Visualizing Cell Processes

A Series of Five Programs
produced by BioMEDIA ASSOCIATES

Content Guide for Program 2

The Genetic Code and Its Translation

www.eBioMEDIA.com

Each of the five programs in this series consists of a set of short, narrated, full-motion modules 1-3 minutes long. Each module conveys an essential process or concept of cellular biology. The modules are organized around national standards for teaching biology.

With the discovery of the structure of DNA it became immediately apparent that the rearrangement of DNA's four building blocks (nucleotides) could be an instruction code for building proteins. Proteins have up to 20 amino acid units, so it was reasoned that three DNA nucleotides would be needed (64 possible code words could be made using combinations of the four DNA nucleotides in suits of three). Two letters code words would give only 16 combinations, not enough to code for all of the 20 amino acids found in proteins.

Breaking the DNA code, and the explosion of molecular genetics that followed, have defined biology for a half-century. Modifying the code to create special traits in work-horse organisms such as the bacterium *E.coli* is now a major biotechnology industry that has led to modifying genes, and splicing genes between species with often re-markable, if controversial, results.

All of this work rests on a simple principle. The genetic code is a code that specifies the precise structure for the millions of kinds of proteins found in living things.

1: Proteins, The Molecules of Life
Proteins are a major component of living things. Skin and hair are made from structural proteins, as is the internal framework of cells.

Motor proteins are found in muscles. Motor proteins produce the bending action of flagella and cilia and they are responsible for the movement of material within cells. Refer to Program 2, *Cell Movement and Transport*, to see how motor proteins work.

Enzymes (the biological catalysts that control virtually all of the chemical processes of life) are proteins.

The shape of a protein is determined by the sequence of its amino acids, for that will determine the location of the chemical cross-links that give a protein its three-dimensional form. If a key amino acid is changed, the protein's shape will change, and this will probably change its function. A striking example of an amino acid substitution is found in hemoglobin, the red pigment in blood cells. This mutation causes normal red blood cells to change their shape from bi-concave disks to crescent shaped cells, or "sickle cells."

A change from the amino acid glutamic acid to the amino acid valine at a specific site in the hemoglobin molecule causes red blood cells to sickle under low concentrations of oxygen, producing a condition known as sickle-cell anemia.

In the malaria belt of Africa, the sickle cell mutation is maintained in the population because it reduces the incidence of malaria (sickling kills the malarial parasites that have invaded the red blood cells). It's an evolutionary trade off—malaria is simply a greater threat to survival than is having the sickle cell gene.

2: Building Proteins According to Specifications

In nature, proteins can only be manufactured by living cells. The instructions for building them are encoded on the cell's DNA where the amino acid sequence is spelled out on one strand of DNA, the "instructing strand." This instructing strand serves as a template, or linear code for the amino acids that will make up the protein.

To make a protein, the three-lettered DNA codes are first transcribed to a new template made from RNA nucleotides. The process of transcription involves the enzyme RNA polymerase. RNA polymerase opens the double-stranded DNA and begins pairing RNA nucleotides to the instruction strand of DNA. This forms a single strand of RNA called messenger RNA or mRNA. The RNA polymerase enzyme transcribes the DNA instructing strand until it reaches a termination sequence of nucleotides and releases the mRNA transcript.

In the next step, the string of mRNA triplet codes will be translated into a string of amino acids. This process occurs on ribosomes, tiny organelles found on the membranes surrounding the nucleus

Translating the mRNA requires an intermediate form of RNA—transfer RNA. Each transfer RNA has an attachment site for a particular amino acid and another site containing the anticodon for that amino acid. Matching tRNA anticodon with mRNA codon positions the correct amino acid in the growing chain. The translation animation should help in understanding this fundamental process.

Translation stops when a termination signal encoded on the mRNA is reached. As translation is completed the chain of amino acids called a polypeptide twists into its final shape.

In many cases, two or more polypeptides come together to form the final protein.

3: Gene Regulation in Prokaryotes

Discovering how genes are translated into proteins was a giant step in our understanding of cell processes. The next question to confront cell researchers was—How are genes turned on when the protein products they specify are needed by the cell? As usual, bacteria, the simplest experimental subjects were called on.

from the script . . .
The bacterium Eschericia coli can be grown on a medium containing lactose, a disaccharide sugar. E. coli produces the enzyme beta galactosidase that splits lactose into two simple sugars—molecules the bacterium can metabolize. The question is how does the presence of lactose trigger E. coli to turn on the gene for beta galactosidase?

Controlling the production of the enzyme is a section of DNA known as the lactose operon—or "lac operon." The lac operon is composed of three DNA sections: a promoter, an operator and genes for beta galactosidase and its companion enzymes. The promoter is a landing site for RNA polymerase—the enzyme that transcribes the nucleotide sequence into messenger RNA. At another location on the DNA is a gene that codes for repressor protein that binds to the operator section of the lac operon preventing transcription of the beta galactosidase gene. This is the normal condition of the resting bacterium when no lactose is available.

