ve some picture of the DNA molecule as a whole. You see. DNA molecules arc so huge that details of their organ ization and functioning can. with present technology. be learned only by studying them in pieces, and preferably not pieces generated by random breaks in the molecules but pieces cut at the same points along each molecule. That's what restriction enzymes do. and i t's d ifficult to overstate **bow** absolutely essential they are to molecular genetics.

Getting back to the question. in most applications a library is used as a source of probes, A probe is a radioactively or fluorescently labeled DNA fragment that is used to detect and locate, in a sample of D A being investigated. portions of DNA with a complementary base sequence. Of course, any fragment in a library can be labeled and used as a probe. but it may not be detecting anything very exciting. The challenge is to

What Is a Chromosome?

The hereditary information in each cell of an organism is awesome in extent, complex in organization. and beautiful in the dynamics of its expression. Chromosomes are the packages housing this information. Most of' us think of chromosomes as the *X*-shaped objects viewed through a microscope in a high-school biology course, This familiar manifestation of chromosomes exists. owever, only during a certain brief' stage in the life of a cell. Since chromosomes—in particular, human chromosomes—arc the essential raw material for the Gene Library Project. a brief review of what is now known about these carriers of heredity is worthwhile.

Located in the nucleus, each chromosome includes. as a single molecule, some fraction of' the cell's complement of DNA and hence of' the cell's genes. The number of chromosomes in a normal cell depends on the species of origin of' the cell. Although all organisms of a species have the same chromosome number, that number is not unique to the species. For example, the evening bat (*Nycticeius humeralis*), the red squirrel (*Tamiasciurus hudsonicus streatori*), and the human all have the same chromosome number. The dependence of' chromosome number on species is not straightforward but exhibits a general downward trend with progression up the evolutionary ladder.

In higher organisms another variation in chromosome number occurs: all gametes (cells tbat participate in sexual reproduction, such as egg and sperm cells) have the same "haploid" number of chromosomes, and all other cells are "diploid," with exactly twice that number. Human gametes, f'or example, possess twenty-three chromosomes, and all other human cells possess forty-six.

How is the chromosome number of a species maintained constant from one cellular and organismal generation to another? In the case of humans, a ncw organism begins with the formation of an ancestral diploid cell by fusion of two haploid gametes: an egg carrying an X sex chromosome and a sperm carrying an X or a Y sex chromosome. (The new organism is a female (male) if an X-bearing (Y-bearing) sperm participated in the fusion.) Other cells are tben formed by successive waves of mitosis. a type of cell division resulting in two diploid cells, each containing duplicates of the chromosomes of both the maternal and the paternal gametes. (The maternal and paternal -versions of each chromosome are referred to jointly as a homologous pair.) The organism grows and develops as mitosis continues and certain of the cells become differentiated in function. At some stage or stages in the organism's Ii fe, special diploid cells (germ cells) undergo meiosis. a type of cell division resulting in formation of four haploid gametes. Each chromosome in these gametes is a duplicate of either the

maternal or the paternal chromosome or. more usual, is some combination of the two arrived at by exchange of hereditary material between homologous chromosomes during meiosis. The cycle begins again with fusion of one such gamete with another from an organism of the opposite sex.

A chromosome includes, in addition to a single DNA molecule. a mass of special proteins (histones) roughly equal to that of the DNA molecule (some 15 percent of the total chromosomal mass), small amounts of many other DNA-binding proteins. and RNA. The histones make the DNA molecule more flexible. by neutralizing the negative charges of its phosphate groups, and thus play an essential role in packaging the very long DNA molecules (which, for humans, average 5 centimeters in length) within the confines of the cell nucleus (which is between 5 X 1 $()^{-4}$ and 10^{-3} centimeter in diameter), Many details of this remarkable engineering feat arc as yet unknown. but several levels of packaging are involved (Fig. 1).

Tbe packaging begins with the formation of "nucleosomes," each consisting of-a length of linking DNA and a length of DNA wound around a core of histones. Next, some regular packing of the nucleosomes, also involving histones, leads to the structure known as (be 30-nanometer chromatin fiber (tbc double helical strand of DNA itself is about 2 nanometers in diameter). This fiber in turn is organized into a sequence of "looped domains," each of which is thought to be activated as a unit during gene expression. The sequence of looped domains is further condensed in some fashion to form the "interphase" chromosome, portions of which decondense and recondense as various genes arc activated and deactivated during biosynthesis by the cell.

The ultimate in DNA packaging occurs as a germ cell prepares for meiosis or as a somatic cell (any ccl] other tban a germ cell) prepares for mitosis. Then each interphase chromosome is duplicated, and the two identical "sister chromatics. " which arc joined at some point along their lengths (the "centromere"), form a longitudinally sym-

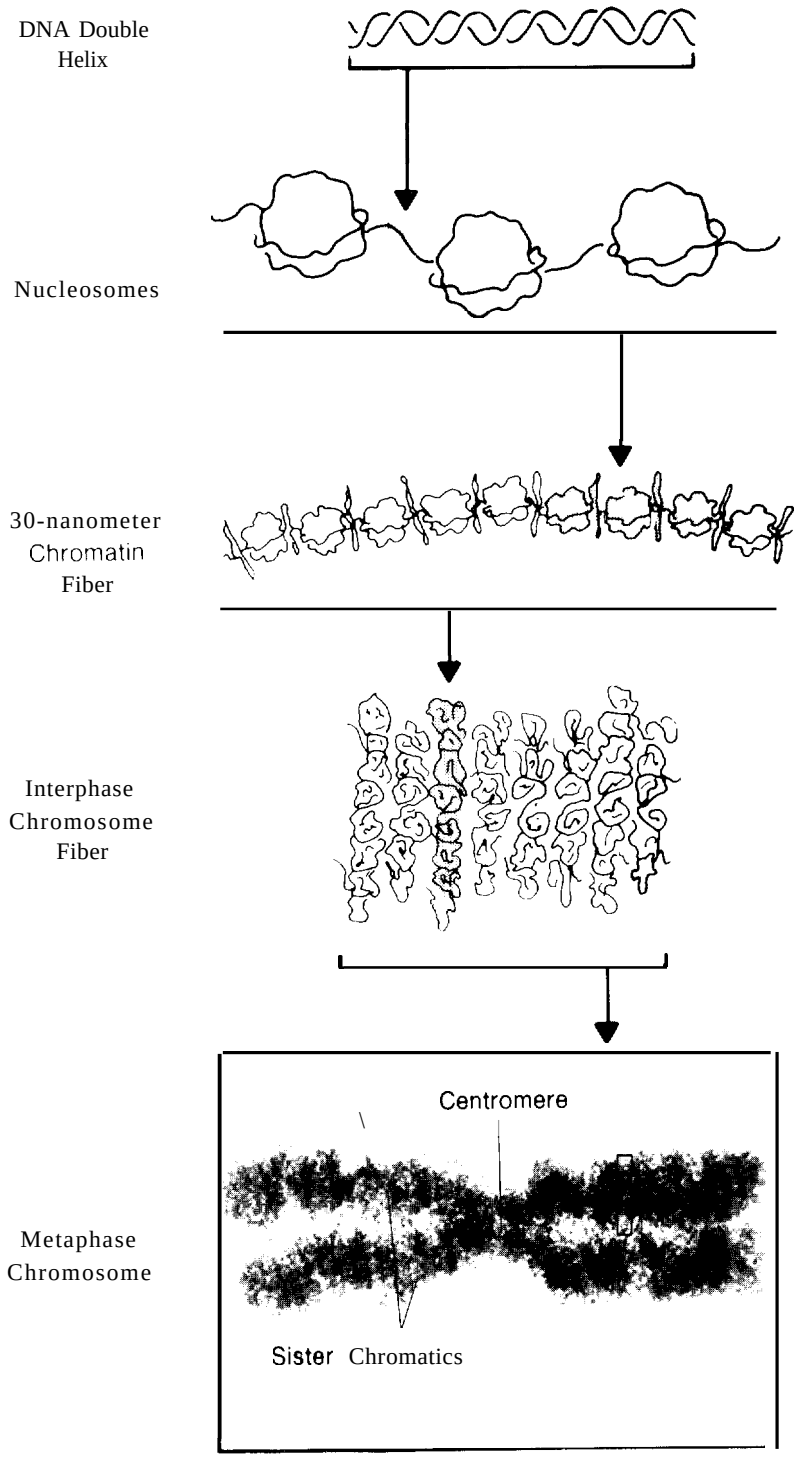


Fig. 1. Many details of the packaging of a DNA molecule and various proteins into the entity known as a chromosome are yet to be determined. These schematic illustrations are merely suggestive of the mechanisms and configurations that may be involved.

