

Researchers used the natural transfer of DNA from *Agrobacterium* to a plant host to introduce DNA fragments of their choice into plant hosts. In nature, the disease-causing *A. tumefaciens* have a set of plasmids, **Ti plasmids** (tumor-inducing plasmids), that contain genes to produce tumors in plants. DNA from the Ti plasmid integrates into the infected plant cell's genome. Researchers manipulate the Ti plasmids to remove the tumor-causing genes and insert the desired DNA fragment for transfer into the plant genome. The Ti plasmids carry antibiotic resistance genes to aid selection and researchers can propagate them in *E. coli* cells as well.

The Organic Insecticide *Bacillus thuringiensis*

Bacillus thuringiensis (Bt) is a bacterium that produces protein crystals during sporulation that are toxic to many insect species that affect plants. Insects need to ingest Bt toxin in order to activate the toxin. Insects that have eaten Bt toxin stop feeding on the plants within a few hours. After the toxin activates in the insects' intestines, they die within a couple of days. Modern biotechnology has allowed plants to encode their own crystal Bt toxin that acts against insects. Scientists have cloned the crystal toxin genes from Bt and introduced them into plants. Bt toxin is safe for the environment, nontoxic to humans and other mammals, and organic farmers have approved it as a natural insecticide.

Flavr Savr Tomato

The first GM crop on the market was the Flavr Savr Tomato in 1994. Scientists used antisense RNA technology to slow the softening and rotting process caused by fungal infections, which led to increased shelf life of the GM tomatoes. Additional genetic modification improved the tomato's flavor. The Flavr Savr tomato did not successfully stay in the market because of problems maintaining and shipping the crop.

17.2 Mapping Genomes

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

- Define genomics
- Describe genetic and physical maps
- Describe genomic mapping methods

Genomics is the study of entire genomes, including the complete set of genes, their nucleotide sequence and organization, and their interactions within a species and with other species. **Genome mapping** is the process of finding the locations of genes on each chromosome. The maps that genome mapping create are comparable to the maps that we use to navigate streets. A **genetic map** is an illustration that lists genes and their location on a chromosome. Genetic maps provide the big picture (similar to an interstate highway map) and use genetic markers (similar to landmarks). A **genetic marker** is a gene or sequence on a chromosome that co-segregates (shows genetic linkage) with a specific trait. Early geneticists called this linkage analysis. Physical maps present the intimate details of smaller chromosome regions (similar to a detailed road map). A **physical map** is a representation of the physical distance, in nucleotides, between genes or genetic markers. Both genetic linkage maps and physical maps are required to build a genome's complete picture. Having a complete genome map of the genome makes it easier for researchers to study individual genes. Human genome maps help researchers in their efforts to identify human disease-causing genes related to illnesses like cancer, heart disease, and cystic fibrosis. We can use genome mapping in a variety of other applications, such as using live microbes to clean up pollutants or even prevent pollution. Research involving plant genome mapping may lead to producing higher crop yields or developing plants that better adapt to climate change.

Genetic Maps

The study of genetic maps begins with **linkage analysis**, a procedure that analyzes the recombination frequency between genes to determine if they are linked or show independent assortment. Scientists used the term *linkage* before the discovery of DNA. Early geneticists relied on observing phenotypic changes to understand an organism's genotype. Shortly after Gregor Mendel (the father of modern genetics) proposed that traits were determined by what we now call genes, other researchers observed that different traits were often inherited together, and thereby deduced that the genes were physically linked by their location on the same chromosome. Gene mapping relative to each other based on linkage analysis led to developing the first genetic maps.

Observations that certain traits were always linked and certain others were not linked came from studying the offspring of crosses between parents with different traits. For example, in garden pea experiments, researchers discovered, that the flower's color and plant pollen's shape were linked traits, and therefore the genes encoding these traits were in close proximity on the same chromosome. We call exchanging DNA between homologous chromosome pairs **genetic recombination**, which occurs by crossing over DNA between homologous DNA strands, such as nonsister chromatids. Linkage analysis involves studying the recombination frequency between any two genes. The greater the distance between two genes, the higher the chance that a

recombination event will occur between them, and the higher the recombination frequency between them. [Figure 17.11](#) shows two possibilities for recombination between two nonsister chromatids during meiosis. If the recombination frequency between two genes is less than 50 percent, they are linked.

