

CHAPTER 18

MICROBIAL MODELS: THE GENETICS OF VIRUSES AND BACTERIA

OUTLINE

- I. The Genetics of Viruses
 - A. Researchers discovered viruses by studying a plant disease: *science as a process*
 - B. A virus is a genome enclosed in a protective coat.
 - C. Viruses can reproduce only within a host cell: *an overview*
 - D. Phages reproduce using lytic or lysogenic cycles
 - E. Animal viruses are diverse in their modes of infection and of replication
 - F. Plant viruses are serious agricultural pests
 - G. Viroids and prions are infectious agents even simpler than viruses
 - H. Viruses may have evolved from other mobile genetic elements
- II. The Genetics of Bacteria
 - A. The short generation span of bacteria facilitates their evolutionary adaptation to changing environments
 - B. Genetic recombination produces new bacterial strains
 - C. The control of gene expression enables individual bacteria to adjust their metabolism to environmental change

OBJECTIVES

After reading this chapter and attending lecture, the student should be able to:

1. Recount the history leading up to the discovery of viruses and discuss the contributions of A. Mayer, D. Ivanowsky, Martinus Beijerinck, and Wendell Stanley.
2. List and describe structural components of viruses.
3. Explain why viruses are obligate parasites.
4. Describe three patterns of viral genome replication.
5. Explain the role of reverse transcriptase in retroviruses.
6. Describe how viruses recognize host cells.
7. Distinguish between lytic and lysogenic reproductive cycles using phage T₄ and phage λ as examples.
8. Outline the procedure for measuring phage concentration in a liquid medium.
9. Describe several defenses bacteria have against phage infection.
10. Using viruses with envelopes and RNA viruses as examples, describe variations in replication cycles of animal viruses.
11. Explain how viruses may cause disease symptoms, and describe some medical weapons used to fight viral infections.
12. List some viruses that have been implicated in human cancers, and explain how tumor viruses transform cells.
13. Distinguish between horizontal and vertical routes of viral transmission in plants.

14. List some characteristics that viruses share with living organisms, and explain why viruses do not fit our usual definition of life.
15. Provide evidence that viruses probably evolved from fragments of cellular nucleic acid.
16. Describe the structure of a bacterial chromosome.
17. Describe the process of binary fission in bacteria, and explain why replication of the bacterial chromosome is considered to be semiconservative.
18. List and describe the three natural processes of genetic recombination in bacteria.
19. Distinguish between general transduction and specialized transduction.
20. Explain how the F plasmid controls conjugation in bacteria.
21. Explain how bacterial conjugation differs from sexual reproduction in eukaryotic organisms.
22. For donor and recipient bacterial cells, predict the consequences of conjugation between the following: 1) F^+ and F^- cell, 2) Hfr and F^- cell.
23. Define transposon, and describe two essential types of nucleotide sequences found in transposon DNA.
24. Distinguish between an insertion sequence and a complex transposon.
25. Describe the role of transposase and DNA polymerase in the process of transposition.
26. Explain how transposons can generate genetic diversity.
27. Briefly describe two main strategies cells use to control metabolism.
28. Explain why grouping genes into an operon can be advantageous.
29. Using the *trp* operon as an example, explain the concept of an operon and the function of the operator, repressor, and corepressor.
30. Distinguish between structural and regulatory genes.
31. Describe how the *lac* operon functions and explain the role of the inducer allolactose.
32. Explain how repressible and inducible enzymes differ and how these differences reflect differences in the pathways they control.
33. Distinguish between positive and negative control, and give examples of each from the *lac* operon.
34. Explain how CAP is affected by glucose concentration.
35. Describe how *E. coli* uses the negative and positive controls of the *lac* operon to economize on RNA and protein synthesis.

KEY TERMS

capsid	provirus	transformation	operator
viral envelope	retrovirus	transduction	operon
bacteriophage (phage)	reverse transcriptase	conjugation	repressor
host range	HIV	F factor	regulatory gene
lytic cycle	AIDS	episome	corepressor
virulent virus	vaccine	F plasmid	inducer
lysogenic cycle	virion	R plasmid	cyclic amp (cAMP)
temperate virus	prion	transposon	cAMP receptor protein (CRP)
prophage	nucleoid	insertion sequence	

LECTURE NOTES

Scientists discovered the role of DNA in heredity by studying the simplest of biological systems—viruses and bacteria. Most of the molecular principles discovered through microbe research applies to higher organisms, but viruses and bacteria also have unique genetic features.

- Knowledge of these unique genetic features has helped scientists understand how viruses and bacteria cause disease.
- Techniques for gene manipulation emerged from studying genetic peculiarities of microorganisms.

I. The Genetics of Viruses

A. Researchers discovered viruses by studying a plant disease: *science as a process*

The discovery of viruses resulted from the search for the infectious agent causing tobacco mosaic disease. This disease stunts the growth of tobacco plants and gives their leaves a mosaic coloration (see Campbell, Figure 18.8a).

1883: A. Mayer, a German scientist demonstrated that the disease was contagious and proposed that the infectious agent was an unusually small bacterium that could not be seen with a microscope.

- He successfully transmitted the disease by spraying sap from infected plants onto the healthy ones.
- Using a microscope, he examined the sap and was unable to identify a microbe.

1890s: D. Ivanowsky, a Russian scientist proposed that tobacco mosaic disease was caused by a bacterium that was either too small to be trapped by a filter or that produced a filterable toxin.

- To remove bacteria, he filtered sap from infected leaves.
- Filtered sap still transmitted disease to healthy plants.

1897: Martinus Beijerinck, a Dutch microbiologist proposed that the disease was caused by a reproducing particle much smaller and simpler than a bacterium.

- He ruled out the theory that a filterable toxin caused the disease by demonstrating that the infectious agent in filtered sap could reproduce.

Plants were sprayed with filtered sap from diseased plant.



Sprayed plants developed tobacco mosaic disease.



Sap from newly infected plants was used to infect others.

- This experiment was repeated for several generations. He concluded that the pathogen must be reproducing because its ability to infect was undiluted by transfers from plant to plant.
- He also noted that unlike bacteria, the pathogen:
 - Reproduced only within the host it infected
 - Could not be cultured on media
 - Could not be killed by alcohol

1935: Wendell M. Stanley, an American biologist, crystallized the infectious particle now known as *tobacco mosaic virus (TMV)*.

B. A virus is a genome enclosed in a protective coat

In the 1950s, TMV and other viruses were finally observed with electron microscopes. Viral structure appeared to be unique from the simplest of cells.

- The smallest viruses are only 20 nm in diameter.
- The virus particle, consists of nucleic acid enclosed by a protein coat and sometimes a membranous envelope.

1. Viral genomes

Depending upon the virus, viral genomes:

- May be double-stranded DNA, single-stranded DNA, double-stranded RNA, or single-stranded RNA
- Are organized as single nucleic acid molecules that are linear or circular
- May have as few as four genes or as many as several hundred

2. Capsids and envelopes

Capsid = Protein coat that encloses the viral genome

- Its structure may be rod-shaped, polyhedral, or complex
- Composed of many *capsomeres*, protein subunits made from only one or a few types of protein.

Envelope = Membrane that cloaks some viral capsids

- Helps viruses infect their host
- Derived from host cell membrane which is usually virus-modified and contains proteins and glycoproteins of viral origin

The most complex capsids are found among *bacteriophages* or bacterial viruses.

- Of the first phages studied, seven infected *E. coli*. These were named types 1 – 7 (T1, T2, T3, ... T7).
- The T-even phages – T2, T4, and T6—are structurally very similar.
 - The icosohedral head encloses the genetic material.
 - The protein tailpiece with tail fibers attaches the phage to its bacterial host and injects its DNA into the bacterium.

