Chapter 13: The Genetics of Viruses and Prokaryotes

I Using Prokaryotes and Viruses to Probe the Nature of Genes

- These organisms have advantages for the study of genetics.
  - It is easier to work with small amounts of DNA.
  - A virus typically has 1/100 of the DNA of a bacteria.
  - Bacteria have 1/1000 of the DNA of a human cell.
  - Data on large numbers are easy to obtain.
  - *E. Coli* doubles every 20 minutes.
  - Viruses replicate far more quickly than bacteria.
- Use of bacteria and viruses in genetic research propelled the science of genetics for the last century.
- Prokaryotes still continue to play a central role in the understanding of genetics, both prokaryotic and eukaryotic.
  - Eukaryotic genes are inserted into bacterial and viral genomes.
  - Newly discovered bacterial enzymes provide powerful tools for studies.

II Viruses: Reproduction and Recombination

- Most viruses are just a nucleic acid polymer and a few proteins.
- Some infect cells but postpone reproduction until certain cellular conditions are met.
- Some reproduce shortly after infecting.
- The simplest are viroids, which are made of just nucleic acid polymers.

A. Scientists studied viruses before they could see them

- The tobacco mosaic virus was the first to be discovered.
- Russian botanist Dmitri Ivanovsky provided evidence for the existence of this virus in 1892, but failed to believe his own evidence.
  - He passed extract from infected plants through a filter fine enough to capture even small bacteria.
  - The infectious agent passed through the filter.
  - Martinus Beijerinick repeated the experiment in 1898 by showing the agent could pass through agar. He called the infectious agent contagium virum fluidum.
  - This long name was later shortened to virus.
- In the late 1930's, the infectious agent was crystallized by Wendell Stanley, who won the Nobel Prize for his success.
- In the 1950's, electron microscopes showed how much viruses differed from bacteria.

B. Viruses reproduce only with the help of living cells

- Viruses are acellular (non-cellular).
- Viruses never directly arise from preexisting viruses.
- Viruses are obligate intracellular parasites.
- When they reproduce, viruses usually destroy the host cell, releasing viruses.
- Many diseases of humans, other animals, and plants are caused by viruses.
- Viruses are unaffected by antibiotics.

- Outside the cell, the viral particles are called virons.
- The genetic material is either DNA or RNA.
- This is generally surrounded by a capsid, or coat composed of protein.
- Many animal viruses have a lipid and protein membrane acquired from the host cell plasma membrane.

C. There are many kinds of viruses

- One classification is whether they are RNA or DNA viruses.
- Another is if they are double or single stranded.
- Some RNA viruses have more than one RNA molecule.
- The DNA of many viruses is linear; one family is circular.
- They can be classified by the shape of their capsid (protein coat); or on whether they have a membranous envelope around the virion.
- Another way to classify viruses is based on the type of host.

D. Bacteriophages reproduce by a lytic cycle or a lysogenic cycle

- Viruses that infect bacteria are called bacteriophages.
- These recognize their host by specific receptor proteins on the host's cell.
- The virons must get their genetic material into the cell and many do so by attaching with the tail assemblies and then injecting the DNA into the cell.
- The lytic cycle is when the virus infects the cell, takes over the cellular machinery, and then lyses the cell releasing its phage progeny.
- The lysogenic cycle is when the infected cell harbors a quiescent virus.
- *See Figure 13.2.*
- Phage that only have lytic cycles are called virulent.

The Lytic Cycle:

  - After infection, viral early genes are transcribed. Early gene products often include proteins that shut down host transcription, and stimulate viral genome replication.
  - Late stage genes code for the protein coat and an enzyme that cause cell lysis.

- The whole process takes around 30 minutes.
- On rare occasions, two viruses infect the same cell at the same time, providing the opportunity for recombination.
- *See Figure 13.3.*

The Lysogenic Cycle:

  - Bacteria harboring phages that are quiescent are called lysogenic.
  - The phage that infect in this manner are called temperate viruses.
  - The lysogenic bacteria have the *prophage* inserted into their bacterial chromosome.
  - Temperate phage have better assurance of long term survival by not killing their host, replicating with their host and occasionally spreading to other hosts.