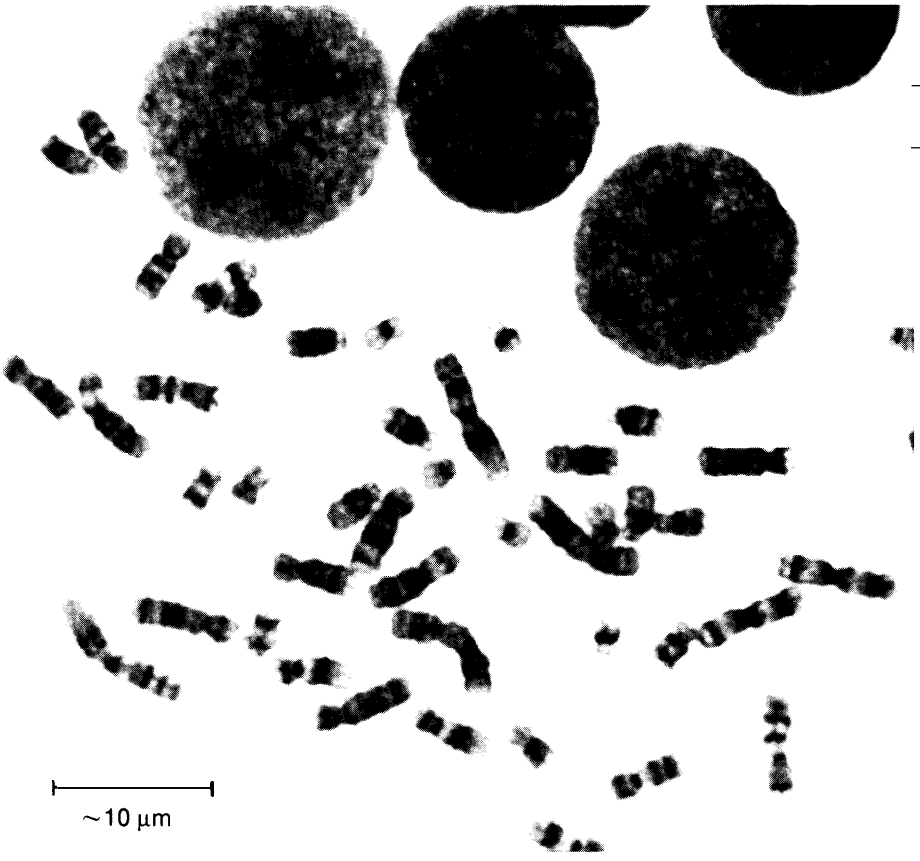


Fig. 2. Photomicrograph of metaphase chromosomes from human leukocytes (white blood ceils). The cells were cultured for 48 hours, blocked at metaphase by exposure to the drug Colcemid for 2 hours, swollen in a hypotonic solution, and treated with a fixative. The chromosomes were released from the metaphase cells simply by dropping the swollen cells onto microscope slides, which breaks the cellular membranes. (The dark, round objects are nuclei released from interphase cells.) The chromosomes were then partially digested with trypsin and treated with Giemsa stain. Since trypsin #referentially digests chromosomal proteins bound to DNA rich in guanine-cytosine pairs, those portions of a chromosome containing such DNA are less readily stained. The result is a pattern of heavily stained (dark) and lightly stained bands along the length of each chromosome. The patterns of these bands permit unambiguous identification of chromosomes.

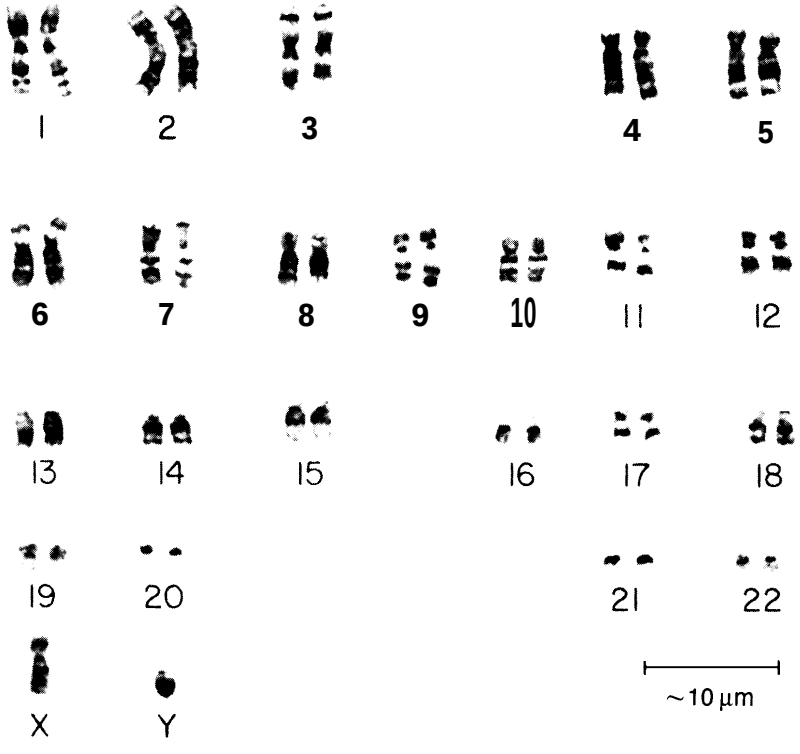


Fig. 3. A display of the metaphase chromosomes of an organism, banded as in Fig. 2 and arranged in order of decreasing size, is known as the karyotype of the organism. Shown here is the karyotype of a human male. By convention the pair of sex chromosomes is displayed separately from the homologous pairs of autosomes. Chromosomes are designated by a size-related number and, historically, by a morphology-related group: group A includes numbers 1 through 3; group B, numbers 4 and 5; group C, numbers 6 through 12; group D, numbers 13 through 15; group E, numbers 16 through 18; group F, numbers 19 and 20; group G, numbers 21 and 22. The similar morphology of the members of a group made them extremely difficult to distinguish before discovery in 1970 of the banding method described in Fig. 2.

WHAT ISA CHROMOSOME'!

metric double structure with the familiar *X* shape. These "metaphrase," or "mitotic," chromosomes are easily visible through an optical microscope (Fig. 2) and were first observed by German cell biologists in the early 1880s.

The metaphase chromosomes of a given species can be distinguished on the basis of a number of properties, such as overall length, DNA content, or location of the centromere. (Except in the case of an X-Y pair, the differences between the two members of a homologous pair of chromosomes are much less pronounced than the differences among homologous pairs.) The most definitive identifying property is the pattern of bands produced on metaphase chromosomes by selective digestion of the chromosomal proteins and subsequent staining. Figure 3 displays such banding patterns for the human chromosomes.

Which genes are packaged in which chromosomes, and exactly where, is known only for some 700 of the estimated 100,000 human genes. Examples of those for which this information is known include the gene for Huntingtons's chorea (chromosome 4), for interferon (chromosome 9), for β-globin (chromosome 11), and for hemophilia and red-green colorblindness (the X chromosome).

The marvelous choreographies of mitosis and meiosis normally transmit intact and unchanged the time-tested chromosomes of an organism from one cellular and organismal generation to another. But sometimes errors. such as rearrangements of chromosomes or departures from the normal number, do occur. The consequences of errors in meiosis are usually grave, if not fatal, to the organism that inherits them. For example, an abnormal number of chromosomes are found in the cells of about half of the human embryos that are spontaneously aborted, and the presence of three copies of human chromosome 21 results in Down's syndrome. Errors in mitosis also have serious consequences, as evidenced by the association of chromosome rearrangements with various types of cancer. ■

find the interesting probes, ones that reveal something unusual about the DNA, such as an alteration related to a genetic disease or a mutation.

SCIENCE: *Have libraries been produced before?*

CRAM: Whole-genome libraries, ones containing DNA fragments from all the human chromosomes, are already available, and so are a few chromosome-specific libraries, ones for the chromosomes that are easier to sort. But this project will be the very first to provide chromosome-specific libraries for all twenty-four types of human chromosomes.

VAN DILLA: With a chromosome-specific library a user can begin his research by fishing around among a pool of DNA fragments from a particular chromosome rather than the whole sea of human DNA. This is a tremendous advantage, and the libraries should increase the productivity of research in molecular genetics just as separated isotopes increase the productivity of research in nuclear physics,

DEAVEN: I have some data to support that expectation, at least as far as one application of the libraries—gene mapping—is concerned. By gene mapping is meant localizing a gene on a particular region of a particular chromosome. In 1958. when molecular genetics was still in its infancy. only about four hundred human genes had been identified. These had been classified as dominant or recessive and autosomal or X-linked but. except for the X-linked genes, could not be mapped. That had to await, first of all, the discovery some ten years later of chromosome banding, a technique for unambiguously distinguishing all twenty-four types of human chromosomes. This discovery made it possible to map a limited number of genes by associating chromosome abnormalities and genetic diseases. At about the same time human-rodent hybrid cell lines became available, and information about the proteins synthesized by various of these cell lines, each carrying only a few human chromosomes, made it possible to map more genes. Still, by 1973 only 64 genes had been assigned to particular autosomcs and 15510 the X chromosome. By 198 I the total number of mapped genes had increased to only about four hundred and fifty, but in that year molecular genetics provided a powerful new tool for mapping. It was discovered that what are known as restriction fragment length polymorphisms could pinpoint the locations of genes about which practically nothing is known except the observable evidence of their expression, such as the symptoms of diseases.

A restriction enzyme cuts a DNA molecule into fragments of various lengths by breaking bonds at every occurrence of a cleavage site, a certain sequence of bases. Mutations can generate new cleavage sites and destroy existing ones, and these changes cause differences in the sets of restriction fragments from different individuals. For example, in the fragments produced by the restriction enzyme *Hpa*I from the DNA of American Blacks, the β-globin gene usually appears on a 7.6-kilobase fragment but sometimes on a 7.0- or 13.0-kilobase fragment. Differences like these arc called restriction fragment length polymorphisms, or RFLPs. An RFLP can be detected by using electrophoresis to separate, according to length, the restriction fragments from, say, two individuals and seeing whether a probe that includes the RFLP sticks to different places in the two lineups.

An RFLP within a gene is obviously inherited along with the gene, but so also is an RFLP so close to a gene that it is not separated from it during meiotic recombination. So the location of such an RFLP on a chromosome is a very close approximation to that of the gene. This method of mapping, say- for a gene that causes some disease, involves testing many probes to find one that reveals an RFLP unique to people with and without the disease—that's the hard part-and then mapping the probe on the chromosome—that's the easy part.

