

GENETIC BASICS OF VARIATIONS IN BACTERIA

Background knowledge: Basic biochemistry of DNA; fundamentals of gene structure, expression, and regulation; structure and growth of bacteria; concept of antibiotics

Suggested reading: *Medical Microbiology*, 5th ed., Murray et al., Chapter 5

Introduction

A. Significance.

The study of bacterial genetics has provided much of the conceptual foundation for understanding the structure, function, and expression of genes. The detailed knowledge of genetic mechanisms in bacteria has also resulted in immensely powerful and sophisticated tools for studying the molecular biology of a wide variety of prokaryotic and eukaryotic organisms. Because many of these tools make it possible to do detailed genetic studies on previously intractable bacterial species, there has been considerable interest and recent exciting progress in studying the genetic basis of antibiotic resistance and pathogenesis.

We know quite a bit about the molecular basis of genetic variation in bacteria. The purpose of these lectures is to provide you with a basic understanding of the ongoing variation and evolution of bacteria in nature. Sometimes, with appropriate selective pressure, new genes and elements can evolve and spread rapidly. One of the most deadly examples is the development of genetic elements that encode resistance to several antibiotics and transfer easily from one bacterial cell to another. Such elements have caused severe problems in the treatment of infectious bacterial disease. In other cases, the genetic changes are programmed by the bacterial cell, as in the case of antigenic variation of certain pathogens. We are just beginning to understand the relevance of these mechanisms to disease. There is no question that the progress will be rapid and that the ultimate objective of using this knowledge to provide new approaches for the control of infectious disease will be realized during the course of your careers.

B. An overview of genetic variation in bacteria.

When a bacterial cell divides, the two daughter cells are generally indistinguishable. Thus, a single bacterial cell can produce a large population of identical cells or clone. On solid medium, a clone is manifested as an easily isolated colony. Occasionally, a spontaneous genetic change occurs in one of the cells. This change (**mutation**) is heritable and passed on to the progeny of the variant cell to produce a **subclone** with characteristics different from the original (**wild type**) parent. This is termed **vertical inheritance**. If the change is detrimental to the growth of the cell, the subclone will quickly be overrun by the healthy, wild type population. However, if the change is beneficial, the subclone may overtake the wild type population. This is an example of how evolution is directed by **natural selection**.

Spontaneous mutations are of two classes: (1) **point mutation**, or change of a single nucleotide, and (2) **DNA rearrangement**, or shuffling of the genetic information to produce insertions, deletions, inversions, or changes in structure. DNA rearrangements can affect a few to several thousand nucleotides. Both types of mutations generally occur at a low frequency (roughly once in 10^6 to 10^8 cells for any particular gene) and lead to a continuous, slow evolution of bacterial populations.

Bacterial variation can also occur by **horizontal transfer** of genetic material from one cell to another. Consider two cells from different populations: bacterium B has features distinct from those of bacterium A. There are three possible mechanisms for transferring a trait from B to A: (1) **transformation**, release and uptake of naked DNA; (2) **transduction**, packaging and transfer of bacterial DNA by viruses, and (3) **conjugation**, bacterial mating in which cells must be in contact. For all three processes, the transferred DNA must be stably incorporated into the genetic material of the recipient bacterium. This can occur in two ways: (1) **recombination**, or integration of the transferred DNA into the bacterial chromosome; or (2) **establishment of a plasmid**, i.e., the transferred material essentially forms a minichromosome capable of autonomous replication.

Mutation and gene transfer work together to accelerate the rate of bacterial evolution. The spontaneous changes required to produce a new function (e.g. antibiotic resistance) may occur at a low frequency. However, once the function has developed, it can readily spread to other bacterial populations. The limitation is the probability and efficiency of gene exchange between different bacteria. Under certain conditions, gene exchange is very efficient.

C. Bacteria as a genetic system.

Most bacterial cells contain a single chromosome. The chromosome of *Escherichia coli* is a double stranded DNA circle of about 5 million base pairs encoding approximately 5,000 genes. Because there is only one chromosome, each gene (with occasional exceptions) is present in only one copy. Thus, bacteria are **haploid**. One important consequence of having a haploid genome is that genetic changes have an immediate effect on the phenotype or properties of the bacterial cell. This is a highly desirable characteristic for the isolation of mutants in the laboratory. The ability of bacterial cells to produce colonies on solid medium makes it possible to physically separate and identify mutant clones of bacteria. The short generation time and ability to produce large numbers of progeny make it possible to isolate virtually any kind of mutation. In nature, these properties mean that evolution is rapid.

D. Genetic changes occurring within a single cell.

Point mutations, single nucleotide changes in the DNA, can have a number of consequences. In coding regions they may alter an amino acid in a polypeptide. The effect may be deleterious (inactivation or lower activity) or beneficial (enhanced or new activity). Changes in the targets for several antibiotics can result in functional proteins that are no longer sensitive to the antibiotic (e.g. certain mutants of DNA gyrase are resistant to quinolones). In non-coding regions, point mutations can affect a variety of signals for expression and regulation of a gene. Often a gene from one bacterial species is not expressed after transfer to another bacterial species because of differences in promoters, ribosome binding sites, codon usage, etc. Spontaneous single nucleotide changes can result in the generation of a functional gene. Such a process may account for the relationship of cholera toxin of *Vibrio cholerae* and enterotoxin of *E. coli*. The toxin proteins are highly homologous, but the genes are only expressed in the bacterial species from which they were isolated. Point mutations can occur spontaneously through errors in DNA replication, or they can be induced by environmental **mutagens** that act directly on the DNA. Chemicals that damage DNA can also induce mutation indirectly. Bacteria encode several genes to help repair damage to the DNA. If the amount of damage is considerable, repair genes that are normally repressed become active. This phenomenon is known as the **SOS response**. One of the induced genes causes a reduction in the proofreading of DNA polymerase and leads to an increase in mutation. This system may have evolved as a mechanism for **hyperevolution** to increase the possibility of generating mutants that can survive in a dangerous environment.