E. Animal viruses have diverse reproductive cycles

- Almost all vertebrates are susceptible to viral infections.
- Among invertebrates, only arthropods commonly get viral infections.
- *Arborviruses* infect both insects and mammals.
- The mammal gets infected from an insect bite.

- The arthropod is called a vector for the disease transmission.
- Animal viruses include those that are just particles of protein surrounding a nucleic acid core to those with membrane. Some have DNA and some RNA.
- Animal viruses enter cells either by endocytosis of a naked virion, endocytosis of a membrane encased virus, or by fusion of a membrane encased virus with the cells membrane.
- *See Figure 13.4*
- Retroviruses such as HIV have a more complex reproductive cycle.
- *See Figure 13.5.*

F. Many plant viruses spread with the help of vectors

- Plant viruses spread horizontally or vertically.
  - Horizontally is one plant to another.
  - Vertically is parent to offspring.
- A virus has to get through the cell wall.
- Insect vectors are one way.
- Another is through damaged tissue.
- Vertical transmission is through vegetative or sexual reproduction.
- Once inside a plant cell, viruses can spread by moving through plasmodesmata.
- They code for special proteins that expand the size of the pores of the plasmodesmata.

G. Viroids are infectious agents consisting entirely of only RNA

- Theodore Diener of the U.S. Department of Agriculture discovered them and reported their existence in 1971.
- They consist of a circular, single-stranded, RNA molecule.
- They are just a few hundred nucleotides in length.
- Each is just 1/1000 the size of the smallest virus.
- No evidence yet exists that viroid RNA is translated.
- They cause a variety of plant diseases, but not much is yet known about them.
- They bare similarities to introns and have some catalytic activities.

III Prokaryotes: Reproduction, Mutation, and Recombination

A. The reproduction of prokaryotes gives rise to clones

- If a number of cells are spread on a semi-solid medium, agar, individual cells give rise to clearly visible colonies.
- If a large number of cells are spread, a confluent lawn develops after cell division.
- *See Figure 13.6.*

B. Some bacteria conjugate, recombining their genes

- *See Figure 13.7.*
- In 1946, Joshua Lederberg and Edward Tatum demonstrated the exchange of DNA between two living bacteria.
- Two strains of auxotrophic, one $met^- \ bio^- \ thr^+ \ teu^+$ and the other met^+ bio^+ thi^- leu^- were grown together and then plated on minimal medium.

- Prototrophic bacteria were recovered from these mixtures.
- Later experiments showed that the exchange of hereditary information was by direct contact and was then named conjugation.
- One bacterial cell – the recipient – had received DNA from the donor.
- Recombination then created a genotype with four wild-type alleles.
- Physical contact requires a pilus, which is a fine projection produced by the donor cell.
- *See Figure 13.8.*
- The DNA actually transfers through a conjugation tube.
- A linear portion of the genome transfers.
- Once inside, the DNA fragment recombines with a homologous region.
- Enzymes cut the DNA in two places and insert the donor region.
- *See Figure 13.9.*

C. In transformation, cells pick up genes from the environment

- More than 75 years ago, Griffith obtained the first evidence for transfer of genes between bacteria.
- This was demonstrated by genetically changing nonvirulent pneumonococci bacteria to the virulent with transforming substance from killed virulent bacteria.
- The transforming substance was later found to be DNA.
- Transformation of bacteria is when bacteria take up extracellular DNA and incorporate it.
- *See Figure 13.10.*

D. In transduction, viruses carry genes from one cell to another

- During the lytic cycle, some bacteriophage heads package the host bacteria's DNA.
- Cells infected by such viruses get a segment of another bacteria's DNA, not the viral DNA.
- This DNA sometimes recombines with the chromosomal DNA and alters the genetic composition.
- *See Figure 13.10.*

E. Plasmids are extra chromosomes in bacteria

- Plasmids are small circular extrachromosomal DNA.
- They are found in many plasmids.
- Plasmids replicate separately from the chromosomes and have an origin of replication.
- There are different types called factors:
  - Metabolic factors carry genes for unusual metabolic functions.
  - F factors carry genes for conjugation.
    - Around 35 genes, including the ones responsible for the pilus, are on the F factor plasmid.
    - Those with this plasmid are called F^+.
    - On occasion, this F plasmid inserts into the genome.
    - When this occurs, chromosomal genes are transferred during conjugation.
    - *See Figure 13.11.*

- R factors are resistance factors.
  - These carry genes that code for proteins that protect the bacteria.
  - Antibiotic resistance genes breakdown or modify antibiotics, or produce components that interfere with antibiotic activity or prevent their transport.
  - These plasmids were discovered in 1957 in Shigella bacteria, which were resistant to several antibiotics.

