Microbial Genomics and Chromosome Organization

Prokaryotic Chromosome Organization

Generation of full genome sequences

Genomic Structure & Functional Genomics

Genome size vs. No. of orfs

Minimal genome concept

Lessons from full genomes

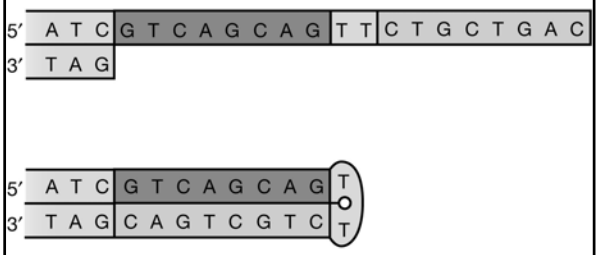
TABLE 7.1 Kinds of genetic elements

Element	Description
Prokaryote	
Chromosome	Extremely long, usually circular, double-stranded DNA molecule
Plasmid	Typically a relatively short, usually circular, double-stranded DNA molecule which is extrachromosomal
Viral genome	Single- or double-stranded DNA or RNA molecule
Transposable element	Double-stranded DNA molecule always found within another DNA molecule
Eukaryote	
Chromosome	Extremely long, linear, double-stranded DNA molecule
Plasmid[a]	Typically a relatively short circular or linear double-stranded DNA molecule which is extrachromosomal
Mitochondrion or chloroplast	Intermediate-length DNA molecules, usually circular
Viral genome	Single- or double-stranded DNA or RNA molecules
Transposable element	Double-stranded DNA molecule always found within another DNA molecule

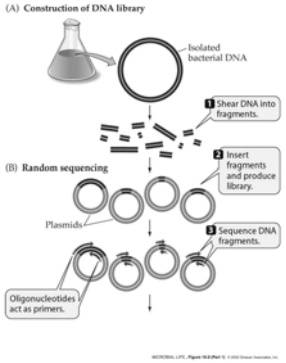
[a]Plasmids are very uncommon in eukaryotes.

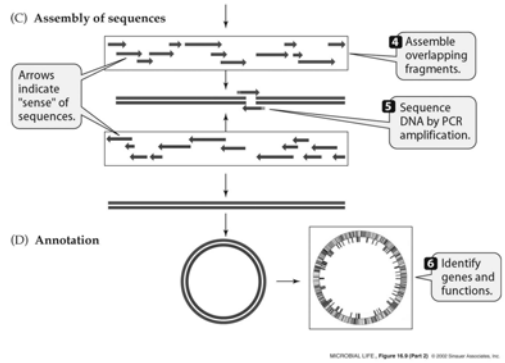
Chromosomal & nonchromosomal genetic elements

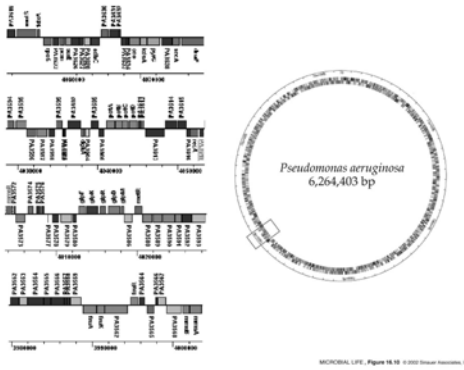
A hairpin structure on end of linear DNA molecule