Other spontaneous events have greater consequence for the structure of the chromosome. Some mutants were found to have large insertions, deletions, or inversions in the chromosomal DNA. Of primary importance to these types of mutations are transposable genetic elements known as Insertion Sequences, or **IS elements**. Scattered throughout the chromosome are active sequences of DNA with the remarkable ability to jump into other regions of the chromosome. Different bacteria have different numbers and different types of IS elements. Typically these elements are 700-3000 bp in length, have inverted repeat sequences at their ends, and encode 1 or 2 proteins responsible for translocating the element to a new location. Transposition is spontaneous and occurs at a frequency comparable to that of point mutations. (About 1 in 10^6 - 10^8 organisms has acquired a mutation in any particular gene.) Most IS elements have very low target specificity and can insert virtually anywhere in the DNA. After insertion of the transposable element into a new location, depending on the element, a copy may be left behind in the original position (**replicative transposition**), or the original copy may be excised and transposed to the new location (**conservative transposition**). Insertion of an element into a gene destroys that gene and can have drastic consequences for the expression of other genes in the same transcriptional unit. Some IS elements carry promoters and can activate a quiescent gene in a single step. The movement of IS elements is often associated by formation of large deletions, inversions, and generation of small circular DNA molecules. These rearrangements can provide the raw material for new genes or new operons. The process is finished by the more subtle fine tuning of point mutations. One important property of IS elements, particularly relevant to the

spread of antibiotic resistance genes, is the ability of two identical IS elements flanking a gene to move that gene to another position in the chromosome or any other DNA molecule in the cell. Such an arrangement is called a **composite transposon**. The spontaneous and random formation of transposable genes and their subsequent insertion into plasmids or bacteriophages generate the potential for rapid dissemination of these genes to other bacteria.

Gene exchange between bacteria

A. Transformation.

The uptake of naked DNA molecules and their stable maintenance in bacteria is called transformation. The phenomenon was discovered in 1928 by Griffith, who was studying the highly virulent pathogen *Streptococcus pneumoniae*. He showed that injecting into mice a mixture of heat killed virulent (smooth) *S. pneumoniae* with a live attenuated (rough) strain led to the development of a live virulent strain, which ultimately killed the mouse. Avery, MacCleod, and McCarty purified the transforming substance and identified it as DNA. This experiment was the first to demonstrate that DNA was the genetic material. It was also the first discovery of gene transfer between bacteria. Since then, other bacteria, including certain species of *Haemophilus*, *Bacillus*, *Actinobacillus*, and *Neisseria*, have been found to be naturally transformable. These bacteria have developed highly specialized functions that will bind DNA fragments and transport them into the cell. These mechanisms can be quite distinct. In the case of *Bacillus subtilis*, any DNA can be taken up. *Bacillus* and *Streptococcus* unwind the DNA and transport only a single strand. In contrast, *Haemophilus*, *Actinobacillus*, and *Neisseria* require a specific sequence to be present on the DNA fragments and transport double-stranded DNA fragments. Transformable organisms take up DNA when they are in a **competent** state. In *Bacillus*, this state is triggered by small diffusible molecules whose concentration indicates when the culture has reached a certain density. In *Haemophilus*, competence is induced by nutritional starvation. These signals somehow trigger the expression of proteins that enable the cells to bind and take up DNA. In nature, the DNA to be taken up is thought to be released into the environment by lysis of bacterial cells. Transformation is probably the least efficient mechanism of gene transfer because naked DNA is sensitive to nucleases in the environment. In the laboratory, mutant strains can be transformed to wild type by the addition of purified DNA extracted from a wild type strain. The process depends on the DNA and is sensitive to the addition of DNase. Some organisms that are not naturally transformable, like *E. coli*, can be made competent for transformation by treating the cells with CaCl₂ or placing them in an electric field (electroporation). The ability to introduce DNA into bacterial cells in the laboratory is the basis for "reverse genetics," in which a gene is first cloned, mutated in vitro, and reintroduced into the bacterial cell to study the resulting phenotype.

B. Homologous recombination.

How is the DNA fragment stably incorporated into the bacterial genome once it is taken up by the cells? One very efficient mechanism is **homologous recombination**. Two molecules of nearly identical sequence can readily exchange segments. In bacteria, the key protein for homologous recombination is **RecA**.

