The Genetics of Viruses and Bacteria

I. A virus is a genome enclosed in a protective coat.
   A. 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.
   B. 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.
   C. The smallest viruses have only four genes, while the largest have several hundred.
   D. 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.
   E. A membranous envelope surrounds the capsids of some viruses.
      1. These **viral envelopes** are derived from the membrane of the host cell.
      2. They also have some host cell viral proteins and glycoproteins, as well as molecules of viral origin.

II. Viruses can reproduce only within a host cell.
   A. An isolated virus is merely a packaged set of genes in transit from one host cell to another.
   B. Each type of virus can infect and parasitize only a limited range of host cells, called its **host range.**
   C. 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).
   D. Some viruses have a broad enough host range to infect several species, while others infect only a single species. Most viruses of eukaryotes attack specific tissues.
   E. A viral infection begins when the genome of the virus enters the host cell.
      1. Once inside, the viral genome commandeers its host, reprogramming the cell to copy viral nucleic acid and manufacture proteins from the viral genome.
      2. The host provides nucleotides, ribosomes, tRNAs, amino acids, ATP, and other components for making the viral components dictated by viral genes.
      3. Most DNA viruses use the DNA polymerases of the host cell to synthesize new genomes along the templates provided by the viral DNA.
      4. RNA viruses use special virus-encoded polymerases that can use RNA as a template.
   F. The nucleic acid molecules and capsomeres then self-assemble into viral particles and exit the cell.
   G. 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.

III. Phages reproduce using lytic or lysogenic cycles.
   A. In the **lytic cycle,** the phage reproductive cycle culminates in the death of the host.
      1. In the last stage, the bacterium lyses (breaks open) and releases the phages produced within the cell to infect others.
      2. Each of these phages can infect a healthy cell.
      3. **Virulent phages** reproduce only by a lytic cycle.
      4. While phages have the potential to wipe out a bacterial colony in just hours, bacteria have defenses against phages.
         a. Natural selection favors bacterial mutants with receptor sites that are no longer recognized by a particular type of phage.

b. Bacteria produce *restriction endonucleases,* or restriction enzymes, that recognize and cut up foreign DNA, including certain phage DNA.
c. Chemical modifications to the bacteria's own DNA prevent its destruction by restriction nucleases.
d. Natural selection also favors phage mutants that are resistant to restriction enzymes.

B. In the **lysogenic** cycle, the phage genome replicates without destroying the host cell.
1. **Temperate phages,** like phage lambda, use both lytic and lysogenic cycles.
a. The lambda phage that infects *E. coli* demonstrates the cycles of a temperate phage.
b. Infection of an *E. coli* by phage lambda begins when the phage binds to the surface of the cell and injects its DNA.
c. What happens next depends on the reproductive mode: lytic or lysogenic cycle.
(1) 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.
(2) During a lysogenic cycle, the viral DNA molecule is incorporated by genetic recombination into a specific site on the host cell's chromosome.
(a) 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.
(b) A few other prophage genes may also be expressed during lysogenic cycles.
(c) Every time the host divides, it copies the phage DNA and passes the copies to daughter cells so the viruses propagate without killing the host cells on which they depend.
(d) The term *lysogenic* implies that prophages are capable of giving rise to active phages that lyse their host cells.
(e) That happens when the viral genome exits the bacterial chromosome and initiates a lytic cycle.

IV. Animal viruses are diverse in their modes of infection and replication.
A. One key variable is the type of nucleic acid that serves as a virus's genetic material.
B. Another variable is the presence or absence of a membranous envelope derived from the host cell membrane.
C. Viruses equipped with an outer envelope use the envelope to enter the host cell.
1. Glycoproteins on the envelope bind to specific receptors on the host's membrane.
2. The envelope fuses with the host's membrane, transporting the capsid and viral genome inside.
3. The viral genome duplicates and directs the host's protein synthesis machinery to synthesize capsomeres with free ribosomes and glycoproteins with bound ribosomes.
4. 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.
5. The viral envelope is thus derived from the host's plasma membrane, although viral genes specify some of the molecules in the membrane.
6. These enveloped viruses do not necessarily kill the host cell.
D. The viruses that use RNA as the genetic material are quite diverse, especially those that

infect animals.