Since 1983 about three hundred RFLP probes for genes have been identified and mapped per year. Even that rate, though, is painfully slow considering how many more genes remain to be mapped. But our chromo-

some-specific libraries will make the search for probes less difficult, and we estimate that about two thousand RFLP probes for genes will be identified and mapped in the first year the libraries are available, and even more in succeeding years.

HILDEBRAND: Once a probe for, say, a defective gene has been identified, it has applications in addition to gene mapping. It can be used to detect the gene in unborn children and in unknowing carriers of the gene—unknowing because the gene may be recessive, not clearly expressed. or, like the gene for Huntington's chorea, expressed only late in life. People from families with histories of genetic disease usually want such information to help them make decisions about the course of pregnancies or of their own lives,

RFLP probes for genes are also very valuable to basic research in molecular genetics. For instance, by studying the inheritance patterns of probes for various genes, we can learn about the relative locations of the genes along the DNA molecules,

SCIENCE: *What is the history of the project?*

DEAVEN: About three years ago a group of us in the Life Sciences Division at Los Alamos decided to look into the possibility of establishing a new program oriented around the structure and functioning of mammalian chromosomes. We envisioned a small group of related projects to be supported by some federal agency, Among the ideas we came up with, that of producing DNA libraries for each of the human chromosome types seemed particularly worthwhile and, furthermore, related to two established strengths of the Laboratory -GenBank, which is a repository for DNA sequence data, and the National Flow Cytometry Resource, which organizes our capability in the technology required for sorting chromosomes.

The DOE's reaction to our idea was quite favorable. They thought that the libraries would provide the means for new research capabilities at the national laboratories and elsewhere and that production of the libraries would be a unique contribution to the biomedical community. Of course, onc ultimate payoff of the project will probably be a better method of detecting mutations in humans caused by exposure to radiation or other mutagens, and this is a basic interest of the DOE's Office of Health and Environmental Research. In any case, the DOE decided to fund the project as a service to the biological and medical communities. In discussions with personnel at the DOE, the suggestion was made that we initiate the project jointly with a group of biomedical scientists at Livermore, which also was strong in flow cytometry,

People at Livermore were enthusiastic, and together we began by meeting, in October and November of 1983, with a group of molecular biologists and human geneticists who advised us about what was needed and the relative priorities to assign to those needs. This advisory group suggested tbat wc construct, as quickly as possible, a set of libraries to be used as sources of potential probes for gene mapping and diagnosis of genetic diseases. For those applications small DNA fragments are best, since a small fragment that includes a sequence coding for a gene is not likely to include in addition one of the many noncoding repetitive sequences present in DNA, So that's what we're doing now—making libraries of relatively small fragments. When these Phase I libraries are completed, we'll go on to constructing the Phase II libraries, with larger human DNA inserts, to be used for basic research in gene structure, arrangement, and expression.

This is a good opportunity for me to acknowledge the foresight of our technical representatives at the DOE. Usually, the most important products of research are

L. Scott Cram of Los Alamos, a biophysicist, is a specialist in the application of flow cytometry to identification and sorting of chromosomes. His research interests include the role of karyotype instability in tumorigenesis.

Larry L. Deaven, a cytogeneticist, is principal investigator for the Gene Library Project at Los Alamos. His research interests include the chromosome constitution of evolving cell populations and the chromosome damage induced by agents associated with energy production.

Carl E. (Ed) Hildebrand of Los Alamos, a biophysicist turned molecular biologist, is a specialist in recombinant DNA technology. His research interests include DNA replication, chromatin structure, and gene expression.

publications, and to some extent research is funded on the basis of the expected number and quality of' the publications. But for this project the DOE set that criterion aside in favor of a different goal—the production of libraries, Although we expect important publications from our work on the project, our major aim is 10 produce high-quality libraries for use by interested scientists throughout the world.

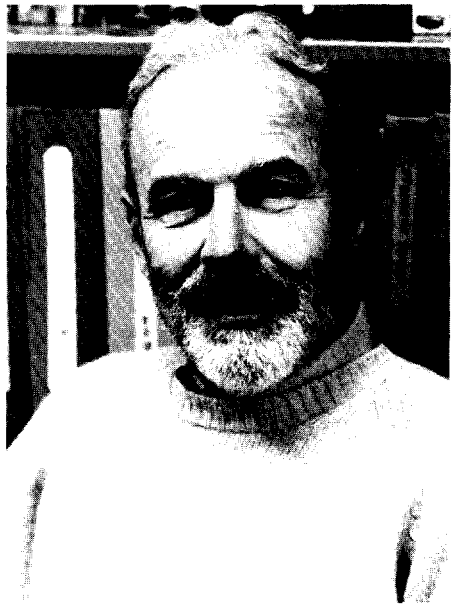
In addition, the Gene Library Project wouldn't be as far along as it is if' it hadn't been funded by the DOE. which can act very quickly when something regarded as worth supporting comes along, We have now completed approximately 75 percent of-the Phase 1 libraries. If our technical representatives at the DOE had not understood the significance and timeliness of the project. we might well still be in the organizational stages.

SCIENCE: ***How*** *does it happen that two weapons laboratories have the personnel and facilities required, for such a project?*

DEAVEN: That's not so much of an anomaly as you might think. Tbe AEC had a natural interest in the biological effects of radiation and realized. as have its successors, that understanding those effects in detail had to be based on knowledge of cellular processes at a molecular level. So in addition to programmatic research on the biological effects of radiation—and now of' various chemicals associated with other means of producing energy—basic research in cellular biology has received strong support at the national laboratories since their beginnings.

SCIENCE: *W 'hat kind of collaboration between the two laboratories is involved, and how well is it working?*

Robert K. Moyzis, a molecular biologist, is Leader of the Genetics Group at Los Alamos. His research interests include the organization of human DNA and the regulation ofgene expression.



Marvin Van Dilla, a biophysicist, is principal investigator for the Gene Library Project at Livermore. His research interests include the development of flow instrumentation and its application to biological problems.

VANDILLA: We collaborate in the sense of' drawing on each other's expertise rather than in the sense of parceling out various aspects of the project to one or the other laboratory. Both of us have gone through the whole process of' making Phase I libraries. but our products are not identical—by edict rather than by chance. The Livermore Ii braries contain fragments of DNA cut by onc restriction enzyme, and the Los Alamos I ibraries contain fragments cut by a different restriction enzyme. So a fragment missing from onc version, because the recombinant DNA molecule containing that fragment is too large to fit into the bacteriophage, will quite likely bc found in the other.

In my judgment the collaboration has been very effective, with each laboratory contributing particular strengths. For example, we at Livermore had greater experience in

sorting human chromosomes with flow cytometers and had been developing one capable of doing the job at a rate compatible with the scope of the project, We need something on the order of a million sorted chromosomes for each library. (-emmercially available flow cytometers can sort that many in a few tens of hours—and that's operating time. so the job actual l>' takes much longer—but the high-speed sorter reduces the sorting time to just a few hours. We made copies of the basic components of our sorter and sent them here. I understand the Los Alamos machine is now up and running.

As for Los Alamos. the> brought to the project greater experience and expertise in recombinant DNA technology and in fact had made a few chromosome-specific libraries for the Chinese hamster before the project even started. The Los Alamos group had also explored the use of hybrid cells as sources of human chromosomes.

Our first joint meeting, at Los Alamos in December of '84. was proof to me of a successful collaboration. Almost everyone involved in the project at Livermore, from Ph.D.-s to technicians. attended the meeting and took the opportunity to discuss common problems and exchange tricks of the trade with their counterparts at Los Alamos. Our chief cell farmer, for instance, had never been to Los Alamos to talk to the chief cell farmer here. They both do the same thing with somewhat different techniques and have developed somewhat different ways of getting around the many nitty-gritty problems that crop up in the busines of cell culture. Together they did two experiments during the meeting. Onc day they grew cells and isolated the chromosomes using the technique that's favored at Livermore, and on the next day they went through the same steps using the technique favored at Los Alamos. Each learned something from the other and seemed very happy about the whole experience.

SCIENCE: *How much interest has been shown in the libraries?*

Making the Libraries

Supplying the Chromosomes

As raw material for each chromosome-specific library, we must collect between half a million and a few million of a particular chromosome by sorting through tens to hundreds of millions with a flow cytometer. Our original hope was to obtain such large numbers of chromosomes for sorting by culturing, or multiplying *in vitro*, human cells. namely, fibroblasts from human foreskin tissue. But. as discussed below, these cells proved to bean acceptable source for only the smaller chromosomes, and another source had to be found for the larger.

We culture fibroblasts in three stages. beginning by placing small pieces of' the tissue sample, together with nutrients essential for growth, in a small plastic container. (Despite the translation of *in vitro*, plastic has largely supplanted glass as a surface for cell culture.) The fibroblasts migrate from the tissue, adhere to the bottom of the container, and multiply by mitosis. (This process is similar to that by which fibroblasts *in vivo* repair a break in the skin.) Mitosis stops when the bottom of the container is covered with a monolayer of cells (Fig. I). The monolayer is then partially digested with the enzyme trypsin to yield a suspension of single cells. This suspension is divided among several containers. fresh growth medium is added. and the culturing is repeated. again until a monolayer has formed.