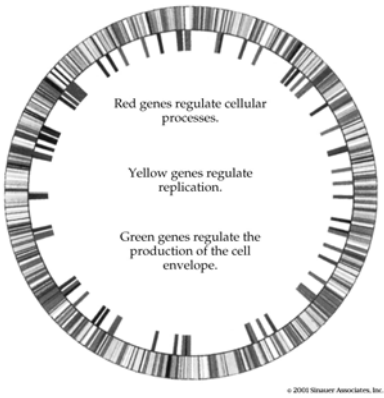


Whole-genome shotgun sequencing





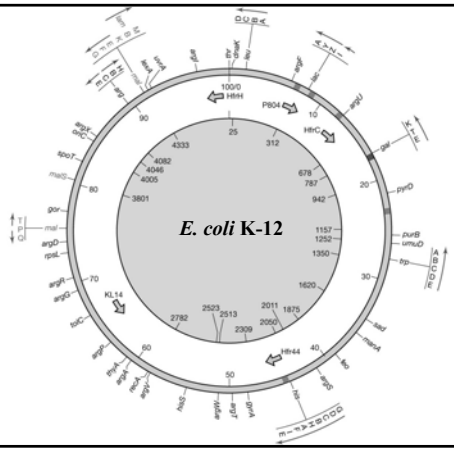




Functional Organization of *H. influenzae*: 1.8 Megabases

Table 16.2 Comparison of regulatory genes in selected bacterial genomes

Microorganism	# Genes in the Genome	# Regulatory Proteins	% of Total
Pseudomonas aeruginosa	5570	468	8.4
Escherichia coli	4289	250	5.8
Bacillus subtilis	4100	217	5.3
Mycobacterium tuberculosis	3918	117	3.0
Helicobacter pylori	1566	18	1.1



E. coli K-12 Genome Organization Lessons

Cured of F-plasmid and Lambda phage, 4.6 Mbs

88% of genome comprised in 4200 orfs
1% tRNAs and rRNAs
0.5% noncoding repetitive sequences
10% regulatory sequences, including promotors, operators, origin and terminus of DNA replication

70% single copy genes
Some gene clusters, e.g., lac, trp, his, and NOT arg
~6% polycistronic mRNAs

E. coli K-12 Genome Organization Lessons (cont.)

Gene orientation can be in both directions, however, highly expressed genes in the same direction as DNA replication forks travel.

For example, all 7 rrn operons
Rem: transcription always 5' to 3'

Also found were several different cryptic, defective prophages and IS elements.
18% of genome from horizontal gene transfer, including large regions (~40Kb) known as **islands of pathogenicity**.

TABLE 15.2 Gene function in bacterial genomes

Functional categories	Percentage of genes on chromosome[a] in that category: *Escherichia coli*	*Haemophilus influenzae*	*Mycoplasma genitalium*
Metabolism	21.0	19.0	14.6
Structural	5.5	4.7	3.6
Transport	10.0	7.0	7.3
Regulation	8.5	6.6	6.0
Translation	4.5	8.0	21.6
Transcription	1.3	1.5	2.6
Replication	2.7	4.9	6.8
Other, known	8.5	5.2	5.8
Unknown	38.1	43.0	32.0

[a]For the size of the chromosome of each of these species and the number of open reading frames that each contains, see Table 15.1

Genome size vs. ORFs

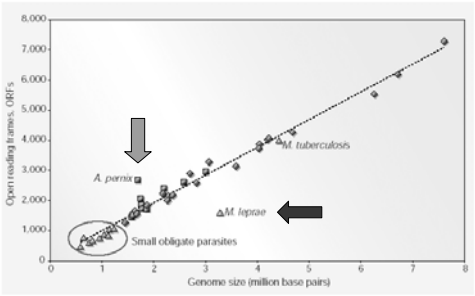


Figure 1 **Number of genes (ORFs) plotted against genome size for 44 fully sequenced genomes, including ten Archaea (squares) and 34 Bacteria. Obligate bacterial parasites are denoted by triangles; all other bacteria are shown as diamonds. *Mycobacterium leprae* is a genome 'in decay' that has a large number of pseudogenes. The archaeon *Aeropyrum pernix* is unusual in having an excessive number of duplicated ORFs.**

Organism (number of genes)	Glycolysis	Tricarboxylic-acid cycle	Amino-acid biosynthesis	Purine biosynthesis	Pyrimidine biosynthesis	Ancestral stock
Mycoplasma genitalium (470)	+	–	–	–	–	*Bacillus–Clostridium*
Buchnera species (588)	+	–	+	+	+	Gamma-proteobacteria
Rickettsia prowazekii (834)	–	+	–	–	–	Alpha-proteobacteria
Chlamydia trachomatis (894)	+	–	+	–	–	Main line
Treponema pallidum (1,041)	+	–	–	–	–	Main line
Mycobacterium leprae (1,604)	Partial	In decay	+	+	+	*Bacillus–Clostridium*