Several molecules of RecA bind to single stranded DNA and allow it to search another double stranded DNA molecule for closely related sequences. If such a region is found, RecA promotes precise alignment of the molecules, which is then followed by breakage and rejoining of the individual strands. A similar process in another region of the fragment results in an exchange of DNA between the fragment and the chromosome: on the incoming fragment, the segment between the two recombination events (crossovers) has replaced the original segment in the chromosome. In this way, incoming DNA encoding an altered gene can replace the original gene in the chromosome. This is known as **allelic exchange**. The transformed bacterial cell will then express a new phenotype (e.g. virulence, antibiotic resistance). Homologous recombination completely depends on the RecA protein; it does not occur in a *recA*⁻ mutant. The probability of recombination increases with increasing lengths of the homologous regions flanking the segment to be inserted. Homologies of several hundred base pairs can yield high frequencies of recombination.

Analysis of penicillin resistance in *Streptococcus pneumoniae* has revealed variant penicillin-binding proteins (PBPs) with lower affinity for penicillin. Analysis of the sequences of these variant PBP genes shows that they are mosaics resulting from substitution of segments of PBP protein genes from other distantly related streptococci (*S. mitis*, *S. oralis*) acquired by DNA transformation and homologous recombination.

If two genes are close together, they will be inserted together at a high frequency. Sometimes homology can be found only in distinct small segments of the incoming DNA. The two crossovers would then result in a **deletion-substitution**: the chromosomal DNA between the crossovers is replaced by an unrelated piece of the incoming DNA. This latter process permits the incorporation of a new gene not previously found in the chromosome (e.g. antibiotic resistance or toxin).

C. Transformation and pili antigenic variation in *Neisseria*.

The pathogenic *Neisseria* species are *Neisseria gonorrhoeae* (gonococci) and *Neisseria meningitidis* (meningococci), which are responsible for gonorrhea and meningitis, respectively. These organisms are Gram negative, non-motile, diplococci that prefer to grow aerobically in an atmosphere slightly enriched with CO₂. Both the gonococci and the meningococci exist only in the human host; there is no other reservoir. Not surprisingly, they have evolved highly sophisticated mechanisms for colonization and invasion of human host tissues and evasion of host defenses. Consequently, untreated infections, especially those of the meningococci, can lead to a variety of serious complications, including death. Advances in DNA technology and the ability to genetically manipulate *Neisseria* in the last decade have allowed explosive progress in our understanding of the molecular basis of pathogenicity in these organisms. While there is much left to learn, these studies have already revealed some fascinating details.

The first step for gonococci to establish an infection is attachment to mucosal epithelial cells. The gonococci have thin, hairlike protein fibers, known as **pili** (or **fimbriae**), on their surfaces. Pili are composed primarily of a single protein subunit (pilin) and possibly minor accessory proteins in low amounts.

These structures are responsible for tight binding of the bacteria to non-ciliated mucosal cells of the urethra or vagina. The tight binding prevents the gonococci from being washed away by vaginal discharge or urine. The presence of pili has also been shown to inhibit phagocytosis by neutrophils. Pili are essential for virulence. Infection is not established in human volunteers inoculated with pil mutants of *N. gonorrhoeae*.

Another contributor to colonization by the gonococci is a set of cell surface proteins, or adhesins, known as Opa or PII. These proteins also promote adherence of gonococci to each other to form aggregates and confer resistance to the bactericidal activity of serum. Recently, it has been shown that specific Opa proteins enable the gonococci to enter the epithelial cells in a vacuole. Once inside, the bacteria multiply and are transported to the base of the non-ciliated cells, where they are released by a "reverse-phagocytosis" into the subepithelial connective tissue. There the organisms continue to multiply, causing damage to tissue and in some cases entering the bloodstream.

One might predict that an immune response mounted against the pili proteins would be effective in preventing or ending infection. Indeed, pilin proteins are highly immunogenic. However, the gonococci have evolved effective ways of evading the immune response. First, they secrete an IgA protease which cleaves the heavy chain of the IgA1 isotype and inactivates it. While it has not been proved that this is an important virulence factor, it seems reasonable that it has an effect on the survival of gonococci in the host. A strategy that is clearly important is the ability to switch production of pilin proteins on or off (**phase variation**) and to express different antigenic forms of pilin proteins (**antigenic variation**). A single clone of *N. gonorrhoeae* can give rise to variants expressing pili antigenically distinct from that of the original clone. The variants arise at a high frequency in the population (approximately 10^4).

The mechanisms of phase variation and antigenic variation rely on multiple copies of the pilin genes in the *Neisseria* chromosome. The different pilin proteins expressed from a single clone can vary from 18 to 24 kilodaltons (kD) in mass. All the proteins have a highly conserved aminoterminal sequence thought to be crucial for transport of the pilin through the membrane and interaction of individual pilin subunits with each other. The carboxyterminal domain is characterized by two conserved Cys residues that form a disulfide bond. This domain is the most immunogenic region of the pilin and is highly variable in amino acid sequence. In addition, there are five other small regions of variable amino acid sequences within the protein. These regions can be thought of as mini-cassettes of variation. Once a pilin gene was cloned, hybridization analysis showed that there are 10 or more copies of related genes in the *Neisseria* chromosome. These copies were generated by gene duplication and point mutations. However, only one of the genes is expressed. The expression locus is called *pilE*. The other non-expressed copies are scattered around the chromosome. These silent loci are called *pilS1*, *pilS2*, etc. Only the *pilE* locus has a promoter and the complete gene, including the conserved aminoterminal coding region for pilin. The silent loci have no promoter and are missing the coding sequences for the aminoterminal domain. When a *Neisseria* clone expresses a different pilin, all or part of a *pilS* gene replaces the homologous region of the *pilE* gene, thus generating a new variant (antigenic variation). Six variable domains in 10 *pilS*

structural genes gives a theoretical probability of about 10^4 different pilin proteins. Some cassettes result in the formation of a pilin gene whose product is processed differently and cannot form pili. These variants have a Pil⁻ phenotype (phase variation). In addition, genes such as *pilC* that regulate pilin expression are also subject to phase variation and lead to Pil⁻ strains. Because the presence of pili discourages phagocytosis, the Pil⁻ forms are probably important for the invasive stage of gonococcal infection.