Our aim is to maximize the number of fibroblasts at metaphase (on the verge of dividing), since only those cells contain the metaphase chromosomes that can be sorted. Obviously, the mom rapidly a population of cells is multiplying, the larger is the fraction of that population at metaphase. But the fibroblasts begin to multiply less rapidly after about fifteen to twenty mitoses and cease multiplying altogether after about fifty. Therefore, after the second culture some of the fibroblasts are suspended in a medium containing glycerol (or some other ice-crystal inhibitor) and frozen in 1 iquid nitrogen. These cells, which can be thawed and recultured at any time up to several years later, serve as a reservoir of fibroblasts at the height

I. SUPPLYING THE CHROMOSOMES

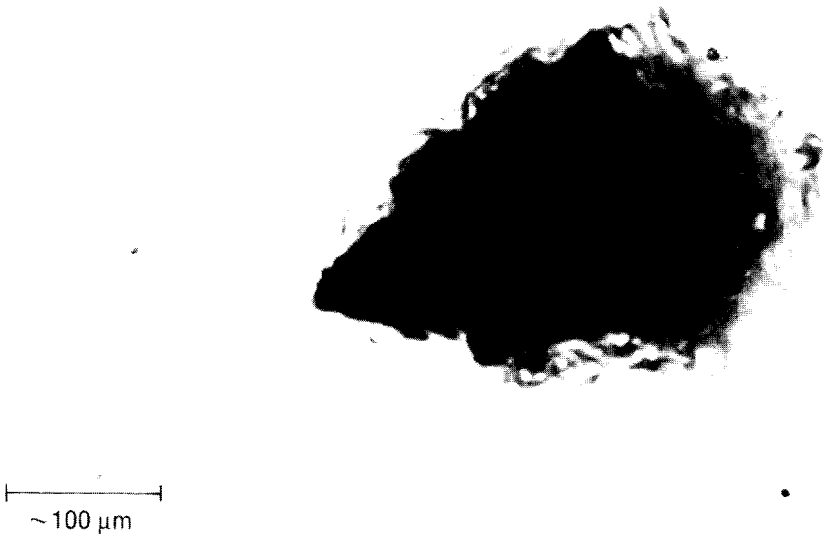


Fig. 1. Cells in culture multiply by mitosis and spread over the surface of the culture dish until a confluent monolayer of cells is formed, Shown here is an optical micrograph of a confluent monolayer of fibroblasts cultured from apiece of human skin tissue (dark area).

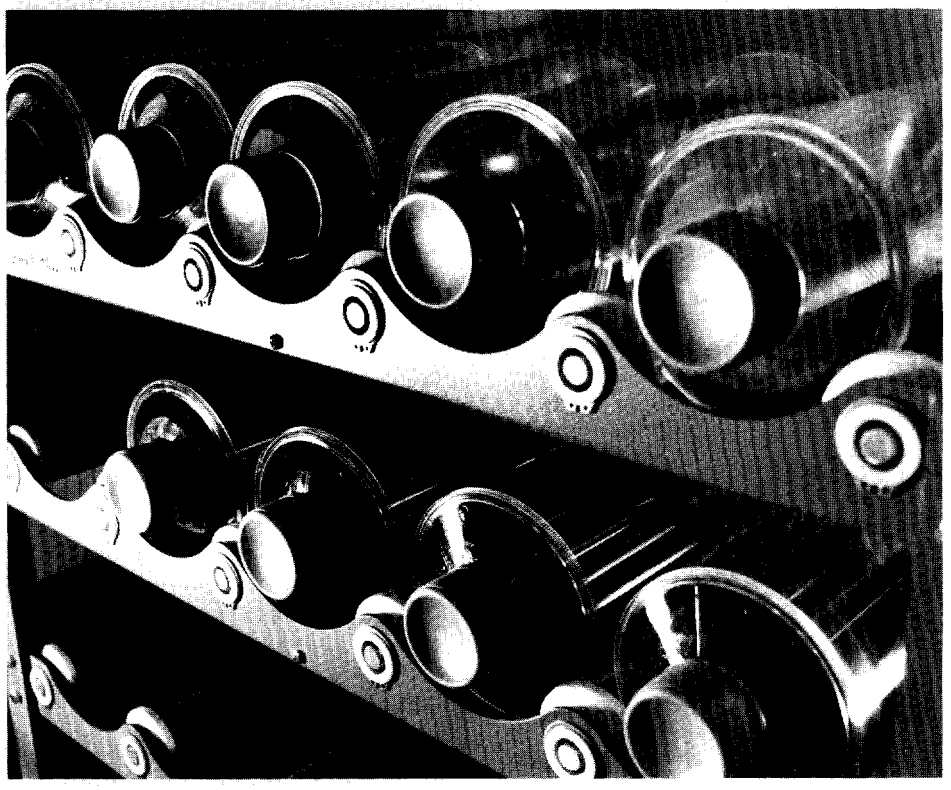


Fig. 2. Cells can multiply manyfold in large, slowly rotating cylinders ("roller bottles") since a large surface area is kept in contact with the culture medium.

of their multiplication rate.

The final culture of the remaining fibroblasts takes place in large "roller bottles" (Fig. 2). When a monolayer has almost formed, the drug Colcemid is added to arrest the cell cycle at metaphase. After exposure to this drug for an appropriate time. the fraction of cells at metaphase increases from the usual maximum of 0.5 percent to between 30 and 50 percent. Since the metaphase fibroblasts "round up" from the culture surface because of their more spherical shape, they are easily separated by gentle shaking of the roller bottles.

The next step is to isolate the chromosomes from the cells. first, the cells are swollen in an "isolation buffer" with an osmotic pressure lower than that of physiological saline. The buffer is then forced through a fine-gauge hypodermic needle to break the cell membranes (the nuclear membranes are already dissipated in metaphase cells) and release the chromosomes. The buffer solution also contains materials. such as proteins that coat chromosomes or dyes that intercalate with DNA, to stabilize the morphology of the chromosomes. Although this isolation step sounds simple in principle. difficulties that plague even the best methods are responsible for our yield of only 2 10 6 sorted chromosomes from every 100 cells.

The final step in preparing the chromosomes for sorting is addition to the buffer of two fluorescent dyes, One of these binds preferentially to DNA sequences rich in adenine-thymine base pairs and the other to sequences rich in guanine-cytosine base pairs. The fluorescence intensities excited in the dyes serve as the basis for flow-cytometric analysis and sorting of the chromosomes,

Analysis of chromosome preparations from human fibroblasts indicated that all the chromosomes except numbers 9 through 12 were sufficiently well resolved for sorting (see Fig. 2 in "Sorting the Chromosomes"). In practice, however, sorts for the larger chromosomes were very time-consuming, apparently because the shearing action that breaks the cell membranes also breaks up many of the larger chromosomes and greatly reduces their numbers. More important. the purities of these sorts were reduced to less than the acceptable level of 90 percent by many doublets of smaller chromosomes, Evidently the shearing action is not sufficient to break up these doublets. which form during isolation and cannot be distinguished by the flow cytometer from the larger chromosomes. Furthermore, reducing the shearing force to avoid breaking the larger chromosomes only increased the number of doublets.

As a result, we have had to turn to human-hamster hybrid cells as a source of the larger chromosomes. Such cells result from fusion of human and hamster cells and initially contain a complete set of both human and hamster chromosomes. However, after several weeks to several months in culture, the hybrid cells spontaneously and randomly lose some of their complement of human chromosomes. These losses make it possible to clone hybrid cells carrying selected human chromosomes. For our purpose an ideal hybrid cell is one with no more than three human chromosomes, each separable in a flow cytometer from each other and from the hamster chromosomes.

Since hybrid cells with a small number of human chromosomes are useful in other areas of research, various cell lines have been established and frozen at a number of Laboratories. We are indeed grateful to those who have shared their cell lines with us for the Gene Library Project.

Culture of the hybrid cells is similar to, and no more difficult than, that of human fibroblasts. However, as sources of chromosomes for the libraries, hybrid cells are far from perfect, Their habit of unexpectedly losing human chromosomes, for instance, requires frequent monitoring for the continued presence of the desired chromosomes. Even more troubling is the fairly common occurrence of DNA exchange between the human and the hamster chromosomes. Such exchanges are difficult to detect but, undetected, contaminate the libraries with hamster genes. ■

DEAVEN: Well, we recently filled requests for over two hundred and fifty Phase I libraries from users in the United States and seven other countries. We expect many more requests as word of their availability gets around.

MOYZIS: I would be very surprised if'ever] research group in the world working on human genes wouldn't want at least one of the libraries, expecially after the Phase II libraries arc available.

SCIENCE: *Do you exercise any control over the uses to which the libraries are put?*

DEAVEN: No. The request form for a library includes a question about the nature of the proposed research. but we ask that primarily because we want to keep a record of the various applications. (Users funded by government agencies have already agreed to the guidelines established by the National Institutes of Health for recombinant DNA research. and industry voluntarily abides by those guidelines.

Incidentally the plan for the future is that the NIH will establish a repository for the libraries we produce and will handle the distribution. They will also collect information determined by the users about the libraries, such as purity and completeness data and characteristics of probes isolated from the libraries, This in information will serve as feedback to us for improving future libraries and will of course be valuable to other users. NIH also plans to establish a repository for probes pulled from our libraries and others.

MOYZIS: One thing we do request of users is that they don't pass the libraries on to other investigators. We want the libraries to originate directly from us, or from the repository when that comes about. One reason for this is so that wc can keep in touch with all the users. A more important reason for not wanting the libraries passed around is to preserve their characteristics. Each amplification of a library unavoidably introduces changes. For example, since phages carrying human DNA fragment A do not multiply at exactly the same rate as those carrying fragment B. after several amplifications thc rela-

tive numbers of fragments A and B will be quite different, Some fragments may disappear altogether.