Figure 2 **Many routes to intracellular adaptation. The differing presence (+) or absence (–) of certain metabolic pathways in the streamlined genomes of parasitic bacteria shows how variable the process may be.**

Global Transposon Mutagenesis and a Minimal Mycoplasma Genome

Clyde A. Hutchison III,[1,2*] Scott N. Peterson,[1*†] Steven R. Gill,[1] Robin T. Cline,[1] Owen White,[1] Claire M. Fraser,[1] Hamilton O. Smith,[1‡] J. Craig Venter[1‡§]

Mycoplasma genitalium with 517 genes has the smallest gene complement of any independently replicating cell so far identified. Global transposon mutagenesis was used to identify nonessential genes in an effort to learn whether the naturally occurring gene complement is a true minimal genome under laboratory growth conditions. The positions of 2209 transposon insertions in the completely sequenced genomes of *M. genitalium* and its close relative *M. pneumoniae* were determined by sequencing across the junction of the transposon and the genomic DNA. These junctions defined 1354 distinct sites of insertion that were not lethal. The analysis suggests that 265 to 350 of the 480 protein-coding genes of *M. genitalium* are essential under laboratory growth conditions, including about 100 genes of unknown function.

265 to 350 genes are the minimum necessary genome

The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*

Gerard Deckert*†, Patrick V. Warren*‡, Terry Gaasterland‡, William G. Young*, Anna L. Lenox*, David E. Graham§, Ross Overbeek‖, Marjory A. Snead*, Martin Keller*, Monette Aujay*, Robert Huber¶, Robert A. Feldman*, Jay M. Short*, Gary J. Olsen§ & Ronald V. Swanson*

*Diversa Corporation, 10665 Sorrento Valley Road, San Diego, California 92121, USA
‡Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Illinois 60439, USA
§Department of Microbiology, University of Illinois, Urbana, Illinois 61801, USA
‖Lehrstuhl für Mikrobiologie, Universität Regensburg W-8400, Regensburg W-8400, Germany

Aquifex aeolicus **was one of the earliest diverging, and is one of the most thermophilic, bacteria known. It can grow on hydrogen, oxygen, carbon dioxide, and mineral salts. The complex metabolic machinery needed for *A. aeolicus* to function as a chemolithoautotroph (an organism which uses an inorganic carbon source for biosynthesis and an inorganic chemical energy source) is encoded within a genome that is only one-third the size of the *E. coli* genome. Metabolic flexibility seems to be reduced as a result of the limited genome size. The use of oxygen (albeit at very low concentrations) as an electron acceptor is allowed by the presence of a complex respiratory apparatus. Although this organism grows at 95 °C, the extreme thermal limit of the Bacteria, only a few specific indications of thermophily are apparent from the genome. Here we describe the complete genome sequence of 1,551,335 base pairs of this evolutionarily and physiologically interesting organism.**

Metabolic pathways and transport systems for *Thermotoga maritima*

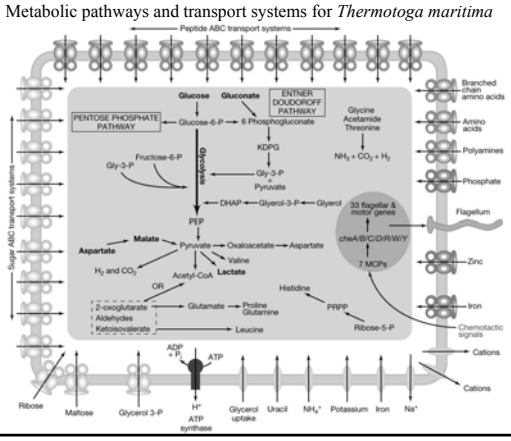


Table 16.3 **Distribution of genes of unknown function among selected bacterial genomes** *(Part 1)*

