Chapter 18

The Genetics of Viruses and Bacteria

Lecture Outline

Overview: Microbial Model Systems

- Viruses and bacteria are the simplest biological systems—microbial models in which scientists find life's fundamental molecular mechanisms in their most basic, accessible forms.
- Molecular biology was born in the laboratories of microbiologists studying viruses and bacteria.
  - Microbes such as E. coli and its viruses are called model systems because of their use in studies that reveal broad biological principles.
  - Microbiologists provided most of the evidence that genes are made of DNA, and they worked out most of the major steps in DNA replication, transcription, and translation.
  - Techniques enabling scientists to manipulate genes and transfer them from one organism to another were developed in microbes.
- In addition, viruses and bacteria have unique genetic features with implications for understanding the diseases that they cause.
- Bacteria are prokaryotic organisms, with cells that are much smaller and more simply organized than those of eukaryotes, such as plants and animals.
- Viruses are smaller and simpler still, lacking the structure and metabolic machinery of cells.
  - Most viruses are little more than aggregates of nucleic acids and protein—genes in a protein coat.

Concept 18.1 A virus has a genome but can reproduce only within a host cell

Researchers discovered viruses by studying a plant disease.

- The story of how viruses were discovered begins in 1883 with research on the cause of tobacco mosaic disease by Adolf Mayer.
  - This disease stunts tobacco plant growth and mottles plant leaves.
  - Mayer concluded that the disease was infectious when he found that he could transmit the disease by rubbing sap from diseased leaves onto healthy plants.
  - He concluded that the disease must be caused by an extremely small bacterium.
  - Ten years later, Dimitri Ivanovsky demonstrated that the sap was still infectious even after passing through a filter designed to remove bacteria.
- In 1897, Martinus Beijerinck ruled out the possibility that the disease was due to a filterable toxin produced by a bacterium by demonstrating that the infectious agent could reproduce.

- The sap from one generation of infected plants could be used to infect a second generation of plants that could infect subsequent generations.
- Beijerinck also determined that the pathogen could reproduce only within the host, could not be cultivated on nutrient media, and was not inactivated by alcohol, generally lethal to bacteria.

- In 1935, Wendell Stanley crystallized the pathogen, the tobacco mosaic virus (TMV).

A virus is a genome enclosed in a protective coat.

- Stanley's discovery that some viruses could be crystallized was puzzling because not even the simplest cells can aggregate into regular crystals.
- However, viruses are not cells.
- They are infectious particles consisting of nucleic acid encased in a protein coat and, in some cases, a membranous envelope.
  - The tiniest viruses are only 20 nm in diameter—smaller than a ribosome.
- The genome of viruses may consist of double-stranded DNA, single-stranded DNA, double-stranded RNA, or single-stranded RNA, depending on the kind of virus.
  - A virus is called a DNA virus or an RNA virus, according to the kind of nucleic acid that makes up its genome.
  - The viral genome is usually organized as a single linear or circular molecule of nucleic acid.
  - The smallest viruses have only four genes, while the largest have several hundred.
- The **capsid** is the protein shell enclosing the viral genome.
- Capsids are built of a large number of protein subunits called *capsomeres.*
  - The number of different *kinds* of proteins making up the capsid is usually small.
  - The capsid of the tobacco mosaic virus has more than 1,000 copies of the same protein.
  - Adenoviruses have 252 identical proteins arranged into a polyhedral capsid—as an icosahedron.
- Some viruses have accessory structures to help them infect their hosts.
- A membranous envelope surrounds the capsids of flu viruses.
  - These **viral envelopes** are derived from the membrane of the host cell.
  - They also have some host cell viral proteins and glycoproteins, as well as molecules of viral origin.
  - Some viruses carry a few viral enzyme molecules within their capsids.
- The most complex capsids are found in viruses that infect bacteria, called **bacteriophages** or **phages.**
- The T-even phages (T2, T4, T6) that infect *Escherichia coli* have elongated icosahedral capsid heads that enclose their DNA and a protein tailpiece that attaches the phage to the host and injects the phage DNA inside.

Viruses can reproduce only within a host cell.

- Viruses are obligate intracellular parasites.
- They can reproduce only within a host cell.