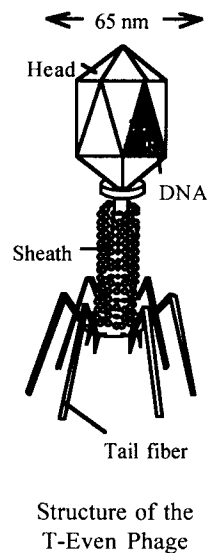
Campbell, Figure 18.2 shows the structure of various viruses.

C. Viruses can reproduce only within a host cell: an overview

Viral reproduction differs markedly from cellular reproduction, because viruses are *obligate intracellular parasites* which can express their genes and reproduce only within a living cell. Each virus has a specific *host range*.

Host range = Limited number or range of host cells that a parasite can infect

- Viruses recognize host cells by a complementary fit between external viral proteins and specific cell surface *receptor sites*.
- Some viruses have broad host ranges which may include several species (e.g., swine flu and rabies).
- Some viruses have host ranges so narrow that they can:
 - Infect only one species (e.g., phages of *E. coli*)
 - Infect only a single tissue type of one species (e.g., human cold virus infects only cells of the upper respiratory tract; AIDS virus binds only to specific receptors on certain white blood cells)



There are many patterns of viral life cycles, but they all generally involve:

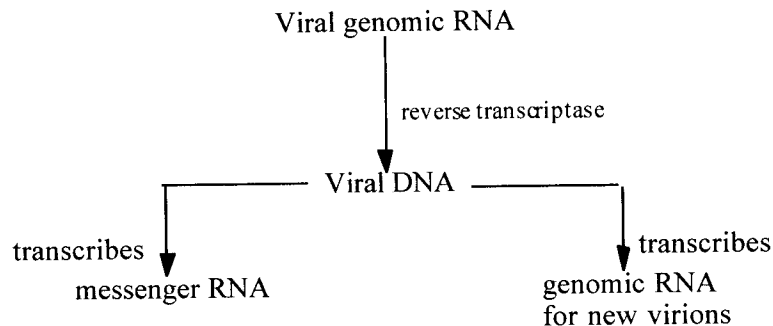
- Infecting the host cell with viral genome
- Co-opting host cell's resources to:
 - Replicate the viral genome
 - Manufacture capsid protein
- Assembling newly produced viral nucleic acid and capsomeres into the next generation of viruses (see Campbell, Figure 18.3)

There are several mechanisms used to infect host cells with viral DNA.

- For example, T-even phages use an elaborate tailpiece to inject DNA into the host cell.
- Once the viral genome is inside its host cell, it commandeers the host's resources and reprograms the cell to copy the viral genes and manufacture capsid protein.

There are three possible patterns of viral genome replication:

1. *DNA* → *DNA*. If viral DNA is double-stranded, DNA replication resembles that of cellular DNA, and the virus uses DNA polymerase produced by the host.
2. *RNA* → *RNA*. Since host cells lack the enzyme to copy RNA, most RNA viruses contain a gene that codes for *RNA replicase*, an enzyme that uses viral RNA as a template to produce complementary RNA.
3. *RNA* → *DNA* → *RNA*. Some RNA viruses encode *reverse transcriptase*, an enzyme that transcribes DNA from an RNA template.



Regardless of how viral genomes replicate, all viruses divert host cell resources for viral production.

- Viral genes use the host cell's enzymes, ribosomes, tRNAs, amino acids, ATP, and other resources to make copies of the viral genome and produce viral capsid proteins.
- These viral components—nucleic acid and capsids—are assembled into hundreds or thousands of virions, which leave to parasitize new hosts.

Viral nucleic acid and capsid proteins assemble spontaneously into new virus particles, a process called *self-assembly*.

- Since most viral components are held together by weak bonds (e.g., hydrogen bonds and Van der Waals forces), enzymes are not usually necessary for assembly.
- For example, TMV can be disassembled in the laboratory. When mixed together, the RNA and capsids spontaneously reassemble to form complete TMV virions.

D. Phages reproduce using lytic or lysogenic cycles

Bacteriophages are the best understood of all viruses, and many of the important discoveries in molecular biology have come from bacteriophage studies.

- In the 1940s, scientists determined how the T phages reproduce within a bacterium; this research:
- Demonstrated that DNA is the genetic material
- Established the phage-bacterium system as an important experimental tool
- Studies on lambda (λ) phage of *E. coli* showed that some double-stranded DNA viruses reproduce by two alternative mechanisms: the *lytic cycle* and the *lysogenic cycle*.

1. The lytic cycle

Virulent bacteriophages reproduce only by a *lytic* replication cycle.

Virulent phages = Phages that lyse their host cells

Lytic cycle = A viral replication cycle that results in the death or lysis of the host cell

The lytic cycle of phage T4 illustrates this type of replication cycle (see Campbell, Figure 18.4):

1. Phage attaches to cell surface.
 - T4 recognizes a host cell by a complementary fit between proteins on the virion's tail fibers and specific receptor sites on the outer surface of an *E. coli* cell.
2. Phage contracts sheath and injects DNA.
 - ATP stored in the phage tailpiece is the energy source for the phage to: a) pierce the *E. coli* wall and membrane, b) contract its tail sheath, and c) inject its DNA.
 - The genome separates from the capsid leaving a capsid "ghost" outside the cell.
3. Hydrolytic enzymes destroy host cell's DNA.
 - The *E. coli* host cell begins to transcribe and translate the viral genome.
 - One of the first viral proteins produced is an enzyme that degrades host DNA. The phage's own DNA is protected, because it contains modified cytosine not recognized by the enzyme.
4. Phage genome directs the host cell to produce phage components: DNA and capsid proteins.
 - Using nucleotides from its own degraded DNA, the host cell makes many copies of the phage genome.
 - The host cell also produces three sets of capsid proteins and assembles them into phage tails, tail fibers, and polyhedral heads.
 - Phage components spontaneously assemble into virions.
5. Cell lyses and releases phage particles.
 - Lysozymes specified by the viral genome digest the bacterial cell wall.
 - Osmotic swelling lyses the cell which releases hundreds of phages from their host cell.
 - Released virions can infect nearby cells.
 - Lytic cycle takes only 20 to 30 minutes at 37°C. In that period, a T4 population can increase a hundredfold.

Bacteria have several defenses against destruction by phage infection.

- Bacterial mutations can change receptor sites used by phages for recognition, and thus avoid infection.

- Bacterial *restriction nucleases* recognize and cut up foreign DNA, including certain phage DNA. Bacterial DNA is chemically altered, so it is not destroyed by the cell's own restriction enzymes.

Restriction enzymes = Naturally occurring bacterial enzymes that protect bacteria against intruding DNA from other organisms. The enzymes also catalyze restriction, the process of cutting foreign DNA into small segments.

Bacterial hosts and their viral parasites are continually *coevolving*.

- Most successful bacteria have effective mechanisms for preventing phage entry or reproduction.
- Most successful phages have evolved ways around bacterial defenses.
- Many phages check their own destructive tendencies and may coexist with their hosts.

2. The lysogenic cycle

Some viruses can coexist with their hosts by incorporating their genome into the host's genome.

Temperate viruses = Viruses that can integrate their genome into a host chromosome and remain latent until they initiate a lytic cycle

- They have two possible modes of reproduction, the lytic cycle and the lysogenic cycle.
- An example is phage λ , discovered by E. Lederberg in 1951 (see Campbell, Figure 18.5)

Lysogenic cycle = A viral replication cycle that involves the incorporation of the viral genome into the host cell genome

Details of the lysogenic cycle were discovered through studies of phage λ life cycle:

1. Phage λ binds to the surface of an *E. coli* cell.
2. Phage λ injects its DNA into the bacterial host cell.
3. λ DNA forms a circle and either begins a lytic or lysogenic cycle.
4. During a lysogenic cycle, λ DNA inserts by genetic recombination (crossing over) into a specific site on the bacterial chromosome and becomes a prophage.

