Chapter 2

Genetic Exchange in Bacteria and the Modular Structure of Mobile DNA Elements

JAMES W. WILSON

Department of Microbiology and Immunology, Tulane University Health Sciences Center, 1430 Tulane Avenue, New Orleans, LA 70112; Present address: Center for Infectious Diseases and Vaccinology, BioDesign Institute, Arizona State University, PO Box 875401, Tempe, AZ 85287

Historical Landmarks

1915–1917 Bacteriophages are discovered independently by Frederick W. Twort in Great Britain and Félix d'Hérelle in France (reviewed in Duckworth, 1976).

1928 Griffith shows that avirulent, nonencapsulated pneumococcus can be "transformed" to become virulent and encapsulated by incubating with heat-killed virulent pneumococcus. He hypothesizes that a substance called transforming principle is transferred from the virulent to the avirulent pneumococcus (Griffith, 1928).

1944 Avery, MacLeod, and McCarty demonstrate that DNA is the transforming principle (Avery et al., 1944).

1946 Lederberg and Tatum discover conjugation when they observe that strains of bacteria can exchange genetic markers. Lederberg later coins the term "plasmid" in a review paper on cellular genetics (Lederberg and Tatum, 1946).

1951 Freeman demonstrates that a bacteriophage can transfer virulence traits from pathogenic to nonpathogenic strains of *Corynebacterium diphtheriae* (Freeman, 1951).

1952 Zinder and Lederberg discover generalized transduction of genetic markers using *Salmonella* bacteriophage (Zinder and Lederberg, 1952).

1952 Using bacteriophages labeled with radioactive protein and DNA, Hershey and Chase provide undisputable proof that DNA is the genetic material (Hershey and Chase, 1952).

1959 First R-plasmid discovered when multiple antibiotic resistance is transferred from *Escherichia coli* to *Shigella* (reviewed in Mitsuhashi, 1977).

1967–1971 Pathogenic properties are observed to be transferred at high frequency between strains of *E. coli*, and the concept of the virulence plasmid is established (reviewed in Miller et al., 1994).

1974–1977 The *Agrobacterium tumefaciens* Ti plasmid is discovered and subsequently shown to insert into the host plant genome (reviewed in Zhu et al., 2000).

1980 Plasmid-mediated tissue invasiveness by *Yersinia enterocolitica* is demonstrated by Zink et al., (1980).

1980–1981 The first conjugal transposons are discovered: Tn*916* from *Enterococcus faecalis* and Tn*5253* from *Streptococcus pneumoniae* (Buu-Hoi and Horodniceanu, 1980; Franke and Clewell, 1981).

1998 A large chromosomal pathogenicity island is shown to be transferred from one bacterial strain to another by a helper bacteriophage (Lindsay et al., 1998).

1. Introduction

Evolution demands that genetic traits be passed on to other members of a given population. The transmission of genetic information in bacteria can be divided into two main modes: vertical and horizontal. Vertical transfer of genes occurs when the bacterial chromosome replicates and each daughter cell receives a chromosomal copy upon cell division, analogous to mitosis and cell division in eukaryotic cells. By contrast, the horizontal transfer of genes occurs between individual members of a population via mechanisms that are not based on chromosomal replication and cell division. Horizontal transmission can be thought of as the bacterial form of sexual genetic transfer and involves mechanisms such as transformation, plasmid conjugation, and bacteriophage transduction. Horizontal transfer allows a genetically unrelated bacterial strain or species to enter a population and transmit its genes into the other cells that are present. In other words, horizontal transfer allows diverse types of bacteria to exchange genes and therefore represents a powerful force to drive evolution.

The mechanisms that have evolved to allow bacteria to transfer genetic information between cells via horizontal transfer are true marvels of nature. The evidence that these genetic exchange mechanisms have played major roles in the evolution and dissemination of virulence genes is vast and very difficult to dispute. The following observations serve as testament to the role of horizontal genetic transfer in the evolution of pathogens:

- A large number of virulence genes are contained on mobile genetic elements such as plasmids, bacteriophages, and DNA regions that are substrates for recombinases that allow their excision from a particular genome. For a number of pathogenic species, such as *Vibrio cholera*, *Staphylococcus aureus*, pathogenic *Escherichia coli*, and others, the presence of a single virulence function on a mobile DNA element is the essential factor for a particular disease.
- There are several examples of large (10–50+ kb) contiguous blocks of chromosomal virulence genes that function together for a specific process, typically for the secretion of virulence factors. Through genomic sequence analysis, it is clear that these regions have been acquired by horizontal

transfer and are absent from genomes of related, nonpathogenic strains or species. These regions have been termed "pathogenicity islands." Frequently, pathogenicity islands are associated with integrases and other genes that are related to mobile DNA elements like plasmids and bacteriophages.

- Large blocks of *related* virulence genes (5–25 genes), such as type III and type IV secretion systems, have been found in the genomes of many different pathogenic species (frequently as part of pathogenicity islands). The genes encoding these systems from different species are clearly homologous and ancestrally related. For such large regions of related genes to have all evolved independently in these different species is highly unlikely. Instead, these genes were most likely disseminated over the course of evolution via horizontal transfer mechanisms.
- The emergence of antibiotic resistance in response to the widespread use of antibiotics worldwide is one of the most impressive examples of the power of microbes to adapt to selective pressure. Analysis of the spread of antibiotic resistance genes among different bacterial species has revealed a central role for mobile DNA elements such as plasmids and transposons in this process.

Put simply, horizontal transfer of genetic information allows bacteria to evolve in ways that vertical transmission of random mutations from mother to daughter cells will not facilitate. Therefore, the study of these genetic exchange mechanisms will reveal a set of strategies that bacteria use to evolve and adapt, frequently in ways that allow them to interact with eukaryotic hosts and cause disease. Table 1 lists numerous examples of bacterial toxins that are contained on mobile genetic elements or large sections of horizontally transferred DNA (Novick, 2003).

The main objective of this chapter is to describe the major genetic exchange mechanisms in bacteria and how different parts of these mechanisms can be utilized in various combinations to create different genetic elements. In addition to learning the basic conceptual themes and mechanisms of genetic exchange, it is becoming increasingly important to be able to combine different parts of the basic concepts in novel ways in order to explain the functioning of several newly discovered genetic elements. As an increasing number of bacterial genomes become sequenced and new genetic elements get discovered, it is apparent that the lines between plasmid, bacteriophage, and transposon are becoming blurred when describing these new elements.

The field of bacterial horizontal gene transfer is vast and includes many concepts that have been shaped by data obtained from highly talented scientists worldwide. Some of the references for individual sections of this chapter are reviews that contain much of the presented information for that section. Additional, specific references (if not provided here) can be found within those reviews.

TABLE 1. Examples of bacterial toxins that are contained on mobile genetic elements.[a]

Disease	Toxin	Organism	Element	Features
Diphtheria	Diphtheria toxin	*Corynebacterium diphtheriae*	β-Prophage	~35 kb
Anthrax	Anthrax toxin	*Bacillus anthracis*	Plasmid	~110 kb
Cholera	Cholera toxin	*Vibrio cholerae*	Prophage	Filamentous
Dysentery	Shiga toxin	*Shigella dysenteriae*	Prophage	Lambdoid
Botulism	Botulin toxin	*Clostridium botuliman*	Prophage	110–165 kb
Enterocolitis	Cpe toxin	*Clostridium perfringens*	Transposon	$IS1470_2$::cpe ~6 kb
Gangrene	α-toxin	*Clostridium perfringens*	Chromosome[b]	
Tetanus	Tetanus toxin	*Clostridium tetani*	Plasmid	"Large"
Enterocolitis	Enterotoxin	*Clostridium difficile*	Path Islet	~3 kb
Diarrhea	Labile toxin (LT)	*Escherichia coli*	Plasmid	90 kb F-like
Diarrhea	Stable toxin (ST)	*Escherichia coli*	Transposon	$IS1_2$::cat ~3 kb
Toxic shock syndrome	TSST-1	*Staphylococcus aureus*	Path island	~15 kb ϕ-related
Food poisoning, TSS	Enterotoxin A	*Staphylococcus aureus*	Prophage	~45 kb
Food poisoning, TSS	Enterotoxin B, C	*Staphylococcus aureus*	Path island	~15 kb ϕ-related
Food poisoning, TSS	Enterotoxin D	*Staphylococcus aureus*	Plasmid	~30 kb
Scalded skin syndrome	Exfoliatin A	*Staphylococcus aureus*	Prophage	~45 kb
Scalded skin syndrome	Exfoliatin B	*Staphylococcus aureus*	Plasmid	~30 kb
Necrotizing pneumonia	Leukocidin	*Staphylococcus aureus*	Prophage	~45 kb
Scarlet fever	SPEA, C	*Streptococcus pyogenes*	Prophage	~45 kb

[a] Data from Novik (2003). Reprinted with permission from Elsevier.
[b] *C. perfringens* α-toxin is the single exception to the rule that toxinosis-causing toxins are encoded by mobile genetic elements.

1.1. Vehicles that Mediate Horizontal Gene Transfer

Horizontal transfer is driven by DNA elements that are mobile; they are able to spread from one bacterium to another and can serve as vehicles that can carry large sections of genes between hosts. Certain mobile DNA elements, such as transposons, can "jump" from one DNA element to another; this type of genetic transfer is very important because it allows a section of DNA that is located on a vertically transferred element (a chromsome) to be moved to a horizontally transferred element (like a bacteriophage or plasmid). The transposed DNA section can then be potentially transferred to a variety of other different strains or species. Below is a list of mobile DNA elements that we will refer to in this chapter:

- *Plasmids*—DNA elements that autonomously replicate apart from the host chromosome.
- *Bacteriophages*—viruses that infect bacterial cells using a protein package that contains genetic information, usually DNA though some RNA phages do exist.
- *Transposons*—sections of DNA located between repeated sequences that can be excised and moved to a separate DNA element via a recombination enzyme called a transposase.

It is helpful to think of these genetic elements as combinations of genetic modules or cassettes. Genetic modules contain genes and/or DNA sites that allow mobile DNA elements to move between hosts, jump between genomes, and be stably maintained in the population. Various combinations of certain genetic modules that we will discuss in this chapter give rise to different mobile DNA elements and different mechanisms for their propagation and transfer. However, the descriptions of different mobile DNA elements are based on the three main vehicles: plasmids, bacteriophages, and transposons.

1.2. The Four Major Horizontal Transfer Paradigms: Transformation, Conjugation, Transduction, and Transposition

The study of bacterial genetic exchange has revealed three processes that allow horizontal gene transfer between cells to occur and a fourth process that allows vertically transferred DNA to jump onto horizontally transferred elements and vice versa. *Transformation* is the uptake of naked DNA that has been released into the environment. As part of their life cycles, it is common for bacterial cells within a population to lyse and release their DNA, and bacteria have evolved mechanisms to bind this DNA and import it into their cytoplasm. *Conjugation* is the directed transfer of DNA through a membrane pore from a donor cell to a recipient cell. A single strand of DNA is transferred and is copied in the recipient, while the other single strand is copied in the donor. This process is commonly mediated by plasmid systems that can either mediate their own transmission ("self-transmissible") or transfer other plasmids that express the appropriate functions (termed "mobilizable plasmids"). *Transduction* is genetic exchange that is mediated by bacteriophages. Bacteriophages package certain DNA molecules present in a host bacterium in a protein sheath termed a capsid and can then transfer this DNA to another bacterium upon release of the phage particles from the host. *Transposition* is the movement of a section of DNA from one location in a particular genome, such as a chromosome or plasmid, to another separate genome. This is a critical concept in horizontal gene transfer because it provides an important way for vertically and horizontally transferred genetic elements to exchange genes.

In the following sections, we will discuss the concepts outlined below:

(1) Transformation mechanisms
(2) Plasmid replication, conjugation, and maintenance
(3) Bacteriophages and transduction
(4) Transposons and the transposition of DNA
(5) The modular structure of mobile genetic elements
(6) A world of genetic modules

2. Transformation Mechanisms

Both Gram-positive and Gram-negative bacteria have evolved mechanisms to allow the uptake of naked DNA from the extracellular environment (Chen and Dubnau, 2004). In these systems, the extracellular DNA is recognized and bound by receptor proteins on the outer surface. Another complex of membrane-associated proteins then mediates the physical transport of the DNA across the cell envelope. During this transport, one strand of the DNA is degraded and a single-stranded DNA molecule is introduced into the recipient cytoplasm. The newly transported, single-stranded DNA molecule can then serve as a substrate for the host recombination machinery, which integrates the DNA piece into the recipient's genome via recombination mechanisms. Several of the proteins in both Gram-positive and Gram-negative bacteria that are involved in the DNA uptake process belong to a family of proteins also involved in other processes such as type IV pilus formation, type II and III secretion, and twitching motility (the pilus/secretion/twitching/competence or PSTC family). In the well-studied transformation systems, there is a strong correlation between the expression of type IV–related pilus proteins and the ability to import DNA from the environment.

2.1. Gram-positive Transformation

Most of the information on Gram-positive transformation is derived from experiments with *Bacillus subtilis* and *Streptococcus pneumoniae* (Chen and Dubnau, 2004). Figure 1 illustrates a model of Gram-positive DNA uptake based on the *B. subtilis* system, but the two systems are very similiar. The model can be broken down into two major parts: binding and transport.