Organism	Genome Size (Mbp)	No. of ORFs (% coding)		Unknown Function		Unique ORFs	
Aeropyrum pernix K1	1.67	1,885	(89%)				
A. aeolicus VF5	1.50	1,749	(93%)	663	(44%)	407	(27%)
A. fulgidus	2.18	2,437	(92%)	1,315	(54%)	641	(26%)
B. subtilis	4.20	4,779	(87%)	1,722	(42%)	1,053	(26%)
B. burgdorferi	1.44	1,738	(88%)	1,132	(65%)	682	(39%)
Chlamydia pneumoniae AR39	1.23	1,134	(90%)	543	(48%)	262	(23%)
Chlamydia trachomatis MoP_c	1.07	936	(91%)	353	(38%)	77	(8%)
C. trachomatis serovar D	1.04	928	(92%)	290	(32%)	255	(29%)
Deinococcus radiodurans	3.28	3,187	(91%)	1,715	(54%)	1,001	(31%)
E. coli K-12-MG1655	4.60	5,295	(88%)	1,632	(38%)	1,114	(26%)
H. influenzae	1.83	1,738	(88%)	595	(35%)	237	(14%)
H. pylori 26695	1.66	1,589	(91%)	744	(45%)	539	(33%)
Methanobacterium thermoautotrophicum	1.75	2,008	(90%)	1,010	(54%)	496	(27%)

MICROBIAL LIFE, Table 16.3 (Part 1) © 2002 Sinauer Associates, Inc.

Table 16.3 Distribution of genes of unknown function among selected bacterial genomes *(Part 2)*

Organism	Genome Size (Mbp)	No. of ORFs (% coding)		Unknown Function		Unique ORFs	
Methanococcus jannaschii	1.66	1,783	(87%)	1,076	(62%)	525	(30%)
M. tuberculosis CSU#93	4.41	4,275	(92%)	1,521	(39%)	606	(15%)
M. genitalium	0.58	483	(91%)	173	(37%)	7	(2%)
M. pneumoniae	0.81	680	(89%)	248	(37%)	67	(10%)
N. meningitidis MC58	2.24	2,155	(83%)	856	(40%)	517	(24%)
Pyrococcus horikoshii OT3	1.74	1,994	(91%)	589	(42%)	453	(22%)
Rickettsia prowazekii Madrid E	1.11	878	(75%)	311	(37%)	209	(25%)
Synechocystis sp.	3.57	4,003	(87%)	2,384	(75%)	1,426	(45%)
T. maritma MSB8	1.86	1,879	(95%)	863	(46%)	373	(26%)
T. pallidum	1.14	1,039	(93%)	461	(44%)	280	(27%)
Vibrio cholerae El Tor N1696	4.03	3,890	(88%)	1,806	(46%)	934	(24%)
Totals:	50.60	52,462	(89%)	22.358	(43%)	12,161	(23%)

From Fraser et al., *Nature* 2000, vol. 406. p. 800.

MICROBIAL LIFE, Table 16.3 (Part 2) © 2002 Sinauer Associates, Inc.

Lessons from full genomes

Size range 600Kb to 12Mb

Vast number of putative genes with no known function.

Pathogenicity conferred by "Pathogenicity Islands"
44.5Kb in *Bacillus anthracis*
Symbiotic Island of >600Kb in *Sinorhizobium loti* including genes for nodulation and N-fixation

Adaptive gene losses in parasitic bacteria
Rickettsia and *Chlamidia* are ATP thieves using the same "alien" ADP/ATP translocase

Microbial Genetic Exchange and Plasmids

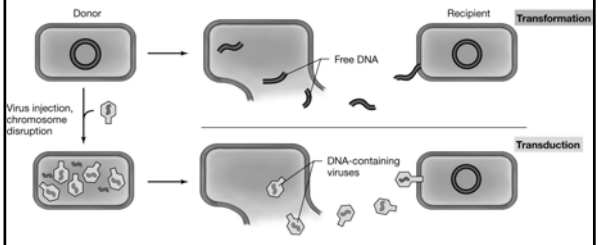
Microbial Genetic Exchange is unidirectional!