Prophage = A phage genome that is incorporated into a specific site on the bacterial chromosome

- Most prophage genes are inactive.
- One active prophage gene codes for the production of *repressor protein* which switches off most other prophage genes.
- Prophage genes are copied along with cellular DNA when the host cell reproduces. As the cell divides, both prophage and cellular DNA are passed on to daughter cells.
- A prophage may be carried in the host cell's chromosomes for many generations.

Occasionally, a prophage may leave the bacterial chromosome.

- This may be spontaneous or caused by environmental factors (e.g., radiation).
- The excision process may begin the phage's lytic reproductive cycle.
- Virions produced during the lytic cycle may begin either a lytic or lysogenic cycle in their new host cells.

Lysogenic cell = Host cell carrying a prophage in its chromosome

- It is called lysogenic because it has the potential to lyse.

- Some prophage genes in a lysogenic cell may be expressed and change the cell's phenotype in a process called *lysogenic conversion*.
- Lysogenic conversion occurs in bacteria that cause diphtheria, botulism, and scarlet fever. Pathogenicity results from toxins coded for by prophage genes.

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

1. Reproductive cycles of animal viruses

Replication cycles of animal viruses may show some interesting variations from those of other viruses. Two examples are the replication cycles of: 1) viruses with envelopes, and 2) viruses with RNA genomes that serve as the genetic material. (See Campbell, Table 18.1 for families of animal viruses grouped by type of nucleic acid.)

a. Viral envelopes

Some animal viruses are surrounded by a membranous envelope, which is unique to several groups of animal viruses. This envelope is:

- Outside the capsid and helps the virus enter host cells.
- A lipid bilayer with glycoprotein spikes protruding from the outer surface.

Enveloped viruses have replication cycles characterized by (see Campbell, Figure 18.6):

1. *Attachment*. Glycoprotein spikes protruding from the viral envelope attach to receptor sites on the host's plasma membrane.
2. *Entry*. As the envelope fuses with the plasma membrane, the entire virus (capsid and genome) is transported into the cytoplasm by receptor-mediated endocytosis.
3. *Uncoating*. Cellular enzymes uncoat the genome by removing the protein capsid from viral RNA.
4. *Viral RNA and protein synthesis*. Viral enzymes are required to replicate the RNA genome and to transcribe mRNA.
 - Some viral RNA polymerase is packaged in the virion.
 - Viral RNA polymerase (transcriptase) replicates the viral genome and transcribes viral mRNA. Note that the viral genome is a strand complementary to mRNA.
 - Viral mRNA is translated into viral proteins including:
 - Capsid proteins synthesized in the cytoplasm by free ribosomes
 - Viral-envelope glycoproteins synthesized by ribosomes bound to rough ER. Glycoproteins produced in the host's ER are sent to the Golgi apparatus for further processing. Golgi vesicles transport the glycoproteins to the plasma membrane, where they cluster at exit sites for the virus.
5. *Assembly and release*. New capsids surround viral genomes. Once assembled, the virions envelop with host plasma membrane as they bud off from the cell's surface. The viral envelope is derived from:
 - Host cell's plasma membrane lipid
 - Virus-specific glycoprotein

Some viral envelopes are not derived from host plasma membrane.

For example, herpesviruses are double-stranded DNA viruses which:

- Contain envelopes derived from the host cell's nuclear envelope rather than from the plasma membrane
- Reproduce within the host cell's nucleus
- Use both viral and cellular enzymes to replicate and transcribe their genomic DNA
- May integrate their DNA into the cell's genome as a *provirus*. Evidence comes from the nature of herpes infections, which tend to recur. After a period of latency, physical or emotional stress may cause the proviruses to begin a productive cycle again.

Provirus = Viral DNA that inserts into a host cell chromosome

b. RNA as viral genetic material

All possible types of viral genomes are represented among animal viruses. Since mRNA is common to all types, DNA and RNA viruses are classified according to the relationship of their mRNA to the genome. In this classification:

- mRNA or the strand that corresponds to mRNA is the *plus (+) strand*; it has the nucleotide sequence that codes for proteins.
- The *minus (-) strand* is a template for synthesis of a plus strand; it is complementary to the sense strand or mRNA.

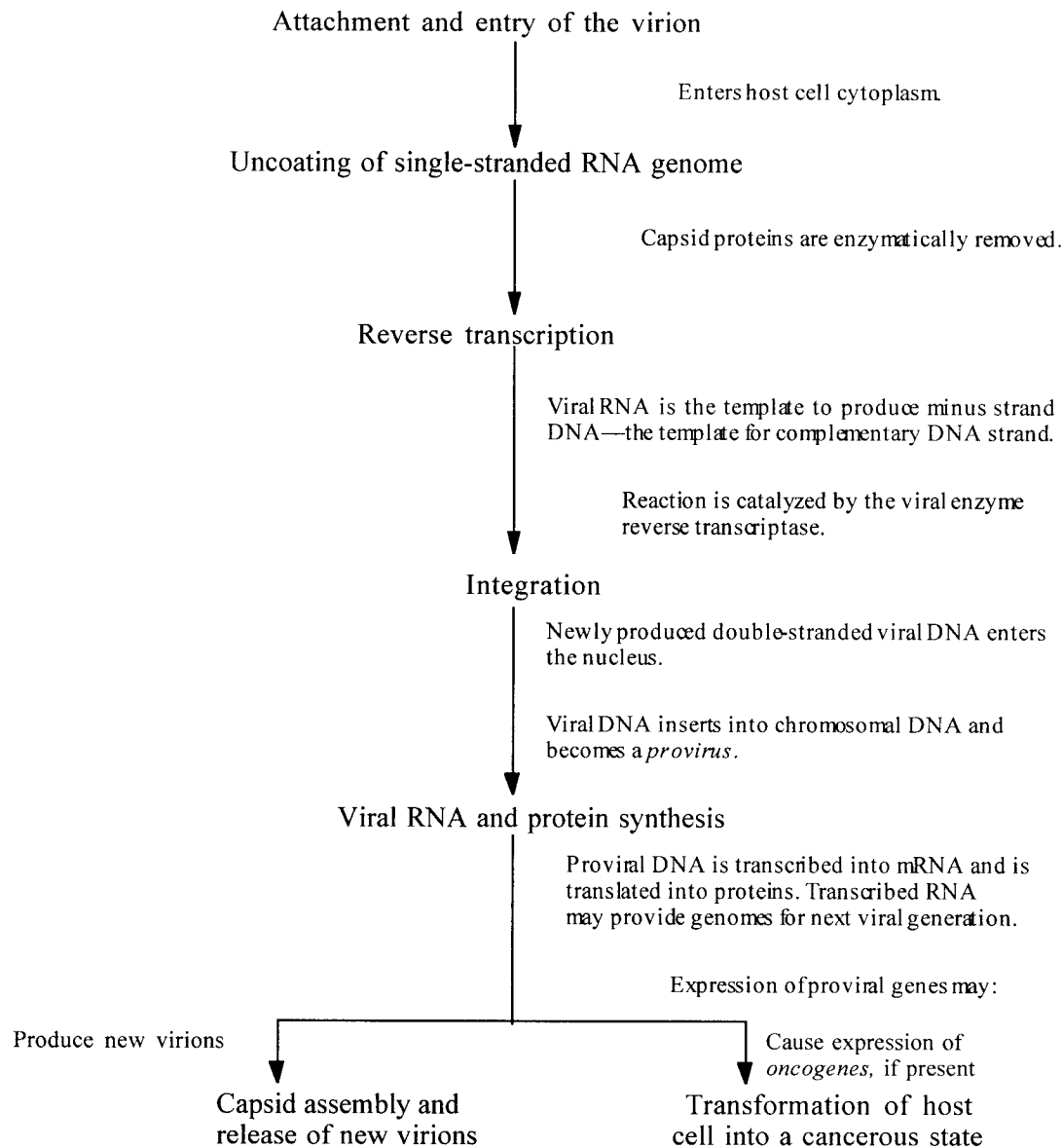
Animal RNA viruses are classified as following:

