

From the Genome to the Proteome

Cells are the fundamental working units of every living system. All the instructions needed to direct their activities are contained within the chemical DNA (deoxyribonucleic acid).

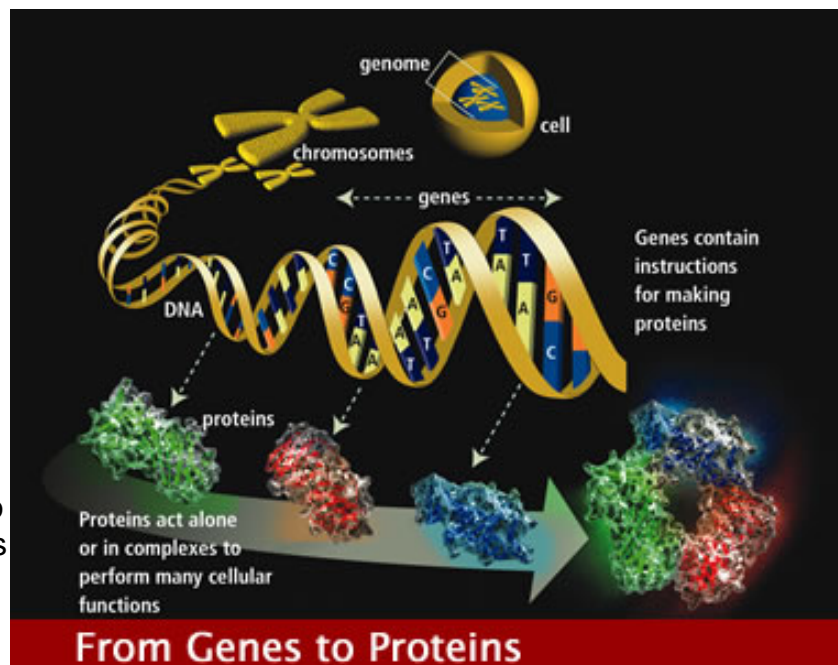
DNA from all organisms is made up of the same chemical and physical components. The DNA sequence is the particular side-by-side arrangement of bases along the DNA strand (e.g., ATTCCGGA). This order spells out the exact instructions required to create a particular organism with its own unique traits.

The **genome** is an organism's complete set of DNA. Genomes vary widely in size: the smallest known genome for a free-living organism (a bacterium) contains about 600,000 DNA base pairs, while human and mouse genomes have some 3 billion. Except for mature red blood cells, all human cells contain a complete genome.

DNA in the human genome is arranged into 24 distinct **chromosomes**--physically separate molecules that range in length from about 50 million to 250 million base pairs. A few types of major chromosomal abnormalities, including missing or extra copies or gross breaks and rejoinings (translocations), can be detected by microscopic examination. Most changes in DNA, however, are more subtle and require a closer analysis of the DNA molecule to find perhaps single-base differences.

Each chromosome contains many **genes**, the basic physical and functional units of heredity. Genes are specific sequences of bases that encode instructions on how to make proteins. Genes comprise only about 2% of the human genome; the

remainder consists of noncoding regions, whose functions may include providing chromosomal structural integrity and regulating where, when, and in what



quantity proteins are made. The human genome is estimated to contain 20,000-25,000 genes.

Although genes get a lot of attention, it's the **proteins** that perform most life functions and even make up the majority of cellular structures. Proteins are large, complex molecules made up of smaller subunits called amino acids. Chemical properties that distinguish the 20 different amino acids cause the protein chains to fold up into specific three-dimensional structures that define their particular functions in the cell.

The constellation of all proteins in a cell is called its **proteome**. Unlike the relatively unchanging genome, the dynamic proteome changes from minute to minute in response to tens of thousands of intra- and extracellular environmental signals. A protein's chemistry and behavior are specified by the gene sequence and by the number and identities of other proteins made in the same cell at the same time and with which it associates and reacts. Studies to explore protein structure and activities, known as proteomics, will be the focus of much research for decades to come and will help elucidate the molecular basis of health and disease.

Source: U.S. Department of Energy Office of Science, Office of Biological and Environmental Research, Human Genome Program

Base URL: www.ornl.gov/hgmis

454 Life Sciences Corporation
20 Commercial St.
Branford, CT 06510
203-871-2300
www.454.com