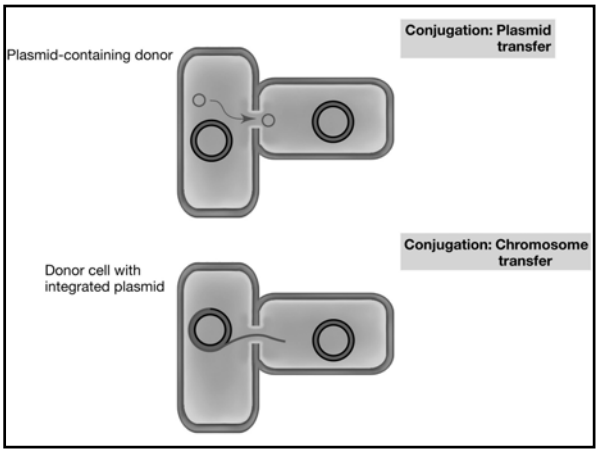
- Transformation
- Transduction
- Conjugation

Each requires Homologous Recombination

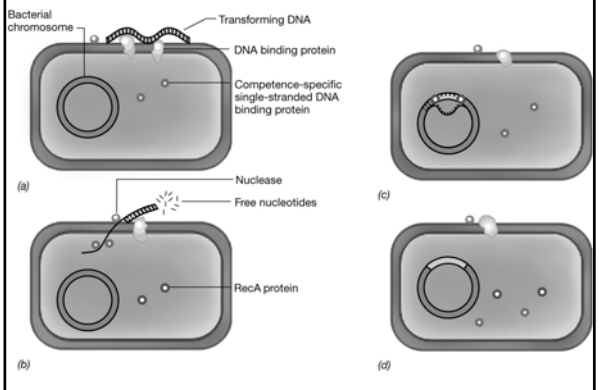
Types of plasmids

Microbial Genetic Exchange

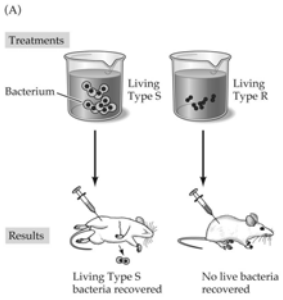




Transformation by a Gram -positive, competent cell

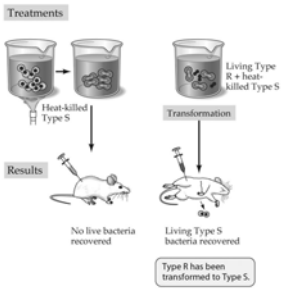


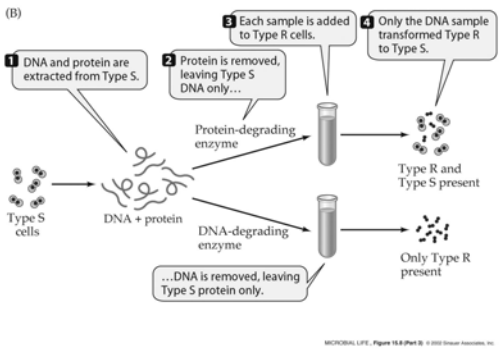
Demonstration of transformation



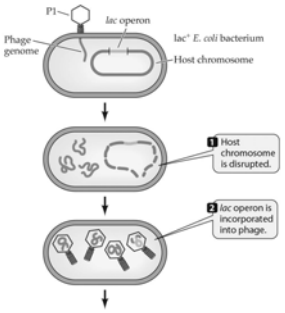
Streptococcus pneumoniae

Demonstration of transformation

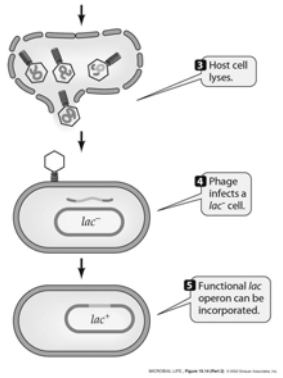




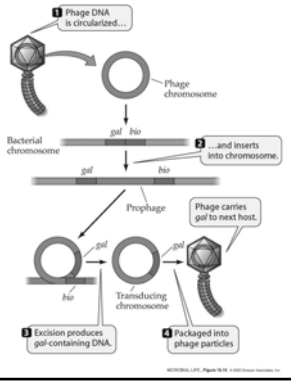