Insertion of *pilS* sequences into the *pilE* locus depends on *recA*, indicating a role for homologous recombination. However, the recombination event usually has the appearance of being **non-reciprocal**; i.e., the replaced *pilE* sequences are not found in the contributing *pilS* locus. One explanation, for which there is convincing evidence, is that the *Neisseria* cell receives new *pilS* information by transformation. The gonococci are known to be continuously competent for transformation. During normal growth of the cells, a portion of the population lyses and releases DNA into the medium. This DNA contains the various *pilS* loci. Thus, *pilS* sequences that have been taken up by a competent cell can recombine with the resident *pilE* gene to generate a new variant. This exchange would have no effect on the *pilS* locus of the transformant, so it would appear that the recombination is non-reciprocal. As predicted by this model, cultures grown in the presence of DNase are far less able to express different pilin variants. The same result is seen when transformation-defective mutants are used.

In summary, pili are important virulence determinants of gonococci. Similar structures have been found for the meningococci, but they confer different tissue specificity. The pili are important for initial attachment. The Opa adhesin is important for tight binding and invasion of the mucosal epithelial cells. Opa undergoes antigenic variation also, but by a different mechanism. The antigenic variation displayed by both sets of proteins has two important consequences. (1) The most obvious is that it helps the organism survive the host immune response. Antibodies specific for one form of pilin or Opa protein are not effective against another form. Volunteers that have been vaccinated with one type of pilin are not protected against naturally varying strains, but they are protected against strains genetically altered to produce the same pilin type (e.g., *recA* mutants). (2) Less obvious, but perhaps equally important, is that stochastic variation allows for the selection of attachment proteins that are most effective for a particular tissue. For example, invasion of cultured epithelial cells selects for bacteria that attach better. Different proteins enhance invasion of other cell lines. Thus, variation may give the bacterial population added flexibility. It allows for selection of variants that produce the colonization proteins most suited for different individuals or other sites of the body.

D. Bacteriophages: Viruses of bacteria.

Viruses are molecules of nucleic acid wrapped in a protein coat (and sometimes lipid). They are found outside the cell, but they can only replicate within a living cell. Viruses that infect bacterial cells are called **bacteriophages** or **phages**. Bacteriophages represent a highly successful strategy for survival in nature. They are ubiquitous - probably all bacteria have phages that can infect and replicate within them. They are medically relevant because they can mediate the efficient transfer of genes between bacteria, including genes for antibiotic resistance and toxins. Some bacteriophages have the ability to convert harmless bacteria into pathogens.

Many varieties of bacteriophages have been identified. The nucleic acid can be DNA or RNA, double stranded or single stranded, circular or linear. The genomes may be small, encoding as few as 5 or 6 genes, or large enough to encode more than 200 genes. The complexity of the genome is usually reflected in the protein coat of the virus particle. Some bacteriophages, like ϕ X174, have simple icosahedral structures; others, like T4, have icosahedral heads attached to contractile tails to provide an elaborate mechanism for attaching to the bacterial cell surface and injecting the DNA. However, all bacteriophages have the same basic life cycle. They infect the bacterial cell, subvert the cell's machinery to replicate themselves, then lyse the cell to release hundreds of new bacteriophage particles into the environment. This process is known as the **lytic cycle**. If other sensitive bacteria are nearby and they happen to come in contact with a newly released particle, the lytic cycle will repeat. A growing bacterial culture can quickly be overwhelmed by the production of bacteriophage and may completely lyse within a short time. If a single bacteriophage particle is added to a densely seeded culture of sensitive bacteria on solid medium, replication of the bacteriophage and lysis of the cells will produce a zone of clearing known as a **plaque**. This is a convenient assay for detecting and quantitating bacteriophages and isolating bacteriophage mutants.

To initiate infection, a bacteriophage particle must adsorb tightly to the surface of the cell. Different bacteriophages have evolved to recognize various specific molecules (proteins or polysaccharides) on the surface of the bacterial cell. In addition to their normal functions for the cell, these molecules act as receptors for the bacteriophages. The bacteriophage-receptor interaction is highly specific, and bacteriophages can be used to discriminate between closely related strains of the same bacterial species. An example of such an assay, called **phage typing**. When a bacteriophage infects a cell, it introduces its nucleic acid into the cytoplasm and leaves its protein coat on the surface. (This finding by Hershey and Chase in 1952 was the first generally accepted evidence that DNA is the genetic material.) Inside the cell, the bacteriophage follows a highly regulated, strictly defined program of development. For bacteriophage T7, which infects *E. coli*, the first genes to be expressed are those involved in DNA replication, shutoff of bacterial genes, degradation of the bacterial chromosome, and exclusive expression of bacteriophage genes. As the cell begins to accumulate hundreds of copies of T7 DNA, middle and late gene products assemble into heads and tails. The heads package DNA and attach to the tails. The last step is lysis of the cell and release of the new bacteriophage particles. All this occurs in less than 20 minutes.