Another kind of change that occurs when a library is amplified is rearrangement of the fragments. One can start, say, with a fragment containing genes A, B, and C in that order and, because cloning is something more than stamping out pennies. end up with the genes in the order C, A, and B, which is misrepresentative of the chromosome. It was to minimize this kind of change that we chose the bacteriophage Charon 21A as the cloning vector for the Phase 1 libraries, since in this phage the tendency for the human DNA insert to get rearranged is minimal. When we move on to the Phase 11 libraries and start using other cloning vectors, we'll have to worry more about this problem.

SCIENCE: *What are some applications of the libraries?*

VANDILLA: We've already discussed gene mapping and mentioned diagnosis of'genetic diseases, detection of mutations. and basic research in molecular genetics. Some other interesting applications are comparisons of DNA sequence organization among the various human chromosomes, pedigree analysis, and comparison of human DNA with those of other species.

MOYZIS: Before we go on to other applications, I'd like to point out that gene mapping is more than just collecting dry facts—it provides basic clues about possible mechanisms for controlling gene expression. For example, some thousands of genes are expressed only in nerve cells. How can that be? If we find that these genes are scattered about all over the genome, we may conclude that some structural feature near each gene regulates its expression. But if we should find that all are located in the same region of just a few chromosomes, then a very different, global regulation mechanism may be at work,

There's also a wealth of data relating various pathological conditions. including cancer, to chromosomal abnormalities such as translocations and deletions. Why is this

Making the Libraries

Sorting the Chromosomes

Before the advent of flow cytometry, human metaphase chromosomes could be coarsely partitioned into at most ten fractions by centrifugation or sedimentation techniques. Now. a state-of~ the-art flow cytometer can separate these chromosomes. with an accuracy of up 95 percent and at a rate of up to 400 per second. into twenty-four fractions, one for each of the two sex chromosomes and the twenty-two homologous pairs of autosomes. Originally developed at Los Alamos in the late 1960s to measure the DNA content of cells,* the flow cytometer is absolutely essential to our goal of producing very pure and highly representative chromosome-specific libraries of DNA fragments.

The necessary accuracy of sorting, and hence purity of the libraries, is achieved with a dual-beam flow cytometer (Fig. 1). This instrument "recognizes" a particular human chromosome on the basis of a property of the DNA it contains, namely, the ratio of adenine-thymine to guanine-cytosine base pairs. This ratio, which varies from one chromosome to another, is translated into measurable fluorescence intensities by staining the chromosomes with two fluorescent dyes. One (Hoechst 33258) binds with high specificity to sequences rich in A-T base pairs, and the other (chromomycin A_3) to sequences rich in G-C' base pairs. As the stained chromosomes intercept the two laser beams of the cytometer, the fluorescence intensities excited in both dyes are measured and used to trigger sorting if they fail within a certain range of the intensities characteristic of the desired chromosome. The data collected during this "bivariate" analysis of a chromosome preparation are monitored in real time and displayed as a "flow histogram" (Fig. 2). Such a histogram is indicative

**See "Flow Cytometry: A New Tool for Quantitative Cell Biology '' Los Alamos Science, Volume 1, Number 1, Summer 1980.*

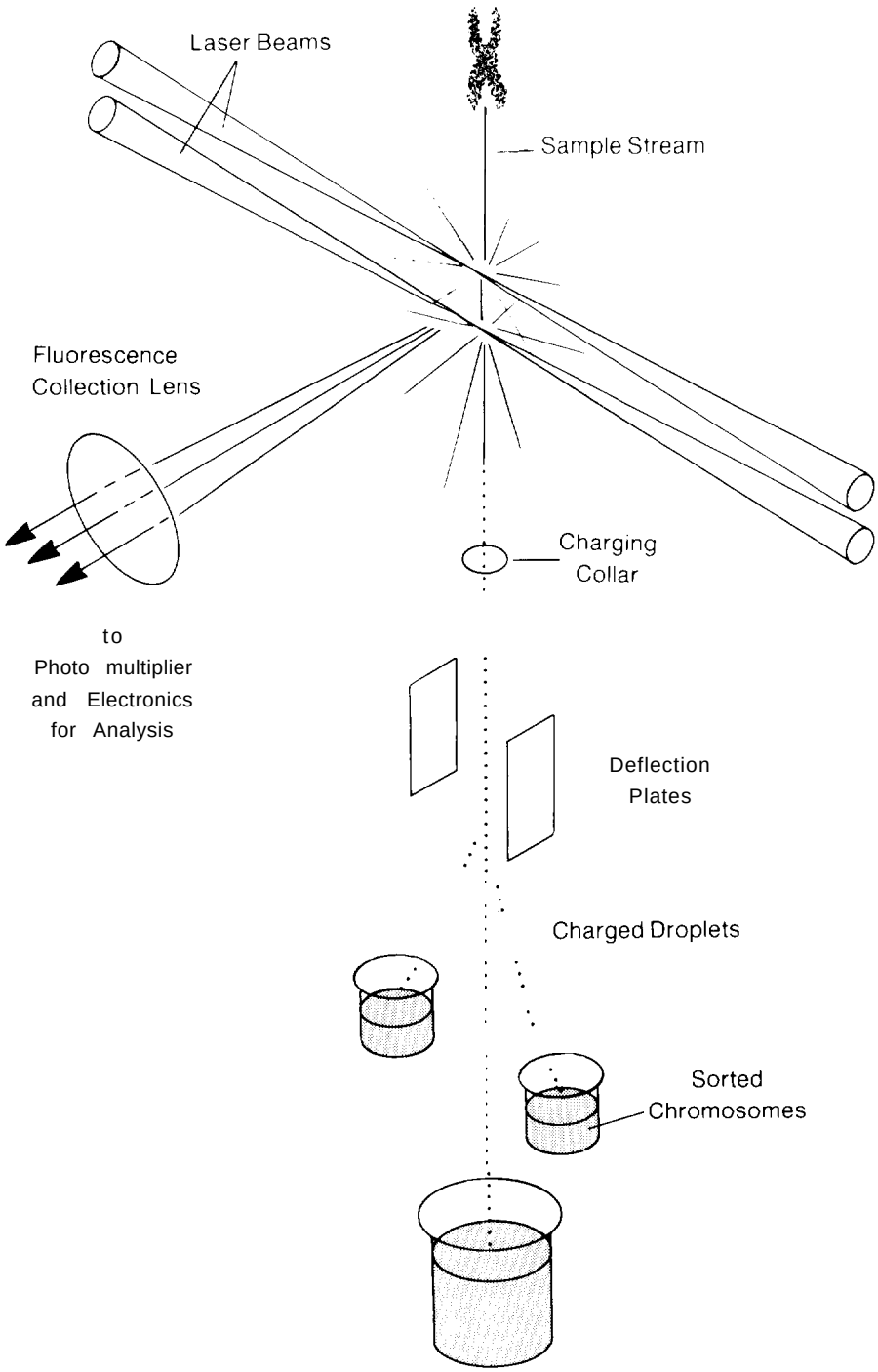


Fig. 1. Operating principles of a flow cytometer for sorting human chromosomes on the basis of the fluorescence intensities of two dyes, each of which binds preferentially to a different type of DNA sequence. A very narrow stream of the chromosome preparation flows rapidly through the focused regions of two lasers, each tuned to the excitation wavelength of one of the dyes, and is then broken into a rapid succession of tiny droplets. If the measured fluorescence intensities fall within a prescribed range of those characteristic of the desired chromosome, an electric charge is imparted to a group of droplets certain to include the chromosome. These droplets are then deflected from the main droplet stream into a collection receptacle. Two different chromosomes can be sorted from the same preparation since either a negative or a positive charge can be imparted to the droplets.

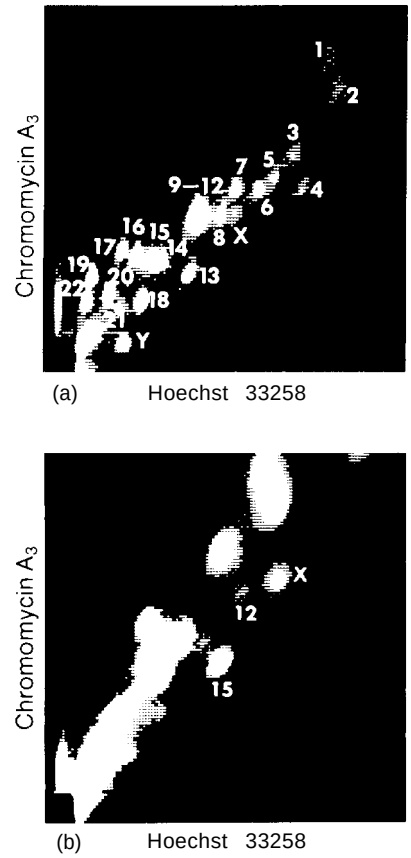


Fig. 2. (a) A flow histogram obtained from analysis of a typical chromosome preparation from human fibroblasts. The histogram is a record of the number of fluorescence events (each presumably due to a single chromosome) versus the fluorescence intensities of the dyes chromomycin A_3 and Hoechst 33258. Peaks in the histogram are attributed to different chromosomes as shown. This histogram indicates that chromosomes 9 through 12 cannot be sorted with the necessary accuracy from such a mixture of all the human chromosomes. In practice, neither can chromosomes 1 through 8, X, 14, and 15 (see "Supplying the Chromosomes"). For these chromosomes the necessary accuracy is achieved by sorting from the chromosomes in human-hamster hybrid cells rather than those in human fibroblasts. (b) A flow histogram of a chromosome preparation from human-hamster hybrid cells carrying human chromosomes X, 12, and 15. Note the excellent resolution of the human chromosomes. Hybrid cells carrying various human chromosomes are used as sources for those chromosomes not sortable at an acceptable purity from fibroblast preparations. Note that the hamster chromosomes lie on a straight line in the histogram; this is evidence of a nearly constant ratio of adenine-thymine base pairs to guanine-cytosine base pairs in the DNA molecules of these chromosomes.