Generalized transduction

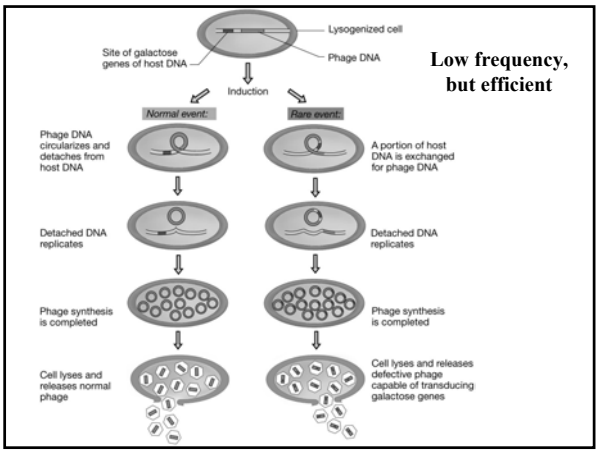


Generalized transduction (cont.)



Specialized transduction





Low frequency, but efficient

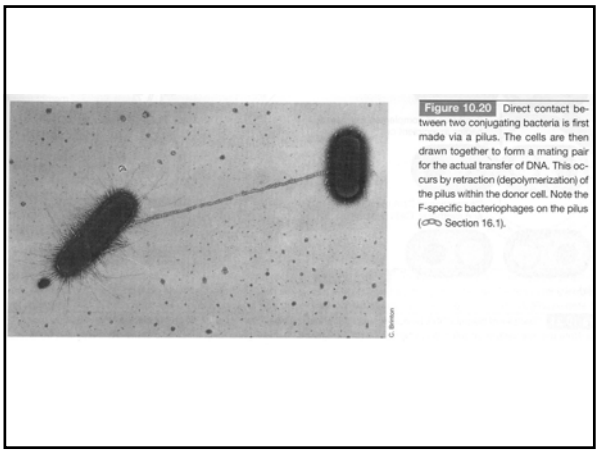
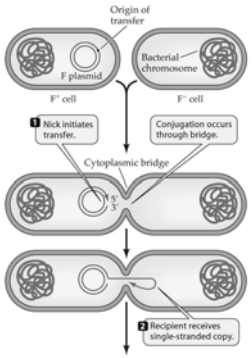
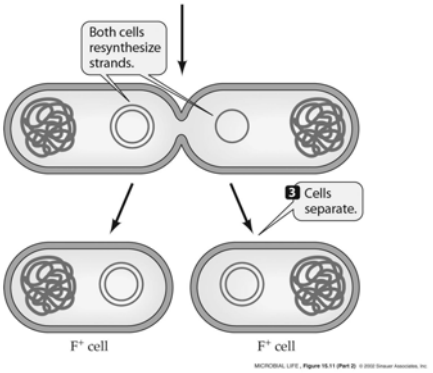


Figure 10.20 Direct contact between two conjugating bacteria is first made via a pilus. The cells are then drawn together to form a mating pair for the actual transfer of DNA. This occurs by retraction (depolymerization) of the pilus within the donor cell. Note the F-specific bacteriophages on the pilus (Section 16.1).