Some bacteriophages, known as **temperate bacteriophages**, have two strategies for replication. One is the lytic cycle described above. The second strategy, known as **lysogeny**, does not result in death of the infected cell. These bacteriophages produce repressors that turn off their lytic genes for replication, production of heads and tails, and lysis. The bacteriophage DNA becomes a benign molecule that coexists with the chromosome known as the **prophage**. In the case of bacteriophage λ , a well studied temperate bacteriophage of *E. coli*, the DNA forms a circle and **integrates** into the chromosome using λ -specified recombination enzymes. The integrated form is replicated passively as part of the bacterial chromosome. The lysogenic state is reversible. When the repressor is inactivated for one of several reasons, including induction of the SOS response caused by DNA damage, the λ DNA is induced to **excise** from the chromosome and commence the lytic cycle. This results in lysis of the cell and the release of newly synthesized bacteriophage λ particles.

E. Transduction.

Bacteriophages have the ability to transfer genes from one bacterial cell to another, a process known as transduction. There are two varieties of bacteriophage-mediated gene transfer: **generalized transduction** and **specialized transduction**.

Generalized transduction occurs as a result of the lytic cycle. In the process of packaging bacteriophage DNA, the head structures of some bacteriophages will package random fragments of the bacterial chromosome. Thus, the lysate contains two kinds of particles that differ only in the kind of DNA they contain. Most of the particles contain viral DNA. When these inject their DNA, the lytic cycle will repeat and new bacteriophage particles will be produced. A small fraction of the particles, possibly as high as 1%, contain fragments of the bacterial chromosome in place of the bacteriophage DNA. When one of these particles injects its DNA into the cell, the cell is not killed. The newly introduced DNA contains only bacterial genes and is free to recombine with the chromosome. Some transducing bacteriophages can introduce 100-200 kilobases of DNA. Because the bacterial fragments that are packaged are essentially random, virtually any bacterial gene of the bacterial chromosome can be transduced (hence, the term "generalized" transduction). Entire plasmids can be transduced by phages. Some plasmids, notably those encoding antibiotic resistance in staphylococci have evolved signals to allow efficient packaging by phage particles and subsequent transfer by transduction. Another element (**pathogenicity island**, see below) in the *Staphylococcus* chromosome which codes for toxic shock toxin senses the presence of an infecting phage, excises, replicates, and is efficiently packaged. **Transducing particles**, like bacteriophage particles, are stable in the environment for long periods of time. Thus, bacterial genes can be stored in the environment and transduction of a bacterial cell may occur long after the original bacterial population was lysed by the bacteriophage infection. Studies on dissemination of antibiotic resistance have revealed generalized transduction to be a significant mechanism of gene transfer in nature.

Specialized transduction requires a temperate bacteriophage. In this class of transduction, a bacterial gene becomes associated with the bacteriophage genome (e.g. by recombination). When such a bacteriophage lysogenizes a new bacterial host, it brings with it the associated bacterial gene. Because it is a bacterial gene, its expression is not turned off by the bacteriophage repressor that inhibits expression of the lytic functions. A well known example is the b phage of *Corynebacterium diphtheriae*, which carries the gene for diphtheria toxin. Cells of *C. diphtheriae* that are non-lysogenic for the bacteriophage are incapable of causing diphtheria. *C. diphtheriae* cells that carry the b phage in the chromosome express the gene for diphtheria toxin and produce disease. Other examples include changes in the O-antigens of *Salmonella*, antibiotic resistance genes, erythrotoxic toxin in *Streptococcus pyogenes*, the tissue destroying α -toxin of *Staphylococcus aureus*, enterotoxin of *E. coli*, cholera toxin of *Vibrio cholerae*, and neurotoxin by *Clostridium botulinum*. Such alterations of the properties of a bacterial cell by lysogeny with a temperate bacteriophage are known as **lysogenic conversion**.

F. Structure and properties of plasmids.

Plasmids are small, double-stranded DNA molecules that are maintained in the cell in addition to the bacterial chromosome. Usually they are circular, but some linear plasmids also exist. These extrachromosomal elements are ubiquitous in the bacterial world. Nearly all bacterial species have one or more plasmids associated with them. All known plasmids are dispensable for basic growth and division of the bacterial cell; yet they are extremely valuable to the success of various bacteria in nature because they are known to encode a wide variety of genes that confer a competitive advantage in certain environments. For example, plasmids have been found to encode bacteriocins (proteins that kill other bacteria), metabolic pathways that allow a cell to use exotic compounds like toluene or pesticide 2,4-D as carbon sources, and resistance to nearly all known antibiotics. It is now recognized that a variety of pathogenicity factors (toxins, lysins, adhesins for attachment to specific cells, iron scavenging systems, resistance to complement, ability to invade tissue cells) are encoded by plasmids. A large part of their success in nature is due to their ability to promote efficient transfer between bacteria by conjugation. This feature has important ramifications with respect to the development, spread, and control of disease-producing bacteria.

