Multiplying Genes by Leaps and Bounds

A new gene amplification method that greatly facilitates DNA analysis is finding a host of applications in medicine, forensic science, and fundamental molecular biology research

TISSUE FROM A WOOLLY MAMMOTH, who last walked the Siberian steppes some 40,000 years ago, would not seem to be a likely source of DNA for sequence analysis. After all, DNA breaks down very readily, and even though the mammoth had been frozen in the Siberian ice until it was chipped out about 10 years ago, its DNA had undergone extensive degradation. With the aid of a new gene amplification method, however, researchers were able to produce enough mitochondrial DNA to determine a nucleotide sequence. The results showed, not surprisingly, that the mitochondrial DNA of this extinct species is very similar to that of its close relative, the modern elephant.

The ability to do the analysis at all nonetheless illustrates one of the great strengths of the gene amplification method, which goes by the name "polymerase chain reaction" (PCR). The technique works with broken down DNA, as well as with intact DNA. Pieces as small as 50 to 100 base pairs can be amplified. "The nice thing about the PCR is that it [the DNA] does not have to be in good shape," says Norman Arnheim of the University of Southern California in Los Angeles.

This opens the way to a variety of applications in addition to analyzing ancient DNA samples. One that is advancing very rapidly is the forensic use of the technology to identify—or exonerate—suspects in rape and murder cases. The biological evidence found at the scene of violent crimes, including bloodstains, semen, and hair, is often of such poor quantity or quality that it can not be analyzed by other forensic techniques. But the DNA in the samples can be amplified and then compared with the suspect's DNA to obtain an identification.

(The forensic applications of PCR-aided DNA analysis and other new DNA-typing technologies that are beginning to be used in court will be discussed more fully in a second article.)

The gene amplification technique is also being applied in clinical medicine. It can be used, for example, to simplify and speed up the prenatal diagnosis of genetic diseases; to detect viral infections, including AIDS; to do the tissue-typing needed for matching the donors and recipients of transplanted organs; and to identify the genes that confer susceptibility to autoimmune diseases.

The PCR, which was originally devised about 3 years ago by Kary Mullis and his colleagues in Henry Erlich's laboratory at the Cetus Corporation in Emeryville, California, is essentially a test-tube method for copying simultaneously the two complementary DNA strands that make up a gene sequence. "It's a way of synthesizing millions of copies of a single sequence in a few hours," Erlich says.

"It's a way of sythesizing millions of copies of a single sequence in a few hours."

The researcher selects the specific DNA segment that he or she wants to amplify by using "primers," short pieces of DNA that have been synthesized to have sequences complementary to the DNA flanking the target region. The primers serve to define the ends of the DNA that will be duplicated. When a DNA sample is heated, the two strands separate, allowing the primers to bind to the flanking sequences, one on each strand.

The primers then initiate the synthesis of two new strands, complementary to the original two, in the presence of a DNA polymerizing enzyme. The PCR was originally performed with a DNA polymerase from the bacterium *Escherichia coli*, but this enzyme is inactivated by the high temperatures needed for strand separation.

More recently, Randall Saiki and David Gelfand of Cetus, Mullis, who is now at Xytronyx in San Diego, and their colleagues have modified the procedure to use a DNA polymerase from the bacterium *Thermus aquaticus*, which normally lives in hot springs at a temperature of 70° to 80°C. The *T. aquaticus* enzyme is heat-stable, a characteristic that gives it a considerable advantage over the *E. coli* polymerase.

Typically, 20 to 30 cycles of heating plus DNA synthesis are run during gene amplification by the PCR. "It's a major technological simplification," Erlich says of the use of the *T. aquaticus* enzyme. "Under the old protocol, you had to add enzyme after every cycle. With this new enzyme that is no longer required." In addition, running the PCR at the higher temperature has made the reaction more sensitive and more specific, resulting in less amplification of nontarget sequences.

Use of the heat-stable enzyme has allowed Cetus to develop, in partnership with the Perkin-Elmer Corporation of Norwalk, Connecticut, an automated instrument, the DNA Thermal Cycler, for running the PCR. Cetus has two patents on the PCR process, and has applied for patents on the DNA Thermal Cycler, the *T. aquaticus* polymerase, and many of the PCR applications.

Because each cycle of the PCR results in a doubling of the DNA sequence being amplified, the amount of material builds up very rapidly. Twenty cycles, for example, can theoretically amplify the DNA by a factor of about 1 million, and 30 cycles can amplify it by a factor of up to 1 billion. Once the amplified DNA is in hand, it can be analyzed by any of the methods of modern molecular biology. The complete nucleotide sequence can be determined, for example.

