Democratizing the DNA Sequence

The polymerase chain reaction has made genetic sequences much more easily available, transforming molecular biology and allowing more investigators to work directly with DNA

SCIENCE IN THE '90s

Fifth in a series

"I KNOW THE TERM IS OVERUSED," says William Rainey of the Museum of Vertebrate Zoology at the University of California at Berkeley, "but this is a revolution."

The force driving Rainey's revolution is the polymerase chain reaction, or PCR. Only a couple of years ago, researchers found it extremely laborious to obtain a specific stretch of DNA in quantity from the mass of genes in a biological sample. By making that task easy, PCR is having profound effects on the work of genome mappers, AIDS researchers, evolutionary biologists like Rainey, developmental biologists, medical diagnosticians, and even forensic scientists. In many areas PCR has cut across the boundaries separating basic research and applied, commercial technology in a way few techniques ever have.

PCR has already turned museum collections into hunting grounds for clues to genetic evolution, made it possible to diagnose HIV infection accurately in newborns, and made substantial contributions to an ever finer grained study of gene regulation. Indeed, as the 1990s begin, PCR has taken its place as a basic tool of molecular biology. By the mid-1990s, some experts predict that most DNA-based diagnosis and screening for common genetic diseases will be done by PCR. And by the decade's end, PCR will be playing a key role in science's most ambitious biological endeavor: the Human Genome Project.

For all its current excitement, the PCR revolution began quietly enough—in the mind of Kary B. Mullis. By his own account, Mullis, then at Cetus Corp., thought up the technique one Friday night in the summer of 1983. The following winter, having made some educated guesses about concentrations and reaction times, he carried out his first experiment. "I was pretty lucky," Mullis says. "The first time I tried the reaction it worked."

What had Mullis discovered? The recipe for revolution.

Mix a DNA sample with a generous helping of DNA's four nucleotides, its chemical building blocks. Add a DNA polymerase (an enzyme that, given one strand of DNA, assembles the complementary strand). Add two primers (strands of DNA perhaps 20 nucleotides long chosen to be complementary to regions at opposite ends of the target sequence) to start the reaction.

This mixture is heated and cooled in cycles of a few minutes each. The first round of heating separates the double-stranded DNA into two single strands. Then, as the mixture cools, the primers find their complementary sequences, and the DNA polymerase gets to work, extending each primer into a new DNA strand.

That much is straightforward. The trick is that the next cycle of heating separates the copies from the original strands—and *both* sets become templates for a new round of DNA synthesis. As a result of this doubling, the target DNA multiplies exponentially—in a chain reaction. Thirty cycles of PCR—no more than an afternoon's work—can amplify a molecular signal that was too small to detect by more than a billionfold.

This was a method whose time had come. Previous techniques for isolating a specific bit of DNA relied on gene cloning—a tedious and slow procedure. PCR, on the other hand, "lets you pick the piece of DNA you're interested in and have as much of it as you want," says Mullis.

Patent applications and further work to refine the technique—much of it in Henry Erlich's laboratory at Cetus—followed Mullis's basic discovery. On the basis of those patents, Cetus now dominates the growing PCR business. In partnership with Perkin-Elmer, Inc., it sells reagents and a PCR machine, and with Hoffmann–La Roche it is developing kits for diagnosis. The only potential threat to this commercial supremacy is a suit filed late last year by the Du Pont Co. challenging Cetus's PCR patents. The basis of the challenge was a set of papers, published in the early 1970s by H. Gobind Khorana of the Massachusetts Institute of Technology and his colleagues, proposing a gene-copying method that shares some of the features of PCR but was never developed into a workable technique.

Although the outcome of that suit might shake Cetus's domination of the business of PCR, it certainly will not stop the scientific juggernaut that is already under way. Even Mullis, who has the natural pride of a new father, confesses to being overwhelmed by how quickly his baby caught on. According to a "PCR Bibliography" published by the Perkin-Elmer–Cetus partnership, 476 papers appeared last year describing work in which PCR played a central role.

From this welter of applications a few themes seem to be emerging that will be of increasing significance during the 1990s. One theme is the transformation of the study of evolution by a wave of DNA data. Allan Wilson of the University of California at Berkeley deciphers evolutionary relations by comparing genetic material from different organisms. Since his laboratory began using PCR, he has been swamped, he says: "With the same number of people in the laboratory, the rate of data production has accelerated so I can't keep up with it."

