九十九學年度 微生物學教材

細菌遺傳學 、細菌基因體學

(11-16-2010, 11-17-2010)

Chapter 14

Bacterial Genetics: **14.4**
Creating genetic variability

11-16-2010

http://life.nctu.edu.tw/~hlpeng/

彭慧玲 分機:56916
(hlpeng@mail.nctu.edu.tw)

課程綱要與評分標準

11/16, 11/17	Bacterial genetics and microbial genomics
11/23, 11/24	Microbial taxonomy, Archaea
11/30	Fungi
12/7	**Exam I- 20%**
12/8	Bacteria (I)
12/14, 12/15	Bacteria (II)
12/21, 12/22	Viruses (I)
12/28, 12/29	Viruses (II), viroids, virusoids, and prions
1/4, 1/5	Microbial diseases and their control
1/11	**Exam II- 20%**

出席率 10.5% (21 h)

NDM-1 (*New Delhi Metallo*-beta-lactamase)

- In December 2009, a carbapenem-resistant *K. pneumoniae* infection acquired in New Delhi, India
  - Resistant to multiple antibiotics
  - Also reported in *E. coli* and *Acinetobacter baumannii*

The Lancet infectious disease 2010

NDM-1 Kp01

超級細菌

Horizontal (Lateral) Gene Transfer (HGT) in bacteria

- Haploid geneome
- Exogenote (Donor)
- Endogenote (Recipient)
- Merozygote
  - Transconjugants
  - Transformants
  - Transductants