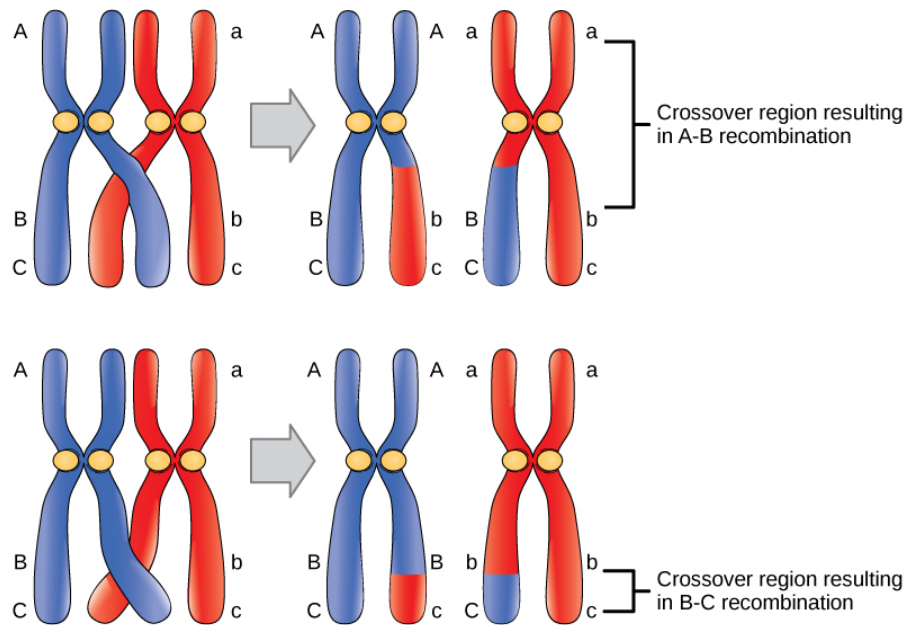


Figure 17.11 Crossover may occur at different locations on the chromosome. Recombination between genes *A* and *B* is more frequent than recombination between genes *B* and *C* because genes *A* and *B* are farther apart. Therefore, a crossover is more likely to occur between them.

The generation of genetic maps requires markers, just as a road map requires landmarks (such as rivers and mountains). Scientists based early genetic maps on using known genes as markers. Scientists now use more sophisticated markers, including those based on non-coding DNA, to compare individuals' genomes in a population. Although individuals of a given species are genetically similar, they are not identical. Every individual has a unique set of traits. These minor differences in the genome between individuals in a population are useful for genetic mapping purposes. In general, a good genetic marker is a region on the chromosome that shows variability or polymorphism (multiple forms) in the population.

Some genetic markers that scientists use in generating genetic maps are **restriction fragment length polymorphisms (RFLP)**, variable number of tandem repeats (VNTRs), **microsatellite polymorphisms**, and the **single nucleotide polymorphisms (SNPs)**. We can detect RFLPs (sometimes pronounced “rif-lips”) when the DNA of an individual is cut with a restriction endonuclease that recognizes specific sequences in the DNA to generate a series of DNA fragments, which we can then analyze using gel electrophoresis. Every individual's DNA will give rise to a unique pattern of bands when cut with a particular set of restriction endonucleases. Scientists sometimes refer to this as an individual's DNA “fingerprint.” Certain chromosome regions that are subject to polymorphism will lead to generating the unique banding pattern. VNTRs are repeated sets of nucleotides present in DNA's non-coding regions. Non-coding, or “junk,” DNA has no known biological function; however, research shows that much of this DNA is actually transcribed. While its function is uncertain, it is certainly active, and it may be involved in regulating coding genes. The number of repeats may vary in a population's individual organisms. Microsatellite polymorphisms are similar to VNTRs, but the repeat unit is very small. SNPs are variations in a single nucleotide.