Until now, researchers who wanted to produce sufficient quantities of a specific gene to do a sequence determination had to clone the gene, which can take weeks, if not longer. But that time-consuming procedure can be avoided by using the PCR to amplify the gene. "We can go from a sample of blood to a sequence in less than a week," says Haig Kazazian of the Johns Hopkins University School of Medicine, who has been using the procedure to study the genetic defects that cause hereditary anemias such as β-thalassemia.

Applications that do not require complete sequence determinations can also be facilitated by the gene amplification procedure. For example, the prenatal diagnoses of some genetic conditions, including β-thalassemia and sickle cell anemia, can be achieved by obtaining samples of fetal cells and analyzing the genetic material with radioactive DNA probes that specifically pick up the gene mutations that cause the diseases.

Before this can be done, however, the cells have to be grown in culture for perhaps a week so that there will be sufficient DNA to carry out the test. But, with the PCR, the

DNA in a single cell is sufficient for amplification, thereby making the culture step unnecessary. Moreover, in the new techniques being developed by the Cetus and Johns Hopkins groups and that of Yuet Wai Kan at the Howard Hughes Institute at the University of California, San Francisco, the subsequent detection of the mutations is also simplified.

In particular, the amounts of DNA produced by the PCR are sufficiently large so that the high sensitivity imparted by the use of radioactive probes is no longer required and the mutant gene segments can be detected by nonradioactive probes or stains. This simplification may be especially valuable in those regions around the Mediterranean and in Africa where the thalassemias and sickle cell disease are endemic and where advanced medical diagnostic laboratories may not always be available.

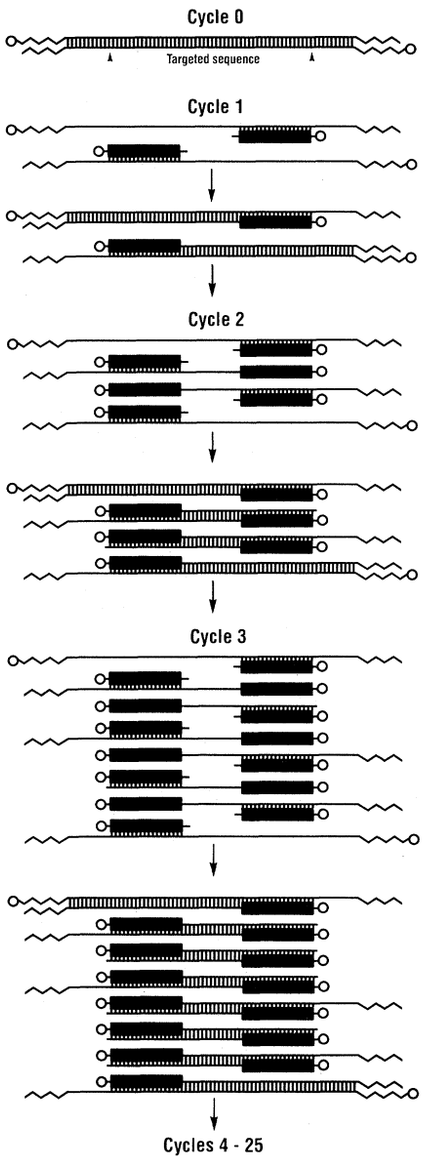
How the cost of the new procedures compares with that of the older methods is not clear at this early stage of development. Kazazian estimates, however, that the PCR techniques will probably be no more expensive than other methods. "What we are getting for the same cost is a quicker answer," he notes.

Autoimmune diseases, which are caused by an abnormal attack of the immune system on the body's own tissues, are not usually classified as genetic diseases. They have a genetic component, however, in that susceptibility to many of the diseases has been linked to particular variants of the class II histocompatibility molecules that mediate the immune cell interactions necessary for mounting normal immune responses.

By allowing the rapid amplification and sequencing of genes, the PCR is also facilitating the studies needed for a better understanding of genetic contributions to autoimmunity. For example, Erlich and his colleagues have identified specific histocompatibility gene variants that confer susceptibility to pemphigus vulgaris, an autoimmune condition characterized by skin blistering.

The high sensitivity of the PCR also makes it well suited for detecting viruses, especially those viruses that go underground in infected persons, often persisting for several years in a small number of cells before causing obvious symptoms of disease. "Most viruses are targets for PCR," says John Sninsky of Cetus, "but viruses associated with latency are particularly attractive for such an approach."