- An isolated virus is unable to reproduce—or do anything else, except infect an appropriate host.
- Viruses lack the enzymes for metabolism and the ribosomes for protein synthesis.
- An isolated virus is merely a packaged set of genes in transit from one host cell to another.
- Each type of virus can infect and parasitize only a limited range of host cells, called its **host range.**
  - This host specificity depends on the evolution of recognition systems by the virus.
  - Viruses identify host cells by a "lock and key" fit between proteins on the outside of the virus and specific receptor molecules on the host's surface (which evolved for functions that benefit the host).
- Some viruses have a broad enough host range to infect several species, while others infect only a single species.
  - West Nile virus can infect mosquitoes, birds, horses, and humans.
  - Measles virus can infect only humans.
- Most viruses of eukaryotes attack specific tissues.
  - Human cold viruses infect only the cells lining the upper respiratory tract.
  - The AIDS virus binds only to certain white blood cells.
- A viral infection begins when the genome of the virus enters the host cell.
- Once inside, the viral genome commandeers its host, reprogramming the cell to copy viral nucleic acid and manufacture proteins from the viral genome.
  - The host provides nucleotides, ribosomes, tRNAs, amino acids, ATP, and other components for making the viral components dictated by viral genes.
- Most DNA viruses use the DNA polymerases of the host cell to synthesize new genomes along the templates provided by the viral DNA.
  - RNA viruses use special virus-encoded polymerases that can use RNA as a template.
- The nucleic acid molecules and capsomeres then self-assemble into viral particles and exit the cell.
  - Tobacco mosaic virus RNA and capsomeres can be assembled to form complete viruses if the components are mixed together under the right conditions.
- The simplest type of viral reproductive cycle ends with the exit of many viruses from the infected host cell, a process that usually damages or destroys the host cell.

Phages reproduce using lytic or lysogenic cycles.

- While phages are the best understood of all viruses, some of them are also among the most complex.
- Research on phages led to the discovery that some double-stranded DNA viruses can reproduce by two alternative mechanisms: the lytic cycle and the lysogenic cycle.
- In the **lytic cycle,** the phage reproductive cycle culminates in the death of the host.
  - In the last stage, the bacterium lyses (breaks open) and releases the phages produced within the cell to infect others.
  - Each of these phages can infect a healthy cell.
- **Virulent phages** reproduce only by a lytic cycle.

- While phages have the potential to wipe out a bacterial colony in just hours, bacteria have defenses against phages.
  - Natural selection favors bacterial mutants with receptor sites that are no longer recognized by a particular type of phage.
  - Bacteria produce *restriction endonucleases,* or restriction enzymes, that recognize and cut up foreign DNA, including certain phage DNA.
    - Chemical modifications to the bacteria's own DNA prevent its destruction by restriction nucleases.
  - Natural selection also favors phage mutants that are resistant to restriction enzymes.
- In the **lysogenic** cycle, the phage genome replicates without destroying the host cell.
  - **Temperate phages,** like phage lambda, use both lytic and lysogenic cycles.
- The lambda phage that infects *E. coli* demonstrates the cycles of a temperate phage.
- Infection of an *E. coli* by phage lambda begins when the phage binds to the surface of the cell and injects its DNA.
  - What happens next depends on the reproductive mode: lytic or lysogenic cycle.
- During a lytic cycle, the viral genes turn the host cell into a lambda phage-producing factory, and the cell lyses and releases its viral products.
- During a lysogenic cycle, the viral DNA molecule is incorporated by genetic recombination into a specific site on the host cell's chromosome.
- In this **prophage** stage, one of the viral genes codes for a protein that represses most other prophage genes.
  - As a result, the phage genome is largely silent.
  - A few other prophage genes may also be expressed during lysogenic cycles.
  - Expression of these genes may alter the host's phenotype, which can have medical significance.
- Every time the host divides, it copies the phage DNA and passes the copies to daughter cells.
  - The viruses propagate without killing the host cells on which they depend.
- The term *lysogenic* implies that prophages are capable of giving rise to active phages that lyse their host cells.
- That happens when the viral genome exits the bacterial chromosome and initiates a lytic cycle.

Animal viruses are diverse in their modes of infection and replication.

- Many variations on the basic scheme of viral infection and reproduction are represented among animal viruses.
  - One key variable is the type of nucleic acid that serves as a virus's genetic material.
  - Another variable is the presence or absence of a membranous envelope derived from the host cell membrane.
  - Most animal viruses with RNA genomes have an envelope, as do some with DNA genomes.
- Viruses equipped with an outer envelope use the envelope to enter the host cell.
  - Glycoproteins on the envelope bind to specific receptors on the host's membrane.

- The envelope fuses with the host's membrane, transporting the capsid and viral genome inside.
- The viral genome duplicates and directs the host's protein synthesis machinery to synthesize capsomeres with free ribosomes and glycoproteins with bound ribosomes.
- After the capsid and viral genome self-assemble, they bud from the host cell covered with an envelope derived from the host's plasma membrane, including viral glycoproteins.