G. Replication and maintenance of plasmids by the bacterial cell.

Plasmids range in size from about 1000 base pairs to greater than 1,000,000 base pairs. Some plasmids have low copy number (1 per bacterial chromosome); others are multicopy (20-30 per chromosome). Multicopy plasmids are generally small. Because they are dispensable, plasmids have evolved highly sophisticated mechanisms to guarantee their replication and survival in the cell. Plasmid-encoded genes are responsible for initiating replication at least once before cell division so that both daughter cells contain the plasmid. For most plasmids, the initiator gene codes for a protein that triggers DNA replication at a specific sequence known as the **origin of replication**. Once the origin is activated, the host cell DNA polymerase replicates the plasmid. In many plasmids, multiple copies of specific binding sites for the initiator (iterons) constitute part of the origin of replication. When the iterons are bound to initiator proteins, a nearby region of the double stranded DNA comes apart to allow the DNA replication

machinery of the host cell to enter the duplex and begin replication. Once replication of the plasmid has occurred, the initiator bound to the iterons of the newly replicated plasmids mediates a coupling of the plasmids. In this configuration, the plasmids are blocked from further replication. They separate at cell division and the cycle repeats. Other plasmids encode a repressor that inhibits transcription of the initiator gene. As the cell grows prior to cell division, the repressor concentration decreases, and the initiator gene is expressed. This event triggers plasmid replication. After replication, the increase in plasmid copies leads to an increase in repressor concentration, which then turns off the initiator gene. Thus, a relatively simple circuit is required to provide all the functions necessary to initiate and control plasmid replication. These functions are referred to as the **replicon** of the plasmid. The idea that a plasmid can control its own replication by controlling its initiation is known as the replicon hypothesis and was proposed in 1963 by Jacob, Brenner, and Cuzin.

Replication of the plasmid is not always sufficient to guarantee that the plasmid will segregate properly to the daughter cells after cell division. This is less of a problem for multicopy plasmids, for which there is a high probability that at least one copy will segregate to each of the daughter cells. For low copy plasmids, it is a serious problem that has been solved by **partition functions**. These plasmid-encoded proteins have a function similar to that of the centromere of a eukaryotic chromosome. The partition proteins bind to the plasmids, assemble them into pairs, and attach them to some as yet unknown structure in the bacterial cell. At cell division, as the septum is formed, the plasmid pairs separate, with each daughter cell receiving one plasmid.

H. Transfer of plasmids between bacterial cells by conjugation.

One of the most important properties of plasmids is their ability to transfer to other bacterial cells. There are two categories of **conjugative plasmids** with respect to transfer: (1) **self-transmissible plasmids**, which encode all the genes necessary to promote cell-to-cell contact and transfer of DNA, and (2) **mobilizable plasmids**, which do not promote conjugation, but can be efficiently transferred when present in a cell that contains a self-transmissible plasmid. The self-transmissible plasmids are usually large. They code for 20-30 proteins specifically required for bacterial cells to form a mating pair, develop a small pore, and transfer plasmid DNA through the pore from one cell to the other. One of the best studied self-transmissible plasmids is the 100 kilobase F plasmid of *E. coli*. About one-third of the F plasmid is devoted to the transfer genes. They encode a long tubular structure, called the sex pilus, which extrudes from the cell. When the sex pilus contacts a cell that does not have the F-plasmid (F⁻ cell, recipient cell), it attaches to its surface. The sex pilus then depolymerizes, draws the two cells together, and creates a small pore that fuses the two cells. One strand of F-plasmid is nicked at a specific sequence, the origin of transfer (*oriT*), by one of the transfer proteins. The nicked single-strand is then unwound from the F-plasmid and threaded through the conjugation pore to the recipient cell. In the donor cell, the transferred single strand of DNA is replaced by synthesis of a new one. Once the single strand of DNA has been transferred to the recipient cell, a complementary strand is synthesized. Thus, when an F⁺ donor cell mates with an F⁻ recipient cell, the result is two F⁺ cells, i.e. the F plasmid has replicated in the process. Conjugation is

efficient. Mixing F^+ donor cells with F^- recipient cells leads to transfer of F to nearly all of the recipient cells in a few minutes.

Mobilizable plasmids are generally small. One example is the 8 kilobase IncQ plasmid, which encodes only three mob genes devoted to its transfer, along with replication control genes and antibiotic resistance genes. When IncQ plasmid is the only plasmid in the cell, it cannot promote conjugation and DNA transfer. However, when it is present along with a self-transmissible plasmid (e.g., RK2), both plasmids can transfer at high efficiency. IncQ plasmid transfers by a mechanism similar to that of RK2. Once the mating pair and conjugation pore has been established by proteins encoded by the RK2 plasmid, IncQ Mob proteins deliver the plasmid to the conjugation pore and nick a single strand of the IncQ plasmid at its origin of transfer. The single strand is transferred and the complementary strands are synthesized.

If a plasmid lacks genes for self-transfer or mobilization, it can still be transferred by recombination. When two circular DNA molecules recombine, they form a single circular DNA molecule known as a **cointegrate**. This can occur by homologous recombination if the two plasmids share a region of identical sequence, such as an IS element. Cointegrate formation is also a common property of transposition by IS elements. If an IS element on a plasmid transposes to another plasmid, one of the possible outcomes is the generation of a cointegrate. It is very common for plasmids to carry IS elements, and the F plasmid is known to have several. When a cell contains the F plasmid and a non-self-transmissible, non-mobilizable plasmid, occasionally (approx. once in every 10^6 cells) the F plasmid will form a cointegrate with the smaller plasmid. When this recombinant F plasmid transfers to a recipient cell, it carries the smaller plasmid along with it. In the recipient cell, the cointegrate can resolve back into two plasmids by recombination. This process is much less efficient than plasmid mobilization because it requires a recombination event first.