Human immunodeficiency virus type 1 (HIV-1)—the AIDS virus—is a case in point. (Also see p. 1407.) AIDS has a latent period of several years and indications are that very few susceptible cells—perhaps 1 in 10,000—actually carry the virus. Nevertheless, Sninsky, Gerald Schochetman of the Centers for Disease Control in Atlanta, and their colleagues have recently shown that the PCR can be used to amplify HIV-1 gene sequences in the DNA of peripheral blood cells from infected individuals, thereby allowing the direct detection of the viral presence. Sninsky estimates that the gene amplification method is capable of picking up one to ten copies of viral DNA per 1 million cells.



The polymerase chain reaction *is a way of amplifying specific gene sequences. The double-stranded DNA is heated and the separated chains are allowed to bind the primers (dark bars) that define the ends of the sequence to be amplified. The primers then initiate the synthesis of two new chains complementary to the originals. This series of events is repeated 20 to 30 times, with each cycle giving a doubling of the DNA.*

The current screening tests for HIV-1 infection identify the viral antibodies that are made by persons who have been exposed to the virus. They do not establish whether symptom-free, but antibody-positive, individuals still carry HIV-1 and are thus at risk of developing the disease themselves and capable of transmitting it to others. Previously the only way to tell whether HIV-1 is present was to isolate it, but the isolation procedure is cumbersome and takes some 3 to 4 weeks to carry out. For comparison, use of the PCR allows HIV-1 detection in just 3 days.

According to Sninsky, the PCR technique, by making it easier to detect HIV-1, may facilitate the epidemiological studies needed to resolve the issue of whether the virus persists in all antibody-positive individuals or whether it may sometimes be eliminated by the immune system. For clinical purposes, it may be possible to use the PCR to estimate how much HIV-1 a person carries. Individuals with the larger virus burdens, who are presumably at high risk of coming down with AIDS, might then be candidates for early drug therapy. In addition, the technique might help to determine whether a candidate antiviral drug actually reduces the amount of virus carried by patients undergoing therapy.

At least for now, however, PCR amplification is too complicated to be used for routine screening for AIDS infections. "ELISA [the current screening test] is just too simple and too inexpensive at this time," Schochetman says.

Sninsky, Bernard Poiesz of the Veterans Administration Medical Center in Syracuse, New York, and their colleagues have also recently used the PCR to detect gene sequences of the human T cell lymphotropic virus 1 (HTLV-1) in cells taken from several persons with a degenerative nerve disease called tropic spastic paraparesis. Although HTLV-1 is best known as a cause of human leukemias and lymphomas, it has also been causally linked to the neurological condition. The PCR work provides further evidence for that connection.

Moreover, researchers would like to know whether the virus that produces the neurological symptoms is identical to the one that causes leukemia. It should now be possible to resolve that issue by determining the sequences of HTLV-I genes amplified from patients with the nerve disease and comparing them with those of the virus obtained from leukemia patients.

The PCR can even allow the simultaneous detection of two viruses, which can be helpful for tracking viruses, such as HIV-1 and HTLV-I, that are sometimes transmitted together. Simultaneous detection can be

achieved by using two sets of primers, one for each virus, to amplify both viral genomes at the same time.

An alternative way of accomplishing simultaneous amplification of more than one virus is to use a primer that recognizes a sequence that is present in the genomes of all of them. Even very distantly related viruses may carry short identical segments in their genomes that are suitable for this purpose. This is true, Sninsky says, for members of the retrovirus family, which includes HIV-1 and HTLV-I.

Such conserved retroviral sequences may also provide handles for amplifying and analyzing as yet unidentified retroviruses. Although the work is controversial, an HTLV-I relative has been implicated in the etiology of multiple sclerosis, and the PCR might aid in the identification of the virus, if it is carried by the patients.

The main problem with using the PCR for virus detection is the possibility of contamination. "The advantage of the technique is also its disadvantage," as Schochetman puts it. The PCR is so sensitive that there is a danger that a virus contaminant might be amplified. Careful quality control will be required to prevent that from happening.

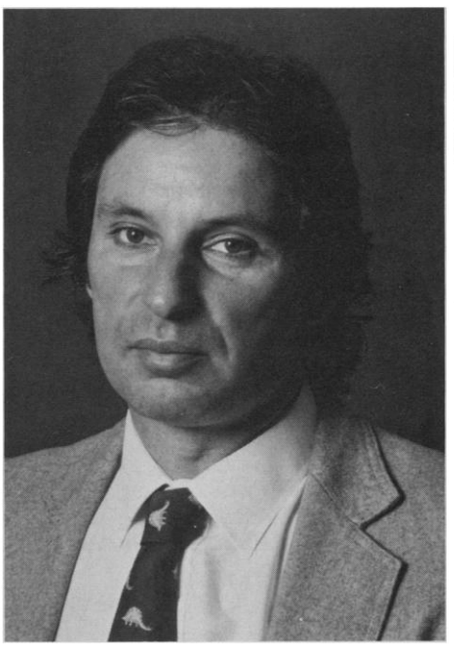
The PCR is also aiding efforts to track down the cellular changes that play a role in cancer development. For example, the normal genes that control cell growth and differentiation may sometimes undergo alterations that convert them to cancer-causing genes, called oncogones. For some genes, such as *ras*, a change in a single base pair is sufficient to do this.