- The viral envelope is thus derived from the host's plasma membrane, although viral genes specify some of the molecules in the membrane.
- These enveloped viruses do not necessarily kill the host cell.
- Some viruses have envelopes that are not derived from plasma membrane.
  - The envelope of the herpesvirus is derived from the nuclear envelope of the host.
  - These double-stranded DNA viruses reproduce within the cell nucleus using viral and cellular enzymes to replicate and transcribe their DNA.
  - In some cases, copies of the herpesvirus DNA remain behind as minichromosomes in the nuclei of certain nerve cells.
  - There they remain for life until triggered by physical or emotional stress to leave the genome and initiate active viral production.
  - The infection of other cells by these new viruses causes cold or genital sores.
- The viruses that use RNA as the genetic material are quite diverse, especially those that infect animals.
  - In some with single-stranded RNA (class IV), the genome acts as mRNA and is translated directly.
  - In others (class V), the RNA genome serves as a *template* for complementary RNA strands, which function both as mRNA and as templates for the synthesis of additional copies of genome RNA.
  - All viruses that require RNA → RNA synthesis to make mRNA use a viral enzyme that is packaged with the genome inside the capsid.
- **Retroviruses** (class VI) have the most complicated life cycles.
  - These carry an enzyme called **reverse transcriptase** that transcribes DNA from an RNA template.
    - This provides RNA → DNA information flow.
  - The newly made DNA is inserted as a **provirus** into a chromosome in the animal cell.
  - The host's RNA polymerase transcribes the viral DNA into more RNA molecules.
    - These can function both as mRNA for the synthesis of viral proteins and as genomes for new virus particles released from the cell.
- Human immunodeficiency virus (HIV), the virus that causes AIDS (acquired immunodeficiency syndrome) is a retrovirus.
- The reproductive cycle of HIV illustrates the pattern of infection and replication in a retrovirus.
- The viral particle includes an envelope with glycoproteins for binding to specific types of red blood cells, a capsid containing two identical RNA strands as its genome, and two copies of reverse transcriptase.
- After HIV enters the host cell, reverse transcriptase molecules are released into the cytoplasm and catalyze synthesis of viral DNA.

- The host's polymerase transcribes the proviral DNA into RNA molecules that can function both as mRNA for the synthesis of viral proteins and as genomes for new virus particles released from the cell.
- Transcription produces more copies of the viral RNA that are translated into viral proteins, which self-assemble into a virus particle and leave the host.

Viruses may have evolved from other mobile genetic elements.

- Viruses do not fit our definition of living organisms.
- An isolated virus is biologically inert, and yet it has a genetic program written in the universal language of life.
- Although viruses are obligate intracellular parasites that cannot reproduce independently, it is hard to deny their evolutionary connection to the living world.
- Because viruses depend on cells for their own propagation, it is reasonable to assume that they evolved *after* the first cells appeared.
- Most molecular biologists favor the hypothesis that viruses originated from fragments of cellular nucleic acids that could move from one cell to another.
  - A viral genome usually has more in common with the genome of its host than with those of viruses infecting other hosts.
  - However, some viruses have genetic sequences that are quite similar to seemingly distantly related viruses.
    - This genetic similarity may reflect the persistence of groups of viral genes that were evolutionarily successful during the early evolution of viruses and their eukaryotic host cells.
- Perhaps the earliest viruses were naked bits of nucleic acids that passed between cells via injured cell surfaces.
  - The evolution of capsid genes may have facilitated the infection of undamaged cells.
- Candidates for the original sources of viral genomes include plasmids and transposable elements.
  - Plasmids are small, circular DNA molecules that are separate from chromosomes.
  - Plasmids, found in bacteria and in eukaryote yeast, can replicate independently of the rest of the cell and are occasionally transferred between cells.
  - Transposable elements are DNA segments that can move from one location to another within a cell's genome.
- Both plasmids and transposable elements are mobile genetic elements.
- The ongoing evolutionary relationship between viruses and the genomes of their hosts is an association that makes viruses very useful model systems in molecular biology.

Concept 18.2 Viruses, viroids, and prions are formidable pathogens in animals and plants

- The link between viral infection and the symptoms it produces is often obscure.
  - Some viruses damage or kill cells by triggering the release of hydrolytic enzymes from lysosomes.
  - Some viruses cause the infected cell to produce toxins that lead to disease symptoms.