I. Integration of F plasmid into the bacterial chromosome and transfer of chromosomal genes.

In addition to recombination with other plasmids, the F plasmid (and other self-transmissible plasmids) can recombine with the bacterial chromosome to form a cointegrate with the chromosome. Now when F transfers to the recipient cell, it also transfers the bacterial chromosome. Thus, it is possible for self-transmissible plasmids to transfer chromosomal genes from one bacterial cell to another. These types of donor cells are called **Hfr** cells because they transfer chromosomal markers at high frequency. Because the chromosome is so large, it requires about 100 minutes to transfer it entirely. However, it is difficult for mating pairs to remain intact for that long, so it is extremely rare to transfer the complete chromosome. Nevertheless, the bacterial genes closest to the integrated plasmid are transferred with high efficiency. Before sequence analysis, it was possible to map the relative position of genes by deliberately breaking mating pairs apart at specific intervals and assaying which genes have been transferred. This type of experiment is known as interrupted mating. Note that Hfr matings do not result in transfer of the complete F plasmid to the recipient cell. The recipients become neither F^+ nor Hfr.

Occasionally the integrated F plasmid in an Hfr cell will excise by recombination. Sometimes it forms the original F plasmid; other times it excises along with adjacent bacterial genes. The resulting plasmid has therefore "captured" some bacterial genes. Such plasmids are called F' (**F-prime**) plasmids. The transfer of bacterial genes to a self-transmissible plasmid allows them to be transferred efficiently to other bacteria. Because the genes are attached to a plasmid replicon, it is not necessary for them to recombine with the chromosome in the recipient cell to be maintained. This process constitutes a highly potent mechanism for bacterial cells to share genetic information.

Transposable elements, antibiotic resistance, and pathogenicity

A. Rapid development of antibiotic resistance.

The use of antibiotics has been essential for the treatment of bacterial infectious disease. Unfortunately, the efficacy of any antibiotic is undermined by the development of bacterial strains that are resistant to it. Bacterial resistance to some antibiotics can occur by point mutations in the gene that encodes the target polypeptide. For example, the quinolones inhibit bacterial DNA gyrases and block DNA replication. Resistance to quinolones can be achieved by mutations that alter DNA gyrase to block binding to the quinolones. Resistance to two different antibiotics would theoretically require mutations in at least two genes and should be relatively rare. Most alarming, however, is the development of highly transmissible plasmids (R-plasmids) that carry resistance genes for several antibiotics. These plasmids are responsible for rapid dissemination of antibiotic resistance to a wide variety of bacteria. The problem is serious enough that unless new varieties of antibiotics are discovered or synthesized, antibiotic therapy to treat certain types of infections may be jeopardized. The study of R-plasmids led to the discovery of a class of transposable elements, called transposons, which have played a critical role in the development of R-plasmids.

One of the most dramatic illustrations of antibiotic resistance was observed in Japan in the 1940's and 50's. Within a few years after beginning the use of sulfonamides (Su) to treat *Shigella dysenteriae* infections responsible for severe diarrheal disease, Su resistant strains began to appear. Within a decade, greater than 90% of the dysentery cases were caused by Su resistant strains. By this time, chloramphenicol (Cm), tetracycline (Tc), and streptomycin (Sm) were in use. The first strain resistant to all four antibiotics was isolated in 1955. By 1970, the large majority of *Shigella* strains were quadruply resistant. Why was it so easy to develop resistance to all four antibiotics? Why did the resistance spread so easily? The answer is that the resistance genes were not present in the chromosome, but rather on a self-transmissible plasmid. The plasmid could transfer freely by conjugation not only to pathogenic *Shigella*, but also to non-pathogenic enteric bacteria like *E. coli*, and vice versa. Thus, it was possible for a member of the normal flora to harbor antibiotic resistance genes and pass them to a pathogen after infection. The fact that the R-plasmid encoded resistance to all four antibiotics means that in a single step, a recipient cell can acquire all the resistances in a single genetic package. This is not an isolated example. Antibiotic resistance plasmids are extremely common. Recently, antibiotic resistant *Salmonella* strains were found in outbreaks in Britain, Saudi Arabia, and India. The *Salmonella* strains were different, but the plasmids in the different outbreaks were found to be

identical. This particular plasmid encoded resistance to tetracycline, streptomycin, sulfonamides, chloramphenicol, kanamycin, spectinomycin, trimethoprim, gentamicin, and ampicillin. A similar problem has occurred in *Pseudomonas aeruginosa* strains that infect the lungs of patients with cystic fibrosis. These strains are often resistant to virtually all antibiotics, leaving no way to treat the infection.