- *Class III RNA viruses*. Double-stranded RNA genome; the minus strand is the template for mRNA. (Reoviruses)
- *Class IV RNA viruses*. Single plus strand genome; the plus strand can function directly as mRNA, but also is a template for synthesis of minus RNA. (Minus RNA is a template for synthesis of additional plus strands.) Viral enzymes are required for RNA synthesis from RNA templates. (Picornavirus, Togavirus)
- *Class V RNA viruses*. Single minus strand genome; mRNA is transcribed directly from this genomic RNA. (Rhabdovirus, Paramyxovirus, Orthomyxovirus)
- *Class VI RNA viruses*. Single plus strand genome; the plus strand is a template for complementary DNA synthesis. *Reverse transcriptase* catalyzes this reverse transcription from RNA to DNA. mRNA is then transcribed from a DNA template. (*Retroviruses*)

Retrovirus = (Retro = backward) RNA virus that uses *reverse transcriptase* to transcribe DNA from the viral RNA genome.

- *Reverse transcriptase* is a type of DNA polymerase that transcribes DNA from an RNA template.
- *HIV (human immunodeficiency virus)*, the virus that causes *AIDS (acquired immunodeficiency syndrome)* is a retrovirus.

RNA viruses with the most complicated reproductive cycles are the retroviruses, because retroviruses must first carry out *reverse transcription*: (See Campbell, Figure 18.7)



2. Important viral diseases in animals

It is often unclear how certain viruses cause disease symptoms. Viruses may:

- Damage or kill cells. In response to a viral infection, lysosomes may release hydrolytic enzymes.
- Be toxic themselves or cause infected cells to produce toxins.
- Cause varying degrees of cell damage depending upon regenerative ability of the infected cell. We recover from colds because infected cells of the upper respiratory tract can regenerate by cell division. Poliovirus, however, causes permanent cell damage because the virus attacks nerve cells which cannot divide.
- Be indirectly responsible for disease symptoms. Fever, aches and inflammation may result from activities of the immune system.

Medical weapons used to fight viral infections include *vaccines* and *antiviral drugs*.

Vaccines = Harmless variants or derivatives of pathogenic microbes that mobilize a host's immune mechanism against the pathogen

- Edward Jenner developed the first vaccine (against smallpox) in 1796. According to the WHO, a vaccine has almost completely eradicated smallpox.
- Effective vaccines now exist for polio, rubella, measles, mumps, and many other viral diseases.

While vaccines can prevent some viral illnesses, little can be done to cure a viral disease once it occurs. Some *antiviral drugs* have recently been developed.

- Several are analogs of purine nucleosides that interfere with viral nucleic acid synthesis (e.g., adenine arabinoside and acyclovir).

3. Emerging viruses

Emerging viruses are viruses that make an apparent sudden appearance. In reality, they are not likely to be new viruses, but rather existing ones that have expanded their host territory.

Emerging viral diseases can arise if an existing virus:

1. Evolves and thus causes disease in individuals who have immunity only to the ancestral virus (e.g., influenza virus)
2. Spreads from one host species to another
 - For example, the 1993 hantavirus outbreak in New Mexico was the result of a population explosion in deer mice that are the viral reservoirs. Humans became infected by inhaling airborne hantavirus that came from the excreta of deer mice.
3. Disseminates from a small population to become more widespread
 - AIDS, once a rare disease, has become a global epidemic. Technological and social factors influenced the spread of AIDS virus.

Environmental disturbances can increase the viral traffic responsible for emerging diseases. For example:

- Traffic on newly cut roads through remote areas can spread viruses among previously isolated human populations.
- Deforestation activities brings humans into contact with animals that may host viruses capable of infecting humans.

4. Viruses and cancer

Some tumor viruses cause cancer in animals.

- When animal cells grown in tissue culture are infected with tumor viruses, they *transform* to a cancerous state.
- Examples are members of the retrovirus, papovavirus, adenovirus and herpesvirus groups.
- Certain viruses are implicated in human cancers:

Viral Group	Examples/Diseases	Cancer Type
Retrovirus	HTLV-1/adult leukemia	Leukemia
Herpesvirus	Epstein-Barr/infectious mononucleosis	Burkitt's lymphoma
Papovavirus	Papilloma/human warts	Cervical cancer
Hepatitis B virus	Chronic hepatitis	Liver cancer

Tumor viruses transform cells by inserting viral nucleic acids into host cell DNA.

- This insertion is permanent as the provirus never excises.
- Insertion for DNA tumor viruses is straightforward.

Several viral genes have been identified as oncogenes.

Oncogenes = Genes found in viruses or as part of the normal eukaryotic genome, that trigger transformation of a cell to a cancerous state.

- Code for cellular growth factors or for proteins involved in the function of growth factors.
- Are not unique to tumor viruses, but are found in the normal cells of many species. In fact, some tumor viruses transform cells by activating cellular oncogenes.

More than one oncogene must usually be activated to completely transform a cell.

- Indications are that tumor viruses are effective only in combination with other events such as exposure to carcinogens.
- Carcinogens probably also act by turning on cellular oncogenes.

F. Plant viruses are serious agricultural pests

As serious agricultural pests, many of the plant viruses:

- Stunt plant growth and diminish crop yields (see Campbell, Figure 18.8a)
- Are RNA viruses
- Have rod-shaped capsids with capsomeres arranged in a spiral

Capsomere = Complex capsid subunit consisting of several identical or different protein molecules

Plant viruses spread from plant to plant by two major routes: horizontal transmission and vertical transmission.

Horizontal transmission = Route of viral transmission in which an organism receives the virus from an external source

- Plants are more susceptible to viral infection if their protective epidermal layer is damaged.
- Insects may be *vectors* that transmit viruses from plant to plant and can inject the virus directly into the cytoplasm.
- By using contaminated tools, gardeners and farmers may transmit plant viruses.

Vertical transmission = Route of viral transmission in which an organism inherits a viral infection from its parent

- Can occur in asexual propagation of infected plants (e.g., by taking cuttings)
- Can occur in sexual reproduction via infected seeds

Once a plant is infected, viruses reproduce and spread from cell to cell by passing through plasmodesmata (see Campbell, Figure 18.8b).

Most plant viral diseases have no cure, so current efforts focus on reducing viral propagation and breeding resistant plant varieties.

G. Viroids and prions are infectious agents even simpler than viruses

Another class of plant pathogens called *viroids* are smaller and simpler than viruses.

- They are small, naked, circular RNA molecules that do not encode protein, but can replicate in host plant cells.
- It is likely that viroids disrupt normal plant metabolism, development, and growth by causing errors in regulatory systems that control gene expression.
- Viroid diseases affect many commercially important plants such as coconut palms, chrysanthemums, potatoes, and tomatoes.

Some scientists believe that viroids originated as escaped introns.

- Nucleotide sequences of viroid RNA are similar to self-splicing introns found within some normal eukaryotic genes, including rRNA genes.

- An alternative hypothesis is that viroids and self-splicing introns share a common ancestral molecule.

As nucleic acids, viroids self-direct their replication and thus are not diluted during transmission from host to host. Molecules other than nucleic acids can be infectious agents even though they cannot self-replicate.