- Others have molecular components, such as envelope proteins, that are toxic.

- In some cases, viral damage is easily repaired (respiratory epithelium after a cold), but in others, infection causes permanent damage (nerve cells after polio).
- Many of the temporary symptoms associated with a viral infection result from the body's own efforts at defending itself against infection.
- The immune system is a complex and critical part of the body's natural defense mechanism against viral and other infections.
- Modern medicine has developed **vaccines,** harmless variants or derivatives of pathogenic microbes that stimulate the immune system to mount defenses against the actual pathogen.
  - Vaccination has eradicated smallpox.
  - Effective vaccines are available against polio, measles, rubella, mumps, hepatitis B, and a number of other viral diseases.
- Medical technology can do little to cure viral diseases once they occur.
- Antibiotics, which can kill bacteria by inhibiting enzymes or processes specific to bacteria, are powerless against viruses, which have few or no enzymes of their own.
  - Most antiviral drugs resemble nucleosides and interfere with viral nucleic acid synthesis.
  - An example is acyclovir, which impedes herpesvirus reproduction by inhibiting the viral polymerase that synthesizes viral DNA.
  - Azidothymidine (AZT) curbs HIV reproduction by interfering with DNA synthesis by reverse transcriptase.
  - Currently, multidrug "cocktails" are the most effective treatment for HIV.

New viral diseases are emerging.

- In recent years, several *emerging viruses* have risen to prominence.
  - HIV, the AIDS virus, seemed to appear suddenly in the early 1980s.
  - Each year new strains of influenza virus cause millions to miss work or class, and deaths are not uncommon.
  - The deadly Ebola virus has caused *hemorrhagic fevers* in central Africa periodically since 1976.
  - West Nile virus appeared for the first time in North America in 1999.
  - A more recent viral disease is *severe acute respiratory syndrome (SARS).*
    - Researchers identified the disease agent causing SARS as a *coronavirus,* a class IV virus with a single-stranded RNA genome.
- The emergence of these new viral diseases is due to three processes: mutation; spread of existing viruses from one species to another; and dissemination of a viral disease from a small, isolated population.
- Mutation of existing viruses is a major source of new viral diseases.
  - RNA viruses tend to have high mutation rates because replication of their nucleic acid lacks proofreading.
  - Some mutations create new viral strains with sufficient genetic differences from earlier strains that they can infect individuals who had acquired immunity to these earlier strains.
    - This is the case in flu epidemics.
- Another source of new viral diseases is the spread of existing viruses from one host species to another.

- It is estimated that about three-quarters of new human diseases originated in other animals.
  - For example, hantavirus, which killed dozens of people in 1993, normally infects rodents, especially deer mice.
  - In 1993, unusually wet weather in the southwestern United States increased the mice's food, exploding the population.
  - Humans acquired hantavirus when they inhaled dust-containing traces of urine and feces from infected mice.
  - The source of the SARS-causing virus is still undetermined, but candidates include the exotic animal markets in China.
  - In early 2004, the first cases of a new bird flu were reported in southeast Asia.
    - If this disease evolves to spread from person to person, the potential for a major human outbreak is great.
- Finally, a viral disease can spread from a small, isolated population to a widespread epidemic.
  - For example, AIDS went unnamed and virtually unnoticed for decades before spreading around the world.
  - Technological and social factors, including affordable international travel, blood transfusion technology, sexual promiscuity, and the abuse of intravenous drugs allowed a previously rare disease to become a global scourge.
- These emerging viruses are generally not new. Rather, they are existing viruses that mutate, spread to new host species, or expand their host territory.
- Changes in host behavior and environmental changes can increase the viral traffic responsible for emerging disease.
  - Destruction of forests to expand cropland may bring humans into contact with other animals that may host viruses that can infect humans.

Plant viruses are serious agricultural pests.

- More than 2,000 types of viral diseases of plants are known.
  - These diseases account for an annual loss of $15 billion worldwide.
- Plant viruses can stunt plant growth and diminish crop yields.
- Most are RNA viruses with rod-shaped or polyhedral capsids.
- Plant viral diseases are spread by two major routes.
- In *horizontal transmission,* a plant is infected with the virus by an external source.
  - Plants are more susceptible if their protective epidermis is damaged, perhaps by wind, chilling, injury, or insects.
  - Insects are often carriers of viruses, transmitting disease from plant to plant.
- In *vertical transmission,* a plant inherits a viral infection from a parent.
  - This may occur by asexual propagation or in sexual reproduction via infected seeds.
- Once a virus starts reproducing inside a plant cell, viral particles can spread throughout the plant by passing through plasmodesmata.
  - These cytoplasmic connections penetrate the walls between adjacent cells.
  - Proteins encoded by viral genes can alter the diameter of plasmodesmata to allow passage of viral proteins or genomes.