F. Transposable elements move genes among plasmids and chromosomes

- This is gene transport within a cell.
- It involves a segment of chromosome or plasmid DNA that can insert at new locations.
- Their movement into other genes disrupts normal function.
- Long transposable elements (about 5000 bp), which include one or more several genes, are called tranposons.
- *See Figure 13.12.*

IV Regulation of Gene Expression in Prokaryotes

- Bacteria can synthesize needed compounds more than one way.
  - For example, amino group for amino acids can come from either N_2 then ammonia, which requires a lot of energy and many enzymes or it can take an amino group from glutamine (see Table 3.2)
  - Cells must regulate how they synthesize molecules to suit their condition, environment and needs.
- Cells can control synthesis by regulating or controlling the amount of enzymes.
  - Cells can block transcription of the gene that codes for a protein.
  - Cells can hydrolyze the RNA.
  - Cells might prevent translation of mRNA.
  - Cells can hydrolyze the protein.
  - Cells can activate or inactivate the enzyme (allosteric enzymes).
  - Most extensively used of the first four is transcriptional control, example 1.

A. Regulation of transcription conserves energy

- *E. coli* prefers glucose, but when glucose availability is low and lactose is available it can use lactose.
- Lactose is a disaccharide, A β linked galactose to glucose
  - Lactose is transported into the cell by a carrier protein called β-galactoside permease.
  - It gets hydrolyzed to glucose and galactose by β-galactosidase.
  - Another enzyme is needed to protect the cell from the side effects of the β-galactosidase. It is called thiogalactoside transacetylase.
- When no lactose is present, levels of all three proteins are very low.
- When glucose is low the lactose high, synthesis of all three proteins occurs rapidly.
- If glucose levels rise again, or lactose levels drop, synthesis ceases.
- Lactose is an inducer for the inducible operon, the *lac* operon. (Constitutively expressed genes constantly produces mRNAs.)

B. A single promoter controls the transcription of adjacent genes

- Structural genes specify the primary structures (amino acid sequence) of the three genes that metabolize lactose.
- They are adjacent to each other on the bacterial chromosome.
- All are transcribed when a single promoter binds RNA polymerase.
- Therefore, their synthesis is coordinated.
- Because there is just a single promoter, all genes are transcribed onto a single mRNA called a polycistronic message.
- The whole system for coordinated control of lactose metabolism is called the *lac operon.*

C. Operons are units of transcription in prokaryotes

- Just down stream from the promoter, between the promoter and structural genes, is a DNA site called the operator.
- If a specific protein, the *lac* repressor, binds to the operator, RNA polymerase is blocked from transcribing the structural genes.
- When the repressor is not attached to the operator, mRNA synthesis proceeds.

D. Operator-repressor control that induces transcription: The *lac* operon

- *See Figure 13.13.*
- The repressor protein has two binding sites: one for the operator and the other for inducers.
- Binding of the repressor by the inducer molecules (as an analog of lactose) changes the shape of the repressor so it fails to bind the operator.
- If the concentration of the inducer drops, functioning repressor binds the operator. *See Figure 13.15.*
- The repressor protein is coded for by the regulatory gene.
- The gene that codes for the *lac* repressor is the *lac* i gene.
- It just happens to be located near the *lac* structural genes. Not all operons' regulatory genes are near the operons they control.
- Regulatory genes like *lac* i have their own promoter. The *lac* i promoter is called Pi.
- The *lac* i gene is expressed constitutively (constantly).
- *See Figure 13.16.*
- Summary of the *lac* operon control:
  - When no inducer is present, *lac* is off.
  - The regulator protein turns the operon off.
  - The *lac* i produces the repressor.
  - The operator and promoter are DNA sequences.
  - Adding inducers turns on the operator.
  - *See Figure 13.17.*

E. Operator-repressor control that represses transcription: The *trp* operon

- The t*rp* operon is repressible.
- *See Figure13.18.*
- The presence of tryptophan prevents transcription of the *trp* operon.
- Tryptophan is the final product of the enzymatic pathway, which is catalyzed by the products of the operon.
- As tryptophan concentration rises, production of transcript drops off.