- *Prions* are pathogens that are proteins, and they appear to cause a number of degenerative brain diseases, such as:
 - Scrapie in sheep
 - "Mad cow" disease
 - Creutzfeldt-Jakob disease in humans
- How can a protein which cannot replicate itself be an infectious pathogen? According to one hypothesis:
 - Prions are defective versions (misfolded) of normally occurring cellular proteins.
 - When prions infect normal cells, they somehow convert the normal protein to the prion version (see Campbell, Figure 18.9).
 - Prions could thus trigger chain reactions that increase their numbers and allow them to spread through a host population without dilution.

H. Viruses may have evolved from other mobile genetic elements

Viruses do not fit our usual definitions of living organisms. They cannot reproduce independently, yet they:

- Have a genome with the same genetic code as living organisms
- Can mutate and evolve

Viruses probably evolved after the first cells, from fragments of cellular nucleic acid that were mobile genetic elements. Evidence to support this includes:

- Genetic material of different viral families is more similar to host genomes than to that of other viral families.
- Some viral genes are identical to cellular genes (e.g., oncogenes in retroviruses).
- Viruses of eukaryotes are more similar in genomic structure to their cellular hosts than to bacterial viruses.
- Viral genomes are similar to certain cellular genetic elements such as plasmids and transposons; they are all *mobile* genetic elements.

II. The Genetics of Bacteria

A. The short generation span of bacteria facilitates their evolutionary adaptation to changing environments

The average bacterial genome is larger than a viral genome, but much smaller than a typical eukaryotic genome.

The major component of the bacterial genome is the *bacterial chromosome*. This structure is:

- Composed of one double-stranded, circular molecule of DNA
- Structurally simpler and has fewer associated proteins than a eukaryotic chromosome
- Found in the *nucleoid* region; since this region is not separated from the rest of the cell (by a membrane), transcription and translation can occur simultaneously.

Many bacteria also contain extrachromosomal DNA in plasmids.

Plasmid = A small double-stranded ring of DNA that carries extrachromosomal genes in some bacteria

Most bacteria can rapidly reproduce by *binary fission*, which is preceded by DNA replication.

- Semi-conservative replication of the bacterial chromosome begins at a single origin of replication.
- The two replication forks move bidirectionally until they meet and replication is complete (see Campbell, Figure 18.10).
- Under optimal conditions, some bacteria can divide in twenty minutes. Because of this rapid reproductive rate, bacteria are useful for genetic studies.

Binary fission is asexual reproduction that produces clones, or daughter cells that are genetically identical to the parent.

- Though mutations are rare events, they can impact genetic diversity in bacteria because of their rapid reproductive rate.
- Though mutation can be a major source of genetic variation in bacteria, it is *not* a major source in more slowly reproducing organisms (e.g., humans). In most higher organisms, genetic recombination from sexual reproduction is responsible for most of the genetic diversity within populations.

B. Genetic recombination produces new bacterial strains

There are three natural processes of genetic recombination in bacteria: *transformation*, *transduction*, and *conjugation*. These mechanisms of gene transfer occur separately from bacterial reproduction, and in addition to mutation, are another major source of genetic variation in bacterial populations.

1. Transformation

Transformation = Process of gene transfer during which a bacterial cell assimilates foreign DNA from the surroundings

- Some bacteria can take up naked DNA from the surroundings. (Refer to Avery's experiments with *Streptococcus pneumoniae* in Chapter 16.)
- Assimilated foreign DNA may be integrated into the bacterial chromosome by recombination (crossing over).
- Progeny of the recipient bacterium will carry a new combination of genes.

Many bacteria have surface proteins that recognize and import naked DNA from closely related bacterial species.

- Though lacking such proteins, *E. coli* can be artificially induced to take up foreign DNA by incubating the bacteria in a culture medium that has a high concentration of calcium ions.
- This technique of artificially inducing transformation is used by the biotechnology industry to introduce foreign genes into bacterial genomes, so that bacterial cells can produce proteins characteristic of other species (e.g., human insulin and human growth hormone).

2. Transduction

Transduction = Gene transfer from one bacterium to another by a bacteriophage (see Campbell, Figure 18.12)

Generalized transduction = Transduction that occurs when random pieces of host cell DNA are packaged within a phage capsid during the lytic cycle of a phage

- This process can transfer almost any host gene and little or no phage genes.
- When the phage particle infects a new host cell, the donor cell DNA can recombine with the recipient cell DNA.

Specialized transduction = Transduction that occurs when a prophage excises from the bacterial chromosome and carries with it only certain host genes adjacent to the excision site. Also known as *restricted transduction*.

- Carried out only by temperate phages
- Differs from general transduction in that:
 - Specific host genes and most phage genes are packed into the same virion.
 - Transduced bacterial genes are restricted to specific genes adjacent to the prophage insertion site. In general transduction, host genes are randomly selected and almost any host gene can be transferred.

3. Conjugation and plasmids

Conjugation = The direct transfer of genes between two cells that are temporarily joined.

- Discovered by Joshua Lederberg and Edward Tatum
- Conjugation in *E. coli* is one of the best-studied examples:

A DNA-donating *E. coli* cell extends
external appendages called *sex pili*
(see Campbell, Figure 18.13)



Sex pili attach to a DNA-receiving cell



A cytoplasmic bridge forms through which
DNA transfer occurs

The ability to form sex pili and to transfer DNA is conferred by genes in a plasmid called the *F plasmid*.

a. General characteristics of plasmids

Plasmid = A small, circular, double-stranded, self-replicating molecule ring of DNA that carries extrachromosomal genes in some bacteria.

- Plasmids have only a few genes, and they are not required for survival and reproduction.
- Plasmid genes can be beneficial in stressful environments. Examples include the F plasmid, which confers ability to conjugate; and the R plasmid, which confers antibiotic resistance.

These small circular DNA molecules replicate independently:

- Some plasmids replicate in synchrony with the bacterial chromosome, so only a few are present in the cell.
- Some plasmids under more relaxed control can replicate on their own schedule, so the number of plasmids in the cell at any one time can vary from only a few to as many as 100.

Some plasmids are episomes that can reversibly incorporate in the cell's chromosome.

Episomes = Genetic elements that can replicate either independently as free molecules in the cytoplasm or as integrated parts of the main bacterial chromosome.

- Examples include some plasmids and temperate viruses such as lambda phage.
- Temperate phage genomes replicate separately in the cytoplasm during a lytic cycle and as an integral part of the host's chromosome during a lysogenic cycle.

While plasmids and viruses can both be episomes, they differ in that:

- Plasmids, unlike viruses, lack an extracellular stage.

- Plasmids are generally beneficial to the cell, while viruses are parasites that usually harm their hosts.

b. The F plasmid and conjugation

The F plasmid (F for fertility) has about 25 genes, most of which are involved in the production of sex pili.

- Bacterial cells that contain the F factor and can donate DNA ("male") are called F^+ cells.
- The F factor replicates in synchrony with chromosomal DNA, so the F^+ factor is heritable; that is, division of an F^+ cell results in two F^+ daughter cells.
- Cells without the F factor are designated F^- ("female").

During conjugation between an F^+ and an F^- bacterium:

- The F factor replicates by *rolling circle replication*. The 5' end of the copy peels off the circular plasmid and is transferred in linear form.
- The F^+ cell transfers a copy of its F factor to the F^- partner, and the F^- cell becomes F^+ (see Campbell, Figure 18.14)
- The donor cell remains F^+ , with its original DNA intact.

The F factor is an episome and occasionally inserts into the bacterial chromosome.

- Integrated F factor genes are still expressed.
- Cells with integrated F factors are called *Hfr cells* (high frequency of recombination).

Conjugation can occur between an Hfr and an F^- bacterium.