- Agricultural scientists have focused their efforts largely on reducing the incidence and transmission of viral disease and in breeding resistant plant varieties.

Viroids and prions are the simplest infectious agents.

- **Viroids,** smaller and simpler than even viruses, consist of tiny molecules of naked circular RNA that infect plants.
- Their several hundred nucleotides do not encode for proteins but can be replicated by the host's cellular enzymes.
- These small RNA molecules can disrupt plant metabolism and stunt plant growth, perhaps by causing errors in the regulatory systems that control plant growth.
- Viroids show that *molecules* can act as infectious agents to spread disease.
- **Prions** are infectious *proteins* that spread disease.
  - They appear to cause several degenerative brain diseases including scrapie in sheep, "mad cow disease," and Creutzfeldt-Jakob disease in humans.
- Prions are likely transmitted in food.
- They have two alarming characteristics.
  - They are very slow-acting agents. The incubation period is around ten years.
  - Prions are virtually indestructible. They are not destroyed or deactivated by heating to normal cooking temperatures.
- How can a nonreplicating protein be a transmissible pathogen?
- According to the leading hypothesis, a prion is a misfolded form of a normal brain protein.
- When the prion gets into a cell with the normal form of the protein, the prion can convert the normal protein into the prion version, creating a chain reaction that increases their numbers.

Concept 18.3 Rapid reproduction, mutation, and genetic recombination contribute to the genetic diversity of bacteria

- Bacteria are very valuable as microbial models in genetics research.
  - As prokaryotes, bacteria allow researchers to study molecular genetics in simple organisms.
  - With the advent of large-scale genome sequencing, information about many prokaryotes has accumulated.
  - The best-studied bacterium is *Escherichia coli*, "the laboratory rat of molecular biology."
- The major component of the bacterial genome is one double-stranded, circular DNA molecule that is associated with a small amount of protein.
  - For *E. coli*, the chromosomal DNA consists of about 4.6 million nucleotide pairs with about 4,400 genes.
  - This is 100 times more DNA than in a typical virus and 1,000 times less than in a typical eukaryote cell.
  - Tight coiling of DNA results in a dense region of DNA, called the **nucleoid,** which is not bound by a membrane.
- In addition, many bacteria have plasmids, much smaller circles of DNA.
  - Each plasmid has only a small number of genes, from just a few to several dozen.

- Bacterial cells divide by binary fission.
  - This is preceded by replication of the bacterial chromosome from a single origin of replication.
- Bacteria proliferate very rapidly in a favorable natural or laboratory environment.
  - Under optimal laboratory conditions, *E. coli* can divide every 20 minutes, producing a colony of 10^7 to 10^8 bacteria in as little as 12 hours.
  - In the human colon, *E. coli* grows more slowly and can double every 12 hours.
  - It does reproduce rapidly enough to replace the 2×10^{10} bacteria lost each day in feces.
- Through binary fission, most of the bacteria in a colony are genetically identical to the parent cell.
  - However, the spontaneous mutation rate of *E. coli* is 1×10^{-7} mutations per gene per cell division.
  - This produces about 2,000 bacteria per day in the human colon that have a mutation in any one gene.
  - About 9 million mutant *E. coli* are produced in the human gut each day.
- New mutations, though individually rare, can have a significant impact on genetic diversity when reproductive rates are very high because of short generation spans.
- Individual bacteria that are genetically well equipped for the local environment clone themselves more prolifically than do less fit individuals.
- In contrast, organisms with slower reproduction rates (like humans) create genetic variation not by novel alleles produced through *new* mutations, but primarily by sexual recombination of existing alleles.

Genetic recombination produces new bacterial strains.

- In addition to mutation, genetic recombination generates diversity within bacterial populations.
- Here, recombination is defined as the combining of DNA from two individuals into a single genome.
- Bacterial recombination occurs through three processes: transformation, transduction, and conjugation.
- Recombination can be observed when two mutant *E. coli* strains are combined.
  - If each is unable to synthesize one of its required amino acids, neither can grow on a minimal medium.
  - However, if they are combined, numerous colonies will be created that started from cells that acquired the missing genes for amino acid synthesis from the other strain.
  - Some of these capable cells may have resulted from mutation. However, most acquired the missing genes by genetic recombination.
- **Transformation** is the alteration of a bacterial cell's genotype by the uptake of naked, foreign DNA from the surrounding environment.
  - For example, harmless *Streptococcus pneumoniae* bacteria can be transformed to pneumonia-causing cells.
  - This occurs when a live nonpathogenic cell takes up a piece of DNA that happens to include the allele for pathogenicity from dead, broken-open pathogenic cells.