A particularly dramatic example of how PCR has speeded things up is the reanalysis Wilson and his colleagues recently did of their 1986 work suggesting all modern humans are descended from a single population in Africa some 200,000 years ago. The original study was laborious and indirect. DNA was extracted from mitochondria in blood samples from 200 individuals representing different populations around the world and analyzed with restriction enzymes, which cut DNA molecules differently if their sequence differs at specific points.

Wilson's group has now retraced its steps using the actual mitochondrial DNA sequences rather than the pattern of restriction fragments—a much more precise kind of data. Before PCR, that work would have meant cloning mitochondrial DNAs for all 200 individuals—a mammoth task. With PCR all it took was synthesizing "versatile primers": short sequences complementary to conserved regions in the mitochondrial DNA. The same pair of primers could pick out the homologous stretch of DNA in all of the nearly 400 samples used in this second study, yielding a product that could be sequenced directly. The results, obtained in only a few months, confirm the original conclusions—and also provide much more

precise dates for divergences between populations.

The value of PCR in this reanalysis was that it enabled Wilson and his co-workers to compare the same gene sequence in many different individuals quickly. For much the same reason, the technique is providing a window on one of the most diverse regions of the human genome: the histocompatibility genes, which code for cell-surface proteins that govern tissue compatibility and regulate the immune response.

Groups led by Erlich and by Hugh O. McDevitt of Stanford University have amplified histocompatibility genes from victims of autoimmune diseases and compared them with the same genes in healthy controls. The comparison enabled them to identify specific DNA sequences associated with such diseases as pemphigus vulgaris (a blistering skin disorder), insulin-dependent diabetes, and a form of rheumatoid arthritis. (Erlich and his Cetus colleagues have also harnessed the diversity of the histocompatibility sequence for forensic purposes in a PCR-based kit for DNA typing that was introduced by Cetus early this year.)

Speed and ease are not the only advantages PCR has over cloning and restriction-fragment analysis. Those methods all require abundant, high-quality DNA. In contrast, PCR can work with almost any genetic starting material, no matter how scanty or degraded. Wilson says that in evolutionary studies with PCR, museum specimens, which rarely have enough intact DNA for cloning, "become valuable resources for molecular analysis." And not just museum specimens: Rainey has amplified DNA for his studies of bats from bits of mummified tissue found in caves and old mines.

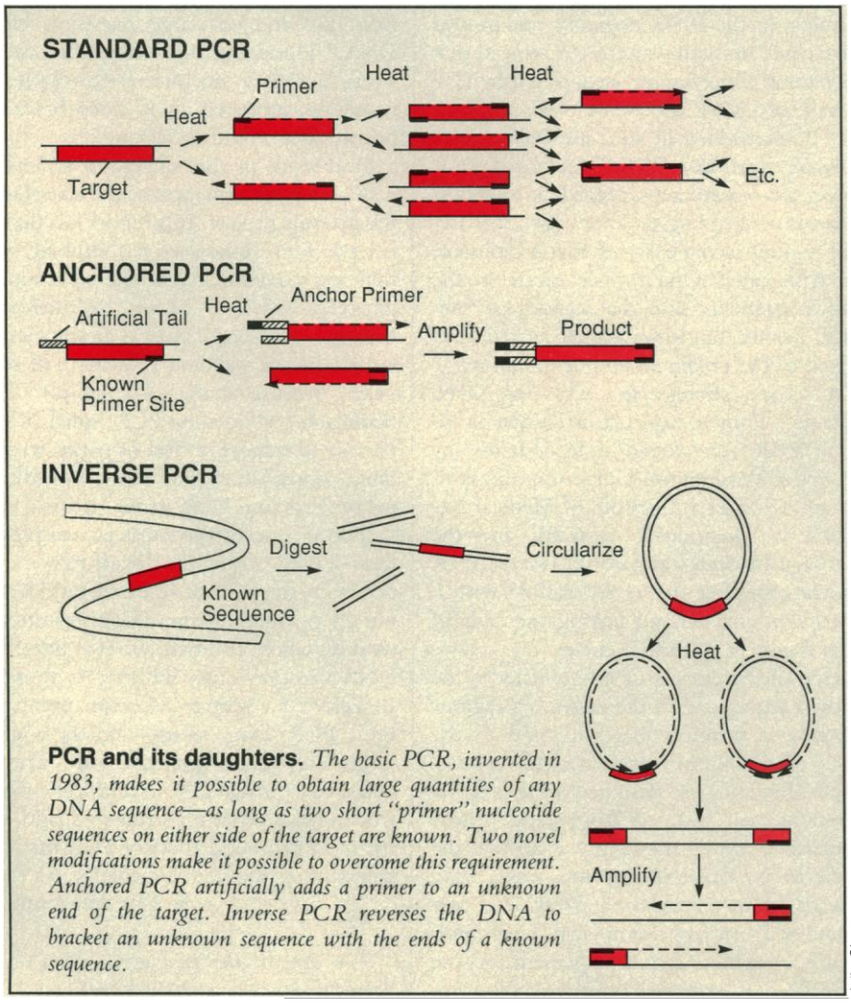
Such economy is making possible entirely new kinds of basic research. A group led by Norman Arnheim of the University of Southern California is using sperm to study genetic recombination, bypassing the extensive family studies that are usually needed. In effect, each sperm serves as an offspring whose genetic makeup can be examined by PCR. The availability of an essentially unlimited number of sperm makes it possible to study certain questions in unprecedented detail. "We can now study some very specific questions about recombination," Arnheim says, "such as the relationship of physical distance [between two loci on a chromosome] to recombination frequency."