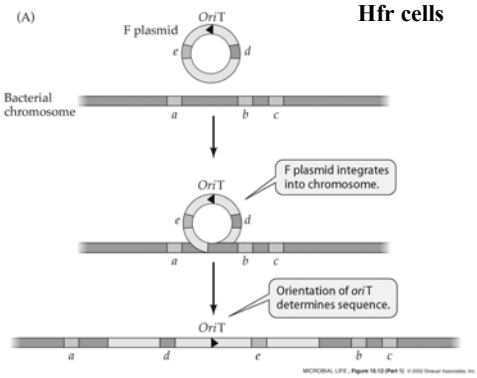
C. Brinton

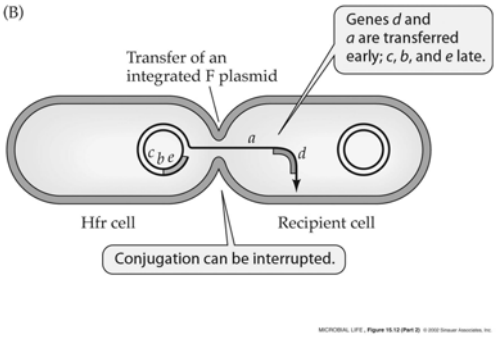
Bacterial conjugation



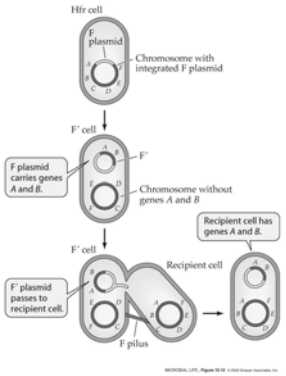


Hfr cells

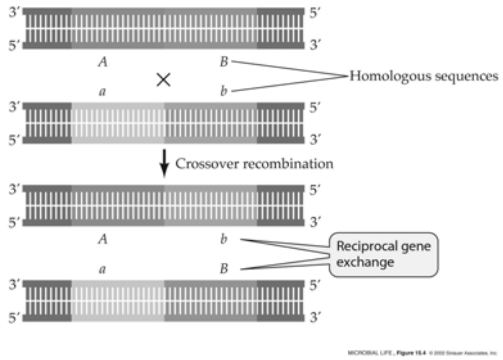


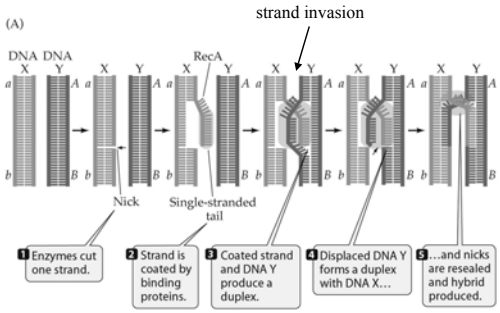


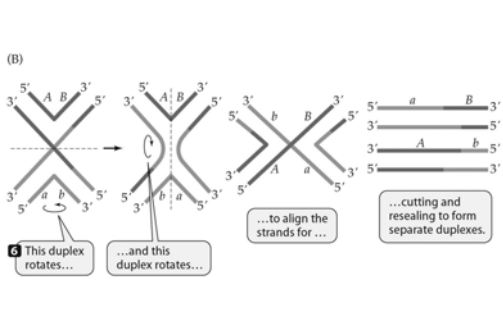
F' cells



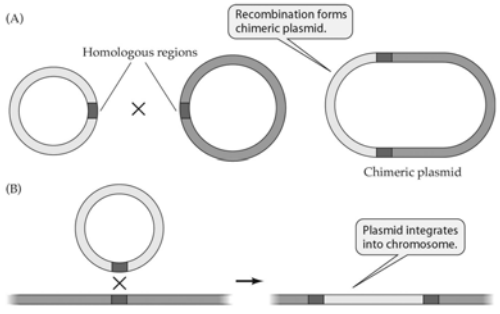
Homologous Recombination



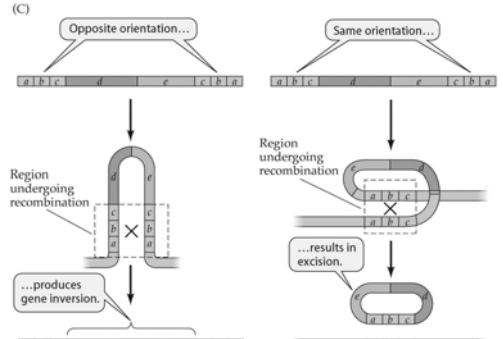




Types of homologous recombination in bacteria



Types of homologous recombination in bacteria



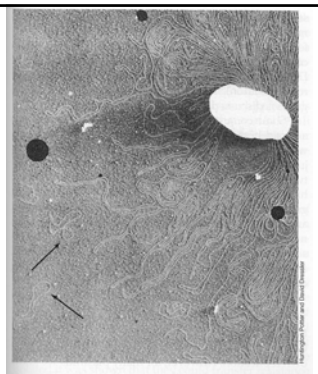
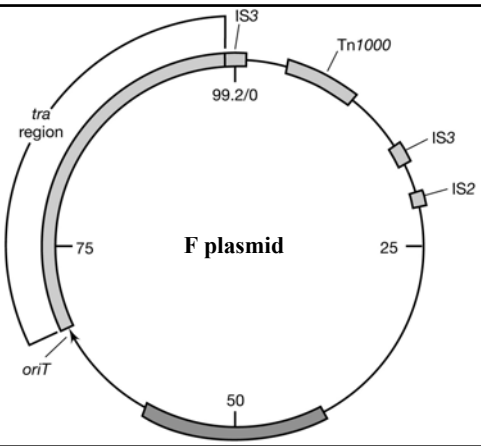
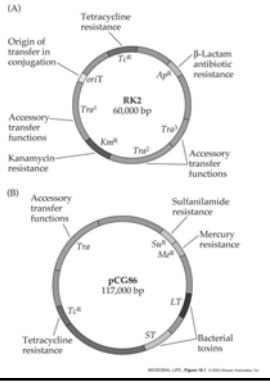


Figure 10.16 The bacterial chromosome and bacterial plasmids, as shown in the electron microscope. The plasmids (arrows) are the circular structures, much smaller than the main chromosomal DNA. The cell (large, white structure) was broken gently so the DNA would remain intact.



Plasmids of pathogenic bacteria



Ti plasmid

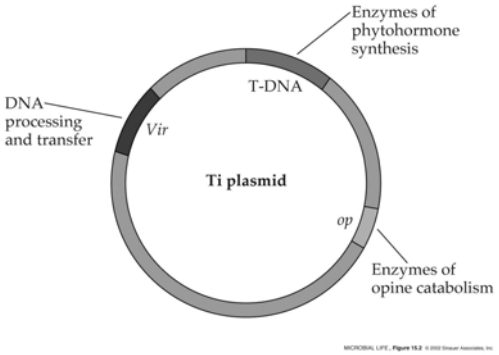


Table 10.3 Some phenotypes conferred by plasmids in prokaryotes

Phenotype class[a]	Organisms[b]
Antibiotic production	*Streptomyces*
Conjugation	*Escherichia, Pseudomonas, Rhizobium, Staphylococcus, Streptococcus, Sulfolobus, Vibrio*