- The foreign allele replaces the native allele in the bacterial chromosome by genetic recombination.
- The resulting cell is now recombinant, with DNA derived from two different cells.

- Years after transformation was discovered in laboratory cultures, most biologists believed that the process was too rare and haphazard to play an important role in natural bacterial populations.
- Researchers have since learned that many bacterial species have surface proteins that are specialized for the uptake of naked DNA.
  - These proteins recognize and transport DNA from closely related bacterial species into the cell, which can then incorporate the foreign DNA into the genome.
  - While *E. coli* lacks this specialized mechanism, it can be induced to take up small pieces of DNA if cultured in a medium with a relatively high concentration of calcium ions.
  - In biotechnology, this technique has been used to introduce foreign DNA into *E. coli*.
- **Transduction** occurs when a phage carries bacterial genes from one host cell to another as a result of aberrations in the phage reproductive cycle.
- In *generalized transduction,* bacterial genes are randomly transferred from one bacterial cell to another.
- Occasionally, a small piece of the host cell's degraded DNA, rather than the phage genome, is packaged within a phage capsid.
  - When this phage attaches to another bacterium, it will inject this foreign DNA into its new host.
  - Some of this DNA can subsequently replace the homologous region of the second cell.
  - This type of transduction transfers bacterial genes at random.
- *Specialized transduction* occurs via a temperate phage.
  - When the prophage viral genome is excised from the chromosome, it sometimes takes with it a small region of adjacent bacterial DNA.
  - These bacterial genes are injected along with the phage's genome into the next host cell.
  - Specialized transduction only transfers those genes near the prophage site on the bacterial chromosome.
- Both generalized and specialized transduction use phage as a vector to transfer genes between bacteria.
- Sometimes known as bacterial "sex," **conjugation** transfers genetic material between two bacterial cells that are temporarily joined.
- The transfer is one-way. One cell ("male") donates DNA and its "mate" ("female") receives the genes.
  - A sex pilus from the male initially joins the two cells and creates a cytoplasmic *mating bridge* between cells.
- "Maleness," the ability to form a sex pilus and donate DNA, results from an **F** (for fertility) **factor** as a section of the bacterial chromosome or as a plasmid.
  - **Plasmids,** including the F plasmid, are small, circular, self-replicating DNA molecules.
- A genetic element that can replicate either as part of the bacterial chromosome or independently of it is called an **episome.**

    - Episomes such as the F plasmid can undergo reversible incorporation into the cell's chromosome.
- Temperate viruses are also episomes.
- Plasmids usually have only a few genes, which are not required for normal survival and reproduction of the bacterium.
    - However, plasmid genes may be advantageous in stressful conditions.
        - The F plasmid facilitates genetic recombination when environmental conditions no longer favor existing strains.
- The F factor or its **F plasmid** consists of about 25 genes, most required for the production of sex pili.
    - Cells with either the F factor or the F plasmid are called F^+ and they pass this condition to their offspring.
    - Cells lacking either form of the F factor, are called F^-, and they function as DNA recipients.
- When an F^+ and F^- cell meet, the F^+ cell passes a copy of the F plasmid to the F^- cell, converting it.
- The plasmid form of the F factor can become integrated into the bacterial chromosome.
- A cell with the F factor built into its chromosome is called an *Hfr* cell (for *H*igh *f*requency of *r*ecombination).
    - Hfr cells function as males during conjugation.
- The Hfr cell initiates DNA replication at a point on the F factor DNA and begins to transfer the DNA copy from that point to its F^- partner.
- Random movements almost always disrupt conjugation long before an entire copy of the Hfr chromosome can be passed to the F^- cell.
- In the partially diploid cell, the newly acquired DNA aligns with the homologous region of the F^- chromosome.
- Recombination exchanges segments of DNA.
- The resulting recombinant bacterium has genes from two different cells.
- In the 1950s, Japanese physicians began to notice that some bacterial strains had evolved antibiotic resistance.
    - Mutations may reduce the ability of the pathogen's cell-surface proteins to transport antibiotics into the bacterial cell.
    - Some of these genes code for enzymes that specifically destroy certain antibiotics, like tetracycline or ampicillin.
- The genes conferring resistance are carried by plasmids, specifically the **R plasmid** (*R* for resistance).
- When a bacterial population is exposed to an antibiotic, individuals with the R plasmid will survive and increase in the overall population.
- Because R plasmids also have genes that encode for sex pili, they can be transferred from one cell to another by conjugation.
- The DNA of a single cell can also undergo recombination due to movement of **transposable genetic elements** or **transposable elements** within the cell's genome.

