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.*
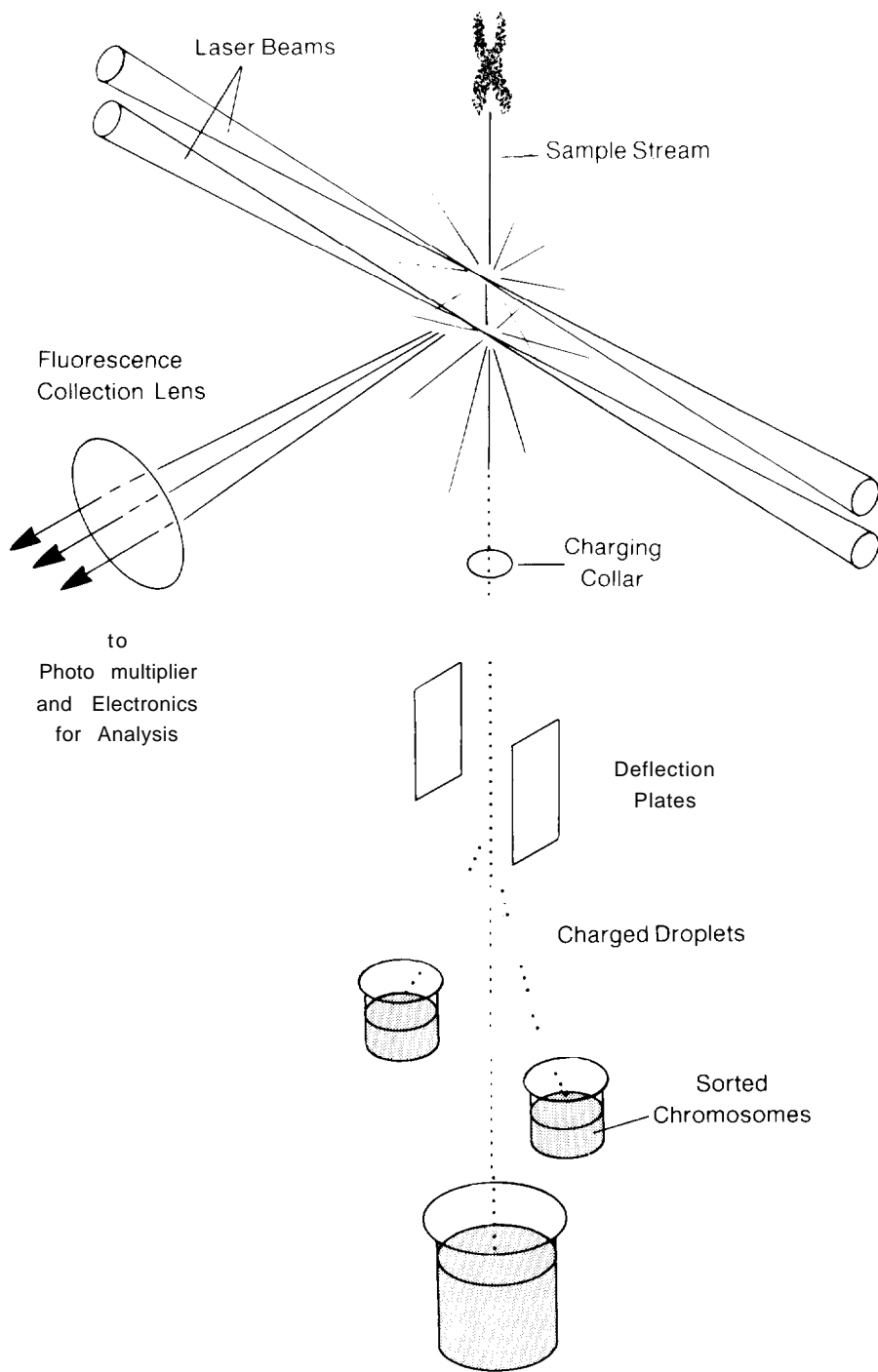
Fig. 1. Operating principles of a flow cytometer for sorting human chromosomes on the basis of the fluorescence intensities of two dyes, each of which binds preferentially to a different type of DNA sequence. A very narrow stream of the chromosome preparation flows rapidly through the focused regions 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.
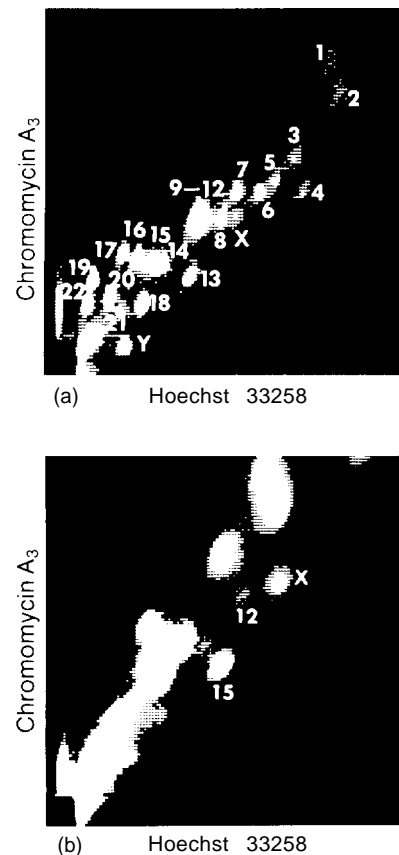

Fig. 2. (a) A flow histogram obtained from analysis of a typical chromosome preparation from human fibroblasts. The histogram is a record of the number of fluorescence events (each presumably due to a single chromosome) versus the fluorescence intensities of the dyes chromomycin A_3 and Hoechst 33258. 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