- *Binding*—the ComEA protein is a surface receptor that binds the DNA non-specifically and delivers the DNA to the machinery that transports it across the membrane. ComEA is homologous to the C-terminal domain of a kinesin-like DNA-binding protein called Kid. In addition, there is a flexible peptide sequence adjacent to the DNA-binding domain of ComEA that may allow the protein to bend. It is thought that after ComEA binds the DNA, it conformationally changes to hand off the DNA to the

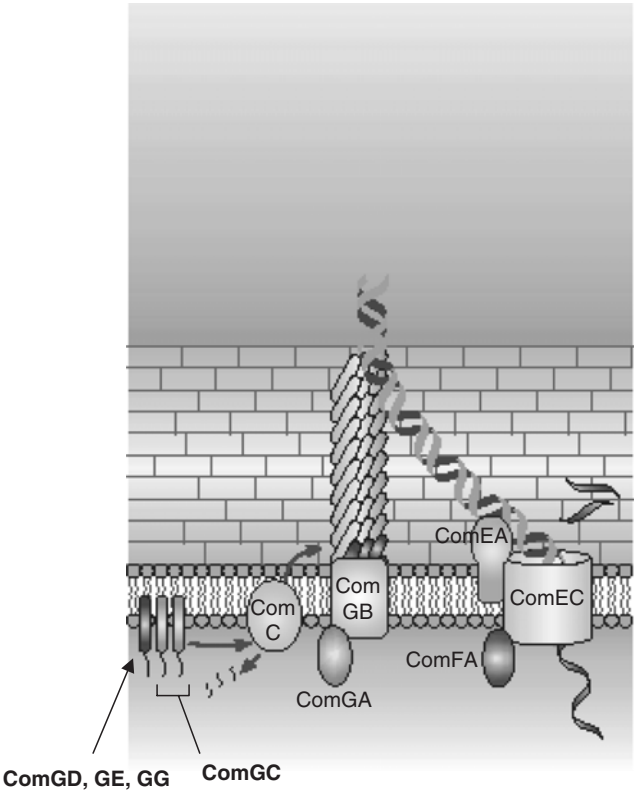

FIGURE 1. Model of DNA uptake during natural transformation of Gram-positive bacteria. On the surface of Gram-positive cells expressing the Com DNA uptake system, ComGB and ComGA aid in the formation of a pilus-like structure consisting of the major subunit (ComGC) and minor subunits (ComGD, -GE, and -GG). The pilus-like structure extends from the ComGB protein in the diagram. These subunits are processed by the ComC prepilin peptidase. This "pseudopilus" is thought to allow the DNA to access the membrane-bound receptor, ComEA, which delivers the DNA to the ComEC channel in the cytoplasmic membrane. The ATP-binding protein ComFA aids in the process of single-stranded DNA transport across the membrane. This model is based on the nomenclature of the *Bacillus subtilis* DNA uptake system. (From Chen and Dubnau, 2004.)

membrane DNA transporter. The ComG proteins (ComGC, ComGD, ComGE, and ComGG), which belong to a class of pilus assembly proteins, form a structure in the outer peptidoglycan layer that provides access of the DNA to ComEA. These pilus-like proteins are processed by the prepilin peptidase ComC. The ComGA and ComGB proteins also aid in the process of forming the ComG pilus-like structure on the cell surface.

- *Transport*—the ComEC protein is a polytopic membrane protein, and ComFA is a membrane-associated protein, which is a member of the DEAD

helicase family (named for its conserved sequence of Asp-Glu-Ala-Asp amino acids), and contains a consensus nucleotide-binding motif. These two proteins together form a transport complex that moves the DNA across the membrane. In addition, a nuclease (termed NucA in *B. subtilis* and EndA in *S. pneumoniae*) serves to degrade one of the DNA strands during transport, possibly providing energy for the ComEC/ComFA-mediated process.

2.2. Gram-negative Transformation

The model for Gram-negative DNA uptake is derived from the *Neisseria gonorrheae* and *Haemophilus influenzae* systems, and can be divided into the same parts as the Gram-positive model above (Chen and Dubnau, 2004). This model uses the nomenclature from *N. gonorrheae* (Figure 2), but the mechanism can be applied to the *H. influenzae* (and other) systems.

- *Binding*—in both the *N. gonorrheae* and *H. influenzae* systems, there are specific uptake sequences (USs) on the naked DNA that are recognized by a receptor on the outer membrane. This is a major difference from the Gram-positive model. For *N. gonorrheae*, the US is GCCGTCTGGA; in *H. influenzae*, it is AAGTGCGGT. Each of these sequences is found at a significantly greater-than-random frequency in each of the respective genomes, indicating that these bacteria preferentially import DNA from their own species. In *H. influenzae*, the DNA is bound at specific locations on the cell surface called "transformasomes," which form a bleb-like subcellular structure in which the DNA is sequestered until transport.
- *Transport*—the following components are involved with DNA uptake in *N. gonnorheae*: (1) PilQ, an outer membrane protein belonging to the secretin family; (2) PilC, a pilus-associated protein; (3) PilE, a pilin-like protein thought to provide a channel through the cell wall like the *B. subtilis* ComG proteins above; (4) ComA, a polytopic inner-membrane protein, which is an ortholog of the *B. subtilis* ComEA protein; (5) Tpc and ComL, which are predicted murein hydrolases; (6) PilF, a traffic NTPase involved in pilus formation; (7) ComP, a prepilin-like protein; and (8) ComE, a periplasmic protein similar to *B. subtilis* ComEA. Together, these proteins allow transport of the naked DNA from the outer surface across the outer membrane, cell wall, and inner membrane to the cytoplasm. There is also thought to be a nuclease, similar to the Gram-positive model, that degrades one of the DNA strands during transport.

2.3. *The* Helicobacter pylori *DNA Uptake Mechanism Is Related to Type IV Secretion*

Helicobacter pylori is a Gram-negative gastric pathogen that uses a transformation mechanism with components related to type IV secretion systems, most notably the plant pathogen *Agrobacterium tumefaciens* T-DNA export system

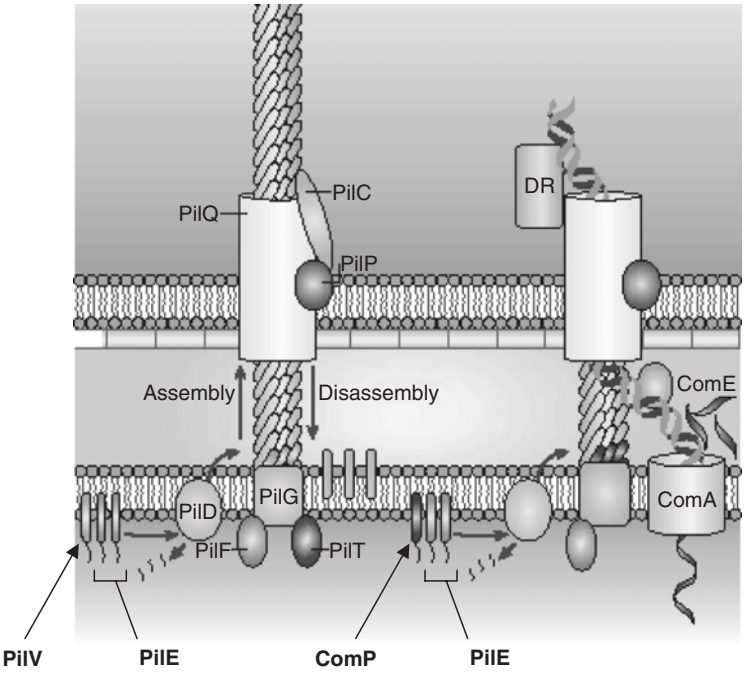

FIGURE 2. A model for DNA uptake during natural transformation of Gram-negative bacteria. This model of DNA uptake in Gram-negative bacteria is based on the system found in *Neisseria gonorrheae*. In *N. gonorrheae*, a type IV pilus is formed on the surface using the Pil proteins. The Pil proteins used to form this surface pilus also function with other Com proteins to facilitate DNA uptake. The left side of this diagram depicts the apparatus involved with type IV pilus formation. The prepilin peptidase (PilD) processes the major (PilE) and minor (PilV) subunits for assembly into the polymerized structure by the PilG membrane protein and the PilF and PilT traffic NTPases. The pilus-like structure extends from the PilG protein in the diagram. The outer membrane secretin protein PilQ, aided by the proteins PilC and PilP, allows the pilus structure to cross the outer membrane. This pilus is able to extend and retract by alternating between assembly and disassembly modes.

The right side of the diagram depicts the model of DNA uptake that involves the Pil and Com proteins. The PilE protein, together with a separate minor pilus subunit (ComP), forms a pseudopilus via the same mechanism as for type IV pilus assembly. This pseudopilus is thought to extend across the periplasm to the PilQ secretin. An as-yet-to-be-indentified DNA receptor (DR) located at the surface binds the DNA and shuttles it into the secretin pore in the outer membrane. The DNA is recognized in the periplasm by the ComE protein and is delivered to a DNA channel in the plasma membrane formed by the ComA protein. A single strand of the DNA enters the cytosol, while the other strand is degraded. (From Chen and Dubnau, 2004.)

(Chen and Dubnau, 2004). In *A. tumefaciens*, the VirB7, B8, B9, and B10 proteins are encoded by colinear genes as part of the T-DNA system, and these proteins are major components of the transmembrane pore through which the T-DNA is transported. In *H. pylori*, four genes of the DNA uptake system, *com*B7–B10, are similiarly arranged, and the protein products (ComB7–B10) resemble their *A. tumefaciens* counterparts. The ComB7–B10 proteins most likely form a membrane apparatus in a manner similar to the *A. tumefaciens* (and other type IV) system(s). ComB7 is an outer membrane protein that is likely bound to ComB9 by a disulfide bond. ComB8, ComB9, and ComB10 form a membrane complex that spans from the inner membrane to the periplasmic space and binds to ComB7. ComH, a protein predicted to be exported to the cell surface (and not a type IV system homologue), is a candidate to be a DNA-binding protein that serves to recognize the DNA substrate and provide nuclease activity that is involved in the processing of the DNA for transport (Chen and Dubnau, 2004).

2.4. Competence Induction

In certain bacterial species, transformation can only take place after the cells in the population have been made "competent" for the uptake of DNA—a process termed "competence induction." Competence induction is an example of a quorum-sensing mechanism: when the cells of the population get to a certain density, a signal molecule is produced in sufficient quantities to induce a regulatory cascade in the cells that changes their physiology. In the case of competence induction, the result of quorum sensing is that the cells express the genes necessary for DNA uptake. One of the best studied competence induction systems is that of *S. pneumoniae*, and this system can serve as an analogous example for other competence systems (Claverys and Havarstein, 2002). In the *S. pneumoniae* system, the competence inducer molecule is expressed from the *comC* gene as a large ComC precursor protein. The ComC precursor is recognized by an ATP-binding cassette (ABC) transport apparatus formed by the membrane-bound ComA and ComB proteins. The ComC precursor is cleaved upon its secretion by ComA and ComB and a small competence-stimulating peptide (CSP) derived from the ComC precursor is exported to the extracellular environment. The CSP is then recognized by other *S. pneumoniae* cells in the population via a surface histidine kinase receptor termed ComD. ComD is a member of a two-component regulatory system with ComE, a transcriptional regulator that gets phosphorylated by ComD upon CSP binding. Phosphorylated ComE then initiates a regulatory cascade that serves to turn on the genes required for DNA uptake. By using competence induction, the cells of a population optimize the control of DNA uptake genes and orchestrate the DNA uptake system to function when conditions are most favorable for DNA exchange.

3. Plasmid Replication, Conjugation, and Maintenance

3.1. Plasmid Replication

The study of plasmid replication has revealed that there are conserved mechanisms used for this process (del Solar et al., 1998; Khan, 1997; Kobryn and Chaconas, 2001; Qin and Cohen, 1998). Most plasmids are covalently closed, circular DNA molecules, but linear plasmid DNA elements do exist as well. These two different molecules have replication mechanisms that are conceptually distinct and are therefore described separately. The major concept of plasmid replication is that of the *replicon* consisting of a cis-acting origin of replication and an initiator protein (most commonly plasmid-encoded) that recognizes the origin and initiates replication. Plasmid replication mechanisms are based on this important scheme. Another important plasmid-related concept is that of *incompatibility*. Plasmids that share common replication and maintenance functions typically cannot coexist in the same bacterial cell and this leads to the loss of one of the plasmids. This is based on the fact that common functions on the plasmids can interact with each other and interfere with normal replication and maintenance of each plasmid. In this case, one plasmid typically "outcompetes" the other and is maintained while the other plasmid is lost. Plasmid replication determinants are common mediators of incompatibility.

The different plasmid replication mechanisms are: (1) theta-type; (2) strand displacement; (3) rolling circle; and (4) linear. We will give brief descriptions of these mechanisms here; however, the reader is directed to several excellent reviews for more details (del Solar et al., 1998; Khan, 1997; Kobryn and Chaconas, 2001; Qin and Cohen, 1998). The overviews of replication presented here serve as conserved models that may be utilized when characterizing the replication modes of newly discovered genetic elements.

3.1.1. Theta-type Replication

For theta-type replication (Figure 3), the origin of replication is bound by the initiator (or "Rep" protein), and the two DNA strands are opened (or "melted") so that two single DNA strands are exposed (del Solar et al., 1998). A DNA polymerase and DNA helicase (these are typically both host-encoded) then start replication at the replication forks that are formed upon melting of the origin. Typically, DNA synthesis is continuous on one strand of the replication fork (and results in the formation of the "leading strand") and discontinous on the other (and results in the formation of "lagging strands"). If only one replication fork is used during DNA synthesis, the replication is termed "unidirectional" and terminates when the replication fork returns to the origin. If both replication forks are used, replication is "bidirectional" and terminates somewhere on the DNA molecule opposite the origin. The term "theta" was used to describe this mode of circular DNA

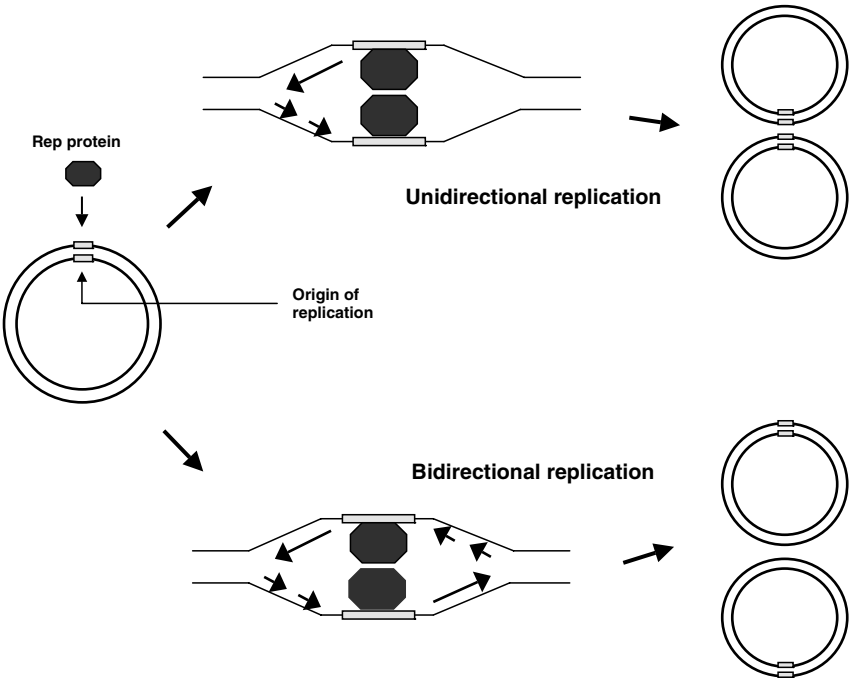

FIGURE 3. Unidirectional and bidirectional theta-type plasmid replication. A circular plasmid is depicted and its origin of replication is indicated. The replication initiator protein (or "Rep protein") recognizes the origin, binds to it, and facilitates origin melting. This leads to host factors being recruited to the origin and commencement of DNA synthesis. Note the formation of leading and lagging strands at the replication forks. In unidirectional replication, only one replication fork is used and replication terminates when the DNA synthesis machinery returns to the origin. In bidirectional replication, both replication forks are used and replication terminates somewhere on the DNA molecule opposite the origin.

replication because when these origins were first visualized via electron microscopy, they were shaped like the Greek character "theta" (θ).