- Unlike plasmids or prophages, transposable elements never exist independently but are always part of chromosomal or plasmid DNA.
  - During transposition, the transposable element moves from one location to another in a cell's genome.
  - In bacteria, the movement may be within the chromosome, from a plasmid to a chromosome (or vice versa), or between plasmids.
  - Transposable elements may move by a "copy and paste" mechanism, in which the transposable element replicates at its original site, and the copy inserts elsewhere.
  - In other words, the transposable element is added at a new site without being lost from the old site.
- Most transposable elements can move to many alternative locations in the DNA, potentially moving genes to a site where genes of that sort have never before existed.
- The simplest transposable elements, called **insertion sequences,** exist only in bacteria.
- An insertion sequence contains a single gene that codes for transposase, an enzyme that catalyzes movement of the insertion sequence from one site to another within the genome.
- The insertion sequence consists of the transposase gene, flanked by a pair of *inverted repeat* sequences.
  - The 20 to 40 nucleotides of the inverted repeat on one side are repeated in reverse along the opposite DNA strand at the other end of the transposable element.
- The transposase enzyme recognizes the inverted repeats as the edges of the transposable element.
- Transposase cuts the transposable elements from its initial site and inserts it into the target site.
- Insertion sequences cause mutations when they happen to land within the coding sequence of a gene or within a DNA region that regulates gene expression.
- Insertion sequences account for 1.5% of the *E. coli* genome, but a mutation in a particular gene by transposition is rare, occurring about once in every 10 million generations.
  - This is about the same rate as spontaneous mutations from external factors.
- Transposable elements longer and more complex than insertion sequences, called **transposons,** also move about in the bacterial genome.
- In addition to the DNA required for transposition, transposons include extra genes that "go along for the ride," such as genes for antibiotic resistance.
- In some bacterial transposons, the extra genes are sandwiched between two insertion sequences.
- While insertion sequences may not benefit bacteria in any specific way, transposons may help bacteria adapt to new environments.
  - For example, a single R plasmid may carry several genes for resistance to different antibiotics.
  - This is explained by transposons, which can add a gene for antibiotic resistance to a plasmid already carrying genes for resistance to other antibiotics.
  - The transmission of this composite plasmid to other bacterial cells by cell division or conjugation can spread resistance to a variety of antibiotics throughout a bacterial population.

- In an antibiotic-rich environment, natural selection factors bacterial clones that have built up R plasmids with multiple antibiotic resistance through a series of transpositions.

- Transposable elements are also important components of eukaryotic genomes.

Concept 18.4 Individual bacteria respond to environmental change by regulating their gene expression

- An individual bacterium, locked into the genome that it has inherited, can cope with environmental fluctuations by exerting metabolic control.
  - First, cells can vary the number of specific enzyme molecules they make by regulating gene expression.
  - Second, cells can adjust the activity of enzymes already present (for example, by *feedback inhibition*).
- The tryptophan biosynthesis pathway demonstrates both levels of control.
  - If tryptophan levels are high, some of the tryptophan molecules can inhibit the first enzyme in the pathway.
  - If the abundance of tryptophan continues, the cell can stop synthesizing additional enzymes in this pathway by blocking transcription of the genes for these enzymes.
- The basic mechanism for this control of gene expression in bacteria, the *operon model,* was discovered in 1961 by François Jacob and Jacques Monod.
- *E. coli* synthesizes tryptophan from a precursor molecule in a series of steps, with each reaction catalyzed by a specific enzyme.
  - The five genes coding for these enzymes are clustered together on the bacterial chromosome, served by a single promoter.
  - Transcription gives rise to one long mRNA molecule that codes for all five enzymes in the tryptophan pathway.
  - The mRNA is punctuated with start and stop codons that signal where the coding sequence for each polypeptide begins and ends.
- A key advantage of grouping genes of related functions into one transcription unit is that a single "on-off switch" can control a cluster of functionally related genes.
- When an *E. coli* cell must make tryptophan for itself, all the enzymes are synthesized at one time.
- The switch is a segment of DNA called an **operator.**
- The operator, located between the promoter and the enzyme-coding genes, controls the access of RNA polymerase to the genes.
- The operator, the promoter, and the genes they control constitute an **operon.**
- By itself, an operon is on and RNA polymerase can bind to the promoter and transcribe the genes.
- However, if a **repressor** protein, a product of a **regulatory gene,** binds to the operator, it can prevent transcription of the operon's genes.
  - Each repressor protein recognizes and binds only to the operator of a certain operon.
  - Regulatory genes are transcribed continuously at low rates.
- Binding by the repressor to the operator is reversible.