However, if a lactose molecule bonds with the repressor it will release its chemical grip and transcription of the beta galactosidase enzyme can proceed. When as lactose molecules are all broken down, no more can bond with the represser and it goes back to work stopping transcription, creating a simple feedback regulation system for this gene.

4: Exons and Introns

Prokaryotes and eukaryotes differ in the way the genetic information for specifying protein structure is distributed along their DNA molecules. This difference showed up when biologists experimenting with the protein ovalbumin, egg white, made a startling discovery.

When they tried matching DNA from the egg nucleus with messenger RNA from the egg cytoplasm only short sections of the two strands matched up. Surprisingly, the DNA blueprint for a protein is broken up in sections, separated by long sections of DNA code words that are not transcribed.

The coded sections that specify the protein are called "exons." The interspersed silent sections, much longer than the exons, are called "introns." Surprisingly, all of the DNA in a gene, including the introns, is transcribed into messenger RNA. Then enzymes cut the introns out of the mRNA transcript and splice the exons into a corrected strand of genetic code, which can then be translated into a protein.

This situation is similar to a computer's badly fragmented hard drive. The information you wish to retrieve has been fit here, there, and everywhere, but although scattered all over the disc, it can still be retrieved to form a coherent message (most of the time).

One of the great questions of cellular genetics is: What is the possible function of the vast sections of silent DNA found in eukaryotic cells? One hypothesis is that introns may constitute a store of alternate protein instructions that may be occasionally accessed by a mutation. If so, introns may be an important source of raw material of evolution. Or, are introns just junk we eukaryotes can't seem to get rid of? Stay alert as the intron story unfolds.

5: Mutations

The DNA code-book for a frog or mammal contains millions of code words, each handed down from generation to generation. If a publisher were given the task of copying the instruction blood with 100% accuracy, they would turn down the job as simply impossible. Life generally does a little better than publishing technology, but there is still ongoing modification of the genes through mutagenic processes.

Organisms tend to be quite well tuned to their lives, so any genetic change is most likely detrimental, and some single nucleotide changes are fatal to the individual. However, a lot of mutations tend to be neutral—so what if a protein's shape changes a little when one of its 5,000 amino acids is swapped for another... it may have no effect on survival, at least under the organism's present conditions.

Mutations are caused by radiation, by reactive chemicals that find their way into the nucleus, or by viruses that snip their DNA into the host's DNA, where they will be replicated as a hidden guest.

Cells contain repair systems that correct such changes but some slip by. If they occur in a gamete, or in a cell in an early embryo, then all of that cell's descendants will carry the mutation.

One disastrous change is a frame shift mutation. Frame shifts occur when the loss of a DNA nucleotide causes the reading frame to shift, throwing off the entire triplet code downstream from that point in the gene. Consequently most of the amino acids specified after the frame shift will be different resulting in a garbled protein that will fail to function.

The most common mutation—a point mutation—occurs when a nucleotide is replaced. The resulting protein may still perform its function or it may function in a slightly different way, creating new possibilities for biological evolution.

6: The HIV Virus

A good way to review the way DNA specifies proteins is to take a look at a current health problem and follow the HIV virus as it goes about replicating itself in human cells.

from the script . . .

HIV researchers knew that many viruses carry surface proteins often held out on short stems. These proteins are used by certain viruses including HIV to find the proper type of cell to infect. They are looking for a cell with matching receptor proteins on its surface.

HIV specifically targets a type of white blood cell known as the "helper T-lymphocyte"(one of the many types of cells that make up our immune system. A helper T-cell carries a unique protein on its surface and HIV surface proteins latch on fitting like a key in a lock.

Next, the core which contains the viral genes encoded on RNA (not DNA) enters the cell. HIV is a retrovirus— a type that will insert its genes directly into the host cell's DNA. To do this, the retrovirus carries reverse transcriptase, the enzyme that transcribes RNA into DNA strands.

The DNA transcript enters the nucleus and snips into the helper T-cell's own DNA. This foreign DNA can now be transcribed into messenger RNA just as though it were another set of T-cell genes. At the same time, the retrovirus RNA genes instruct the cell to manufacture virus proteins.

Assembly takes place on the inside of the helper T-cell's plasma membrane and the

retroviruses are budded from the cell ready to tumble through the bloodstream in search of other cells carrying the appropriate surface receptors.

At some point, the lymphocyte, having produced many infective HIV particles, dies. As more lymphocytes are infected the host's immune system begins to break down leaving the person defenseless against even the most common germs.

So, viruses are little more than sets of genes that reproduce themselves by taking over the cell machinery for replication and protein synthesis, a way of life that probably began with the earliest cellular life forms. See *The Biology of Viruses* for details.