

## 17.5 Genomics and Proteomics

*By the end of this section, you will be able to do the following:*

- Explain systems biology
- Describe a proteome
- Define protein signature

Proteins are the final products of genes, which help perform the function that the gene encodes. Amino acids comprise proteins and play important roles in the cell. All enzymes (except ribozymes) are proteins that act as catalysts to affect the rate of reactions. Proteins are also regulatory molecules, and some are hormones. Transport proteins, such as hemoglobin, help transport oxygen to various organs. Antibodies that defend against foreign particles are also proteins. In the diseased state, protein function can be impaired because of changes at the genetic level or because of direct impact on a specific protein.

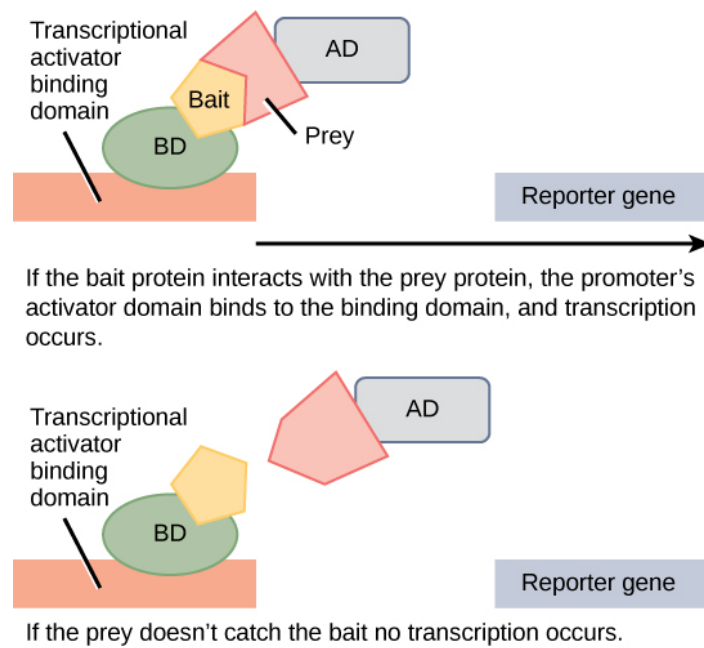
A **proteome** is the entire set of proteins that a cell type produces. We can study proteomes using the knowledge of genomes because genes code for mRNAs, and the mRNAs encode proteins. Although mRNA analysis is a step in the right direction, not all mRNAs are translated into proteins. **Proteomics** is the study of proteomes' function. Proteomics complements genomics and is useful when scientists want to test their hypotheses that they based on genes. Even though all multicellular organisms' cells have the same set of genes, the set of proteins produced in different tissues is different and dependent on gene expression. Thus, the genome is constant, but the proteome varies and is dynamic within an organism. In addition, RNAs can be alternately spliced (cut and pasted to create novel combinations and novel proteins) and many proteins modify themselves after translation by processes such as proteolytic cleavage, phosphorylation, glycosylation, and ubiquitination. There are also protein-protein interactions, which complicate studying proteomes. Although the genome provides a blueprint, the final architecture depends on several factors that can change the progression of events that generate the proteome.

Metabolomics is related to genomics and proteomics. **Metabolomics** involves studying small molecule metabolites in an organism. The **metabolome** is the complete set of metabolites that are related to an organism's genetic makeup. Metabolomics offers an opportunity to compare genetic makeup and physical characteristics, as well as genetic makeup and environmental factors. The goal of metabolome research is to identify, quantify, and catalogue all the metabolites in living organisms' tissues and fluids.

### Basic Techniques in Protein Analysis

The ultimate goal of proteomics is to identify or compare the proteins expressed from a given genome under specific conditions, study the interactions between the proteins, and use the information to predict cell behavior or develop drug targets. Just as scientists analyze the genome using the basic DNA sequencing technique, proteomics requires techniques for protein analysis. The basic technique for protein analysis, analogous to DNA sequencing, is mass spectrometry. Mass spectrometry identifies and determines a molecule's characteristics. Advances in spectrometry have allowed researchers to analyze very small protein samples. X-ray crystallography, for example, enables scientists to determine a protein crystal's three-dimensional structure at atomic resolution. Another protein imaging technique, nuclear magnetic resonance (NMR), uses atoms' magnetic properties to determine the protein's three-dimensional structure in aqueous solution. Scientists have also used protein microarrays to study protein interactions. Large-scale adaptations of the basic two-hybrid screen ([Figure 17.17](#)) have provided the basis for protein microarrays. Scientists use computer software to analyze the vast amount of data for proteomic analysis.

Genomic- and proteomic-scale analyses are part of **systems biology**, which is the study of whole biological systems (genomes and proteomes) based on interactions within the system. The European Bioinformatics Institute and the Human Proteome Organization (HUPO) are developing and establishing effective tools to sort through the enormous pile of systems biology data. Because proteins are the direct products of genes and reflect activity at the genomic level, it is natural to use proteomes to compare the protein profiles of different cells to identify proteins and genes involved in disease processes. Most pharmaceutical drug trials target proteins. Researchers use information that they obtain from proteomics to identify novel drugs and to understand their mechanisms of action.



**Figure 17.17** Scientists use two-hybrid screening to determine whether two proteins interact. In this method, a transcription factor splits into a DNA-binding domain (BD) and an activator domain (AD). The binding domain is able to bind the promoter in the activator domain's absence, but it does not turn on transcription. The bait protein attaches to the BD, and the prey protein attaches to the AD. Transcription occurs only if the prey “catches” the bait.

Scientists are challenged when implementing proteomic analysis because it is difficult to detect small protein quantities. Although mass spectrometry is good for detecting small protein amounts, variations in protein expression in diseased states can be difficult to discern. Proteins are naturally unstable molecules, which makes proteomic analysis much more difficult than genomic analysis.

## Cancer Proteomics

Researchers are studying patients' genomes and proteomes to understand the genetic basis of diseases. The most prominent disease researchers are studying with proteomic approaches is cancer. These approaches improve screening and early cancer detection. Researchers are able to identify proteins whose expression indicates the disease process. An individual protein is a **biomarker**; whereas, a set of proteins with altered expression levels is a **protein signature**. For a biomarker or protein signature to be useful as a candidate for early cancer screening and detection, they must secrete in body fluids, such as sweat, blood, or urine, such that health professionals can perform large-scale screenings in a noninvasive fashion. The current problem with using biomarkers for early cancer detection is the high rate of false-negative results. A **false negative** is an incorrect test result that should have been positive. In other words, many cancer cases go undetected, which makes biomarkers unreliable. Some examples of protein biomarkers in cancer detection are CA-125 for ovarian cancer and PSA for prostate cancer. Protein signatures may be more reliable than biomarkers to detect cancer cells. Researchers are also using proteomics to develop individualized treatment plans, which involves predicting whether or not an individual will respond to specific drugs and the side effects that the individual may experience. Researchers also use proteomics to predict the possibility of disease recurrence.

The National Cancer Institute has developed programs to improve cancer detection and treatment. The Clinical Proteomic Technologies for Cancer and the Early Detection Research Network are efforts to identify protein signatures specific to different cancer types. The Biomedical Proteomics Program identifies protein signatures and designs effective therapies for cancer patients.