- The number of active repressor molecules available determines the on or off mode of the operator.

- Repressors contain allosteric sites that change shape depending on the binding of other molecules.
  - In the case of the *trp,* or tryptophan, operon, when concentrations of tryptophan in the cell are high, some tryptophan molecules bind as a **corepressor** to the repressor protein.
  - This activates the repressor and turns the operon off.
  - At low levels of tryptophan, most of the repressors are inactive, and the operon is transcribed.
- The *trp* operon is an example of a *repressible* operon, one that is *inhibited* when a specific small molecule binds allosterically to a regulatory protein.
- In contrast, an *inducible* operon is *stimulated* when a specific small molecule interacts with a regulatory protein.
  - In inducible operons, the regulatory protein is active (inhibitory) as synthesized, and the operon is off.
  - Allosteric binding by an **inducer** molecule makes the regulatory protein inactive, and the operon is turned on.
- The *lac* operon contains a series of genes that code for enzymes that play a major role in the hydrolysis and metabolism of lactose (milk sugar).
  - In the absence of lactose, this operon is off, as an active repressor binds to the operator and prevents transcription.
- Lactose metabolism begins with hydrolysis of lactose into its component monosaccharides, glucose and galactose.
- This reaction is catalyzed by the enzyme ß-galactosidase.
  - Only a few molecules of this enzyme are present in an *E. coli* cell grown in the absence of lactose.
  - If lactose is added to the bacterium's environment, the number of ß-galactosidase increases by a thousandfold within 15 minutes.
- The gene for ß-galactosidase is part of the *lac* operon, which includes two other genes coding for enzymes that function in lactose metabolism.
- The regulatory gene, *lacI*, located outside the operon, codes for an allosteric repressor protein that can switch off the *lac* operon by binding to the operator.
- Unlike the *trp* operon, the *lac* repressor is active all by itself, binding to the operator and switching the *lac* operon off.
  - An **inducer** *inactivates* the repressor.
- When lactose is present in the cell, allolactose, an isomer of lactose, binds to the repressor.
  - This inactivates the repressor, and the *lac* operon can be transcribed.
- Repressible enzymes generally function in anabolic pathways, synthesizing end products from raw materials.
  - When the end product is present in sufficient quantities, the cell can allocate its resources to other uses.
- Inducible enzymes usually function in catabolic pathways, digesting nutrients to simpler molecules.

- By producing the appropriate enzymes only when the nutrient is available, the cell avoids making proteins that have nothing to do.

- Both repressible and inducible operons demonstrate *negative* control because active repressors switch off the active form of the repressor protein.
- Positive gene control occurs when an activator molecule interacts directly with the genome to switch transcription on.
- Even if the *lac* operon is turned on by the presence of allolactose, the degree of transcription depends on the concentrations of other substrates.
  - If glucose levels are low, then **cyclic AMP (cAMP)** accumulates.
- The regulatory protein *catabolite activator protein (CAP)* is an **activator** of transcription.
  - When cAMP is abundant, it binds to CAP, and the regulatory protein assumes its active shape and can bind to a specific site at the upstream end of the *lac* promoter.
- The attachment of CAP to the promoter directly stimulates gene expression.
- Thus, this mechanism qualifies as positive regulation.
- The cellular metabolism is biased toward the use of glucose.
- If glucose levels are sufficient and cAMP levels are low (lots of ATP), then the CAP protein has an inactive shape and cannot bind upstream of the *lac* promoter.
  - The *lac* operon will be transcribed but at a low level.
- For the *lac* operon, the presence/absence of lactose (allolactose) determines if the operon is on or off.
- Overall energy levels in the cell determine the level of transcription, a "volume" control, through CAP.
- CAP works on several operons that encode enzymes used in catabolic pathways.
  - If glucose is present and CAP is inactive, then the synthesis of enzymes that catabolize other compounds is slowed.
  - If glucose levels are low and CAP is active, then the genes that produce enzymes that catabolize whichever other fuel is present will be transcribed at high levels.