- Tryptophan is called a corepressor.
- It binds to the repressor and changes the repressor's shape so that together, the repressor can bind to the operator.

F. Protein synthesis can be controlled by increasing promoter efficiency

- When glucose is high, even when lactose is available, the *lac* operon fails to transcribe frequently.
- When glucose is low, and lactose is available, *lac* structure genes are transcribed.
- Low glucose levels cause elevated intracellular cyclic AMP levels (cAMP). The molecule AMP is the monophosphate form of the familiar ATP. Cyclic AMP has a phosphodiester linkage between its own 5'and 3' carbon.
- The Cyclic AMP is a common intracellular signaling molecule.
- When glucose is low and cAMP high, cAMP binds to a protein called CRP. CRP is short for cAMP receptor protein.
- The CRP-cAMP complex binds the DNA just upstream of the promoter.
- Binding of this site *increases* rates of transcription.
- Glucose decreases the synthesis of *lac* enzymes.
- *See Table 13.2.*
- This is called catabolite repression.
- *See Figure 13.19.*

V Control of Transcription in Viruses

- Viruses have gene regulation mechanisms.
- Early genes must be transcribed before later ones.
- Temperate viruses need to regulate when to undertake a lytic cycle.
- Lambda is a temperate phage.
  - When host bacteria are growing in rich medium, lambda enters a lytic cycle.
  - When the host is less healthy, lambda lays low.
  - The decision is done by a genetic switch.
    - Two regulatory proteins compete for two operator/promoter sites on the phage DNA.
    - One operator controls the lytic gene activities, the other, lysogenic cycles.
    - The two regulatory proteins have opposite effects on the two operators.
    - cI represses the lytic operator/promoter and activates the lysogenic operator/promoter.
    - Cro activates the lytic operator/promoter and represses the lysogenic operator/promoter.
    - The relative amounts of cI versus Cro determine the outcome.
    - *See Figure 13.20.*

VI Prokaryotic Genomes

- Viral genomes were first to be sequenced.

- *Haemophilus influenzae* was the first free-living organism to be completely sequenced. *See Figure 13.21*.
- This has revealed details of how bacteria allocate and organize their genes.
- Three types of information can be obtained from a genomic sequence.
  - Open reading frames can be recognized by promoter regions and start and stop codons.
  - Amino acid sequences can be deduced.
  - Gene control sequences of promoters and terminators can be identified.

A. Functional genomics relates gene sequences to functions.

- *Haemophilus influenzae*, which infects humans, has a circular chromosome of 1,830,137 base pairs.
- It has 1,743 protein coding regions.
- When it was finished being sequenced, only 58% of the proteins coded for were previously known ones.
- Roles for most of the unknown proteins have now been identified.
- This learning process is called annotation.
- Functional genomics is the assignment of roles to genes and the description of how they work in the organism.
- In addition to *H. influenzae*, *Mycoplasma genitalium* (580,070 base pairs) and *E. coli* (4,639,211) base pairs) have been completed.

B. The sequencing of prokaryotic genomes has medical applications

- Scientists are discovering genes for proteins in infectious disease causing bacteria – potential targets for new drugs.
- New vaccines might be possible as cell surface antigen coding genes are discovered.

C. What genes are required for cellular life?

- There are some universal genes needed by all organisms.
- There are some universal gene segments, like those coding for an ATP binding site.
- *M. genitalium* has just 470 genes.
- Some genes are dispensable under certain conditions. For example, those for lactose utilization are not needed for *E. coli* grown on glucose.
- Using mutagens to knock out genes, it has been determined that *M. genitalium* can survive in the laboratory with just 337 genes!
- *See Figure 13.22.*