II. SORTING

of the accuracy with which the chromosomes in the preparation can be distinguished.

One problem we faced in sorting the chromosomes was finding a suitable stabilizer, a material added to the isolation buffer to stabilize the morphology of'the metaphase chromosomes. Some of these materials appeared to decrease the specificity of the dyes (thus decreasing the sorting accuracy) and to degrade the chromosomal DNA (thus producing fragments that cannot be incorporated in recombinant DNA molecules), These undesirable side effects have now been minimized by using the polyamides spermine and spermidine as stabilizers.

The speed of sorting is crucial to producing highly representative libraries, which requires cloning DNA fragments from very large numbers of sorted chromosomes. For example, about two million were needed for each library produced during the first phase of the project. (These Phase I libraries contain relatively small DNA fragments.) With the commercial flow cytometer then available, sorting that many chromosomes took about twenty hours, a not unreasonable time. But ten times more sorted chromosomes will be needed for each of the Phase 11 libraries (which will contain relatively large DNA fragments), and two hundred hours is indeed an unreasonable time. With this in mind. our collaborators at Livermore developed the high-speed sorter now in use at both laboratories. Its approximately tenfold greater sorting speed is due to more rapid formation of droplets, a modification that permits sorting from more concentrated chromosome preparations. We are also investigating the feasibility of simultaneously sorting four, rather than two, different chromosomes by charging the selected droplets to one of two levels of either polarity.

The successful application of flow cytometry to sorting chromosomes is evidence of the many advances made in its practice. But the technology still has its drawbacks. It is expensive and, more important. very labor-intensive. Computerized monitoring and control may help reduce the human labor involved. ■

so? A good first question to answer is what genes are located near those abnormalities.

HILDEBRAND: The chromosome translocation associated with Burkitt's lymphoma, a cancer of the immune system, is known to put a cancer-causing gene near a regulatory sequence for an immunoglobulin gene. But we don't know for sure whether this observation has general significance, For example, are there many such rearrangements that lead to pathological states?

DEAVEN: The same argument holds for mapping close to the translocations known to be involved in many hereditary diseases. If we had documentation of what genes were located on the chromosomes that participate in a translocation, then we would have insight into the mechanisms at work in expression of the disease.

MOYZIS: It wasn't too many years ago that the immense amount of DNA in a human cell was regarded as having been just thrown into the nucleus like a mess of spaghetti. Now more and more evidence is accumulating that this mess has a very defined order and that any perturbation of that order results in some alteration in the individual. It's now known, for instance, that approximately one-half of all pregnancies end in spontaneous abortions at a very early stage of embryonic development, and the chromosomes of about one-half of all spontaneously aborted embryos are badly scrambled. So the genetic diseases that we know about are only a very small fraction of what can possibly go wrong. Most alterations in embryonic DNA lead to embryonic death.

DEAVEN: Between 2 and 3 percent of all newborns are afflicted with one of the more than three thousand known genetic diseases. And these numbers don't include some relatively common adult disorders. such as Alzheimer's disease, that have a genetic component. So genetic disease is a major biomedical and social problem. I think that the libraries will accelerate the identification of probes for rapid, inexpensive prenatal diagnosis and genetic counseling. Biotech companies should find this a profitable field-finding and marketing batteries of such probes.

The ultimate goal, of course, is not simply to predict the occurrence of a genetic disease but to uncover the biochemical defect involved. It's amazing how little we understand about something as common as Down's syndrome. We know that it's associated with three copies of chromosome 21, but why an extra copy of that chromosome should have such devastating effects is not known. Down's syndrome may be an extremely difficult disease to tackle, though, since a whole chromosome is involved. Diseases caused by defects in single genes should be a lot easier.

HILDEBRAND: Once a disease has been traced to a defect in a gene, copies of the normal gene can be produced by recombinant DNA techniques, and this opens the possibility of' gene therapy, of treating the disease by introducing the normal gene into the patient. Gene therapy is still very much in the experimental stage, and it's a sensitive issue because of the ethical and legal implications of intervention with the human genome.

MOYZIS: Yes, but probably no one is going to complain about attempts to apply gene therapy to really terrible diseases—say, Lesch-Nyhan syndrome with its symptoms of severe mental retardation, kidney damage, cerebral palsy and self-destructive behavior. What is worrisome to many is the specter of eugenics, of using such techniques to alter mental or physical characteristics. Many people feel it's one thing to genetically engineer plants or animals for desirable traits, and quite another to genetically engineer humans. In the first place. who's to say what are desirable traits? I strongly believe, however, that the types of genetic tampering that arc the most feared are exactly the ones that arc the least likely to be performed in the immediate future, mainly because we don't know how. It's extremely unlikely that anyone alive today will ever see genetic engineering applied to so complex a trait as intelligence. We haven't the foggiest notion

Making the Libraries

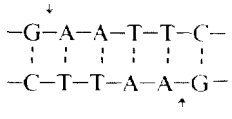
Cloning the DNA

The cells have been cultured, the chromosomes isolated and sorted. Now can begin the final act in the making of a library—that of multiplying the DNA in the sorted chromosomes by a factor of more than a million. Some tricks of recombinant DNA technology make such great amplification possible, and the magicians who perform these tricks are lowly but accommodating bacteria and viruses.

Briefly, four steps are involved in cloning the DNA: preparation of DNA fragments from the sorted chromosomes and from bacterial viruses (bacteriophages, or simply phages); formation of recombinant DNA molecules including both chromosomal and viral DNA fragments: assembly of phages containing the recombinant DNA molecules: and cloning of the phages (and hence replication of the chromosomal DNA fragments) in a bacterial host. For concreteness we describe, and illustrate in the accompanying figure, the making of what we call a Phase I library, one based on relatively small chromosomal DNA fragments. (Our Phase 11 libraries, which will serve a different purpose, will be based on larger fragments.)

First the DNA molecules in the sorted chromosomes must be extracted from the other chromosomal components. For reasons that will become clear later, this extraction must be performed very gently to minimize breaking the very long and very fragile DNA molecules.

The extracted DNA is then digested with the restriction enzyme *EcoRI*, which recognizes the sequence of six bases –G–A–A–T–T–C– and, as shown below, cuts the DNA at each occurrence of' this sequence by breaking one covalent bond on each strand:

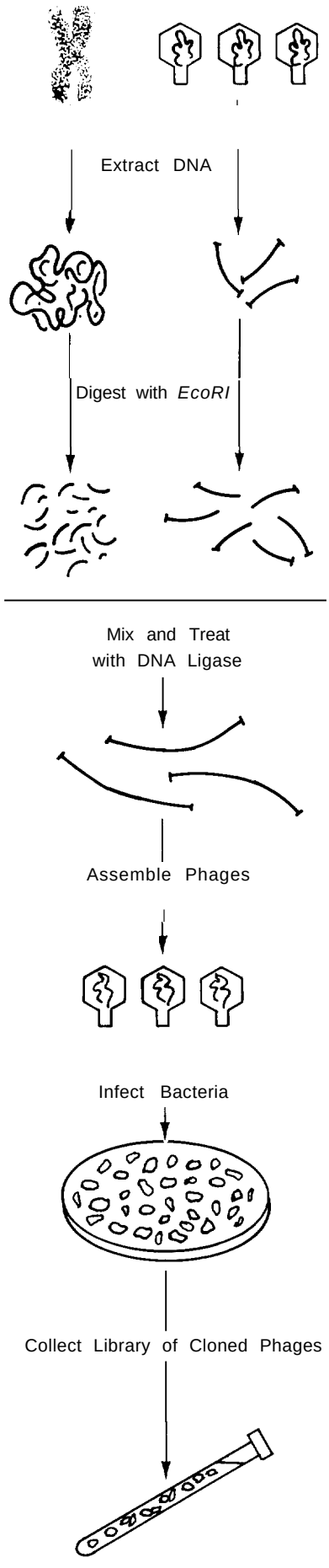


EcoRI is produced by certain strains of the bacterium *Escherichia coli* as a defense against infection by the DNA of bacteriophages. The enzyme severs the viral DNA and thus prevents its propagation in the bacterium. (To avoid self-destruction, the bacterium also produces another enzyme that modifies its own DNA so that it cannot be severed by *EcoRI.)* other bacterial species produce restriction enzymes that recognize other sequences. For example, *HindIII,* the restriction enzyme used by our collaborators at Livermore for their phase I libraries, is produced by the bacterium *Hemophilus influenzae.* Like *EcoRI, HindIII* recognizes a six-base sequence (–A–A–G-T–T--T–) and makes staggered cuts.

Complete digestion of the chromosomal DNA with *EcoRI* (or *HindIII)* produces a set of relatively small flagments. We can estimate the average length of these fragments by assuming a random distribution along the chromosomal DNA of equal numbers of the four bases A, T, G, and C. Then the probability of any six successive bases being the *EcoRI* (or *HindIII)* recognition sequence is $(1/4)^6$, which implies an average of about 4100 base pairs between cleavage sites, (An intact DNA molecule in the largest human chromosome (number 1) is about 2.5×10^8 base pairs in length.)

We, and our collaborators at Livermore, use DNA from the bacteriopbage Charon 2 I A as a cloning vector. that is. as a medium capable of accommodating foreign DNA segments for replication. Charon 21A is a modified version of the λ bacteriophage; its linear DNA molecule is about 42,000 base pairs long, terminates with short stretches of single-stranded DNA, and includes a single *EcoRI* cleavage site.) The DNA molecules extracted from the phages are cut by *EcoRI* into two "arms." The cut ends of these arms are then stripped of a phosphate group (for a reason to be explained later) by treatment with the enzyme alkaline phosphatase.