1. In some with single-stranded RNA, the genome acts as mRNA and is translated directly.
2. In others, 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.
3. All viruses that require RNA → RNA synthesis to make mRNA use a viral enzyme that is packaged with the genome inside the capsid.
4. **Retroviruses** have the most complicated life cycles.
    a. These carry an enzyme called **reverse transcriptase** that transcribes DNA from an RNA template.
    b. This provides RNA → DNA information flow.
    c. The newly made DNA is inserted as a **provirus** into a chromosome in the animal cell.
    d. The host's RNA polymerase transcribes the viral DNA into more RNA molecules.
    e. These can function both as mRNA for the synthesis of viral proteins and as genomes for new virus particles released from the cell.
        (1) Human immunodeficiency virus (HIV), the virus that causes AIDS (acquired immunodeficiency syndrome) is a retrovirus.
        (2) 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.
        (3) After HIV enters the host cell, reverse transcriptase molecules are released into the cytoplasm and catalyze synthesis of proviral DNA.
        (4) 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.

V. Viruses may have evolved from other mobile genetic elements.

A. Viruses do not fit our definition of living organisms.

B. 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.

C. In general, a viral genome usually has more in common with the genome of its host than with those of viruses infecting other hosts.

D. 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.

1. Perhaps the earliest viruses were naked bits of nucleic acids that passed between cells via injured cell surfaces.
2. The evolution of capsid genes may have facilitated the infection of undamaged cells.
3. Candidates for the original sources of viral genomes include mobile genetic elements such as plasmids and transposable elements.
    a. Plasmids are small, circular DNA molecules that are separate from chromosomes.
    b. Plasmids, found in bacteria and in eukaryote yeast, can replicate independently of the rest of the cell and are occasionally transferred

between cells.

c. Transposable elements are DNA segments that can move from one location to another within a cell's genome.

VI. Viruses are formidable pathogens in animals and plants.

A. Some viruses damage or kill cells by triggering the release of hydrolytic enzymes from lysosomes.

B. Some viruses cause the infected cell to produce toxins that lead to disease symptoms.

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

D. Many of the temporary symptoms associated with a viral infection result from the body's own efforts at defending itself against infection.

E. Modern medicine has developed **vaccines,** harmless variants or derivatives of pathogenic microbes that stimulate the immune system to mount defenses against the actual pathogen.

1. Medical technology can do little to cure viral diseases once they occur.
2. 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.
3. Most antiviral drugs resemble nucleosides and interfere with viral nucleic acid synthesis.

VII. 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.

A. Mutation of existing viruses is a major source of new viral diseases.

1. RNA viruses tend to have high mutation rates because replication of their nucleic acid lacks proofreading.
2. 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.

B. Another source of new viral diseases is the spread of existing viruses from one host species to another.

C. Finally, a viral disease can spread from a small, isolated population to a widespread epidemic.

1. Technological and social factors, including affordable international travel, blood transfusion technology, sexual promiscuity, and the abuse of intravenous drugs can allow a previously rare disease to become global.

D. These emerging viruses are generally not new. Rather, they are existing viruses that mutate, spread to new host species, or expand their host territory.

VIII. Viroids and prions are the simplest infectious agents.

A. **Viroids,** smaller and simpler than even viruses, consist of tiny molecules of naked circular RNA that infect plants.

1. Their several hundred nucleotides do not encode for proteins but can be replicated by the host's cellular enzymes.
2. These small RNA molecules can disrupt plant metabolism and stunt plant growth, perhaps by causing errors in the regulatory systems that control plant growth.
3. Viroids show that *molecules* can act as infectious agents to spread disease.

B. **Prions** are infectious *proteins* that spread disease.

1. They appear to cause several degenerative brain diseases including scrapie in sheep, "mad cow disease," and Creutzfeldt-Jakob disease in humans.
2. Prions are likely transmitted in food.
3. They have two alarming characteristics.

a. They are very slow-acting agents. The incubation period is around ten

years.
b. Prions are virtually indestructible. They are not destroyed or deactivated by heating to normal cooking temperatures.