The origin of replication of theta-type plasmids have some common features (del Solar et al., 1998): (1) repeated DNA sequences termed "iterons" that serve as the binding sites for the initiator protein (these plasmids are also described as "iteron-containing" plasmids); (2) A/T-rich sequences that facilitate the opening of DNA origin; and (3) binding sites for host replication factors that aid the replication process, such as DnaA, IHF, FIS, and others. There are some deviations from this scheme, such as plasmid ColE1, which does not use a plasmid-encoded protein initiator but rather a small RNA molecule to prime DNA polymerase I-intitated replication; however, these general features are remarkably conserved among iteron-containing, theta-type plasmids.

3.1.2. Strand Displacement Replication

The strand displacement mechanism of replication is used by the broad-host-range plasmids of the IncQ group (del Solar et al., 1998). It is similar to theta-type replication, but is unique in the following ways: (1) the early stages of replication do not require host-encoded replication factors typically involved in theta-type replication such as RNA polymerase, DnaA, DnaB, DnaC, and DnaG; (2) three plasmid-encoded proteins are utilized for replication: the RepA helicase, the RepB primase, and the RepC origin-binding protein; (3) DNA synthesis is bidirectional, but is exclusively continuous (proceeding as the leading DNA strand is displaced by the RepA helicase). The fact that this mechanism of initiation is independent of host replication factors may help to explain the broad-host-range nature of the IncQ plasmids.

3.1.3. Rolling Circle Replication

The model for rolling circle replication is presented in Figure 4, and it differs significantly from the theta-type and strand displacement mechanisms (del Solar et al., 1998; Khan, 1997). The general mechanistic features of rolling circle replication are as follows: (1) there are two replication origins: a double-stranded origin and a single-stranded origin; (2) the replication initiator protein, in a dimer form for most systems, recognizes the double-stranded origin, nicks a single DNA strand at this site, and becomes covalently attached to this strand at a conserved tyrosine residue; (3) replication is initiated at the exposed 3′-OH DNA group at the nick site via other factors including DNA polymerase III, helicase, and single-stranded DNA-binding protein; (4) termination of this replication cycle results in a newly copied, double-stranded molecule and a displaced, single-stranded molecule; (5) replication of the single-stranded molecule is initiated by host factors at the single-stranded origin and results in a second, newly copied double-stranded molecule.

Rolling circle replication is a highly conserved mode of replication. It is used by single-stranded and double-stranded bacteriophages and many plasmids found in Gram-positive bacteria. In addition, the conjugative transfer of DNA is basically a rolling circle replication mechanism where the displaced single strand is transferred from a donor to a recipient bacterial host.

3.1.4. Linear Replication

The best-known examples of linear plasmids are found in *Streptomyces* spp. and in the Lyme disease agent *Borrelia borgdorferi* (Kobryn and Chaconas, 2001; Qin and Cohen, 1998). The linear *Streptomyces* plasmids, typified by the well-studied pSLA2, have terminal proteins bound to the 5′ ends of the linear molecule, similar to adenovirus or the *B. subtilis* phage ϕ29. In contrast, the *Borrelia* plasmids are covalently closed at each end by an interstrand DNA bond. Both strategies serve to protect the linear plasmids from degradation by exonucleases.

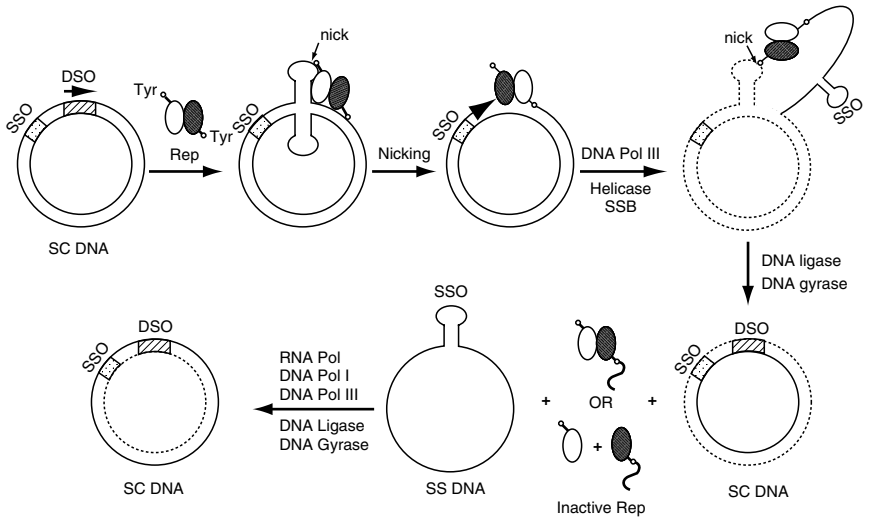

FIGURE 4. Rolling circle replication of a bacterial plasmid. At the double-stranded origin (DSO), the replication initiator protein (Rep protein, shown here as a dimer) forms a nick at a hairpin structure and becomes covalently attached to a single strand of the DNA via a tyrosine residue on one of the monomers. DNA replication is initiated at the exposed 3′-OH DNA group at the nick site via host factors including DNA polymerase III, helicase, and single-stranded DNA-binding protein. Replication of the new strand displaces the strand that is bound by the Rep protein dimer. DSO-mediated replication is terminated when the replication machinery returns to the DSO site, the Rep dimer once again nicks the DSO site (via the tyrosine residue on the other monomer), and the newly replicated strand is ligated together via host DNA ligase. This process also results in the displaced strand being religated (via the Rep protein activity) and the formation of inactive Rep protein molecules, which are bound to small pieces of plasmid DNA that are byproducts of the religation reaction. This DNA byproduct can be initially observed at the nick in the newly replicated DSO where it joins at the tyrosine residue of one of the Rep monomers. The displaced strand is replicated from the single-stranded origin (SSO) via the activity of a number of host factors. This model results in the formation of two double-stranded copies of the plasmid. (From Khan, 1997.)

The replication of *Streptomyces* pSLA2 initiates from an origin located internally within the linear genome and proceeds bidirectionally toward each end. In this model, lagging strand synthesis leads to the formation of a 3′ overhang of between 200 and 300 bp at the ends of the newly replicated DNA molecules. The repair (or "filling-in") of these overhangs is likely to be primed by a "fold-back" mechanism, where inverted repeats at the ends of the genome allow a hairpin structure to form that is recognized by the terminal protein (Figure 5). The terminal protein would aid in the initiation of repair synthesis and become covalently bound to the 5′ end of this new molecule.

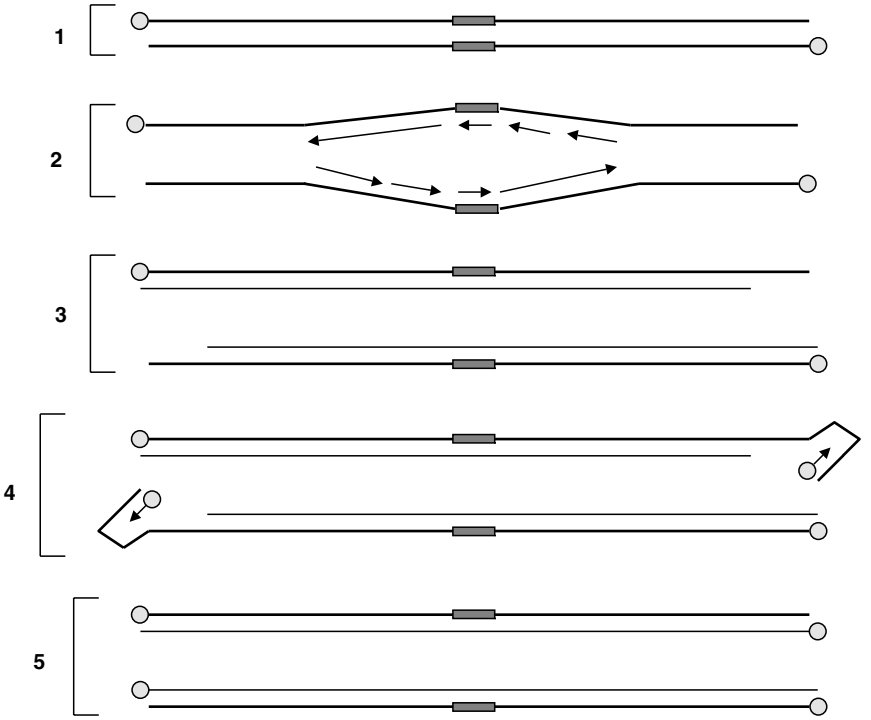

FIGURE 5. Model for replication of linear plasmids that have terminal proteins bound to their 5-primed ends. (1) A linear plasmid with covently bound terminal proteins at each 5′ end is depicted. (2) Bidirectional replication commences at an origin located in the middle of the molecule. Note the formation of leading and lagging strands during DNA synthesis. (3) Lagging strand synthesis results in 3′ overhangs that must be repaired (or "filled in"). (4) A "folded-back" structure forms at the 3′ ends of the overhanging strands. This structure is recognized by the terminal protein, which catalyzes the repair synthesis to fill in the terminal gap. (5) Fill-in repair of the overhangs results in completely replicated molecules that have terminal proteins bound to each 5′ end.

Borrelia plasmids also initiate replication from an internally located origin, but a "head-to-head", "tail-to-tail" dimer replication intermediate is predicted to form (Figure 6). This dimer is resolved into two full linear genomes by a DNA breakage and rejoining reaction catalyzed by the ResT telomere resolvase. This reaction produces the convalently closed telomeric ends of the linear genome.

3.2. Plasmid Conjugation

As described above, conjugation is the replicative, horizontal transfer of DNA from a donor cell to a recipient cell (Cascales and Christie, 2003; Christie, 2001; Llosa et al., 2002). Typically, conjugation involves the

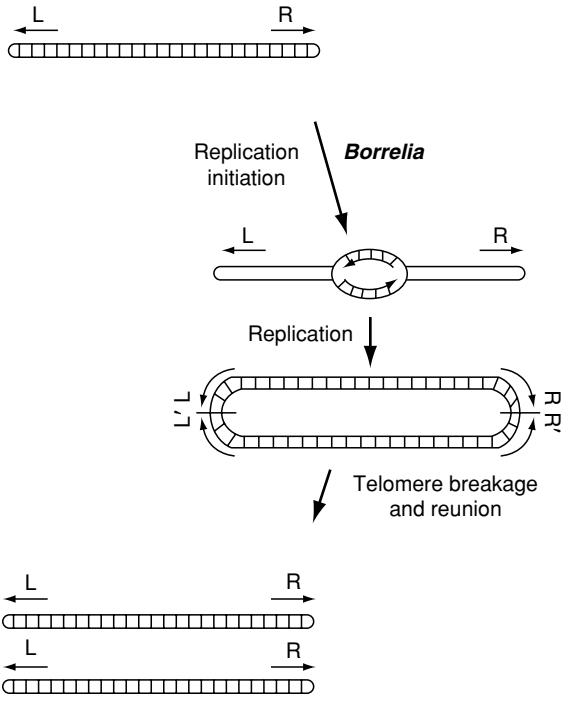

FIGURE 6. Replication of a linear plasmid with covalently closed hairpin ends. A linear plasmid with covalently closed hairpin ends is depicted, and direct repeats present at the left and right telomeric ends of the genome are indicated. This particular example is based on linear plasmids from *Borrelia* spp., but may be applicable to similar linear plasmids that exist in nature. Replication is believed to proceed from internal, bidirectional origins. The replicated telomeres are processed by a DNA cleavage and reunion event that is catalyzed by a telomere resolvase enzyme. (From Kobryn and Chaconas, 2001. Reprinted with permission from Elsevier.)

transfer of plasmid molecules, but the conjugative transfer or mobilization of transposons and bacterial chromosomes also occurs. An important distinction when discussing conjugation is that between self-transmissible and mobilizable elements. Self-transmissible elements encode all the conjugation functions needed to mediate their own transfer from a donor to a recipient (the *oriT*, Dtr, and Mpf components described below). Mobilizable elements are able to conjugatively transfer, but do not encode all the functions necessary for self-transmission. Typically, these elements contain *oriT* and Dtr functions, but not a mating pair formation (Mpf) system. They rely on the Mpf functions from another self-transmissible element for their transfer from donor to recipient.

3.2.1. Conjugation Systems Are Composed of Three Major Components

(1) A cis-acting DNA site termed the origin of transfer or *oriT*;
(2) DNA transfer and replication proteins (Dtr) (also known as DNA-processing functions) that recognize the *oriT* and create a single-stranded nick at this site. The protein that binds the *oriT* and creates the nick is termed the *relaxase* protein. There is also a protein that couples the nicked *oriT*/relaxase complex to the mating pair formation system (see below) termed the *coupling protein*. The Dtr functions are also referred to as "Mob" genes.
(3) Mpf proteins form any pili or fimbrial structures associated with conjugation and also form the conjugal pore that serves as the conduit for DNA transfer. The Mpf components of almost all conjugation systems belong to the type IV secretion family that is described in Chapter 9 on bacterial protein secretion mechanisms. This family of secretion systems has evolved to transport DNA and proteins substrates and has become widely disseminated among bacterial species including pathogens. The Mpf functions are also referred to as "Tra" genes.

3.2.2. Conjugation Is Essentially Rolling Circle Replication Coupled to Type IV Secretion

The mechanism of conjugation is shown in Figure 7. There are four steps in this mechanism, which is very similar to rolling circle replication and type IV secretion (Cascales and Christie, 2003; Christie, 2001; Llosa et al., 2002):

(1) The relaxase protein binds to the *oriT* on the DNA substrate and cleaves one of the DNA strands to create a single-stranded nick. The nicking reaction results in the relaxase becoming covalently attached to the 5′ end of the cleaved DNA strand at a conserved tyrosine residue in the *N*-terminal region of the protein.
(2) The coupling protein, which forms a homo-hexamer and is associated with the Mpf apparatus at the cytoplasmic membrane, facilitates interaction of the *oriT*/relaxase with the conjugation pore. This is how the DNA is "hooked up" to the transfer apparatus.
(3) The coupling protein is able to drive the transfer of the *oriT*/relaxase complex through the Mpf type IV secretion pore and into the recipient cell.
(4) After the initial transfer of the oriT/relaxase complex, the remainder of the single-stranded DNA is pumped through the Mpf pore into the recipient. After this transfer is completed, the ends of the single-stranded molecule are rejoined by the catalytic ligase activity of the relaxase protein, which is released in this reaction.

