What Controls the Genes?

Molecular biologists are using a virus to work out the nitty-gritty details explaining which genes in a cell will function and when

BY JULIE ANN MILLER

Brain cells, skin cells, liver cells have different shapes, make different enzymes, perform different functions. Yet they, and almost all other cells in the body, contain the same genetic information.

Molecular biologists have described the structure of the genetic material DNA and the basic mechanisms by which it directs construction of proteins. Researchers are now making progress toward understanding the processes that direct which genes will be active in a particular cell at a particular time. The controversial technique of recombinant DNA construction is one of the methods used to address this problem.

To examine control of gene expression, several laboratories are studying a virus called lambda, which infects the intestinal bacteria *E. coli*. Because of its dual lifestyle, lambda must control expression of its genes very closely.

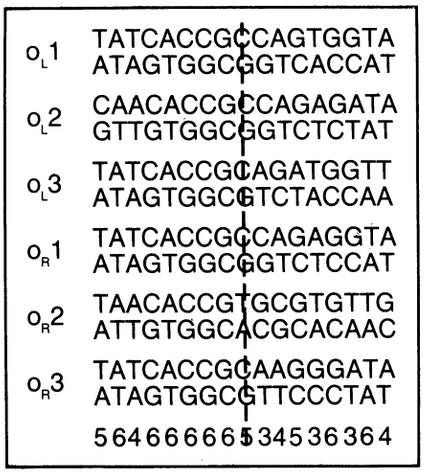
Sometimes when the virus infects a bacterium it takes over protein production, and the bacterium begins manufacturing viral proteins. After about 45 minutes the bacterial cell bursts, releasing hundreds of new lambda viruses.

Under different conditions, the virus does not immediately harm the infected cell. Rather, the viral DNA attaches to the bacterial DNA, and the lambda genes are quietly reproduced and passed on to the daughter cells each time the bacterium divides. When conditions change, the viral DNA can again take charge to manufacture new lambda viruses and destroy the bacterial host.

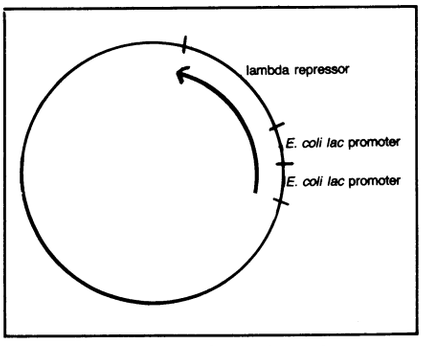
When the viral DNA is attached to the bacterial chromosome, it is crucial to the virus that its own genes remain silent while the bacterial genes run the cell operations.

Fifteen years ago two French scientists, François Jacob and Jacques Monod, proposed a theory of how cells turn off expression of parts of their DNA. Jacob and Monod were studying a group of genes in *E. coli* bacteria. Those genes specify proteins involved in metabolism of the sugar lactose.

Jacob and Monod hypothesized that one of the genes was responsible for a type of protein that was not an enzyme or a structural part of the cell, but rather that regulated the expression of DNA. When that hypothetical protein, called repressor, bound to a specific site on the DNA molecule, it would prevent formation of the products of a group of nearby genes.



Sequences of the six DNA sites where repressor protein binds. The number of sequences where the same base pair appears in each position is tabulated.



Recombinant DNA plasmid used to make large amounts of lambda repressor.

Repressor protein has since been actually identified and isolated, both for the lactose genes *(lac)* that Jacob and Monod studied and for genes of the virus lambda. The exact sequences of the *lac* repressor's submolecular components, 347 amino acids, were determined a few years ago by Konrad Beyreuther and his colleagues in Cologne, Germany (SN: 1/19/74, p. 40). Analysis of the sequence of lambda repressor's 272 amino acids is nearing completion in the laboratory of Mark Ptashne at Harvard University. Knowledge of these sequences may eventually help to explain exactly how the repressor interacts with DNA.

Using an electron microscope, researchers can now see both lambda and *lac* repressor bound to DNA. The photograph on page 349, of lambda repressor and DNA, was taken by Jay Hirsh and Robert Schleif at Brandeis University. Demonstration that such micrographs show repressor on DNA and the experimental methods are contained in a paper by these authors in JOURNAL OF MOLECULAR BIOLOGY, in press.