The next step is fairly simple, at least in theory. We mix the two DNA preparations and let nature take its course—nature in this case meaning the tendency for single strands of DNA to "stick together" (become linked by hydrogen bonds) at regions with complementary bases, Since the cutting action of *EcoRI* has equipped both the chromosomal DNA fragments and the vector arms with just such sticky ends. aggregates consisting of' two arms linked by a (single) chromosomal DNA fragment will form. (Herein lies the reason for gentle extraction of the DNA from the chromosomes: physical breaks do not result in sticky ends.) These aggregates arc then converted into intact DNA molecules by treatment with the enzyme DNA ligase, which establishes a covalent bond at the break along each strand. The result is a set of recombinant DNA molecules—that is, molecules of phage DNA carrying human DNA fragments. (To increase the probability of forming recombinant aggregates. the chromosomal fragments are mixed with an excess of arms. Therefore, many of the aggregates formed will consist simply of two arms. But DNA ligase cannot repair the



III. CLONING THE: DNA

DNA in such aggregates because of the missing phosphate groups, So our foresight in the previous step reduces the background of nonrecombinant DNA molecules in the final library.)

Next we assemble Charon 21A phages, drawing on proteins produced by mutant varieties of the phage for the necessary structural components and on our supply of recombinant DNA molecules for the genetic component. This step unavoidably introduces a deficiency in the resulting library: those recombinant DNA molecules containing fragments longer than about 9100 base pairs are too large to be encased in the heads of the phages, and therefore the library will lack that portion (about 33 mass percent) of the chromosomal DNA residing on such fragments. It is for this reason that we and our collaborators at Livermore use different restriction enzymes and thus produce libraries based on different sets of chromosomal DNA fragments. The expectation is that the DNA missing from one laboratory's version of a library will be found in the other,

The final step in making the libraries, that of cloning the recombinant DNA molecules, is accomplished by seeding the phages onto a "lawn," or monolayer, of an appropriate strain of *E. coli,* When a phage injects its recombinant DNA molecule into such a cell, the biological apparatus of the cell is subverted to production of new phages. The cell ruptures when between one and two hundred new phages have been produced, and these then infect surrounding cells. The process continues. resulting in formation of a plaque, or clear area, in the translucent lawn of bacteria. In each plaque arc some one to ten million infective phages, each containing an exact copy of the recombinant DNA molecule carried by the original infecting phage.

These vastly multiplied recombinant DNA molecules, then, constitute a library-conveniently and safely packaged within the proteinaceous armor of phages. This construction of a library puts us at the starting line. Its application to problems of clinical and fundamental significance lies ahead. ■

what genes or combinations of genes are involved; in fact, we haven't come close to settling even the nature-versus-nurture issue. And even when a single defective gene is known to be the culprit—as it is in certain kinds of mental retardation—developing the techniques for introducing the normal gene and getting it expressed properly will take a long time.

HILDEBRAYD: Yes, routine clinical application of gene therapy is a long way off, especially in light of the ethical and legal questions it raises, But in the process of ironing out those questions, society should focus on the extraordinary benefits to be derived from the technology rather than its "Brave New World" implications.

MOYZIS: I think other developments in molecular genetics pose more immediate concerns. For instance. the availability of probes specific to the X and Y chromosomes has made it very easy to determine the sex of a fetus. Since even in "sexually liberatcd" western countries the preference for having a male child first is still strong. routine determination of fetal sex could conceivably lead to dramatic changes in the ratio of the sexes. The ethical. social, and economic implications of this possibility need immediate consideration.

SCIENCE: *Are issues such as these heavy on your mind as you go about your research?*

CRAM: I don't think so. The goals of basic research arc information and understanding. It's up to society to determine how those are applied. Take the photoelectric effect, for example. Its applications are mainly viewed as beneficial, although a few may be regarded by some as deplorable. But a scientist doesn't hide his head in the sand in the face of information that can possibly be applied in a harmful way.

MOYZIS: I would second that. You can't be paralyzed because a bit of knowledge can be used inappropriately. And 1 repeat that most of the worrisome aspects of gene therapy are a long way down the road. A major problem that must he solved before it can be applied to many diseases is how to introduce the good genes into only the appropriate cells of

a patient. If the defect is in a gene that is expressed, say. only in liver cells. extra copies of this gene in other cells might cause a great deal of mischief.

I am somewhat amazed, however, at the speed with which the medical community is moving toward gene therapy. I'm aware, though, that clinicians confronting patients have a different viewpoint than we do. It's no particular virtue on the part of people in basic research to be more cautious.

I think the patient and the patient's family should have a large say in deciding about experimental treatments. If a child of mine had Lesch-Nyhan syndrome, I'd be willing to take a chance on gene therapy. Terminal cancer patients are taking similar chances right now with other experimental therapies.

HILDEBRAND: An indirect form of gene therapy has already been tested clinically. Patients with a form of thalassemia were treated with a drug that would—it was hoped—reactivate the fetal hemoglobin genes. The drug worked, but since it is a potential carcinogen, the ultimate result may be initiation of tumors or other defects through inappropriate activation of other normally silent genes. Of course, all therapeutic regimes have the risks of unwanted side effects. Gene therapy presents the special risk of possibly altering the human gene pool. But so do chemotherapy and radiation therapy.

Direct gene therapy seems imminent. and public interest is high, even though the genetic diseases now being considered as candidates for this method of treatment are extremely rare. The Office of Technology Assessment has published a background paper on gene therapy, and the NIH's Recombinant DNA Advisory Committee has appointed a Working Group on Human Gene Therapy. The legal problems may take a long time to clear up, though.

DEAVEN: Yes, but that can change rapidly. About six years ago Margery Shaw, Director of the Institute of Medical Genetics at the University of Texas in Houston, gave a colloquium here at Los Alamos on the legal aspects of genetics. At that time the fetus had absolutely no legal status in this country or in other western countries. Last year she gave an update of that talk at the annual meeting of the American Society of Human Genetics. She reported that in parallel with scientific developments in recognizing forms of fetal abuse, such as overconsumption of alcohol or other drugs by the mother, principles are being established concerning the legal status of the fetus and its right to be born in a state of good health. My feeling is that a similar thing will happen with gene therapy, and that by the time gene therapy is a reality we will have arrived at legal guidelines for its application.

SCIENCE: *Do you think society turns to scientists-for ethical guidance as well as technical expertise?*

CRAM: Yes, and I think most of us accept that role, But the scientific community can lose credibility because even scientists can interpret the same technical data differently, This happened when policies for radiation exposure were being established.

MOYZIS: That's because we're human beings just like everybody else. Although we may be more aware of the possible applications of our research than the average person, most of us are not particularly qualified to spout off about the ethics of those applications. Whether or not I favor abortion or genetic engineering of humans is in some ways based on my scientific expertise but in other ways is not. We're often called upon at cocktail parties to discuss these sticky issues, but I tend to feel that most scientists should be asked for nothing more than scientific input. Most scientists aren't any more objective than anybody else when it comes to judgments that involve emotions and prejudices.

DEAVEN: Yes, but I do think scientists should spend some time thinking about the implications of their research and should actively participate in the formulation of ethical and legal guidelines for society. Some have altered their careers in this direction, I'll mention Margery Shaw again. After acquiring an M.D. and having a successful career in human genetics, she went back to school and got a degree in law because she saw the legal implications of new developments in human genetics. We need more people like that.

HILDEBRAND: This is all very interesting, but perhaps we should retreat to a more direct application of the libraries—that of developing more sensitive methods for detecting mutations.

MOYZIS: Yes, we badly need a means of detecting damage to human DNA caused by low-level exposure to mutagens, It's not hard to detect the damage caused by high-level exposure. Ionizing radiation, for example, causes hundreds to thousands of double- and single-strand breaks in the DNA molecules of a single cell at exposures over 1000 rads. But very few people ever get that kind of exposure. Most of the concern, medical and legal, is about the effects of low dose rates. Is exposure to fallout from the test of a nuclear weapon in Nevada some thirty years ago, or to Agent Orange in Vietnam, the cause of some of today's cancers? Does living in Los Alamos at an altitude of 7500 feet cause a higher incidence of spontaneous abortions? Who knows? It's very difficult to make these assessments.

Currently, the incidence of mutations is determined from the observable changes within a small number of genes, most commonly a single gene. Since a single gene is Only about one-millionth of a DNA molecule, the assays are not very sensitive. But with a battery of hundreds or perhaps thousands of gene probes pulled from the libraries, a much larger fraction of the DNA can be assayed for damage. A wealth of cytogenetic data indicates that some regions of certain chromosomes are more susceptible to damage. Obviously, probes for those regions would be of particular value.

DEAVEN: Los Alamos is already working on detecting mutations induced by low-level radiation by assaying what are known as tandem repeats for RFLPs.

MOYZIS: Yes, and although our results on irradiated cells are very preliminary. they're very encouraging. We're working with the repetitive sequences in human DNA. which

make up about one-quarter of the total. Many of these repetitive sequences are thought to be involved in defining chromosome structure or in switching genes on and off. They often jump to new locations in the DNA, altering the gene sequence they move into. We've cloned and characterized over 75 percent of the repetitive DNA sequences present in human DNA. Most of these sequences occur about a hundred thousand times per genome. and about a third of them occur in tandem, one right after another. With a single RFLP probe for such a tandem repeat, we can sample many orders of magnitude more DNA for mutations than with probes for a few genes. With this increase in sensitivity Susan Atchley, a graduate student in the Laboratory's Genetics Group, has been able to detect damage to DNA at a much lower level of radiation exposure than has been reported previously. But it appears that most of the observed changes in the DNA disappear with time following irradiation. Presumably the changes arc being repaired. These preliminary results have led us to facetiously throw around the idea—although maybe it's not so facetious-that a little bit of' radiation may be a good thing in the sense that it keeps the repair enzymes induced to such a level that they can more readily cope with larger doses.