Conjugation can be thought of as rolling circle replication (where the relaxase plays the role of the initiator protein) that is coupled to type IV secretion (where the coupling protein and Mpf pore are the type IV apparatus). The mechanism described above has been termed the "shoot and pump" model

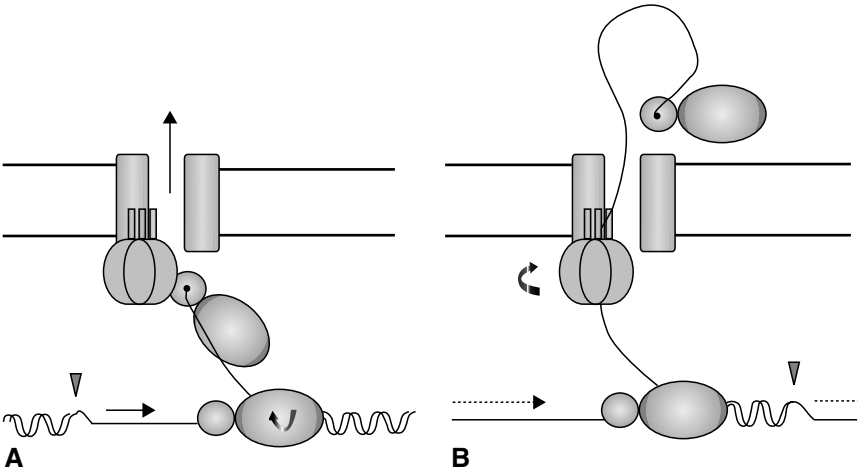

FIGURE 7. A model for DNA transfer during bacterial conjugation. This diagram depicts a conjugative plasmid being processed at its origin of transfer (*oriT*) and its subsequent interaction with the mating pore apparatus that allows transfer of the plasmid to the recipient. The straight, dark lines indicate bacterial envelopes of the donor and recipient (the donor contains the plasmid and is at the bottom of the diagram). A pore is formed between the two cells via the mating pair formation (Mpf) proteins, and the coupling protein is shown to be associated with the mating pore on the donor side. The wavy line indicates the plasmid DNA, and the vertical arrow indicates the location of the single-stranded DNA nick that is produced at the *oriT* via the activity of the relaxase protein (depicted as the two-part oval shapes). The relaxase protein forms a dimer that binds to the oriT region. One subunit of the dimer becomes covalently attached to one strand of the plasmid DNA as a result of the nicking reaction. (A) After the relaxase-catalyzed nicking reaction at *oriT*, the relaxase monomer that is bound to the single-strand of plasmid DNA is recognized by the coupling protein at the mating pore. The relaxase enzyme is then "shot" through the mating pore into the recipient cell via the activity of the coupling protein. Replication of the plasmid DNA in the donor cell proceeds from the *oriT* site (this is indicated by the horizontal arrow) and is aided by the helicase activity of the relaxase monomer that remains associated with the donor plasmid DNA. (B) After shooting the DNA-bound relaxase monomer into the recipient cell, the coupling protein pumps the rest of the single DNA strand into the recipient. When replication of the plasmid DNA in the donor returns to the *oriT*, the relaxase monomer in the donor ligates it together at the nick site and liberates the transferred single strand to complete its transfer to the recipient. Inside the recipient, the relaxase monomer rejoins the transferred strand, and this strand gets replicated into a double-stranded molecule via the activity of a primase (often plasmid-encoded) and host DNA synthesis functions. (From Llosa et al. 2002. Reprinted with permission from Blackwell Publishing.)

because the *oriT*/relaxase complex is initially shot through the Mpf apparatus and the rest of the DNA is pumped into the recipient cell. Conjugation is a beautiful example of two molecular processes (DNA replication and macromolecular secretion) that have been combined to produce a powerful means of genetic exchange.

3.3. Plasmid Maintenance Functions

A plasmid molecule that consists of just a minimal replicon (either with or without a conjugation system) is frequently not maintained efficiently in the cell population and is lost from cells at a high frequency that exceeds the rate of expected random distribution. This could be due to the fact that the plasmid is a metabolic burden on the cells or that the plasmids are not distributed properly upon cell division, among other possible reasons. Therefore, plasmids have evolved mechanisms that ensure their stable maintenance within the bacterial population. The most common way to maintain plasmid vectors in the laboratory is via an antibiotic resistance gene that allows plasmid-containing cells to survive in the presence of a given antibiotic. However, plasmids in nature are often cryptic and do not contain an easily selectable phenotype such as antibiotic resistance. In addition, if a plasmid does contain a resistance marker, the antibiotic is not always present to allow selection for plasmid-containing cells. In these cases, other molecular mechanisms can ensure that all daughter cells inherit a plasmid copy upon cell division. These maintenance mechanisms are extraordinary examples of how extrachromosomal DNA elements have evolved to ensure their persistence within a bacterial population.

3.3.1. Partition Systems

Partition systems, or "Par" systems, serve to actively distribute plasmids to daughter cells upon cell division in a manner that is conceptually similar to centromeres in eukaryotic cell division (Gerdes et al., 2004; Moller-Jensen et al., 2000; Pogliano, 2002). Plasmid molecules are grouped at the midcell division plane, and then copies are moved to areas within each daughter cell so that even distribution is obtained at cell division. This type of distribution has been visualized using plasmid molecules that have been fluorescently tagged, and the distribution has been shown to be dependent on the presence of an active partition system.

The components that form a partition system are remarkably conserved in different bacterial plasmid and chromosomal genomes. A typical partition system consists of the following plasmid-encoded components (with nomenclature based on the well-studied ParAB system of plasmid P1):

(1) A cis-acting plasmid DNA site consisting of several repeated regions (termed ParS in plasmid P1);
(2) A ParA or ParM protein that contains Walker-type or actin-like ATPase domains, respectively;

(3) A ParB protein that binds to the cis-acting site on the plasmid and recruits the ParA or ParM component to this complex.

The ParA or ParM/ParB/ParS complexes that form between copies of plasmid molecules become localized at midcell (possibly via an interaction with a bacterial host cell function) and are thought to be propelled toward the opposite cell poles (or ends) so that cell division compartmentalizes copies of the plasmid in each of the daughter cells. The ATPase activity of the ParA/ParM component is thought to play a critical role in this process. Alternatively, single Par complexes may bind to bacterial host cell functions that are evenly distributed on either side of the midcell but limited in number. There may be just enough of the bacterial host cell functions to bind plasmid copies on either side of midcell to ensure proper plasmid distribution upon cell division. In this way, all the plasmid copies do not bind to one side of the midcell, which would create plasmidless daughter cells.

3.3.2. Multimer Resolution Systems

Frequently, multimers of plasmid copies can form via homologous recombination. This can have the effect of lowering the effective plasmid copy number and decrease the frequency of plasmid inheritance by daughter cells. Plasmids have evolved multimer resolution systems to counteract this phenomenon (Barre and Sherratt, 2002; Sauer, 2002). These systems consist of: (1) a DNA site and (2) a recombinase protein that initiates recombination at the DNA site. Each single plasmid copy contains one DNA site. When two plasmid molecules fuse to form a dimer, there will be two DNA sites contained on this dimer molecule. The recombinase will perform a site-specific recombination reaction between the two DNA sites that will result in the dimer plasmid molecule being resolved back into two separate monomer plasmid copies. The best-known multimer resolution system is the *lox*/Cre system of plasmid P1 and Xer/*cer* of *E. coli* (Barre and Sherratt, 2002; Sauer, 2002). The *lox*/Cre system has been extensively utilized to perform site-specific recombination reactions that allow certain genetic engineering experiments to be performed in a wide array of cell types.

3.3.3. Post-segregational Lethality Systems

Post-segregational lethality systems serve to kill any plasmidless daughter cells that arise in the population (Engelberg-Kulka and Glaser, 1999). These systems are also called "plasmid addiction" or "toxin–antitoxin" systems. A plasmid encoding such a system expresses both a toxin that is lethal to the bacterial cells and an antitoxin that inactivates the toxin. The key point is that the antitoxin component has a very short half-life compared to the toxin component. As long as the plasmid is present in the cell, the antitoxin is produced and the host cell is protected. However, if a plasmidless daughter cell arises, the source of continually expressed antitoxin is gone and any remain-

ing antitoxin degrades very quickly. In this case, the stable toxin is free to mediate killing of the plasmidless host cell.

In the *hok*/*sok* system of plasmid R1, the toxin–antitoxin interaction takes place at the RNA level. The *hok* RNA encodes the toxin protein. The *sok* RNA can hybridize to the *hok* RNA molecule, prevent its translation, and signal for the hybrid to be degraded. Consistent with the model above, the *sok* RNA is quickly degraded in plasmidless cells and translation of the *hok*-encoded toxin commences. In other post-segregational lethality systems (those of F, R100, and RK2), the toxin and antitoxin are proteins that bind to one another in an inactive complex. Degradation of the antitoxin protein in plasmidless host cells releases the toxin to mediate the direct killing of these segregants.

3.4. Example of a Virulence Plasmid from Yersinia pestis

Figure 8 shows the map for virulence plasmid pCD1 from *Y. pestis* (Cornelis et al., 1998). The location of the replication and partition genes are indicated among the large number of genes dedicated to the type III protein secretion system used for virulence of the organism. Very similar plasmids are also found in *Y. enterocolitica* and *Y. pseudotuberculosis* indicating that this organization of genes has been conserved as the plasmid has been transferred to, and/or evolved with, different bacterial hosts.

4. Bacteriophages and Transduction

Bacteriophages (or "phages") are viruses that infect bacterial cells. They consist of a nucleic acid genome (usually DNA, but can also be RNA) that is enclosed within a polymeric protein package called a *capsid* (or "head") (see the phages depicted in Figure 9). Typical phages also possess a multiprotein structure called a *tail* that extends outward from the capsid and serves to interact with the bacterial cell surface during injection of the genome into the host cytosol. Phages bind a host bacterium at the surface via the tail (sometimes with the assistance of tail fibers that extend off the tail), and inject or shoot the genome inside the host where the intracellular part of the phage life cycle commences. As part of this life cycle, the phage DNA may integrate into the host chromosome or remain as an extrachromosomal element. Also, during packaging of the newly replicated phage genomes, the phage may "steal" a large section of the host chromosome and transfer it to a new host bacterium. Thus, bacteriophages are potent machines that can serve to drive the horizontal transfer of genetic material.

4.1. Lytic and Lysogenic Bacteriophages

The life cycle of phages can be divided into two distinct phases or "cycles": the lytic and the lysogenic. In the *lytic cycle*, the phage infects a host bacterium, the phage genes are expressed, new phage genomes are synthesized

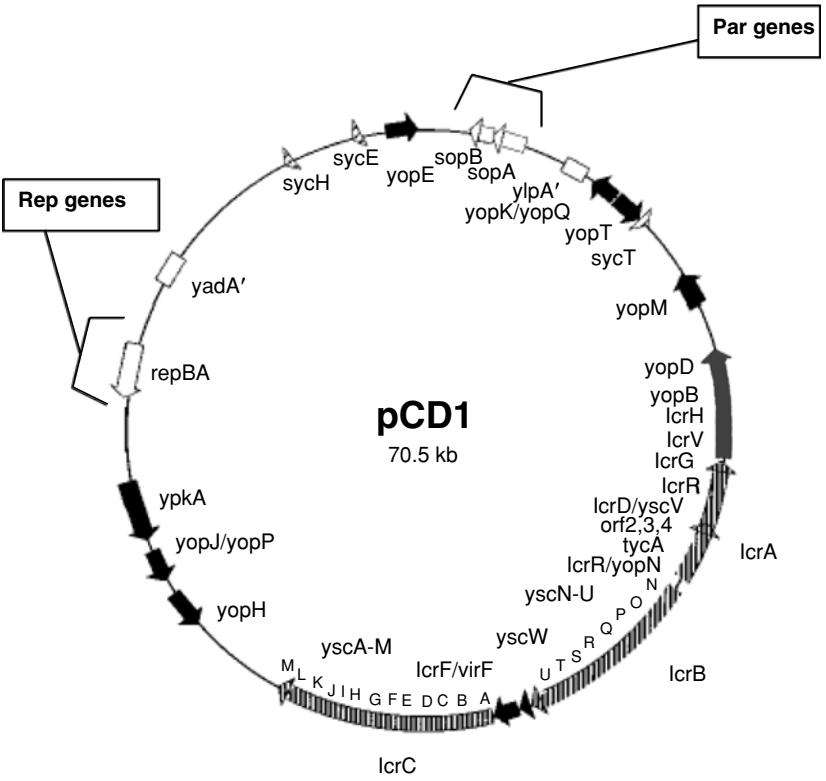

FIGURE 8. Virulence plasmid pCD1 from *Yersinia pestis.* Shown is a map of the virulence plasmid pCD1 from *Yersinia pestis*. The locations of the Rep and Par genes are indicated. The majority of the genes on this plasmid (the *yop*, *ysn*, *syc*, and *lcr* genes) are dedicated to the function of a type III secretion system, which is a major virulence determinant of this bacterium. This genetic structure is conserved on similar plasmids found in *Y. enterocolitica* and *Y. pseudotuberculosis*. (From Cornelis et al, 1998.)

and packaged into newly formed capsids, and the host cell lyses to release progeny phage particles that are able to infect other bacterial cells. This can be thought of as an "infect-replicate-disperse" phase of the phage life cycle. By contrast, the *lysogenic cycle* occurs when the phage genome integrates into the host chromosome, remains there in a "dormant" state (termed a "prophage"), and replicates along with the chromosome. At some point, usually in response to an environmental or cellular signal, the prophage will excise from the host chromosome and initiate a lytic cycle that leads to cell lysis and dispersal of newly formed phage particles. Some bacteriophages are only capable of lytic replication (e.g., the "T-even" phages T2, T4, and T6), while others are able to switch between the lytic and lysogenic cycles (e.g., lambda phage). These two main cycles of phage replication are depicted in Figure 9.