- As the integrated F factor of the Hfr cell transfers to the F^- cell, it pulls the bacterial chromosome behind its leading end.
- The F factor always opens up at the same point for a particular Hfr strain. As rolling circle replication proceeds, the sequence of chromosomal genes behind the leading 5' end is always the same.
- The conjugation bridge usually breaks before the entire chromosome and tail end of the F factor can be transferred. As a result:
 - Only some bacterial genes are donated.
 - The recipient F^- cell does not become an F^+ cell, because only part of the F factor is transferred.
 - The recipient cell becomes a partial diploid.
 - Recombination occurs between the Hfr chromosomal fragment and the F^- cell. Homologous strand exchange results in a *recombinant F^- cell*.
 - Asexual reproduction of the recombinant F^- cell produces a bacterial colony that is genetically different from both original parental cells.

c. R plasmids and antibiotic resistance

One class of nonepisomal plasmids, the *R plasmids* (for resistance), carry genes that confer resistance to certain antibiotics.

- Some carry up to ten genes for resistance to antibiotics.
- During conjugation, some mobilize their own transfer to nonresistant cells.
- Increased antibiotic use has selected for antibiotic resistant bacterial strains carrying the R plasmid.

- Additionally, R plasmids can transfer resistance genes to bacteria of different species including pathogenic strains. As a consequence, resistant strains of pathogens are becoming more common.

4. Transposons

Pieces of DNA called transposons, or transposable genetic elements, can actually move from one location to another in a cell's genome.

Transposons = DNA sequences that can move from one chromosomal site to another.

- Occur as natural agents of genetic change in both prokaryotic and eukaryotic organisms.
- Were first proposed in the 1940s by Barbara McClintock, who deduced their existence in maize. Decades later, the importance of her discovery was recognized; in 1983, at the age of 81, she received the Nobel Prize for her work.

There are two patterns of transposition: a) conservative transposition and b) replicative transposition.

Conservative transposition = Movement of preexisting genes from one genomic location to another; the transposon's genes are not replicated before the move, so the number of gene copies is conserved.

Replicative transposition = Movement of gene copies from their original site of replication to another location in the genome, so the transposon's genes are inserted at some new site without being lost from the original site.

Transposition is fundamentally different from all other mechanisms of genetic recombination, because transposons may scatter certain genes throughout the genome with no apparent single, specific target.

- All other mechanisms of genetic recombination depend upon homologous strand exchange: meiotic crossing over in eukaryotes; and transformation, transduction, and conjugation in prokaryotes.
- Insertion of episomic plasmids into chromosomes is also site specific, even though it does not require an extensive stretch of DNA homologous to the plasmid.

a. Insertion sequences

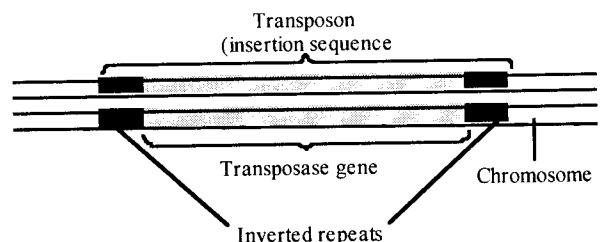
The simplest transposons are insertion sequences (see Campbell, Figure 18.15).

Insertion sequences = The simplest transposons, which contain only the genes necessary for the process of transposition. Insertion sequence DNA includes two essential types of nucleotide sequences:

- Nucleotide sequence coding for *transposase*
- Inverted repeats

Transposase = Enzyme that catalyzes insertion of transposons into new chromosomal sites.

- The transposase gene in an insertion sequence is flanked by inverted repeats.



Inverted repeats (IR) = Short noncoding nucleotide sequences of DNA that are repeated in reverse order on opposite ends of a transposon. For example:

```

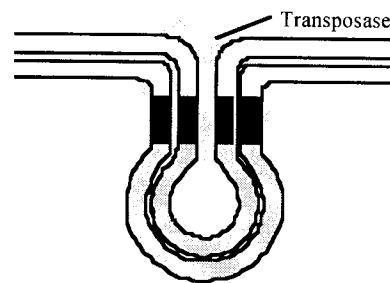
DNA strand #1  ... A T C C G G T ... A C C G G A T ...
DNA strand #2  ... A T C C G G T ... A C C G G A T ...
  
```

Note that each base sequence (IR) is repeated in reverse, on the DNA strand *opposite* the inverted repeat at the other end. Inverted repeats:

- Contain only 20 to 40 nucleotide pairs
- Are recognition sites for transposase

Transposase catalyzes the recombination by:

- Binding to the inverted repeats and holding them close together
- Cutting and resealing DNA required for insertion of the transposon at a new site



Insertion of transposons also requires other enzymes, such as DNA polymerase. For example,

- At the target site, transposase makes staggered cuts in the two DNA strands, leaving short segments of unpaired DNA at each end of the cut.
- Transposase inserts the transposon into the open target site.
- DNA polymerase helps form direct repeats, which flank transposons in their target site. Gaps in the two DNA strands fill in when nucleotides base pair with the exposed single-stranded regions.

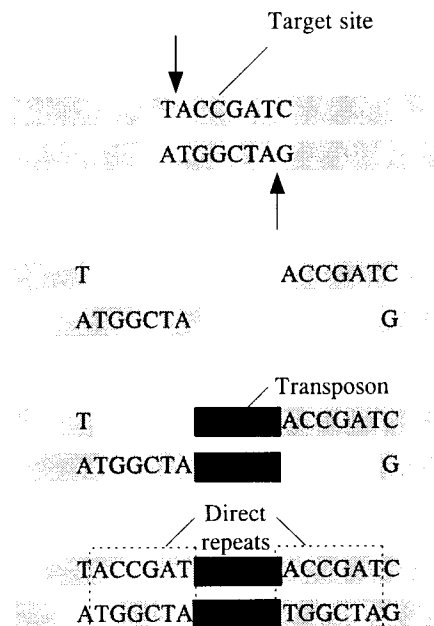
Direct repeats = Two or more identical DNA sequences in the same molecule.

- The transposition process creates direct repeats that flank transposons in their target site (see Campbell, Figure 18.16).

Transposed insertion sequences are likely to somehow alter the cell's phenotype; they may:

- Cause mutations by interrupting coding sequences for proteins.
- Increase or decrease a protein's production by inserting within regulatory regions that control transcription rates.

Transposition of insertion sequences probably plays a significant role in bacterial evolution as a source of genetic variation.

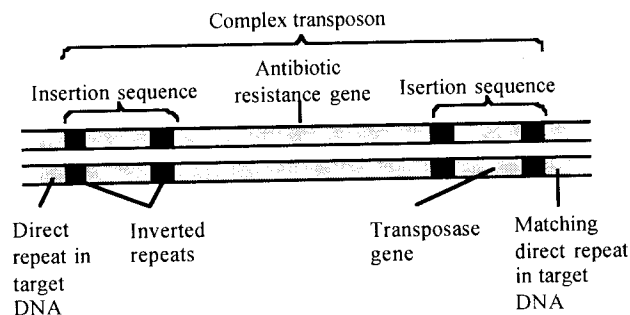


- Though insertion sequences only rarely cause mutations (about one in every 10^6 generations), the mutation rate from transpositions is about the same as the mutation rate from extrinsic causes, such as radiation and chemical mutagens.

b. Composite transposons

Composite (complex) transposons = Transposons which include additional genetic material besides that required for transposition; consist of one or more genes flanked by insertion sequences (see Campbell, Figure 18.17).

- The additional DNA may have any nucleotide sequence.
- Can insert into almost any stretch of DNA since their insertion is not dependent upon DNA sequence homology
- Generate genetic diversity in bacteria by moving genes from one chromosome, or even one species, to another. This diversity may help bacteria adapt to new environmental conditions.