The analysis of lambda repressor required using recombinant DNA molecules to create large amounts of repressor protein. When normal lambda DNA is incorporated into bacterial DNA, only 200 molecules of repressor in each cell suffice to suppress the viral genes. To obtain more repressor, Ptashne and coworkers constructed a small circular piece of DNA, called a plasmid (SN: 6/21/75, p. 404), that contained the gene for the lambda repressor adjacent to two identical pieces of bacterial DNA (see diagram). The bacterial DNA selected was the region that normally initiates expression of the *lac* genes. Therefore in the plasmid the viral repressor gene was regulated by the bacterial promoter region.

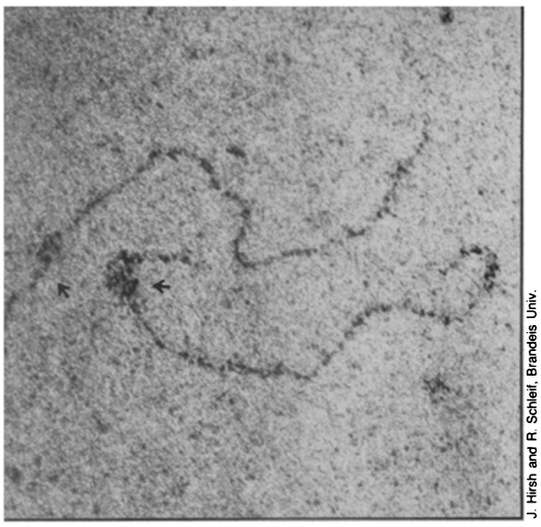
The plasmid of viral and bacterial DNA was introduced into *E.coli* cells. Because more molecules of lactose enzymes than repressor molecules are produced in bacteria, the cells with the plasmid made 50 to 100 times more repressor than did bacteria containing normal lambda DNA. This technique gave the researchers enough repressor protein to determine the exact arrangement of its component amino acids.

"The availability of large amounts of repressor will allow people to isolate and crystallize the protein, to find its shape and to see how it binds to DNA," Ptashne says.

Research on genetic regulation has focused, not only on the repressor protein, but also on the DNA where it binds. Repressor attaches directly to specific sites—one in the group of bacterial genes that control lactose metabolism and six different sites (see diagram) on lambda DNA.

Scientists have discovered the function of some of these DNA control regions. Repressor molecules that bind to sites o_L1 and o_L2 turn off the gene to the left, named *N*. Repressors that bind to sites o_R1 and o_R2 repress the right gene, called *cro*.

The gene name *cro* is an abbreviation for "Control of Repressor and Other things." That name was chosen because geneticist Harvey Eisen predicted that repressor control would depend only on a DNA site and not on a gene. So when Eisen discovered the gene *cro*, he figuratively



J. Hirsh and R. Schleif, Brandeis Univ.

Repressor protein (arrows) bound to specific sites on lambda DNA. Magnification of this electron micrograph is × 360,000.

had to eat it. That gene is also called *ar* (anti-repressor) and *tof* (turn off function).

When repressor molecules are bound to four sites between *N* and *cro*, they turn off almost all the lambda genes, because the protein specified by *N* is required for most of the virus's approximately 50 other genes to function.

"The *N* gene somehow causes the other genes to be expressed," Ptashne says. "The way it works is still a subject of discussion."

Knowing the site where repressor acts has revealed how it turns off gene expression. Repressor bound to DNA gets in the way of RNA polymerase, the enzyme that translates DNA into RNA.

The genetic information of a DNA molecule is coded in the order of paired chemical components called bases. When a gene is expressed, RNA polymerase separates the two strands of DNA along a short distance. The separation occurs between the bases of each pair.

RNA, the messenger between DNA and the cell's protein synthesis machinery, also is composed of building blocks. When the DNA strands have separated, one specific RNA component binds most efficiently to each DNA base. RNA polymerase then links together the RNA components, and the messenger RNA moves into the cell cytoplasm.

To begin the whole process of gene expression, the RNA polymerase must bind to a specific DNA site. If its site is already occupied, the enzyme cannot act. Because the polymerase binding site for gene *N* overlaps with repressor binding sites o_L1 and o_L2, and the polymerase site for *cro* overlaps with o_R1 and o_R2, RNA polymerase cannot begin transcribing the associated genes if repressor has been bound.

"Apparently repressor bound to two sites excludes polymerase more efficiently than does repressor bound to a single site," Ptashne says.

Researchers have determined the exact sequence of DNA subunits of repressor and polymerase sites. The six repressor sites of lambda are similar in their subunit sequences, and the sequence within each site is partially symmetrical around the center pair of bases (see chart, the letters represent the four different DNA bases). The sequence of subunits has so far given few clues as to how repressor protein and DNA interact.