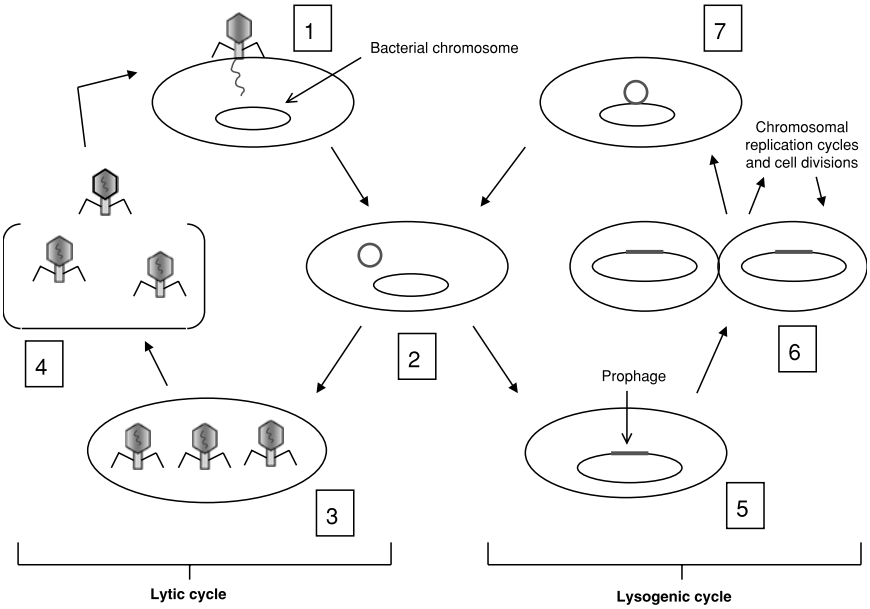

FIGURE 9. The lytic and lysogenic cycles of bacteriophage replication. This diagram shows a single type of phage that can enter either the lytic or lysogenic cycles, but there are phages that are only lytic and do not enter the lysogenic cycle. Also, the phage morphology depicted here shows a capsid, tail, and tail fibers, but other morphologies are common including those with no tail fibers and/or a very short tail. (1) The phage binds to the surface of a host bacterial cell and injects its genome into the cytoplasm. (2) Typically, the phage genome will circularize and the phage enters either the lytic or lysogenic cycle. The expression and stability of the integrase protein at this step is a key determinant of which cycle is entered. If the integrase is expressed and remains stable, this favors the lysogenic pathway. (3) In the lytic cycle, new phage genomes and packaging proteins (i.e. the capsid, tail, and tail fibers) are synthesized and new virions are assembled. (4) Cell lysis occurs and the new phage virions are released. Cell lysis is typically directed by one or more phage-encoded proteins. The newly released virions go on to bind other bacterial cells and initiate more replication cycles. (5) The lysogenic cycle begins with the integration of the phage genome into the host chromosome by the integrase protein (commonly with the assistance of certain host proteins as well). The stably integrated phage genome is termed a "prophage." (6) The prophage is replicated along with the host chromosome. This prophage can remain integrated in the host genome over a very large number of cell generations. (7) At a certain point (usually in response to an environmental or cellular signal), the prophage will excise from the host genome and initiate a lytic cycle.

An important concept to keep in mind when thinking about phages is that upon injection of a DNA phage genome into a host cell, the genome will frequently circularize. This covalently closed circular DNA can then serve as a substrate for a phage-encoded *integrase* protein that catalyzes the insertion of

the phage into the host chromosome (Azaro and Landy, 2002; Van Duyne, 2002). In some cases, the circular phage genome can replicate as a plasmid for several host cell generations before entering the lytic or lysogenic cycle. It is usually the expression and stability of the integrase protein that determines if the phage will enter the lytic or lysogenic cycle. In addition, the DNA site-specificity of the integrase frequently determines where in the host chromosome the circular genome will insert and remain located during the lysogenic cycle (Azaro and Landy, 2002; Van Duyne, 2002).

4.2. Generalized and Specialized Transduction

Transduction is the transfer of bacterial chromosomal DNA between cells via a bacteriophage. There are two main types of transduction: generalized and specialized. In *generalized transduction*, any part of the bacterial chromosome can be packaged and transferred by the phage (Figure 10A). This occurs when pieces of bacterial chromosome are mistakenly packaged into the capsid during phage assembly and are transferred to a new bacterium upon infection by the "chromosomal DNA-carrying" phage particle. The injected chromosomal DNA can be inserted into the new host genome via the endogenous recombination mechanisms. The bacterial host providing the packaged DNA is termed a "donor" and the host which receives that DNA by the phage is termed the "recipient." Usually, the genomes of the donor and recipient in a transduction are highly related. Therefore, the donor chromosomal fragment will integrate at the identical spot in the recipient genome as dictated by the homologous recombination mechanisms. However, any mutation that is present in the donor DNA on that fragment (such as a deletion, point mutation, or transposon insertion) is able to be transferred to the recipient genome.

In *specialized transduction*, only genes that are located near the phage (or "linked") are able to be packaged and transferred (Figure 10B). This happens when a lysogenic prophage excises and, due to a recombination error, carries an adjacent section of the host chromosome with it. The phage genome plus the chromosomal section are packaged in a capsid, transferred to a new host, and both are integrated into the new host chromosome.

The descriptions of generalized and specialized transduction given above show how phages can transfer sections of the host chromosome that have been mistakenly packaged along with the phage genome. It is also important to understand that phages can also transfer genes that have been inserted within the boundaries of their genomes, either by transposons, other phages, or some other kind of recombination mechanism. Several prophages have been identified that contain foreign genes, particularly toxins and virulence factors, which contribute to disease (Table 1) (Novick, 2003). Sometimes, a prophage will become inactivated by mutation over the course evolution and remain inserted at its particular chromosomal location. These prophage "remnants" are becoming increasingly recognized upon sequencing of different bacterial genomes and contribute to the notion that phages play a major role in the horizontal transfer of large blocks of genes (Brussow et al., 2004).

4.3. Regulation of Phage-encoded Toxins by Host-encoded Regulators: Diphtheria Toxin and Cholera Toxin

A striking example of microbial evolution is the regulation of phage-encoded genes by regulator proteins that are encoded outside of the phage genome in the host chromosome. Two classic examples of this phenomenon are illustrated by the regulation of diphtheria toxin encoded by the beta-prophage of

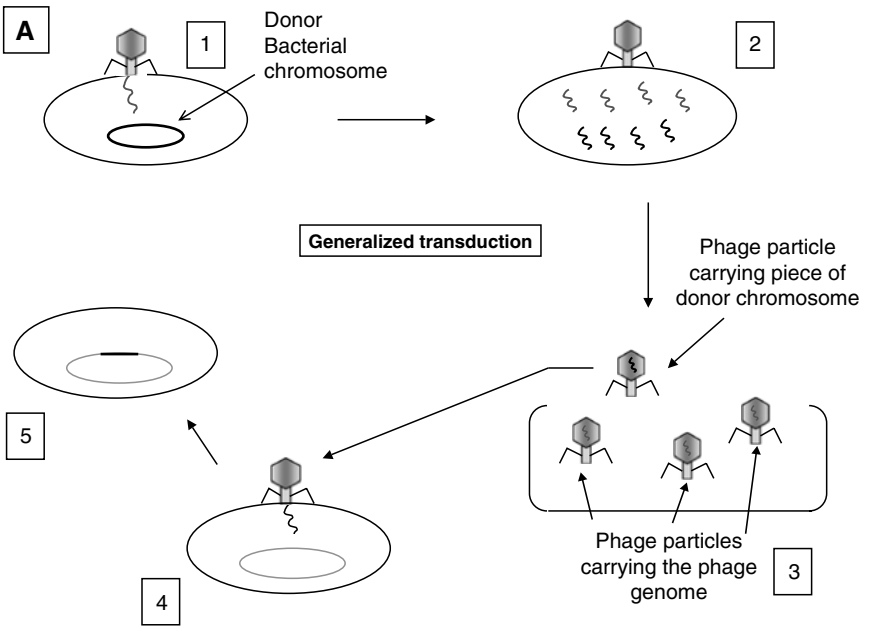

FIGURE 10. Generalized and specialized bacteriophage transduction of bacterial host genes. Panel A: *Generalized transduction*. (1) The phage injects its genome into the host cell, and the lytic cycle of replication is initiated. The infected cells at this point are termed donor cells because genes from their genome will be transferred to a new host. (2) During the lytic phase of certain phages, the host chromosome is fragmented into large pieces prior to cell lysis, and a mixture of newly replicated phage genomes and host chromosomal fragments is produced. (3) The host chromosomal fragments are able to be packaged into phage virions and then transduced to new host cells, termed recipient cells. (4) The phage virion carrying the host chromosomal fragment injects this genomic DNA piece into the new host cell. (5) The injected chromosomal fragment is able to recombine into the recipient host genome by homologous recombination. Usually, the genomes of the donor and recipient in a transduction are highly related. Therefore, the donor chromosomal fragment will integrate at the identical spot in the recipient genome as dictated by the homologous recombination mechanisms. However, any mutation that is present in the donor DNA on that fragment (such as a deletion, point mutation, or transposon insertion) is able to be transfered to the recipient genome. A key feature of generalized transduction is the fact that any donor chromosomal fragment can be packaged by the phage.

(*Continued*)

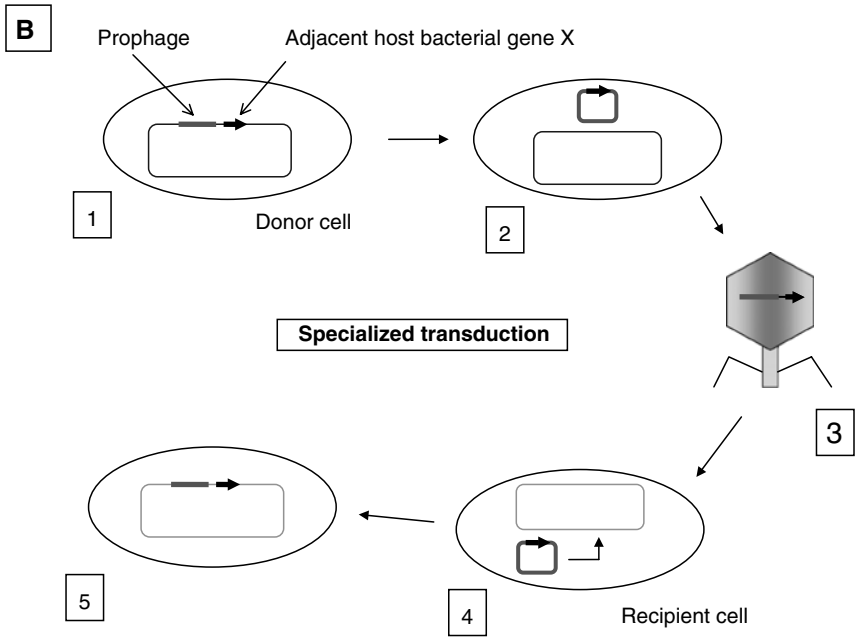

FIGURE 10. (*Cont'd*). Panel B: *Specialized transduction*. (1) In specialized transduction, a host bacterial gene located adjacent to a prophage (termed here "gene X") is transduced to a recipient host. (2) Usually prophage excision from the chromosome is precise and only includes phage DNA with no adjacent host DNA. However, occasional recombination errors occur such that an adjacent host gene(s) is excised with the phage genome. (3) The phage plus host genomic DNA fragment (called a "transducing fragment") is packaged by a phage virion. Typically, excision of the prophage results in a circular DNA molecule, but the phage packages the DNA as a linear molecule. (4) The transducing fragment is transferred to a recipient cell by the phage where it is integrated into the recipient chromosome at its specific integration site. As with generalized transduction, any mutation that is present in the transduced donor gene(s) is able to be transferred to the recipient. (5) The prophage and donor gene *X* are now located in the recipient chromosome.

Corynebacterium diphtheriae and cholera toxin (CT) encoded by the CTX phage of *V. cholerae* (Faruque et al., 1998; Holmes, 2000; Snyder and Champness, 2003). These examples show how genes carried by horizontally transferred elements can evolve to come under the control of host regulatory systems.

The gene encoding diphtheria toxin (termed *tox*) is located on a prophage (beta-prophage), but its expression is regulated by the DtxR protein, which is located elsewhere on the host chromosome (Holmes, 2000; Snyder and Champness, 2003). DtxR is highly related to Fur, the well-characterized iron response regulator in *E. coli*. Not surprisingly, the activity of DtxR is responsive to iron concentrations in a manner similar to Fur. In conditions of high iron concentrations, DtxR binds to ferrous ions (Fe^{2+}) and assumes a

conformation that allows it to bind DNA and act as a repressor of gene transcription. Several genes are repressed by DtxR under these conditions including the *tox* gene on beta-prophage. When iron concentrations drop to low levels (such as in a eukaryotic host environment), DtxR DNA-binding activity drops significantly and repression of the DtxR-regulated genes is lost. These genes, including *tox*, are now able to be expressed under the low-iron conditions where the activity of their products is likely to be advantageous to the bacteria. In this way, the toxin is not expressed outside the host where it is not needed.

Two genes encode the subunits of CT (*ctxA* and *ctxB*), and they are located on the *V. cholerae* CTX prophage. The *ctxAB* genes are regulated by the activity of several proteins (ToxR, ToxS, ToxT, TcpP, and TcpH) that are encoded by genes located elsewhere in the *V. cholerae* genome (Faruque et al., 1998; Snyder and Champness, 2003). The ToxR/ToxS and TcpP/TcpH protein pairs function together to activate the transcription of the gene encoding ToxT as well as other virulence factors. The ToxT protein then goes on to activate the transcription of the *ctxAB* genes. There are DNA-binding sites for ToxR/ToxS upstream of the *ctxAB* genes as well, and the ToxR/ToxS complex also contributes directly to promote *ctxAB* expression. Interestingly, the *toxT*, *tcpP*, and *tcpH* genes are located on another horizontally transferred genetic element, the *V. cholerae* pathogenicity island, termed VPI.

5. Transposons and the Transposition of DNA

Transposons are DNA segments that are able to transpose or "jump" from one genomic location to another. This is accomplished in large part by the activity of a *recombinase* (also called a *transposase*) that catalyzes the insertion of the transposon into a new DNA site. Different transposons may have different insertion site specificities: some insert randomly or near-randomly, while others display insertion site preferences (Craig et al., 2002). Transposable elements are ubiquitous in nature; they are found in cells from all branches of life. In bacterial cells, the transposition of DNA represents a key mechanism by which vertically and horizontally transferred genetic elements can exchange DNA segments. In addition, transposons are frequently used as convenient tools for mutagenesis and have been invaluable in the genetic analysis of many different types of bacterial species (Craig et al., 2002; Hayes, 2003).