An example is a transposon that carries a bacterial gene for antibiotic resistance.

Examples of genetic elements that contain one or more complex transposons include:

- F factor.
- DNA version of the retrovirus genome.

C. The control of gene expression enables individual bacteria to adjust their metabolism to environmental change

This material is difficult to teach. The problem is that there are so many components to track, and the students are not familiar enough with the vocabulary. Figures 18.17 and 18.18 are good teaching aids for the *trp* and *lac* operons. It is also helpful to construct flow charts where there are decision points (e.g., glucose present—glucose absent), so that students can visually follow alternate paths and their respective consequences.

Genes switch on and off as conditions in the intracellular environment change. Bacterial cells have two main ways of controlling metabolism:

1. *Regulation of enzyme activity.* The catalytic activity of many enzymes increases or decreases in response to chemical cues.
 - For example, the end product of an anabolic pathway may turn off its own production by inhibiting activity of an enzyme at the beginning of the pathway (*feedback inhibition*).
 - Useful for immediate short-term response.
2. *Regulation of gene expression.* Enzyme concentrations may rise and fall in response to cellular metabolic changes that switch genes on or off.

- For example, accumulation of product may trigger a mechanism that inhibits transcription of mRNA production by genes that code for an enzyme at the beginning of the pathway (*gene repression*).
- Slower to take effect than feedback inhibition, but is more economical for the cell. It prevents unneeded protein synthesis for enzymes, as well as, unneeded pathway product.

An example illustrating regulation of a metabolic pathway is the tryptophan pathway in *E. coli*. (See Campbell, Figure 18.18) Mechanisms for gene regulation were first discovered for *E. coli*, and current understanding of such regulatory mechanisms at the molecular level is still limited to bacterial systems.

1. Operons: the basic concept

Regulated genes can be switched on or off depending on the cell's metabolic needs. From their research on the control of lactose metabolism in *E. coli*, Francois Jacob and Jacques Monod proposed a mechanism for the control of gene expression, the *operon* concept.

Structural gene = Gene that codes for a polypeptide

Operon = A regulated cluster of adjacent *structural genes* with related functions

- Common in bacteria and phages
- Has a single promoter region, so an RNA polymerase will transcribe all structural genes on an all-or-none basis
- Transcription produces a single *polycistronic* mRNA with coding sequences for all enzymes in a metabolic pathway (e.g., tryptophan pathway in *E. coli*)

Polycistronic mRNA = A large mRNA molecule that is a transcript of several genes

- Is translated into separate polypeptides
- Contains stop and start codons for the translation of each polypeptide

Grouping structural genes into operons is advantageous because:

- Expression of these genes can be coordinated. When a cell needs the product of a metabolic pathway, all the necessary enzymes are synthesized at one time.
- The entire operon can be controlled by a single *operator*.

Operator = A DNA segment located within the promoter or between the promoter and structural genes, which controls access of RNA polymerase to structural genes.

- Sometimes overlaps the transcription starting point for the operon's first structural gene
- Acts as an on/off switch for movement of RNA polymerase and transcription of the operon's structural genes

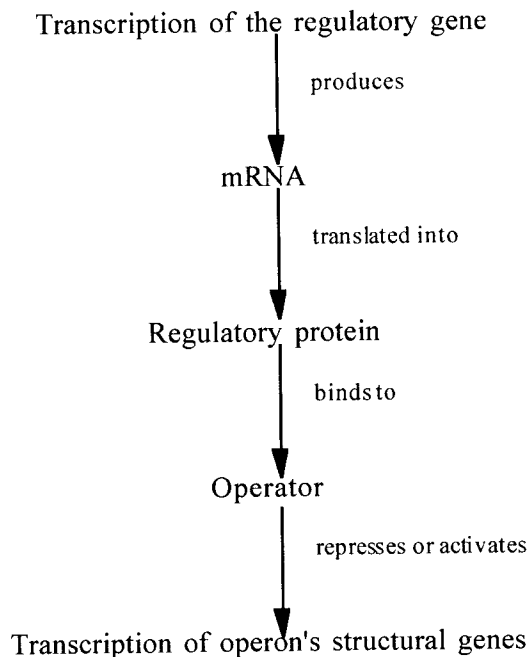
What determines whether an operator is in the "on" or "off" mode? By itself, the operator is on; it is switched off by a protein repressor.

Repressor = Specific protein that binds to an operator and blocks transcription of the operon

- Blocks attachment of RNA polymerase to the promoter
- Is similar to an enzyme, in that it:
 - Has an active site with a specific conformation, which discriminates among operators. Repressor proteins are specific only for operators of certain operons.
 - Binds *reversibly* to DNA
 - May have an allosteric site in addition to its DNA-binding site
- Repressors are encoded by *regulatory genes*.

Regulatory genes = Genes that code for repressor or regulators of other genes

- Are often located some distance away from the operons they control and has its own promotor
- Are involved in switching on or off the transcription of structural genes by the following process:

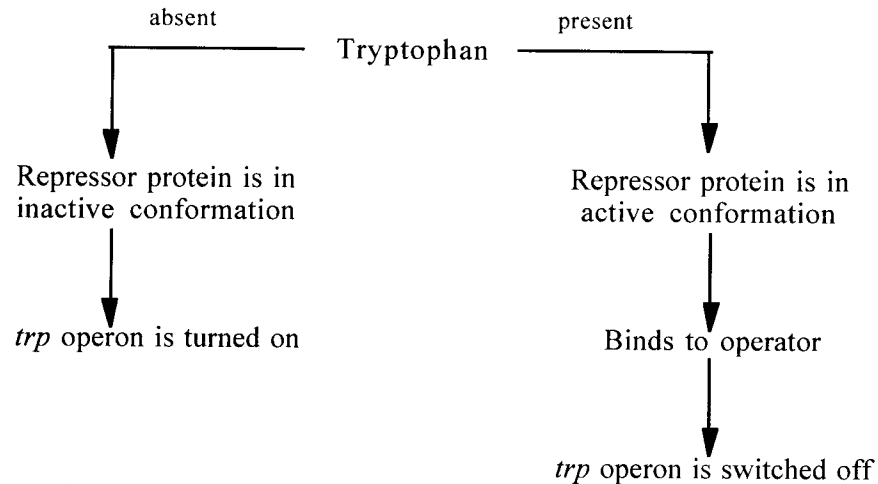


Regulatory genes are continually transcribed, so their activity depends upon how efficient their promoters are in binding RNA polymerase.

- They produce repressor molecules continuously, but slowly.
- Operons are still expressed even though repressor molecules are always present, because repressors are not always capable of blocking transcription; they alternate between inactive and active conformations.

A repressor's activity depends upon the presence of key metabolites in the cell.

- Regulation of the *trp* operon in *E. coli* is an example of how a metabolite cues a repressor (see Campbell, Figure 18.19):
- Repressible enzymes catalyze the anabolic pathway that produces tryptophan, an amino acid.
- Tryptophan accumulation represses synthesis of the enzymes that catalyze its production.



How does tryptophan activate the repressor protein?

- The repressor protein, which normally has a low affinity for the operator, has a DNA binding site plus an allosteric site specific for tryptophan.
- When tryptophan binds to the repressor's allosteric site, it activates the repressor causing it to change its conformation.
- The activated repressor binds to the operator, which switches the *trp* operon off.
- Tryptophan functions in this regulatory system as a *corepressor*.

Corepressor = A molecule, usually a metabolite, that binds to a repressor protein, causing the repressor to change into its active conformation

- Only the *repressor-corepressor complex* can attach to the operator and turn off the operon.
- When tryptophan concentrations drop, it is less likely to be bound to repressor protein. The *trp* operon, once free from repression, begins transcription.
- As concentrations of tryptophan rise, it turns off its own production by activating the repressor.
- Enzymes of the tryptophan pathway are said to be *repressible*.