In all the examples discussed so far, the virtue of PCR is that it enables the investigator to quickly find a known DNA sequence. But PCR is not limited to known bits of DNA. Using PCR, a researcher can, in effect, go fishing for related sequences. Primers based on the ends of a known sequence can pull up other matching sequences from a DNA sample, no matter how much the sequences may differ in the region between the primer targets.

Perhaps the most dramatic fishing expedition so far was made by Mark G. Rush and his colleagues at the New York University School of Medicine. By constructing primers complementary to parts of the gene for reverse transcriptase (the enzyme that copies a retrovirus into the host's DNA), the workers combed human DNA for traces of retroviruses. They found dozens of them, many of them genetically related to known primate retroviruses. These sequences are presumably disabled relics of ancient retroviral infections that are now passed on as parts of the human genome.

Other stretches of DNA are clearly not so benign. Indeed, the role in which PCR is most likely to become familiar to the public is as sleuth for specific pieces of DNA associated with genetic diseases. Among the diseases that can already be detected by PCR are many of the most common threats: beta-thalassemia, hemophilia, Tay-Sachs disease, phenylketonuria, Duchenne muscular dystrophy, and many cases of cystic fibrosis. PCR speeds diagnosis, reduces the size of the required sample, and often eliminates the need for radioactive probes. Considering such benefits, Haig H. Kazazian, Jr., of the Johns Hopkins University School of Medicine has predicted that by about 1995 PCR-based techniques will account for 90% of the DNA diagnosis methods for diseases caused by a single gene.

If DNA can detect genes gone wrong in human beings, it can also detect the DNA of pathogens with exquisite sensitivity. The work on AIDS presents a microcosm of PCR's contribution to medical practice and research. According to John Sninsky of Cetus, the PCR diagnostic assay can give defin-



PCR and its daughters. *The basic PCR, invented in 1983, makes it possible to obtain large quantities of any DNA sequence—as long as two short "primer" nucleotide sequences on either side of the target are known. Two novel modifications make it possible to overcome this requirement. Anchored PCR artificially adds a primer to an unknown end of the target. Inverse PCR reverses the DNA to bracket an unknown sequence with the ends of a known sequence.*

Julie Cherry

itive results when other, antibody-based tests are inconclusive or the infection is too recent for the patient to have developed antibodies. It may be especially valuable for diagnosing HIV infection in babies born to infected mothers.

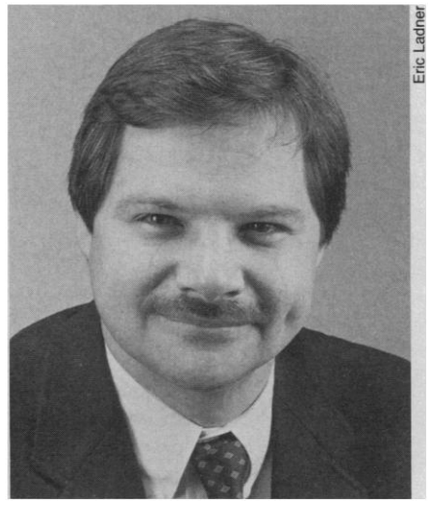
Because newborns carry maternal antibodies, the standard antibody test gives consistently positive results. But only 20 to 30% of such children actually go on to develop AIDS. A group led by Christian Brechot of the Pasteur Institute, Ellen Chadwick of the Children's Memorial Hospital in Chicago, and Martha Rogers of the Centers for Disease Control has shown that PCR can identify the infected children.