Along with the transposase, the activity of a transposon depends on *inverted repeat* sequences that are located on the outside ends of the DNA element. These inverted repeats mark the boundaries of the transposon and are recognized by the transposase during the recombination event that allows transposon movement. Another characteristic of transposons is the presence of *direct repeats* on either end of the transposon, right next to the inverted repeats. The direct repeats represent the site of transposon insertion (or

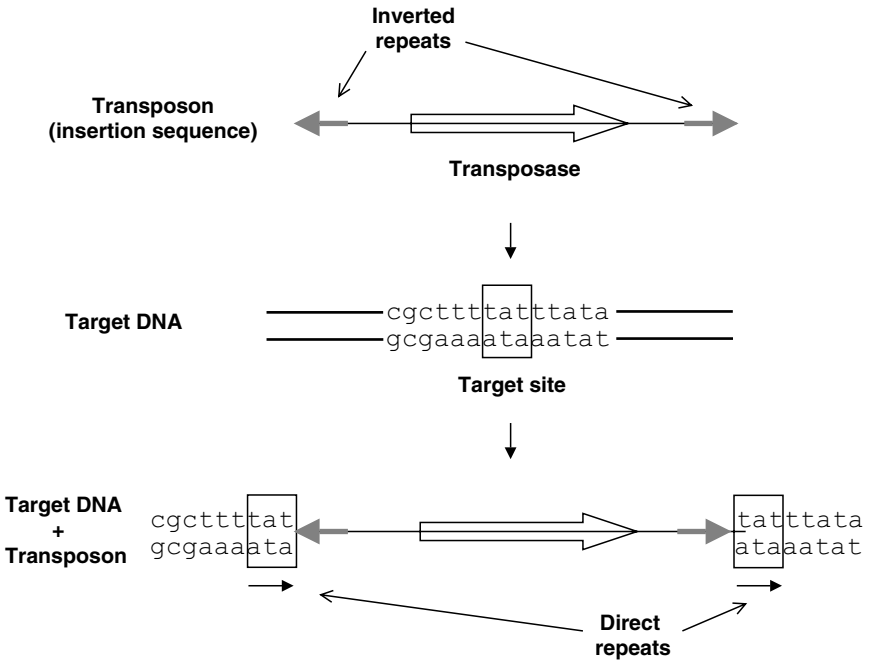

FIGURE 11. Structural features of a transposon. An insertion sequence is depicted as an example of a transposon. Note the inverted repeats formed by each end (that mark the boundary of the transposon) and the internal transposase gene. The inverted repeats vary in size between different transposons but are typically in the range of 15–50 bp. The target DNA contains the target site into which the transposon will insert. This site is duplicated as a result of the transposition reaction.

target site); this site gets duplicated during the recombination reaction. These essential features of transposons are depicted in Figure 11.

5.1. Insertion Sequences, Composite Transposons, and Noncomposite Transposons

The simplest type of transposon is called an *insertion sequence* (or IS element). A typical insertion sequence contains inverted repeats at its ends and a transposase enzyme located in between the inverted repeats (see Figure 12). Insertion sequences do not carry resistance genes and were first discovered because their insertion inactivated certain genes. Although many hundreds of different insertions sequences have been discovered, most of these can be placed in roughly 20 distinct families (Mahillon and Chandler, 1998).

A *composite transposon* is formed when two complete insertion sequences flank other genes, most commonly those encoding resistance to antibiotics.

The activity of the two insertion sequences moves the genes located between them so that the entire composite of two IS elements and associated genes transpose together. Certain composite transposons contain IS elements that are in the same orientation while others have them in inverted or opposite orientation. The composite transposons depicted in Figure 12 have inverted IS elements. An example of a composite transposon with IS elements in the same orientation is Tn9 (not shown in the figure).

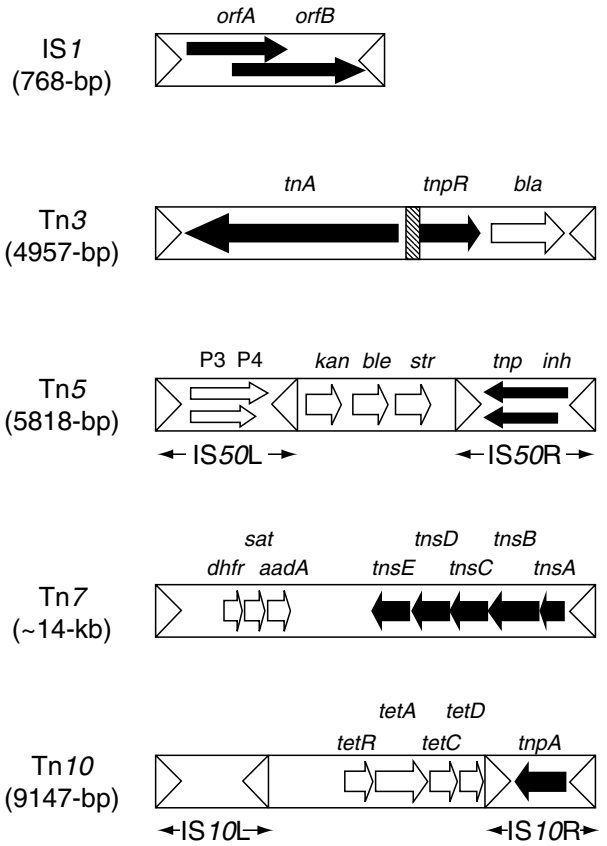

FIGURE 12. Structural organization of selected transposons. Several transposons are depicted as examples of different structural organization of these elements. The inverted repeats are designated as triangles at each end of the transposons, and the genes associated with each element are depicted as arrows. The IS1 element is an insertion sequence and represents the simplest type of transposon. Tn5 and Tn10 represent composite transposons formed by genes that have been "captured" between two insertion sequences. The insertion sequences and associated genes move as one unit during transposition. Tn3 and Tn7 are examples of noncomposite transposons consisting of an insertion sequence that has acquired antibiotic resistance genes between its inverted repeats. (From Hayes, 2003. Reprinted with permission from Annual Reviews, www.annualreviews.org)

Noncomposite transposons can be thought of as a single insertion sequence that contains an antibiotic resistance gene(s) in between its inverted repeats along with the transposase enzymes. Examples of noncomposite transposons are Tn3 and Tn7 as shown in Figure 12.

5.2. Cut-and-paste Versus Replicative Transposition

There are two main mechanisms by which transposons move from one location to another: cut-and-paste and replicative. In *cut-and-paste* transposition, the transposase binds to the ends of the element and brings them together such that the transposon "loops out" of its location. The transposase catalyzes DNA strand cleavage, which releases the transposon as a circular element that is able to integrate into a new DNA site (see Figure 13). A new copy of the transposon is not created during the cut-and-paste mechanism. Examples of cut-and-paste transposons are Tn5, Tn7, and Tn10 (Hayes, 2003; Snyder and Champness, 2003).

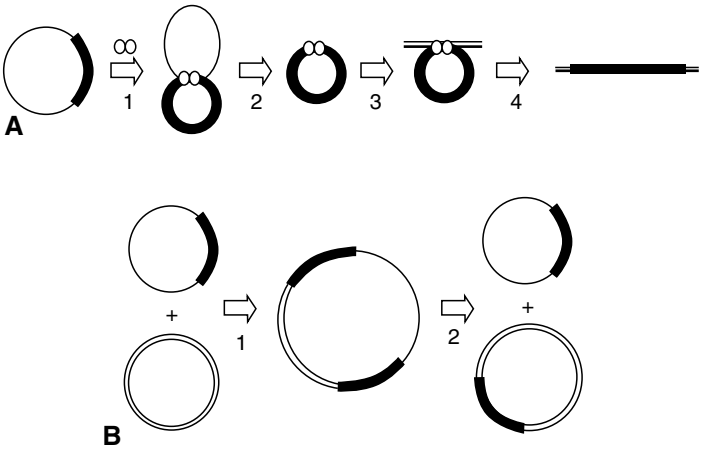

FIGURE 13. Cut-and-paste and replicative transposition. Panel A: *Cut-and-paste transposition*. (1) A transposon residing in a donor DNA molecule is "looped out" by the activity of the transposase enzyme (depicted here as two small circles to represent a dimer). (2) The transposase catalyzes the excision of the transposon from its original location as a circular molecule. (3) The transposase then catalyzes the integration of the transposon into the target DNA molecule. (4) The integrated transposon is depicted. Panel B: *Replicative transposition*. (1) During replicative transposition, a cointegrate is formed between the transposon donor DNA and the target DNA molecules. When the cointegrate is formed, the transposon is duplicated. (2) The activity of a resolvase enzyme catalyzes the resolution of the cointegrate. A copy of the transposon remains in the donor DNA molecule and a new copy now resides in the target DNA. (From Hayes, 2003. Reprinted with permission from Annual Reviews, www.annualreviews.org)

Replicative transposition involves duplication of the transposon and formation of a cointegrate between the donor and target DNA molecules (Figure 13). The resolution of the cointegrate is catalyzed by the activity of a transposon-encoded enzyme called a *resolvase*. This resolution leaves a copy of the transposon in the donor molecule and a new copy in the target molecule. The resolvase is distinct from the transposase; the transposase catalyzes the initial strand cleavage and joining reactions between the donor and target molecules to form the cointegrate, while the resolvase serves to resolve the cointegrate into two separate molecules again. Examples of replicative transposons are Tn3 and the phage Mu transposon (Hayes, 2003; Snyder and Champness, 2003).

6. The Modular Structure of Mobile Genetic Elements

From the descriptions of the major DNA vehicles above, it is clear that their biology involves a number of different functions such as replication, conjugation, phage particle formation, integration, and excision. A useful way to think about these functions is to view them as genetic modules that are combined in different ways to form a given mobile DNA element (Burrus and Waldor, 2004; Osborn and Boltner, 2002; Toussaint and Merlin, 2002). For example, a particular bacteriophage may consist of modules for genome replication, integration, and packaging. Likewise, a given plasmid may contain modules for replication, conjugation, and maintenance. When mobile genetic elements are viewed as combinations of specific modules, it is clear how different elements evolve via the "mixing and matching" of modules. We have explained the basic mobile DNA vehicles above: plasmids, bacteriophages, and transposons. However, nature has evolved many variants of mobile DNA elements. The P1 bacteriophage genome is essentially a large, stably maintained plasmid that encodes packaging functions for movement from host to host. The bacteriophage Mu is a randomly inserting transposon that can be packaged for transfer.

In Section 6.1, we will give some examples of different kinds of mobile genetic elements that represent unique combinations of modules and deviate from typical classification schemes. Below is a list of the different genetic modules that are used in various combinations to create different elements (Burrus and Waldor, 2004; Osborn and Boltner, 2002; Toussaint and Merlin, 2002). These modules are based on functions found in the basic genetic vehicles described above and are illustrated in Figure 14.

6.1. Genetic Modules Found in Mobile DNA Elements

(1) *Replication (Rep)* represents a replicon that drives DNA replication.
(2) *Transfer (Tra)* functions are used for Mpf to allow cell contact and DNA pore generation for conjugative transfer.

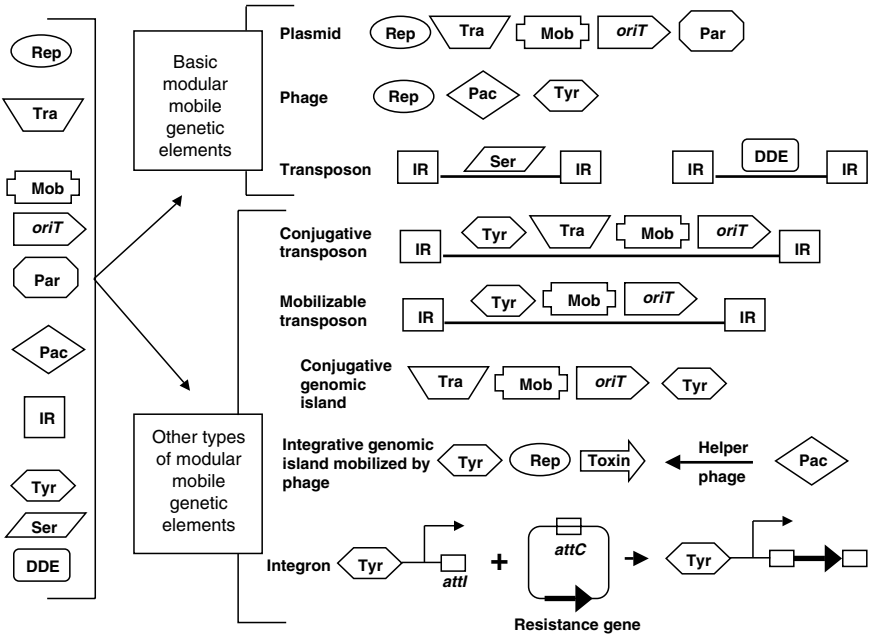

FIGURE 14. Combinations of genetic modules form different mobile DNA elements. This diagram shows the different combinations of genetic modules that make up a variety of mobile DNA elements. In a column on the left side of the diagram, the functions of the different genetic modules are shown: Rep (replication genes), Tra (conjugative transfer genes corresponding to the mating pair formation (Mpf) system), Mob (mobilization genes corresponding to the DNA transfer and replication (Dtr) system that processes the *oriT* for transfer), *oriT* (origin of transfer), Par (genes involved in plasmid maintenance or partition), Pac (genes involved packaging DNA into a phage capsid), IR (inverted DNA repeats), Tyr (tyrosine recombinase typical of bacteriophages and other mobile DNA elements), Ser (serine recombinase found in certain transposons such as Tn3 and gamma-delta), DDE (transposase found in transposons such as Tn7 and Tn10). In different combinations, these genetic modules form the structure of mobile DNA elements that can be divided into two main groups: (1) basic modular mobile genetic elements such as plasmids, bacteriophages, and transposons and (2) other types of modular mobile genetic elements. Details about each mobile element can be found in the text. (Based on concepts presented in Osborn and Boltner, 2002.)

(3) *Mobilization (Mob)* relaxase functions process the *oriT* for conjugative transfer and correspond to the DNA transfer and replication functions (Dtr) indicated above.
(4) *Origin of tranfer (oriT)* is the DNA site on a mobile DNA element recognized by Mob functions for conjugation.
(5) *Maintenance/partition (Par)* functions ensure faithful inheritance and stable maintenance of an autonomously replicating DNA element.

(6) *Packaging (Pac)* genes are necessary for the packaging of mobile DNA into protein capsid particles.
(7) *Inverted repeats (IR)* are DNA sites that are recognized by recombinases for the integration and excision (i.e., transposition) of mobile DNA.
(8) *Recombinases* (*Tyr*, *Ser*, and *DDE*)—*Tyr*: tyrosine recombinase found in several bacteriophages and other mobile DNA elements; *Ser*: serine recombinase (resolvase/invertase) found in certain transposons Tn3 and γδ; *DDE*: typical transposase found in Tn7 and Tn10. Virtually all mobile DNA recombinases can be assigned to one of these families. The reference to tyrosine and serine in the classification of these recombinases refers to conserved amino acids that are involved with the catalytic mechanisms of each enzyme group.

6.2. Other Types of Mobile Genetic Elements that Are Combinations of Modules

These examples of mobile genetic elements consist of different combinations of the modules listed above and in Figure 14. Their discovery and characterization indicate that not all mobile DNA elements fall into standard categories such as plasmid, bacteriophage, and transposon. A common feature of these elements is their ability to integrate into, and excise from, the host bacterial chromosome. In addition, these elements are able to self-transfer or be mobilized from one host to another and then integrate into the genome of the new host. Typically, an extrachromosomal replication phase is not a predominant part of their life cycle; these elements tend to rely on an "excision-transfer-integration" mode of existence.

