17.3B: Use of Whole-Genome Sequences of Model Organisms

Sequencing genomes of model organisms allows scientists to study homologous proteins in more complex eukaryotes, such as humans.

Learning Objectives

• Describe the model organisms used in whole-genome sequencing

Key Points

• The first genome to be completely sequenced was the bacterial virus, bacteriophage fx174, which is 5368 base pairs.
• Scientists utilize genome sequencing from model organisms to study homologous proteins and establish evolutionary relationships.
• Genome annotation is the process of attaching biological information to gene sequences identified using whole-genome sequencing.
• Model organisms include the fruit fly (Drosophila melanogaster), brewers yeast (Saccharomyces cerevisiae), the nematode, Caenorhabditis elegans, and the mouse (Mus musculus).

Key Terms

• genome annotation: the process of attaching biological information to gene sequences.
• model organism: any organism (e.g. the fruit fly) that has been extensively studied as an example of many others and from which general principles may be established
Use of Whole-Genome Sequences of Model Organisms

The first genome to be completely sequenced was of a bacterial virus, the bacteriophage fx174 (5368 base pairs). This was accomplished by Fred Sanger using shotgun sequencing. Several other organelle and viral genomes were later sequenced. The first organism whose genome was sequenced was the bacterium Haemophilus influenzae, which was accomplished by Craig Venter in the 1980s. Approximately 74 different laboratories collaborated on the sequencing of the genome of the yeast Saccharomyces cerevisiae, which began in 1989 and was completed in 1996. It took this long because it was 60 times bigger than any other genome that had been sequenced at that point. By 1997, the genome sequences of two important model organisms were available: the bacterium Escherichia coli K12 and the yeast Saccharomyces cerevisiae. Genomes of other model organisms, such as the mouse Mus musculus, the fruit fly Drosophila melanogaster, the nematode Caenorhabditis elegans, and the human Homo sapiens are now known. Much basic research is performed using model organisms because the information can be applied to the biological processes of genetically-similar organisms. Having entire genomes sequenced aids these research efforts.

The process of attaching biological information to gene sequences is called genome annotation. Annotation aids researchers doing basic experiments in molecular biology, such as designing PCR primers and RNA targets.

Sequencing genomes allows scientists to identify homologous proteins and establish evolutionary relationships. Furthermore, if a newly-discovered protein is homologous to a known protein, through homology, scientists can make an educated guess as to how the new protein functions.

Eukaryotes are organisms containing cells that enclose complex organelles within a well-defined cell membrane. The defining characteristic that sets eukaryotes and prokaryotes apart is the eukaryotes' nucleus, or nuclear envelope, in which an organism's genetic information is contained. The first eukaryotic genome to be sequenced was that of S. cerevisiae, which is the yeast used in baking and brewing. It is the most-studied eukaryotic model organism in molecular and cell biology, similar to E. coli's role in the study of prokaryotic organisms. Research on many proteins that are important to humans is done by examining their homologs in yeasts. For example, signaling proteins and protein-processing enzymes were discovered through the help of yeast genome.
Figure 4.1: *S. cerevisiae: A Model Organism for Eukaryotes*: Saccharomyces cerevisiae, a yeast, is used as a model organism for studying signaling proteins and protein-processing enzymes which have homologs in humans.