B. Transposable antibiotic resistance genes.

There are several questions regarding multiple antibiotic resistance. (1) How do the antibiotic resistance genes get onto the plasmids? (2) Why are there so many on a single plasmid? (3) What is the selective pressure for multiple resistance? One way a chromosomal resistance gene can become associated with a plasmid is by the integration and imprecise excision of bacterial plasmids, as seen above for F' prime plasmids. However, there is a second more efficient way. We now know that the antibiotic resistance genes on plasmids are transposable. One group of scientists discovered that a kanamycin resistance gene (Km) found on a plasmid could jump to a bacteriophage and from then to another plasmid. The structure of this element revealed a Km gene flanked by two identical IS elements. As mentioned previously, this structure is referred to as a **composite transposon**. In this case, it was named Tn5. The Km gene itself has no ability to transpose, but it can be mobilized by the transposition properties of the IS elements. This led to the discovery that any gene can be transposed to a new location if it is flanked by identical IS elements. Numerous complex transposons encoding various resistances and other genes, including those encoding virulence determinants, have been identified. Transposons have gone through their own evolution to make them transpose more efficiently. One example of a more highly evolved transposon is Tn3, which codes for ampicillin resistance (Ap). Its structure is simpler than that of the complex transposons. Rather than two distinct IS elements, Tn3 is smaller and appears similar to an IS element containing the Ap gene within it. How such a structure evolved is not clear, but it transposes at a high frequency (once in 10^3 cells). The large variety of plasmids and bacterial species that harbor this transposon are testimony to its success in nature.

C. The role of plasmids and transposable elements in the spread of antibiotic resistance and pathogenicity determinants.

Because plasmids move freely among various bacteria, particularly some **promiscuous plasmids** with remarkably broad host-range, they are frequent targets for transposons. Once an antibiotic resistance gene has been acquired by a plasmid, that plasmid has a selective advantage. Its host will survive treatment with that particular antibiotic and the population will thrive. This provides even more opportunity for spread of the plasmid and acquisition of yet additional transposons. A plasmid with multiple antibiotic resistances offers enormous selective advantage to organisms that encounter these antibiotics. Note that the presence of a single antibiotic in the environment selects for the plasmid and its entire complement of antibiotic resistance genes by virtue of their linkage to the same DNA molecule.

Other elements promote the spread of multiple antibiotic resistance. Antibiotic resistance-encoding transposons can insert into other antibiotic resistance

transposons to become part of the original. Thus, a single "super transposon" can encode multiple antibiotic resistance genes. In addition, another type of element, called an **integron**, is a highly specialized gene cassette capturing device. It encodes a specialized recombination enzyme that inserts DNA into a specific site located downstream of a transcriptional promoter. The integron can capture antibiotic resistance and other genes and put them into position so that they are all expressed in tandem from the same promoter. Some integrons contain dozens of genes, many of them antibiotic resistance genes. The integrons can also be incorporated into a transposon, so that the large array of genes is able to transpose in a single event to a conjugative plasmid. The combination of these super transposons and integrons with self-transmissible or mobilizable plasmids provides a potent mechanism for rapid spread of multiple antibiotic resistance among bacteria.

In addition, genes other than antibiotic resistance genes may become transposable by flanking IS elements and jump to a plasmid already encoding multiple antibiotic resistance. The survival of these genes will then be selected (indirectly) by the use of any antibiotic whose resistance is encoded by the plasmid. Some plasmids have been found to encode a complete package of antibiotic resistance, colonization factors, and enterotoxins. Because these genes transposed to the plasmid, the same sets of genes, often as a complete package, can be integrated into the chromosome. Some of the genes can integrate via transposition or site-specific recombination similar to that used by integrons and temperate phages. It is also possible for the entire plasmid to integrate into the chromosome, as shown in the production of Hfr strains. If these genes provide a strong selective advantage, they may become permanently incorporated into the chromosome by eliminating plasmid replication genes and elements involved in recombination. Genome sequencing has recognized a number of these elements in the chromosomes of bacteria. They are clearly additions to the basic chromosome. These so called **genomic islands** have been named for their functions. Islands that contain genes for pathogenicity are called **pathogenicity islands**. Comparing several strains of the same species of bacteria has revealed how dynamic the bacterial chromosome is. Different strains often differ by several distinct islands containing hundreds of genes. In addition, some of the genomic islands have been found in distantly related or unrelated bacteria, a potent testimony to the extent of horizontal transfer in bacteria and its effect on bacterial evolution.

Why are antibiotic resistance plasmids so common? The amount of antibiotics used worldwide each year for medical and agricultural purposes can be measured in thousands of tons. This represents enormous selective pressure for resistant bacteria and their resident plasmids. There is an urgent need to take every precaution to reduce unnecessary use of antibiotics. Once selective pressure is reduced, sensitive bacteria begin to reappear because they are not at a disadvantage. There is considerable debate over the use of antibiotics in animal feed, and some countries have attempted to ban certain antibiotics that are used clinically. It is a serious problem because studies have shown that R-plasmids present in animal flora can appear later in human pathogens. As physicians, you can reduce the selective pressure by prescribing antibiotics only when it is clear that they are required, and whenever possible, by choosing a narrow spectrum antibiotic. Hospitals may find it necessary to rotate the use of antibiotics to periodically remove selective pressure for resistance to specific antibiotics.

There is no question that antibiotic therapy of bacterial infectious disease has brought about an extraordinary reduction in human suffering. However, the excessive use of antibiotics, the resultant selective pressure, and the remarkable ability of bacteria to develop and disseminate resistance genes relatively quickly comprise a serious problem that merits continuous attention.