Because genetic maps rely completely on the natural process of recombination, natural increases or decreases in the recombination level given genome area affects mapping. Some parts of the genome are recombination hotspots; whereas, others do not show a propensity for recombination. For this reason, it is important to look at mapping information developed by multiple methods.

Physical Maps

A physical map provides detail of the actual physical distance between genetic markers, as well as the number of nucleotides. There are three methods scientists use to create a physical map: cytogenetic mapping, radiation hybrid mapping, and sequence mapping. **Cytogenetic mapping** uses information from microscopic analysis of stained chromosome sections ([Figure 17.12](#)). It is

possible to determine the approximate distance between genetic markers using cytogenetic mapping, but not the exact distance (number of base pairs). **Radiation hybrid mapping** uses radiation, such as x-rays, to break the DNA into fragments. We can adjust the radiation amount to create smaller or larger fragments. This technique overcomes the limitation of genetic mapping, and we can adjust the radiation so that increased or decreased recombination frequency does not affect it. **Sequence mapping** resulted from DNA sequencing technology that allowed for creating detailed physical maps with distances measured in terms of the number of base pairs. Creating **genomic libraries** and **complementary DNA (cDNA) libraries** (collections of cloned sequences or all DNA from a genome) has sped the physical mapping process. A genetic site that scientists use to generate a physical map with sequencing technology (a sequence-tagged site, or STS) is a unique sequence in the genome with a known exact chromosomal location. An **expressed sequence tag (EST)** and a single sequence length polymorphism (SSLP) are common STSs. An EST is a short STS that we can identify with cDNA libraries, while we obtain SSLPs from known genetic markers, which provide a link between genetic and physical maps.

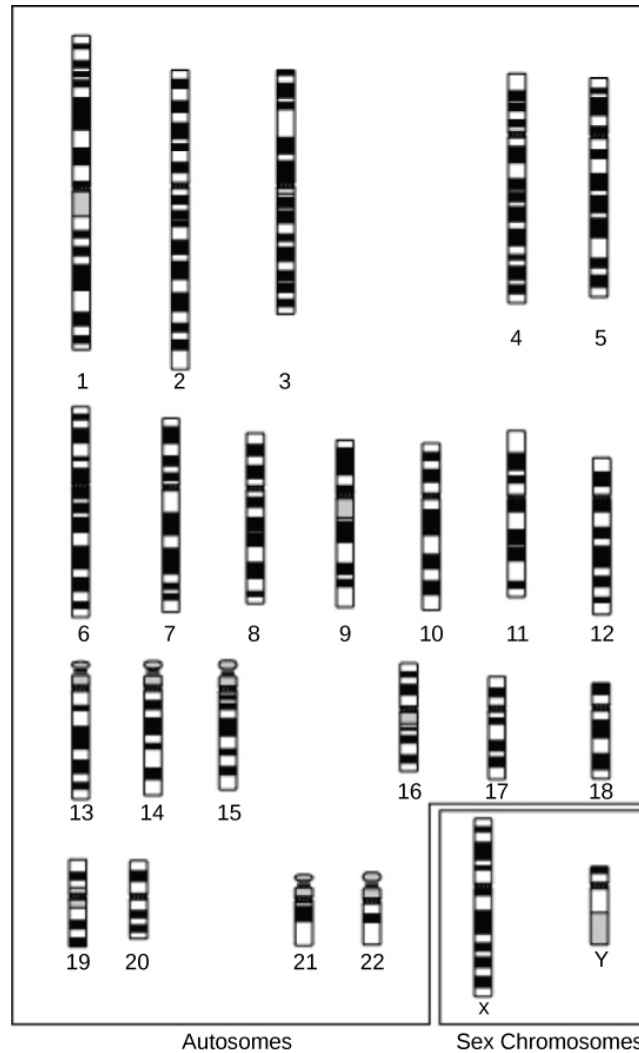


Figure 17.12 A cytogenetic map shows the appearance of a chromosome after scientists stain and exam it under a microscope. (credit: National Human Genome Research Institute)