Repressor bound to the right operator region of DNA controls expression of the repressor's own gene, as well as of *cro*, the research group at Harvard and another led by Vincenzo Pirrotta in Basel, Switzerland, have discovered. Ptashne's group reported in the Oct. 8 SCIENCE both negative and positive mechanisms of this control.

Such extensive regulation is necessary, they believe, to keep repressor concentration within a narrow range. Too much repressor would prevent the virus from taking over protein production in an emergency, and too little repressor would make the virus destroy its host unnecessarily.

The intricate control of repressor relies on the protein binding more tightly to some sites than to others. Of the three right binding sites, the repressor binds most strongly to o_R1 and least strongly to o_R3. Therefore the o_R3 site is likely to be filled only when there is a high concentration of repressor in the cell cytoplasm, and no need to synthesize more. The biologists have determined that repressor protein bound to site o_R3 turns off production of repressor, a negative feedback.

The other means of repressor gene control is positive. Repressor bound to o_R1 enhances expression of the repressor gene, fivefold to tenfold. Bound repressor encourages the binding of RNA polymerase and thus the functioning of the gene, Ptashne suspects, either by slight changes in the local DNA structure or by direct contact with the protein.

Researchers have found yet one more means by which the virus controls repression of gene expression. When the viral DNA first enters a bacterial cell, the immediate rate of repressor synthesis determines which lifestyle the virus will pursue, whether it will rapidly kill the cell or whether its DNA will incorporate into the bacterial chromosome. To insure that incorporation occurs often, the rate of repressor production is five to ten times greater immediately after infection than it is after the viral DNA is incorporated.

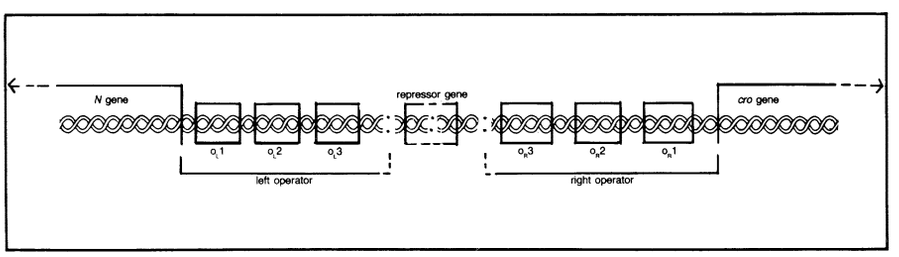
The reason for this difference may be that RNA polymerase binds to different sites in these instances. Pirrotta in the Aug. 19 NATURE and Ptashne reported finding two sites. Immediately after infection, the polymerase binds further from the start of the repressor gene and therefore produces a somewhat longer messenger RNA molecule. Pirrotta and Ptashne suggest that within the extra length of RNA (the leader) there is a strong binding site for ribosomes, the cellular structures that manufacture protein. Ribosomes are therefore most likely to translate the longer RNA into protein.

"It is remarkable that the phage [virus] has developed such elaborate mechanisms both at the level of transcription [DNA to RNA] and translation [RNA to protein]," Pirotta's group writes.

Ptashne feels that even this rather extensive description of repressor regulation is incomplete. When viruses are being made at the expense of the bacterial cell, still another repressor, the product of the *cro* gene, binds to the DNA to turn down the expression of *N*, *cro* and the repressor gene. The mechanism of this repression remains unknown.

"We have not yet elucidated all the functions mediated by these remarkable regulatory sequences," Ptashne's research group concluded.

Although repressor has been thus far identified only in bacteria and viruses, Ptashne believes that similar mechanisms control gene expression in higher organisms. "Everything else has been the same," he says. □



The region of the lambda virus chromosome that controls expression of the genes.

Julie Ann Miller has joined the staff of SCIENCE NEWS as Life Sciences Editor. She holds a B.A. in biochemistry and molecular biology from Harvard University, a Ph.D. in neuroscience from the University of Wisconsin and has published research in a number of scientific journals. Since January she has been a graduate student in journalism at the University of Wisconsin and a writer in the University-Industry Research Program there.

She succeeds Janet L. Hopson, who is on leave to work in the recombinant DNA laboratory of Herbert Boyer at the University of California at San Francisco on a Science-Writer-in-Residence fellowship from the Council for the Advancement of Science Writing. Following that she will remain on the West Coast as a SCIENCE NEWS contributing editor.