6.2.1. The ICE Family

A number of these elements have been ascribed different names such as conjugative transposons, conjugative genomic islands, mobilizable transposons, and conjugal, self-transmissible, integrating element (CONSTIN) (Burrus and Waldor, 2004; Osborn and Boltner, 2002; Toussaint and Merlin, 2002). Recently, it was proposed that all such elements fall under the term "ICE," standing for *i*ntegrative and *c*onjugative *e*lement. The term "ICEland" has also been proposed as a fusion between *ICE* and genomic is*land*. We will describe various ICEs below under the descriptive names they have already been ascribed, but it is helpful to think of the different elements as subclasses of ICEs. Hopefully, these examples of mobile DNA elements illustrate the vast number of mechanistic possibilities that exist when genetic modules are put together in different combinations.

6.2.1.1. Conjugative Transposons

Conjugative transposons utilize excision and integration to jump between different genomes (similar to transposons and integrative phages), but with the

added feature that they encode their own conjugation systems that direct their self-transfer from a donor bacterium to a recipient (Churchward, 2002). These elements utilize a circular intermediate that is produced by excision from the host genome and this intermediate serves as a substrate for conjugative transfer. The element recircularizes in the recipient and is then integrated into the new host genome. The integration and excision of conjugative transposons is mediated by tyrosine recombinases that are more similar to those that perform site-specific recombination (similar to phage integration) than those that catalyze movement of typical transposons (like the Ser and DDE recombinases). Conjugative transposons can integrate at different sites in the genome, although these sites have related sequence features that indicate the existence of preferred targets. Two examples of conjugative transposons are Tn916 from *Enterococcus faecalis* and CTnDOT from *Bacteriodes* spp.

Tn916. Tn916 was first identified in the Gram-positive opportunistic pathogen *E. faecalis* and encodes resistance to tetracycline via the *tetM* determinant (Churchward, 2002). Tn916 integration is mediated by a tyrosine recombinase, Int, while excision is mediated by both Int and another recombinase termed Xis; Int is a member of the lambda integrase family, while Xis is a small, basic protein that is very similar to the lambda Xis protein. The target site specificity of Tn916 integration can be considered semi-random since it displays a clear preference for A/T-rich sites with a consensus target of 5′-TT/ATTTT(*N*6)AAAAAA/TA-3′ (where *N* is any nucleotide).

Tn916 encodes its own conjugation system that contains an *oriT* site, which is similar to the nick sites of plasmids F and R6K and functions to transfer a single DNA strand to the recipient. The product of the Tn916 conjugation gene *orf21* is a member of the SpoIIIE/FtsK family of proteins that can function to drive DNA transport between cells. The Tn916 conjugation system displays a broad host range, capable of mediating transfer between a variety of Gram-positive cells and between Gram-positive and Gram-negative cells.

The excision and insertion of Tn916 involves a novel DNA intermediate where 6 bp regions at the ends of the element form a heteroduplex (a double-stranded DNA structure where both strands are nonhomologous) to form a circular molecule. Heteroduplex DNA is also formed upon insertion of Tn916 into its target site. These heteroduplexes are resolved upon replication of the newly inserted Tn916 and its flanking regions. This mechanism explains why the ends of integrated Tn916 molecules are rarely homologous.

CTnDOT. CTnDOT is part of a family of related mobile elements that are widely distributed among natural isolates of different *Bacteroides* spp. and are responsible for mediating rampant spread of antibiotic resistance in these bacteria (Cheng et al., 2001; Whittle et al., 2002). CTnDOT carries the *tetQ* gene, which confers tetracycline resistance. In addition to directing their own self-transfer, the CTnDOT elements also mobilize other transposons termed non-replicating *Bacteroides* units (NBUs), which are discussed below.

Like Tn916, CTnDOT integration utilizes a tyrosine recombinase, Int, that is a member of the lambda integrase family. Excision of CTnDOT requires both Int and a small, topoisomerase-like protein called Exc. Although CTnDOT can integrate at multiple sites, it displays greater site specificity than Tn916, resulting in a range of preferred sites, and is smaller than that of Tn916. The target site consensus for CTnDOT is a 10 bp sequence: 5′-GTTNNTTTGC-3′. A highly homologous sequence is also found at one end of CTnDOT, providing a possible reason why the element is directed to integrate at the consensus target sites.

A remarkable feature of CTnDOT is the fact that tetracycline can induce its excision and transfer several thousand fold. This induction is dependent on a regulatory system encoded by CTnDOT that consists of three genes: *rteA*, *rteB*, and *rteC*. The *rteA* and *rteB* genes are part of an operon containing the *tetQ* resistance gene, and this operon is induced by the presence of tetracycline. The RteA and RteB proteins then mediate the expression of *rteC*, the product of which goes on to regulate the expression of several CTnDOT genes involved in excision and transfer.

6.2.1.2. Mobilizable Transposons

Some excision and insertion-based mobile elements do not encode their own conjugation systems but are able to be conjugatively-mobilized from a donor to a recipient (Churchward, 2002). These elements have been termed "mobilizable transposons". Three major types of these elements have been found, represented by the examples listed below:

Non-replicating Bacteroides Units (NBUs). Along with the CTnDOT conjugative transposons, NBUs are important vehicles of antibiotic resistance transmission among *Bacteriodes* and possibly other Gram-positive species. NBU1 is the best-studied member of this group, though other related NBUs, NBU2 and NBU3, are also found in *Bacteriodes* (Churchward, 2002; Shoemaker et al., 1996). These elements contain an *oriT* site and a nickase enzyme (Mob) that processes the *oriT* for transfer by conjugative functions provided by a coresident conjugative transposon. In addition, DNA vectors that carry the NBU1 mobilization region can be mobilized by the IncP plasmid conjugation system in *E. coli*, indicating its ability to function with different transfer systems. Integration of NBU1 requires IntN1, a tyrosine recombinase belonging to the lambda integrase family, and the target site is specifically located at the 3′ end of a leucine tRNA gene. This target site contains a 14 bp sequence that is also present at one end of NBU1.

Tn4551 from Clostridia *spp.* This element is distinguished from NBUs by the fact that it encodes a serine recombinase, TnpX, belonging to the resolvase family and can integrate at multiple different genomic sites (Lyras and Rood, 2000). It contains an *oriT* and the Mob-encoding gene *tnpZ*, and has been shown to be mobilized by the IncP plasmid RP4. An interesting feature of Tn4551 is that the site used for integration on the transposon is located in the

promoter for the *tnpX* gene encoding the recombinase. When the circular form of Tn4551 is integrated at the target site, the –10 and –35 regions of the *tnpX* promoter become separated from both ends of the integrated, linear transposon. This serves to turn off *tnpX* expression unless it integrates downstream of another properly positioned –35 region sequence or whole promoter. Excision of Tn4551 to its circular form restores the *tnpX* promoter and allows normal TnpX expression.

Tn5398 from Clostridium difficile. This element is unique in that it possesses an *oriT* sequence that allows it to be mobilized by separate conjugation functions, but does not appear to encode its own Mob protein that would nick the *oriT* for transfer (Farrow et al., 2001). It also does not appear to encode a typical recombinase enzyme that would mediate its integration and excision. Analysis of the entire Tn5398 sequence did not reveal any genes with obvious homology to either Mob protein or recombinase genes. It is thought that the coresident conjugative transposon Tn5397 provides these functions for Tn5398 excision, integration, and mobilization.

6.2.1.3. Conjugative Genomic Islands

The conjugative genomic islands contain Tra, Mob, and *oriT* modules that mediate their self-transfer from donor to recipient (Burrus and Waldor, 2004; Osborn and Boltner, 2002; Toussaint and Merlin, 2002). They also typically possess a tyrosine recombinase of the lambda integrase family that mediates integration and excision. However, they are distinguished from conjugative transposons in that they integrate at a specific site in their host genome, commonly at one end of a tRNA or other house-keeping gene, as opposed to multiple target sites. A distinguishing feature of genomic islands is that they are found at their particular integration site in certain strains and absent from this site in other closely related strains (see Chapter 4). The conjugative genomic islands provide an explanation for how at least some of the genomic islands are able to horizontally transfer between host strains.

It is important to realize that conjugative genomic islands and conjugative transposons are very similar mobile elements. The distinction made between these two elements in the literature based on their integration sites helps to categorize different subclasses of ICEs, but it is possible under certain conditions for a conjugative genomic island to integrate at sites other than the common preferential location. Also, a number of conjugative genomic islands were termed "IncJ" plasmids when they were initially discovered, and this was based on certain incompatibility features shared by these common elements (most likely related to common recombinase functions) (Churchward, 2002). However, these elements lack any discernible plasmid replicon, and strains harboring them do not yield a stable, circular extrachromosomal form. Their classification as plasmids is now considered inappropriate and further study has revealed that they are part of the ICE family.

Analysis of different conjugative genomic islands has revealed a remarkable conservation of modular organization and DNA sequence homology between elements isolated from different species. Comparison of the genomes of the elements SXT (*V. cholera*), R391 (*Providencia rettgeri*), R997 (*Proteus mirabilis*), and pMERPH (*Shewenalla putrefaciens*) showed a highly conserved backbone consisting of integration, conjugation, and regulatory functions punctuated by auxiliary sequences that included transposons and antibiotic resistance determinants (Boltner and Osborn, 2004). In addition, all these elements integrate at the same genomic site, the 5′ end of the *prfC* gene encoding peptide release factor 3 (RF3). These observations strongly suggest that these elements, though isolated from different bacterial genera at vastly different geographic locations, have a common ancestral origin.

The SXT element from *V. cholera* is a 100 kb genomic island that contains genes coding for chloramphenicol, sulphamethoxazole, trimethoprim, and streptomycin resistance. SXT and related elements had not been detected in *V. cholera* strains before 1993, but are now present in the vast majority of current clinical *V. cholera* isolates from Asia. Recently, it was demonstrated that the SOS response pathway promotes conjugative transfer of the SXT element in a manner similar to lambda phage induction (Beaber et al., 2004). The "SOS-activated" form of RecA protein facilitates the autocleavage of the SXT SetR repressor protein, which alleviates repression of another pair of SXT genes, *setC* and *setD*. The SetC and SetD proteins then induce the expression of a number of SXT genes required for SXT excision and transfer. Remarkably, one of the agents used to induce the SOS response in these experiments, and thus induce SXT transfer, was the fluoroquinolone antibiotic ciprofloxacin. This indicates that the clinical and agricultural use of SOS-inducing antibiotics may promote the horizontal transfer of antibiotic resistance via the SXT-type elements.

6.2.2. Non-ICE Modular Genetic Elements

As noted previously, the elements described above are termed integrative and conjugative elements or ICEs. Below, we discuss some other examples of modular genetic elements that seem to fall out of the ICE category because conjugation is not an integral part of their description. For clarity, we will put them under the "non-ICE" category.

6.2.2.1. Integrative Genomic Islands Mobilized by Bacteriophages

There are some examples of genomic islands that are mobilized by a helper bacteriophage. These islands are integrated at a specific site in the host genome, but can be excised and transferred from donor to recipient in the presence of a helper transducing phage. These elements are distinguished from pathogenicity genes and islands that have been inserted into the genome of an active bacteriophage and are specifically transduced by this phage as

part of its packaged genome. Instead, the examples discussed here are able to "hitch a ride" in transducing particles formed by unlinked helper phages.

Staphylococcus aureus *Pathogenicity Island 1 (SaPI1)*. This element encodes the toxic shock syndrome toxin 1 (TSST-1) and two other newly discovered superantigens termed SEK and SEL. It is part of a large family of similar *Staphylococcal* DNA elements that encode superantigens and are integrated at specific sites in the *S. aureus* chromosome. In the presence of phage 80α, SaPI1 can be excised from its genomic site (assisted by the activity of a phage 80α-encoded excision function) and transduced to recipients at a very high frequency (Lindsay et al., 1998; Novick, 2003). There are two populations of phage particles produced in this event: normal-sized particles containing the 80α genome and smaller particles (about 1/3 the size) containing the SaPI1 island. Upon infection of recipient cells, incoming SaPI1 DNA can be detected in a linear form that immediately replicates, forms circular intermediates, and integrates into the genome at its specific *att* site. Another SaPI, termed SaPI2, is similarly excised and transduced by phage 80α. In the absence of 80α, both these islands are very stable and do not detectably excise and transfer. The SaPI elements represent a remarkable example of how a genomic island can cooperate with a bacteriophage to facilitate its transfer.

Vibrio cholera *VPI*. The toxin responsible for severe diarrhea caused by *V. cholera* infection, cholera toxin (CT), is contained on a filamentous bacteriophage that uses the toxin-coregulated pilus (TCP) as its essential receptor during transduction. The TCP is encoded by the large *Vibrio* pathogenicity island (VPI) that is located on chromosome 1 of *V. cholera*, whose genome consists of two circular chromosomes. It has been proposed that VPI is actually another *V. cholera* filamentous phage and that the TCP serves as the coat for this phage (Karaolis et al., 1999). After this initial proposal, several lines of evidence have cast significant doubt on this hypothesis, and other alternative explanations have been used to explain the nature of VPI (Davis and Waldor, 2003; Faruque et al., 2003). Recently, VPI was shown to be transduced via a helper phage (CP-T1) to four different strains of *V. cholera* and insert into its specific integration site in the recipient chromosome (O'Shea and Boyd, 2002). This result strongly suggests that this may be the preferred mechanism for the horizontal transfer of VPI. It has subsequently been shown that VPI excises from its integration site using two VPI-encoded recombinases, Int and VpiT (Rajanna et al., 2003). This excised form is likely the substrate that is packaged by the helper phage.

6.2.2.2. Integrons

Integrons consist of an integrase gene belonging to the tyrosine recombinase family (IntI), a specific integration site termed *attI*, and a promoter that directs transcription through the *attI* site (Komano, 1999; Recchia and Sherratt, 2002). This basic unit then serves as the platform for integration of

genetic elements called "gene cassettes," which consist of a gene (commonly encoding antibiotic resistance) and an integrase-specific site termed *attC*. Site-specific recombination between the integron and the circular form of gene cassettes occurs at the integron *attI* site and the *attC* site. This leads to the insertion of the gene cassette downstream of the *attI*-associated integron promoter and subsequent expression of the inserted gene. Several gene cassettes are able to be inserted into a single integron. "Super-integrons" have been described that contain a very large number (sometimes hundreds) of gene cassettes, and these elements have been shown to be a source of gene cassettes for smaller integrons. Integrons and super-integrons have been found in a diverse range of Gram-negative hosts and are commonly located within transposons and conjugative plasmids (though super-integrons tend to be more stable components of a host genome and not part of the mobile gene pool). Five major classes of integrons have been defined based on the homology of the integrase genes and *att* sites. The ability of integrons to "capture" antibiotic resistance genes and contribute to their dissemination is thought to have played a major factor in the rapid evolution of multiple resistance observed worldwide.