4. How can a nonreplicating protein be a transmissible pathogen?
a. According to the leading hypothesis, a prion is a misfolded form of a normal brain protein.
b. 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.

IX. The short generation span of bacteria helps them adapt to changing environments.
A. The major component of the bacterial genome is one double-stranded, circular DNA molecule that is associated with a small amount of protein.
B. 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.
C. Bacterial cells divide by binary fission which is preceded by replication of the bacterial chromosome from a single origin of replication.
D. Through binary fission, most of the bacteria in a colony are genetically identical to the parent cell.
1. However, the spontaneous mutation rate of *E. coli* is 1×10^{7} mutations per gene per cell division.
2. This produces about 2,000 bacteria per day in the human colon that have a mutation in any one gene.
3. About 9 million mutant *E. coli* are produced in the human gut each day.
4. New mutations, though individually rare, can have a significant impact on genetic diversity when reproductive rates are very high because of short generation spans.
5. 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.

X. In addition to mutation, genetic recombination generates diversity within bacterial populations.
A. Here, recombination is defined as the combining of DNA from two individuals into a single genome.
B. Bacterial recombination occurs through three processes: transformation, transduction, and conjugation.
1. **Transformation** is the alteration of a bacterial cell's genotype by the uptake of naked, foreign DNA from the surrounding environment.
a. Many bacterial species have surface proteins that are specialized for the uptake of naked DNA.
b. These proteins recognize and transport DNA from closely related bacterial species into the cell, which can then incorporate the foreign DNA into the genome.
2. **Transduction** occurs when a phage carries bacterial genes from one host cell to another as a result of aberrations in the phage reproductive cycle.
a. In transduction, bacterial genes are randomly transferred from one bacterial cell to another.
(1) Occasionally, a small piece of the host cell's degraded DNA, rather than the phage genome, is packaged within a phage capsid.
(2) When this phage attaches to another bacterium, it will inject this foreign DNA into its new host.
(3) This type of transduction transfers bacterial genes at random.
(4) This can also occurs when a prophage viral genome is excised from

the chromosome, taking with it a small region of adjacent bacterial DNA.

3. Sometimes known as bacterial "sex," **conjugation** transfers genetic material between two bacterial cells that are temporarily joined.
   a. The transfer is one-way. One cell ("male") donates DNA and its "mate" ("female") receives the genes.
   b. A sex pilus from the male initially joins the two cells and creates a cytoplasmic *mating bridge* between cells.
   c. "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.
      (1) 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.
      (2) When an F^+ and F^- cell meet, the F^+ cell passes a copy of the F plasmid to the F^- cell, converting it.
4. 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.
   a. 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.
   b. 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.
   c. 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.
   d. The simplest transposable elements, called **insertion sequences,** exist only in bacteria.
      (1) 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.
      (2) 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.
5. Transposable elements longer and more complex than insertion sequences, called **transposons,** also move about in the bacterial genome.
   a. In addition to the DNA required for transposition, transposons include extra genes that "go along for the ride," such as genes for antibiotic resistance.
   b. Transposons may help bacteria adapt to new environments.
      (1) For example, a single R plasmid may carry several genes for resistance to different antibiotics.
      (2) This is explained by transposons, which can add a gene for antibiotic resistance to a plasmid already carrying genes for resistance to other antibiotics.

(3) 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.

(4) 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.

c. Transposable elements are also important components of eukaryotic genomes.