Genetic and Physical Maps Integration

Genetic maps provide the outline and physical maps provide the details. It is easy to understand why both genome mapping technique types are important to show the big picture. Scientists use information from each technique in combination to study the genome. Scientists are using genomic mapping with different model organisms for research. Genome mapping is still an ongoing process, and as researchers develop more advanced techniques, they expect more breakthroughs. Genome mapping is similar to completing a complicated puzzle using every piece of available data. Mapping information generated in laboratories all over the world goes into central databases, such as GenBank at the National Center for Biotechnology Information (NCBI).

Researchers are making efforts for the information to be more easily accessible to other researchers and the general public. Just as we use global positioning systems instead of paper maps to navigate through roadways, NCBI has created a genome viewer tool to simplify the data-mining process.



SCIENTIFIC METHOD CONNECTION

How to Use a Genome Map Viewer

Problem statement: Do the human, macaque, and mouse genomes contain common DNA sequences?

Develop a hypothesis.

Go to [this website \(https://openstax.org/l/GeneINSR\)](https://openstax.org/l/GeneINSR) to test the hypothesis.

The web page displays the comparison of the gene sequences of many organisms to the Human Insulin Receptor gene. Explore the type of information provided, select the groups of organisms needed for testing of the hypothesis from the top portion of the displayed data. Focus the attention to the bottom part, the Selected Orthologues. Explore which columns are relevant to the needed information.

On the same page, there are other options to explore, not all are necessary for the task, however it might give more insight to the value of genome/gene comparisons.



LINK TO LEARNING

Online Mendelian Inheritance in Man (OMIM) is a searchable online catalog of human genes and genetic disorders. This website shows genome mapping information, and also details the history and research of each trait and disorder. Click this [link \(http://openstax.org/l/OMIM\)](http://openstax.org/l/OMIM) to search for traits (such as handedness) and genetic disorders (such as diabetes).

17.3 Whole-Genome Sequencing

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

- Describe three types of sequencing
- Define whole-genome sequencing

Although there have been significant advances in the medical sciences in recent years, doctors are still confounded by some diseases, and they are using whole-genome sequencing to discover the root of the problem. **Whole-genome sequencing** is a process that determines an entire genome's DNA sequence. Whole-genome sequencing is a brute-force approach to problem solving when there is a genetic basis at the core of a disease. Several laboratories now provide services to sequence, analyze, and interpret entire genomes.

For example, whole-exome sequencing is a lower-cost alternative to whole genome sequencing. In exome sequencing, the doctor sequences only the DNA's coding, exon-producing regions. In 2010, doctors used whole-exome sequencing to save a young boy whose intestines had multiple mysterious abscesses. The child had several colon operations with no relief. Finally, they performed whole-exome sequencing, which revealed a defect in a pathway that controls apoptosis (programmed cell death). The doctors used a bone-marrow transplant to overcome this genetic disorder, leading to a cure for the boy. He was the first person to receive successful treatment based on a whole-exome sequencing diagnosis. Today, human genome sequencing is more readily available and results are available within two days for about \$1000.

Strategies Used in Sequencing Projects

The basic sequencing technique used in all modern day sequencing projects is the chain termination method (also known as the dideoxy method), which Fred Sanger developed in the 1970s. The chain termination method involves DNA replication of a single-stranded template by using a primer and a regular **deoxynucleotide** (dNTP), which is a monomer, or a single DNA unit. The primer and dNTP mix with a small proportion of fluorescently labeled **dideoxynucleotides** (ddNTPs). The ddNTPs are monomers that are missing a hydroxyl group (–OH) at the site at which another nucleotide usually attaches to form a chain ([Figure 17.13](#)). Scientists label each ddNTP with a different color of fluorophore. Every time a ddNTP incorporates in the growing complementary strand, it terminates the DNA replication process, which results in multiple short strands of replicated DNA