7. Conclusions—A World of Genetic Modules

If evolution is a race, horizontal gene transfer is one of the strategies that give bacteria a huge advantage in that race. Moreover, bacteria do not rely on one plan for this strategy—many types of different modules contribute to horizontal gene transfer. Genetic modules for DNA replication, conjugation, transduction, integration, and excision have all been combined in different ways to create the many types of genetic elements described in this chapter. Armed with this information, molecular microbiologists now have a deeper appreciation of the genetic flexibility of bacteria and the different possibilities that exist when studying the evolution of different microbes. This becomes especially important as we sequence more and more bacterial genomes and strive to develop new ways to combat harmful bacteria.

Questions to Consider

1. During transposition of a cut-and-paste transposon, replication of the transposon does not occur. How do these types of transposons replicate? Is replication of a transposon (or any type of mobile DNA element) important to its survival?

Cut-and-paste transposons are replicated along with the rest of the genome in which it resides and therefore rely on the replication functions the host organism. Therefore, transposition and replication of cut-and-paste transposons are two separate events. Replicative transposons are replicated during both host chromosomal DNA replication and a transposition event.

Replication of a transposon (or any mobile DNA element) is *the most* important thing to its survival. Therefore, learning about how these different DNA elements propagate themselves is key to understanding their existence.

2. Suppose a new mobile DNA element has been discovered that is able to maintain itself as an extrachromosomal circular molecule, but has the ability to integrate into the host chromosome and be transduced to other hosts. List and describe the genetic modules that are likely contained in this new mobile element.

The genetic modules likely present in this element would be: (1) Rep—for its replication as a plasmid-like element; (2) Par—for its stable maintenance as a plasmid; (3) Tyr recombinase—for chromosomal integration (Tyr because it has phage-like features); (4) Pac—for packaging the element for transduction.

3. Describe the three major components of a conjugation system.

The three components of a conjugation system are: (1) an origin of transfer (*oriT*)—the DNA site on the substrate molecule that is recognized and processed by the relaxase; (2) the DNA transfer and replication proteins (Dtr)—refers to the relaxase enzyme that binds and nicks the *oriT* and the coupling protein that joins the *oriT*/relaxase complex to the conjugal pore formed by the mating pair formation proteins; (3) the mating pair formation proteins (Mpf)—these proteins form the sex pilus and conjugal pore.

4. You have discovered a site between two genes in the fully sequenced *Escherichia coli* genome that serves as the target for integration of different mobile genomic islands. You have recently been given a large number of different, uncharacterized clinical isolates of *E. coli* and wish to determine if islands have integrated into this site in these isolates. What strategies could be used to do this?

Some possible strategies could be: (1) with primers designed to hybridize on either side of the target site, use one of the newly improved strategies for directly sequencing bacterial chromosomal DNA using the DNA from the clinical isolates as a template; (2) with the same target site-flanking primers, perform a PCR using the clinical isolate chromosomal DNA as a template to determine if a high–molecular weight product is formed. This would work best if the inserted islands are less than 15 kb in size and a highly processive PCR amplification enzyme is used; (3) perform a Southern blot with clinical isolate chromosomal DNA that has been digested with restriction enzymes that would give an easily identifiable change in restriction pattern if an island is inserted at the target site.

References

Avery, O., Macleod, C., and McCarty, M. (1944). Studies on the chemical nature of the substance inducing transformation of pneumococcal types. *J. Exp. Med.* 79:137–158.

Azaro, M. and Landy, A. (2002). Lambda integrase and the Lambda integrase family. In Craig, N., Craigie, R., Gellert, M., and Lambowitz, A. (eds.) *Mobile DNA II*. Washington, DC: ASM Press.

Barre, F. and Sherratt, D. (2002). Xer site-specific recombination: promoting chromosome segregation. In: Craig, N., Craigie, R., Gellert, M., and Lambowitz, A. (eds.) *Mobile DNA II*. Washington, DC: ASM Press.

Beaber, J. W., Hochhut, B., and Waldor, M. K. (2004). SOS response promotes horizontal dissemination of antibiotic resistance genes. *Nature*. 427(6969):72–74.

Boltner, D. and Osborn, A. M. (2004). Structural comparison of the integrative and conjugative elements R391, pMERPH, R997, and SXT. *Plasmid*. 51(1):12–23.

Brussow, H., Canchaya, C., and Hardt, W. D. (2004). Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol. Mol. Biol. Rev*. 68(3):560–602.

Burrus, V. and Waldor, M. K. (2004). Shaping bacterial genomes with integrative and conjugative elements. *Res. Microbiol*. 155(5):376–386.

Buu-Hoi, A. and Horodniceanu, T. (1980). Conjugative transfer of multiple antibiotic resistance markers in *Streptococcus pneumoniae. J. Bacteriol*. 143(1):313–320.

Cascales, E. and Christie, P. J. (2003). The versatile bacterial type IV secretion systems. *Nat. Rev. Microbiol*. 1(2):137–149.

Chen, I. and Dubnau, D. (2004). DNA uptake during bacterial transformation. *Nat. Rev. Microbiol*. 2(3):241–249.

Cheng, Q., Sutanto, Y., Shoemaker, N. B., Gardner, J. F., and Salyers, A. A. (2001). Identification of genes required for excision of CTnDOT, a *Bacteroides* conjugative transposon. *Mol. Microbiol*. 41(3):625–632.

Christie, P. J. (2001). Type IV secretion: intercellular transfer of macromolecules by systems ancestrally related to conjugation machines. *Mol. Microbiol*. 40(2):294–305.

Churchward, G. (2002). Conjugative transposons and related mobile elements. In Craig, N., Craigie, R., Gellert, M., and Lambowitz, A. (eds.) *Mobile DNA II*. Washington, DC: ASM Press.

Claverys, J. P. and Havarstein, L. S. (2002). Extracellular-peptide control of competence for genetic transformation in *Streptococcus pneumoniae. Front. Biosci*. 7:d1798–1814.

Cornelis, G. R., Boland, A., Boyd, A. P., Geuijen, C., Iriarte, M., Neyt, C., Sory, M. P., and Stainier, I. (1998). The virulence plasmid of *Yersinia*, an antihost genome. *Microbiol. Mol. Biol. Rev*. 62(4):1315–1352.

Craig, N., Craigie, R., Gellert, M., and Lambowitz, A. (2002). *Mobile DNA*. Washington, DC: ASM Press.

Davis, B. M. and Waldor, M. K. (2003). Filamentous phages linked to virulence of *Vibrio cholerae. Curr. Opin. Microbiol*. 6(1):35–42.

del Solar, G., Giraldo, R., Ruiz-Echevarria, M. J., Espinosa, M., and Diaz-Orejas, R. (1998). Replication and control of circular bacterial plasmids. *Microbiol. Mol. Biol. Rev*. 62(2):434–464.

Duckworth, D. H. (1976). Who discovered bacteriophage? *Bacteriol. Rev*. 40(4):793–802.

Engelberg-Kulka, H. and Glaser, G. (1999). Addiction modules and programmed cell death and antideath in bacterial cultures. *Annu. Rev. Microbiol*. 53:43–70.

Farrow, K. A., Lyras, D., and Rood, J. I. (2001). Genomic analysis of the erythromycin resistance element Tn5398 from *Clostridium difficile. Microbiology*. 147(Pt 10):2717–2728.

Faruque, S. M., Albert, M. J. and Mekalanos, J. J. (1998). Epidemiology, genetics, and ecology of toxigenic *Vibrio cholerae. Microbiol. Mol. Biol. Rev*. 62(4):1301–1314.

Faruque, S. M., Zhu, J., Asadulghani, Kamruzzaman, M., and Mekalanos, J. J. (2003). Examination of diverse toxin-coregulated pilus-positive *Vibrio cholerae*

strains fails to demonstrate evidence for *Vibrio* pathogenicity island phage. *Infect. Immun*. 71(6):2993–2999.
Franke, A. E. and Clewell, D. B. (1981). Evidence for a chromosome-borne resistance transposon (Tn916) in *Streptococcus faecalis* that is capable of "conjugal" transfer in the absence of a conjugative plasmid. *J. Bacteriol*. 145(1):494–502.
Freeman, V. J. (1951). Studies on the virulence of bacteriophage-infected strains of *Corynebacterium diphtheriae. J. Bacteriol*. 61(6):675–688.
Gerdes, K., Moller-Jensen, J., Ebersbach, G., Kruse, T., and Nordstrom, K. (2004). Bacterial mitotic machineries. *Cell*. 116(3):359–366.
Griffith, F. (1928). The significance of pneumococcal types. *J. Hyg*. 27:113–159.
Hayes, F. (2003). Transposon-based strategies for microbial functional genomics and proteomics. *Annu. Rev. Genet*. 37:3–29.
Hershey, A.D and Chase, M. (1952) Independent functions of viral protein and nucleic acid in growth of bacteriophage. *J Gen Physiol*. 36:39–56.
Holmes, R. K. (2000). Biology and molecular epidemiology of diphtheria toxin and the tox gene. *J. Infect. Dis*. 181 (Suppl 1):S156–167.
Karaolis, D. K., Somara, S., Maneval, D. R., Jr., Johnson, J. A., and Kaper, J. B. (1999). A bacteriophage encoding a pathogenicity island, a type-IV pilus and a phage receptor in cholera bacteria. *Nature*. 399(6734):375–379.
Khan, S. A. (1997). Rolling-circle replication of bacterial plasmids. *Microbiol. Mol. Biol. Rev*. 61(4):442–455.
Kobryn, K. and Chaconas, G. (2001). The circle is broken: telomere resolution in linear replicons. *Curr. Opin. Microbiol*. 4(5):558–564.
Komano, T. (1999). Shufflons: multiple inversion systems and integrons. *Annu. Rev. Genet*. 33:171–191.
Lederberg, J. and Tatum, E. (1946). Gene recombination in *E. coli. Nature*. 158:558.
Lindsay, J. A., Ruzin, A., Ross, H. F., Kurepina, N., and Novick, R. P. (1998). The gene for toxic shock toxin is carried by a family of mobile pathogenicity islands in *Staphylococcus aureus. Mol. Microbiol*. 29(2):527–543.
Llosa, M., Gomis-Ruth, F. X., Coll, M., and de la Cruz Fd, F. (2002). Bacterial conjugation: a two-step mechanism for DNA transport. *Mol. Microbiol*. 45(1):1–8.
Lyras, D. and Rood, J. I. (2000). Transposition of Tn4451 and Tn4453 involves a circular intermediate that forms a promoter for the large resolvase, TnpX. *Mol. Microbiol*. 38(3):588–601.
Mahillon, J. and Chandler, M. (1998). Insertion sequences. *Microbiol. Mol. Biol. Rev*. 62(3):725–774.
Miller, V., Kaper, J. B., Portnoy, D., and Isberg, R. (1994). *Molecular Genetics of Bacterial Pathogenesis: A Tribute to Stanley Falkow*. Washington, DC: ASM Press.
Mitsuhashi, S. (1977). *R Factor: Drug Resistance Plasmid*. Tokyo: University of Tokyo Press.
Moller-Jensen, J., Jensen, R. B. and Gerdes, K. (2000). Plasmid and chromosome segregation in prokaryotes. *Trends Microbiol*. 8(7):313–320.
Novick, R. P. (2003). Mobile genetic elements and bacterial toxinoses: the superantigen-encoding pathogenicity islands of *Staphylococcus aureus. Plasmid*. 49(2):93–105.
O'Shea, Y. A. and Boyd, E. F. (2002). Mobilization of the Vibrio pathogenicity island between *Vibrio cholerae* isolates mediated by CP-T1 generalized transduction. *FEMS Microbiol. Lett*. 214(2):153–157.
Osborn, A. M. and Boltner, D. (2002). When phage, plasmids, and transposons collide: genomic islands, and conjugative- and mobilizable-transposons as a mosaic continuum. *Plasmid*. 48(3):202–212.

Pogliano, J. (2002). Dynamic cellular location of bacterial plasmids. *Curr Opin Microbiol*. 5(6):586–590.

Qin, Z. and Cohen, S. N. (1998). Replication at the telomeres of the *Streptomyces* linear plasmid pSLA2. *Mol. Microbiol*. 28(5):893–903.

Rajanna, C., Wang, J., Zhang, D., Xu, Z., Ali, A., Hou, Y. M., and Karaolis, D. K. (2003). The *Vibrio* pathogenicity island of epidemic *Vibrio cholerae* forms precise extrachromosomal circular excision products. *J. Bacteriol*. 185(23):6893–6901.

Recchia, G. and Sherratt, D. (2002). Gene acquisition in bacteria by integron-mediated site-specific recombination. In Craig, N., Craigie, R., Gellert, M., and Lambowitz, A. (eds.) *Mobile DNA II*. Washington, DC: ASM Press.

Sauer, B. (2002). Chromosome manipulation by Cre-*lox* recombination. In Craig, N., Craigie, R., Gellert, M., and Lambowitz, A. (eds.) *Mobile DNA II*. Washington, DC: ASM Press.

Shoemaker, N. B., Wang, G. R., and Salyers, A. A. (1996). NBU1, a mobilizable site-specific integrated element from *Bacteroides* spp., can integrate nonspecifically in *Escherichia coli. J. Bacteriol*. 178(12):3601–3607.

Snyder, L. and Champness, W. (2003). *Molecular Genetics of Bacteria*. Washington, DC: ASM Press.

Toussaint, A. and Merlin, C. (2002). Mobile elements as a combination of functional modules. *Plasmid*. 47(1):26–35.

Van Duyne, G. (2002). A structural view of tyrosine recombinase site-specific recombination. In Craig, N., Craigie, R., Gellert, M., and Lambowitz, A. (eds.) *Mobile DNA II*. Washington, DC: ASM Press.

Whittle, G., Shoemaker, N. B., and Salyers, A. A. (2002). Characterization of genes involved in modulation of conjugal transfer of the *Bacteroides* conjugative transposon CTnDOT. *J. Bacteriol*. 184(14):3839–3847.

Zhu, J., Oger, P. M., Schrammeijer, B., Hooykaas, P. J., Farrand, S. K., and Winans, S. C. (2000). The bases of crown gall tumorigenesis. *J. Bacteriol*. 182(14):3885–3895.

Zinder, N. D. and Lederberg, J. (1952). Genetic exchange in *Salmonella. J. Bacteriol*. 64(5):679–699.

Zink, D. L., Feeley, J. C., Wells, J. G., Vanderzant, C., Vickery, J. C., Roof, W. D., and O'Donovan, G. A. (1980). Plasmid-mediated tissue invasiveness in *Yersinia enterocolitica. Nature*. 283(5743):224–226.