XI. Individual bacteria respond to environmental change by regulating their gene expression.

A. First, cells can vary the number of specific enzyme molecules they make by regulating gene expression.

B. Second, cells can adjust the activity of enzymes already present (for example, by *feedback inhibition*).

C. The tryptophan biosynthesis pathway demonstrates both levels of control.

1. If tryptophan levels are high, some of the tryptophan molecules can inhibit the first enzyme in the pathway.
2. 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.
3. The basic mechanism for this control of gene expression in bacteria is called the *operon model*.
   a. *E. coli* synthesizes tryptophan from a precursor molecule in a series of steps, with each reaction catalyzed by a specific enzyme.
   b. The five genes coding for these enzymes are clustered together on the bacterial chromosome, served by a single promoter.
   c. Transcription gives rise to one long mRNA molecule that codes for all five enzymes in the tryptophan pathway.
   d. The mRNA is punctuated with start and stop codons that signal where the coding sequence for each polypeptide begins and ends.
4. 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.
   a. When an *E. coli* cell must make tryptophan for itself, all the enzymes are synthesized at one time.
   b. The switch is a segment of DNA called an **operator.**
5. 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.**
6. By itself, an operon is on and RNA polymerase can bind to the promoter and transcribe the genes.
7. However, if a **repressor** protein, a product of a **regulatory gene,** binds to the operator, it can prevent transcription of the operon's genes.
   a. Each repressor protein recognizes and binds only to the operator of a certain operon.
   b. Regulatory genes are transcribed continuously at low rates.
   c. Binding by the repressor to the operator is reversible.
   d. The number of active repressor molecules available determines the on or off mode of the operator.
   e. Repressors contain allosteric sites that change shape depending on the binding of other molecules.
8. 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.
    a. This activates the repressor and turns the operon off.
    b. At low levels of tryptophan, most of the repressors are inactive, and the operon is transcribed.
9. 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.
10. In contrast, an *inducible* operon is *stimulated* when a specific small molecule interacts with a regulatory protein.
    a. In inducible operons, the regulatory protein is active (inhibitory) as synthesized, and the operon is off.
    b. Allosteric binding by an **inducer** molecule makes the regulatory protein inactive, and the operon is turned on.

D. 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).
1. In the absence of lactose, this operon is off, as an active repressor binds to the operator and prevents transcription.
2. Lactose metabolism begins with hydrolysis of lactose into its component monosaccharides, glucose and galactose.
    a. This reaction is catalyzed by the enzyme ß-galactosidase.
    b. 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.
3. The gene for ß-galactosidase is part of the *lac* operon, which includes two other genes coding for enzymes that function in lactose metabolism.
4. 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.
5. Unlike the *trp* operon, the *lac* repressor is active all by itself, binding to the operator and switching the *lac* operon off.
    a. An **inducer** *inactivates* the repressor.
    b. When lactose is present in the cell, allolactose, an isomer of lactose, binds to the repressor.
    c. This inactivates the repressor, and the *lac* operon can be transcribed.

E. Repressible enzymes generally function in anabolic pathways, synthesizing end products from raw materials.
1. When the end product is present in sufficient quantities, the cell can allocate its resources to other uses.

F. Inducible enzymes usually function in catabolic pathways, digesting nutrients to simpler molecules.
1. By producing the appropriate enzymes only when the nutrient is available, the cell avoids making proteins that have nothing to do.

G. Both repressible and inducible operons demonstrate *negative* control because active repressors switch off the active form of the repressor protein.

H. Positive gene control occurs when an activator molecule interacts directly with the genome to switch transcription on.

I. Even if the *lac* operon is turned on by the presence of allolactose, the degree of transcription depends on the concentrations of other substrates.
1. If glucose levels are low, then **cyclic AMP (cAMP)** accumulates.
2. The regulatory protein *catabolite activator protein (CAP)* is an **activator** of transcription.

3. 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.
4. The attachment of CAP to the promoter directly stimulates gene expression, thus this mechanism qualifies as positive regulation.
5. The cellular metabolism is biased toward the use of glucose.
    a. 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.
    b. The *lac* operon will be transcribed but at a low level.
6. For the *lac* operon, the presence/absence of lactose (allolactose) determines if the operon is on or off.
7. Overall energy levels in the cell determine the level of transcription, a "volume" control, through CAP.
8. CAP works on several operons that encode enzymes used in catabolic pathways.
    a. If glucose is present and CAP is inactive, then the synthesis of enzymes that catabolize other compounds is slowed.
    b. 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.