The repair we are seeing has relevance to the long-standing question of whether or not such a thing as a threshold for damage exists. Perhaps there is a dose rate below which the DNA repair enzymes can handle the damage, even with continuous exposure.

CRAM: If repair is blocked. what levels of damage can be detected?

MOYZIS: We haven't done that experiment yet.

DEAVEN: With the current techniques for assaying genetic damage, large groups of exposed people must be studied to get good statistics, Very few such groups arc available. In fact. despite a massive effort since 1945, the data emerging for the largest group—the Hiroshima and Nagasaki populations-are questionable in terms of genetic significance. The power of the newer techniques is that they can be applied to small populations because they provide so much more information per person.

MOYZIS: The recent incident in Mexico, in which some two hundred people were exposed to cobalt-60 pellets from a junked radiation therapy machine, gave us a new population for study. but those people got horrendous doses.

DEAVEN: One problem with most of the available study populations is that the exposure is not known with certainty. We can estimate the exposure by determining levels of chromosome damage, but this method is not very accurate, especially if the chromosome analysis is not done immediately. Data from laboratory experiments, such as those Bob is talking about, should provide better correlations between damage and dose.

CRAM: Is there so much variation in restriction fragment lengths among individuals that detecting mutations requires a control sample for each individual?

MOYZIS: We've found that the variation

The success of the Los Alamos contribution to the Gene Library Project is due to the competence and enthusiasm of a large number of people. Pictured here are Kevin Albright, Marty Bartholdi, Bill Bentley, Nancy Brown, Evelyn Campbell, Doug Chritton, Scott Cram, Larry Deaven, Ed Hildebrand, Paul Jackson, Ted Lobb, Mary Luedemann, John Martin, Linda Meinke, Julie Meyne, and Bob Moyzis. Jim Jett, Jon Longmire, and Chris Munk **were** ***unavailable for the photograph.***

depends on the sequence. For damage estimates we've concentrated on sequences that seem not to vary in the population at large. Preliminary variation was estimated by isolating DNA from blood samples of any individual who was unfortunate enough to walk past my lab. The Life Sciences Division here has enough people with different ethnic backgrounds to make those estimates meaningful,

Repetitive sequences that do vary in the population are also interesting. Ones that map to particular chromosome locations will make exquisite probes for those chromosomes and will be useful in prenatal diagnosis, since orders of magnitude less fetal DNA will be needed for analysis. Ed and I are currently searching for such sequences in the libraries.

HILDEBRAND: Another application of probes from the libraries is investigating permanent changes in the genome on a larger time scale, for example, in evolution and speciation,

DEAVEN: Yes. Using a battery of probes that represent an ordered arrangement of human genes, one can look for patterns of constancy and of change in the genomes of plants and animals. That sort of information should give some insight as to just what the limits are on rearrangement of the genome. For example, if a group of genes has remained linked closely together over hundreds of millions of years of evolution, that's a pretty good sign that any perturbation of that conserved linkage is deleterious or even lethal, Taxonomists might be able to use such information as the basis for a better definition of speciation.

MOYZIS: It might also help solve the dilemma of cross-species extrapolation. The relevance of evidence from an animal study to humans can always be questioned. But if it is known what genes or organizations of DNA are unique to, say, mice and to humans and common to both, estimates of risk assessment from animal studies will be on firmer ground.

DEAVEN: Pedigree analysis will also benefit from the libraries. Pedigree analysis is the tracing of a genetic trait through generations of a family to determine whether the gene for the trait is dominant or recessive and whether it is located on one of the sex chromosomes or on an automsome. For this job geneticists have been using phenotypic evidence of the presence of a gene. But some genes don't provide unambiguous phenotypic evidence. The use of RFLP probes for such genes will make pedigree analysis easier and more accurate.

MOYZIS: In the near future enough probes for RFLPs will be available that it will be possible to assign paternity with 100 percent certainty.

CRAM: I believe a test like that was done on the pandas at the Washington zoo to find out whether the female was impregnated by her mate or by artificial insemination with sperm from a panda in England.

MOYZIS: I didn't realize that, but similar techniques are used routinely to distinguish different strains of plants. DNA might come to replace fingerprints for certain identification purposes.

SCIENCE: *Has production of the libraries been fairly routine or full of surprises?*

VAN DILLA: Dealing with life processes and living things practically guarantees surprises. An example is our experience with hybrid cells. The first hybrid cell line we worked with was just a dream. The cells were easy to grow and. as advertised, contained three human chromosomes, which were easy to sort. So it came as a rude surprise that other hybrid cell lines didn't perform nearly so well, One was supposed to contain human chromosome 1, but after the cells were grown, we couldn't find any to sort. We had known that hybrid cells have a habit of losing human chromosomes as they multiply but had been lulled into a false sense of security by the behavior of the first cell line.

Even so routine a process as chopping up DNA with a restriction enzyme gave us a surprise. We were getting a cutting efficiency of only about 1 percent even though we were using a standard restriction enzyme whose cutting efficiency was well documented to be close to 100 percent. At first we blamed a bad batch of enzyme, but that idea didn't check out. We still don't know exactly what the problem is, but it's been solved by an additional purification of the DNA extracted from the chromosomes. So we assume the enzyme was being poisoned by some contaminant the chromosomes pick up during sorting,

The logistics of producing the libraries gave us a surprise too. When we switched to sorting with the high-speed flow cytometer in the spring of '84, the poor cell farmers were suddenly faced with keeping up with its ten times greater appetite for chromosomes. That shouldn't have come as a surprise, but it did, probably because we were too busy with other aspects of the project to think hard about the manpower needed to keep the high-speed sorter happy,

SCIENCE: *Have you enjoyed participating in the Gene Library Project?*

HILDEBRAND: Definitely yes, but there has been an element of frustration because it is taking a much longer time than we anticipated. We're still adapting to the reality that this project is a long-term research and development effort with a defined endpoint. The Laboratory has vast experience in dealing with long-term experiments in physics. Life scientists, however, have traditionally been accustomed to experiments that can be completed in a period of days, weeks, or months, To a physicist what we call an experiment would be just one step in a much larger process.

MOYZIS: Just a coffee break. But that does bring up a good point, For the Life Sciences Division the Gene Library Project is a major undertaking in terms of the time necessary to get to the starting point of actually using the libraries. The DOE may have been right in thinking that Livermore and Los Alamos were the only places where enough biologists would talk to each other long enough to get the project going. Physicists are more accustomed to projects involving lots of people. Biology is still very much done by one guy with his test tube at his bench.

DEAVEN: Group research in biology is going to become more and more common, though. As more and more sophisticated technologies come into use, the breadth of expertise needed in a research effort gets greater. This project is certainly an example of that trend. Between twelve and fifteen people, representing a variety of specialties, are involved on a day-to-day or week-to-week basis.

CRAM: The point that Ed made about the difference between biologists and physicists is worth emphasizing. We can do an experiment sometimes in an afternoon, but physicists often spend years on an experiment. A biologist might publish six papers a year, but physicists doing an experiment at an accelerator might publish a single paper with ten authors every couple of years. I think the Gene Library Project is an example that crosses that barrier.

MOYZIS: Biology is still such a wide open field with so much left to learn that we can get significant, publishable results in a relatively short time. A number of people involved in this big project are going to have little holes in their curricula vitae because they felt it was so important. Not many of the biologists at universities would have been willing to devote two years or so to getting these libraries made.

DEAVEN: Working on something so different as this project has certainly been an interesting challenge. Sometimes though, and this is probably true of all of us, I dream of being on the other side of the fence—opening my mail, unwapping a little vial, and actually doing some research with a library. ■

Acknowledgments

The following people have made substantial contributions to the organization and administration of the National Laboratory Gene Library Project: Mark W. Bitensky, Leader of the Life Sciences Division at Los Alamos National Laboratory; Brian D. Crawford, now Director of the Molecular Biology Program at Long Island University; M. Duane Enger, now Head of the Department of Zoology at Iowa State University; Mortimer J. Mendelssohn, Associate Director of the Biomedical and Environmental Research Division at Lawrence Livermore National Laboratory; David A. Smith, Office of Health and Environmental Research, U. S. Department of Energy; Carleton C. Stewart, Leader of the Pathology Group at Los Alamos National Laboratory; Robert Wagner, Consultant at Los Alamos National Laboratory; and Ronald A. Walters, a member of the staff of the Associate Director for Chemistry, Earth and Life Sciences at Los Alamos National Laboratory.

Further Reading

James D. Watson, John Tooze, and David T. Kurtz. *Recombinant DNA: A Short Course.* Scientific American Books, 1983; distributed by W. H. Freeman and Company.

Alan E. H. Emery. *An Introduction to Recombinant DNA.* Chichester, England: John Wiley & Sons, 1984.

Maya Pines. "In the Shadow of Huntington' s." *Science 84,* May 1984, pp. 32-39.

Office of Technology Assessment. *Human Gene Therapy—A Background Paper*. Washington, District of Columbia: U.S. Congress, Office of Technology Assessment, OTA-BP-BA-32, December 1984.

Yvonne Baskin. "Doctoring the Genes." *Science 1984,* December 1984, pp. 52-60.