Later in infection, examining viral DNA can guide therapy by signaling resistance to the drug AZT, according to Brendan Larder and Sharon Kemp of the Wellcome Research Laboratories in England. PCR could also provide a direct gauge of a therapy's effectiveness. David Ho and his colleagues at the University of California at Los Angeles recently found that when AIDS patients are treated with AZT, the amount of virus that can be cultured from their blood plasma falls dramatically; other drugs might be judged by the same criterion. Sninsky thinks PCR could serve for monitoring the viral load directly, without culturing.

That's in the patient. In the lab, PCR is serving as a versatile molecular tool for probing viral genetics. By reverse-transcribing viral messenger RNAs and amplifying the resulting DNA, George N. Pavlakis and his colleagues at the National Cancer Institute have detected "what we believe are all the messenger RNAs of HIV." During its full life cycle HIV seems to make 30 different messenger RNAs—a remarkable number for a retrovirus and far more than any technique before PCR had revealed.

Pavlakis's group has been characterizing these messages by examining the size and sequence of the amplified DNAs and cloning them to get their protein products. Several of them appear to be hybrids of known HIV proteins, such as the *tat* and *rev* proteins. To dissect the functions of the viral proteins and DNA in greater detail, Pavlakis's group and others are harnessing PCR to introduce mutations into the relevant regions of DNA before cloning it. "These things could have been done before [the advent of PCR]," says Pavlakis, who adds that they could have taken days. "Now they can be done overnight."

Remarkable as the feats accomplished with the original version of PCR have been,



AIDS sleuth. *John Sninsky of Cetus is making use of PCR for a more accurate diagnosis of the disease.*

that method did have a fundamental limitation in the need to know a sequence of nucleotides at both ends of the target sequence in order to synthesize the right primers. As the 1990s begin, several groups have come up with ways to get around that constraint. For example, groups led by Gail Martin of UCSF and Mark Davis of Stanford developed a technique the Davis group dubbed "anchored PCR." Anchored PCR halves the requirement, according to Davis: "Instead of needing to know two 20-base-pair regions facing each other you can make do with one." One primer binds to the known sequence and the second, or "anchor" primer, binds to a sequence artificially linked to the unknown end of the target.

A second strategy for exploring DNA outside a known sequence is known as inverse PCR. Developed independently by groups at Washington University, the Walter and Eliza Hall Institute of Medical Research in Melbourne, Australia, and the National Institutes of Health, inverse PCR entails chopping the DNA sample with a restriction enzyme and linking the ends of each fragment to form a circle.

One of the circlets of DNA includes the known sequence and the unknown regions flanking it. Primers that will bind to the ends of the known sequence are synthesized—but they are designed so that DNA synthesis will proceed from each primer "outward," taking the long way around the circle to the other primer site, rather than inward, as in conventional PCR. The first round of DNA replication yields linear molecules in which the arrangement of the sequences is inverted: the unknown sequences are bracketed by the primer sites and ordinary PCR can proceed.

Such novel techniques—along with basic PCR—are likely to figure in the mapping and sequencing of the human genome, which promises to be one of the chief scientific endeavors of the decade. For one thing, what is likely to be the common language for describing landmarks on the chromosomes is predicated on PCR.

In August of last year four key figures in the genome project—Maynard Olson of Washington University, Leroy Hood of Caltech, Charles Cantor of the Lawrence Berkeley Laboratory, and David Botstein of Genentech—proposed that a several hundred base-pair stretch of DNA be sequenced at each landmark and the sequence published. Any investigator could check his material for the presence of a "sequence-tagged site" (STS)—and the corresponding landmark—simply by making the appropriate primers and running a PCR.

"What is now needed," says Hood, "is a really clever way of [establishing new] STSs from the ends of large pieces of cloned DNA." Hood thinks PCR will be the answer—probably an automated technique based on anchored PCR since bracketing primer sites would be lacking.

One result of this change in technology may be a transformation of the sociology of the genome project. In a report last September the four authors of the strategy wrote that as a consequence of the common language provided by STSs, "the dichotomy between 'big' and 'small' laboratories [involved in the genome project] will disappear." Because of the easy access to DNA landmarks provided by PCR, small labs will be able to achieve a kind of parity with big ones, coordinating their work with the major projects unfolding in the national labs.

Indeed, one of the most far-reaching effects of PCR during the decade to come may not be at the level of research results at all, but in opening up new opportunities for workers whose ability to work at the level of DNA was previously limited. As Rainey of Berkeley's Vertebrate Museum points out, until PCR came along, "unless you had biomedical-scale grants, you generally couldn't sequence DNA." The same may be true in fields other than Rainey's, and, if so, the most profound effect of PCR may be to democratize the double helix.

■ **Tim Appenzeller**

Tim Appenzeller is a senior editor at The Sciences *magazine in New York.*