Physiological functions	
Degradation of octane, camphor, naphthalene	*Pseudomonas*
Degradation of herbicides	*Alcaligenes*
Formation of acetone and butanol (☞ Section 30.11)	*Clostridium*
Lactose, sucrose or urea utilization and nitrogen fixation	Enteric bacteria
Nodulation and symbiotic nitrogen fixation (☞ Section 19.22)	*Rhizobium*
Pigment production	*Erwinia, Staphylococcus*
Resistance	
Antibiotic resistance (☞ Section 20.12)	*Campylobacter*, Enteric bacteria, *Neisseria, Staphylococcus*
Resistance to cadmium, cobalt, mercury, nickel, and/or zinc (☞ Section 19.16)	*Acidocella, Alcaligenes, Listeria, Pseudomonas, Staphylococcus*
Bacteriocin resistance (and production)	*Bacillus*, Enteric bacteria, *Lactococcus, Propionibacterium*
Virulence	
Host cell invasion	*Salmonella, Shigella, Yersinia*
Coagulase, hemolysin, enterotoxin (☞ Sections 21.9 and 21.11)	*Staphylococcus*
Enterotoxin, K antigen (☞ Sections 12.11 and 21.11)	*Escherichia*
Tumorigenicity in plants (☞ Section 19.21)	*Agrobacterium*

[a] Only a few of the many phenotypes known to be associated with plasmids are given.
[b] Only a few well-characterized examples are given. All of the organisms given in the list are *Bacteria* except for *Sulfolobus*, which is a member of the *Archaea*.