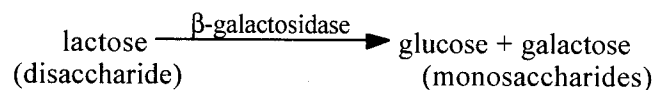
2. Repressible versus inducible operons: two types of negative gene regulation

Repressible operon = Operons which have their transcription inhibited. Usually associated with anabolic processes, (e.g., tryptophan synthesis via *trp* operon).

Inducible operons = Operons which have their transcription stimulated. Usually associated with catabolic processes.

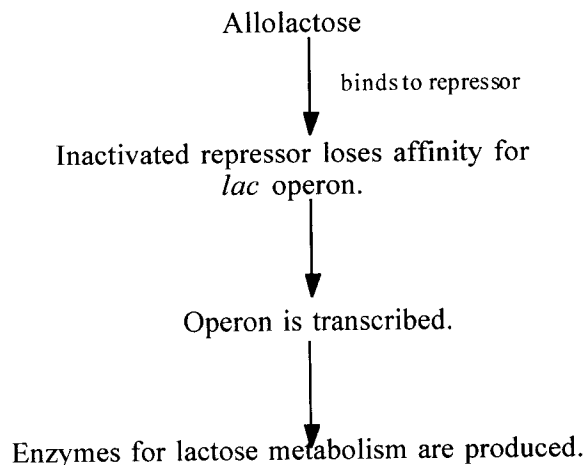
Some operons can be switched on or *induced* by specific metabolites (e.g., *lac* operon in *E. coli*).

- *E. coli* can metabolize the disaccharide lactose. Once lactose is transported into the cell, β -galactosidase cleaves lactose into glucose and galactose:



- When *E. coli* is in a lactose-free medium, it only contains a few β -galactosidase molecules.

- When lactose is added to the medium, *E. coli* increases the number of mRNA molecules coding for β -galactosidase. These mRNA molecules are quickly translated into thousands of β -galactosidase molecules.
- Lactose metabolism in *E. coli* is programmed by the *lac* operon which has three structural genes:
 1. *lac Z* - Codes for β -galactosidase which hydrolyzes lactose
 2. *lac Y* - Codes for a permease, a membrane protein that transports lactose into the cell
 3. *lac A* - Codes for transacetylase, an enzyme that has no known role in lactose metabolism
- The *lac* operon has a single promoter and operator. The *lac* repressor is innately active, so it attaches to the operon without a corepressor.
- Allolactose, an isomer of lactose, acts as an *inducer* to turn on the *lac* operon (see Campbell, Figure 18.20):



Differences between repressible and inducible operons reflect differences in the pathways they control.

Repressible Operons	Inducible Operons
Their genes are switched on until a specific metabolite activates the repressor. They generally function in anabolic pathways. Pathway end product switches off its own production by repressing enzyme synthesis.	Their genes are switched off until a specific metabolite inactivates the repressor. They function in catabolic pathways Enzyme synthesis is switched on by the nutrient the pathway uses.

Repressible and inducible operons share similar features of gene regulation. In both cases:

- Specific repressor proteins control gene expression.
- Repressors can assume an active conformation that blocks transcription and an inactive conformation that allows transcription.
- Which form the repressor assumes depends upon cues from a metabolite.

Both systems are thus examples of *negative control*.

- Binding of active repressor to an operator always turns off structural gene expression.
- The *lac* operon is a system with negative control, because allolactose does not interact directly with the genome. The derepression allolactose causes is indirect, by freeing the *lac* operon from the repressor's negative effect.

Positive control of a regulatory system occurs only if an activator molecule interacts directly with the genome to turn on transcription.

3. An example of positive gene regulation

The *lac* operon is under dual regulation which includes negative control by repressor protein and positive control by cAMP receptor protein (CRP).

CRP (cAMP receptor protein) = An allosteric protein that binds cAMP and activates transcription binding to an operon's promoter region (enhances the promoter's affinity for RNA polymerase) (see Campbell, Figure 18.21)

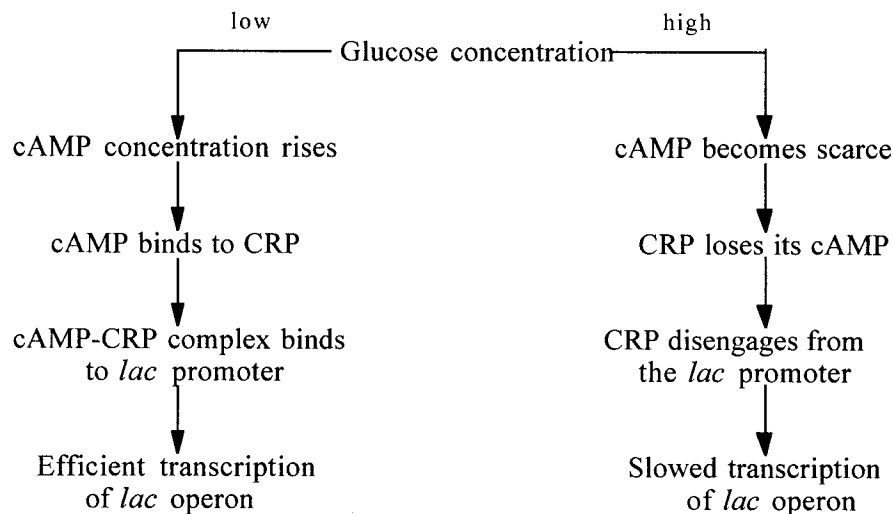
- Exists in two states: inactive (no cAMP bound) and active (cAMP bound). Only the active form of CRP can bind to the promoter to stimulate transcription.
- It is a *positive regulator* because it *directly* interacts with the genome to stimulate gene expression.
- CRP binding to a promoter is dependent on glucose concentration.

E. coli preferentially uses glucose over lactose as a substrate for glycolysis. So normal expression of the *lac* operon requires:

- Presence of lactose
- *Absence* of glucose

How is CRP affected by the absence or presence of glucose?

- When glucose is missing, the cell accumulates *cyclic AMP (cAMP)*, a nucleotide derived from ATP. cAMP activates CRP so that it can bind to the *lac* promoter.
- When glucose concentration rises, glucose catabolism decreases the intracellular concentration of cAMP. Thus, cAMP releases CRP.



In this dual regulation of the *lac* operon:

- Negative control by the repressor determines whether or not the operon will transcribe the structural genes.
- Positive control by CRP determines the rate of transcription.

E. coli economizes on RNA and protein synthesis with the help of these negative and positive controls.

- CRP is an activator of several different operons that program catabolic pathways.
- Glucose's presence deactivates CRP. This, in turn, slows synthesis of those enzymes a cell needs to use catabolites other than glucose.
- *E. coli* preferentially uses glucose as its primary carbon and energy source, and the enzymes for glucose catabolism are coded for by unregulated genes that are continuously transcribed (constitutive).
- Consequently, when glucose is present, CRP does not work and the cell's systems for using secondary energy sources are inactive.

When glucose is absent, the cell metabolizes alternate energy sources.

- The cAMP level rises, CRP is activated and transcription begins of operons that program the use of alternate energy sources (e.g., lactose).
- Which operon is actually transcribed depends upon which nutrients are available to the cell. For example, if lactose is present, the *lac* operon will be switched on as allolactose inactivates the repressor.

REFERENCES

Alberts, B., D. Bray, J. Lewis, M. Raff, K. Roberts and J.D. Watson. *Molecular Biology of the Cell*. 3rd ed. New York: Garland, 1994.

Campbell, N., et al. *Biology*. 5th ed. Menlo Park, California: Benjamin/Cummings, 1998.