Determining how frequently *ras* gene activation might contribute to the development of human cancers has been difficult. Although *ras* oncogenes have been found in several types of human cancers, in most studies, the percentage of tumors carrying the active oncogene was low, 10% or less. This figure may have been so low at least partly because identifying point mutations, such as those causing *ras* gene activation, can be difficult. "You're looking for a 1-base-pair mismatch in the whole human genome," explains Frank McCormick of Cetus.

Application of the PCR can help to simplify the problem, however. For example, Johannes Bos of the State University of Leiden in the Netherlands and his colleagues used the method in a study in which they showed that more than a third of human colon cancers carry oncogenic *ras* mutations. And this year, McCormick and his colleagues applied a similar approach to show that cancer cells from about one-quarter of patients with acute myeloid leukemia also contain *ras* oncogenes. "The PCR in this case reduces the complexity of the DNA so that a single base-pair mismatch is easy to detect," McCormick says.

Infections with some viruses have also been linked to human cancers. These include, in addition to HTLV-I, human papilloma virus, which has been implicated in the etiology of cervical cancer. Arnheim and his colleagues have used the PCR to provide further evidence for this hypothesis, but their work has a wider significance.

When patients undergo surgery for cancer, samples of the tissue removed are embedded in paraffin blocks. This is done so that very thin tissue slices can be made and



Henry Erlich *and his colleagues at Cetus have been developing the polymerase chain reaction and its applications.*

then examined by a pathologist to confirm the cancer diagnosis. The blocks of paraffin-embedded tissue are usually kept so that by now there is a worldwide collection of such preserved tissue specimens, with defined pathologies and known clinical results for the patients.

The Arnheim group has shown that DNA sequences of the human papilloma virus can be detected in such samples of cervical cancer tissue with the aid of the PCR, even though the tissue preparations had been made years ago and the DNA is likely to be degraded. The most important consequence of the work is that it opens up the possibility of using the tissue specimen collection for studies aimed at tracing the contributions of viruses and oncogenes to cancer development.

"Hypotheses linking the presence or absence of specific DNA sequences with a disease or its prognosis may now be immediately tested," Arnheim says. Because information on the clinical outcomes of the patients from whom the tissue specimens were removed is already available, long prospective studies are not required.

Not only does the PCR open up the collection of preserved tissue specimens for DNA analyses, but as already mentioned, it can also be used to study the very highly degraded DNA from archaeological samples. "I've been interested in ancient DNA and how we can get information out of those sequences. We could see immediately that this would be a terrific way of going for very small pieces," says Allan Wilson of the University of California, Berkeley. The analysis of the woolly mammoth DNA was originally performed in Wilson's laboratory by Russell Higuchi, who has since moved to Cetus.

Most of Wilson's work, however, is concerned with using variations in DNA structure to trace the evolution of human populations. In particular, Wilson and his colleagues have been focusing on the mitochondrial DNA, which can be transmitted only by the maternal parent. (Sperm do not have mitochrondria.) By tracing variations in the mitochondrial DNA in populations around the world, the researchers have concluded that all human beings alive today are descended from the same woman—popularized as "mitochondrial Eve"—who lived in Africa about 200,000 years ago (*Science*, 2 October 1987, p. 24).

To facilitate their analysis of mitochondrial DNA, Wilson and his colleagues turned to the PCR about 2 years ago. They were the first to show, in collaboration with the Cetus group, that DNA amplified by the technique could be sequenced directly, without having to clone it. With the method they then identified a particular mutation, a 9-base-pair deletion, in human mitochondrial DNA that apparently originated in East Asia.

Such DNA variations, or polymorphisms as they are called in molecular biology jargon, can aid in tracking the migrations of ancient peoples. The finding of the 9-base-pair deletion in some American Indians and in New Guineans indicates that migrants from East Asia contributed to the founding of these populations.

More recently, Svante Pääbo of the Wilson group has shown that DNA sequences from the dried-out brain of an Egyptian mummy and from a 7000-year-old human brain recovered from a flooded sinkhole in Florida can be amplified by the PCR and sequenced. "We can look at ancient peoples very easily now and can do population studies on extinct peoples," Wilson says.

■ JEAN L. MARX