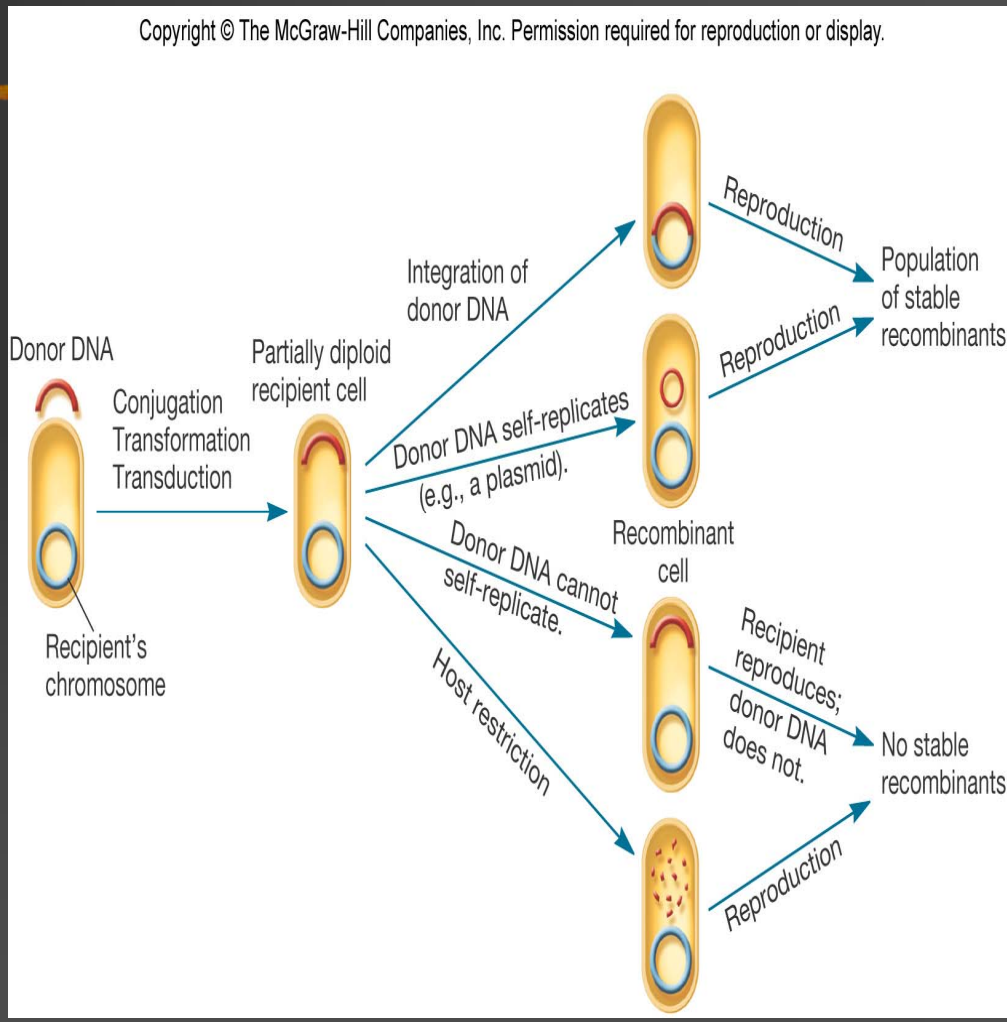


Figure 14.11

Recombination at the molecular level

- Homologous recombination (Reciprocal recombination)
  - Double-Strand Break Model
    - results from <u>DNA strand breakage</u> and <u>reunion</u>, leading to crossing-over

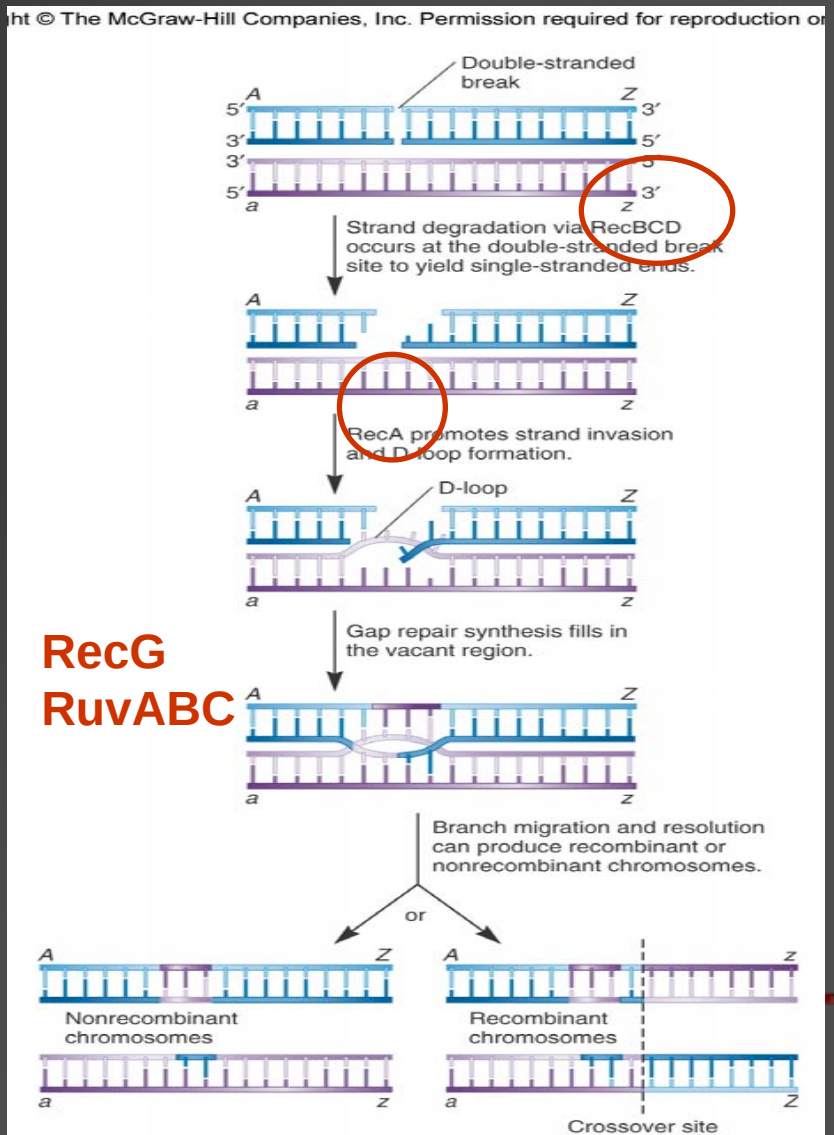


Figure 14.12

Nonreciprocal Homologous Recombination

- incorporation of single strand of DNA into chromosome, forming a stretch of heteroduplex DNA
- proposed to occur during bacterial transformation

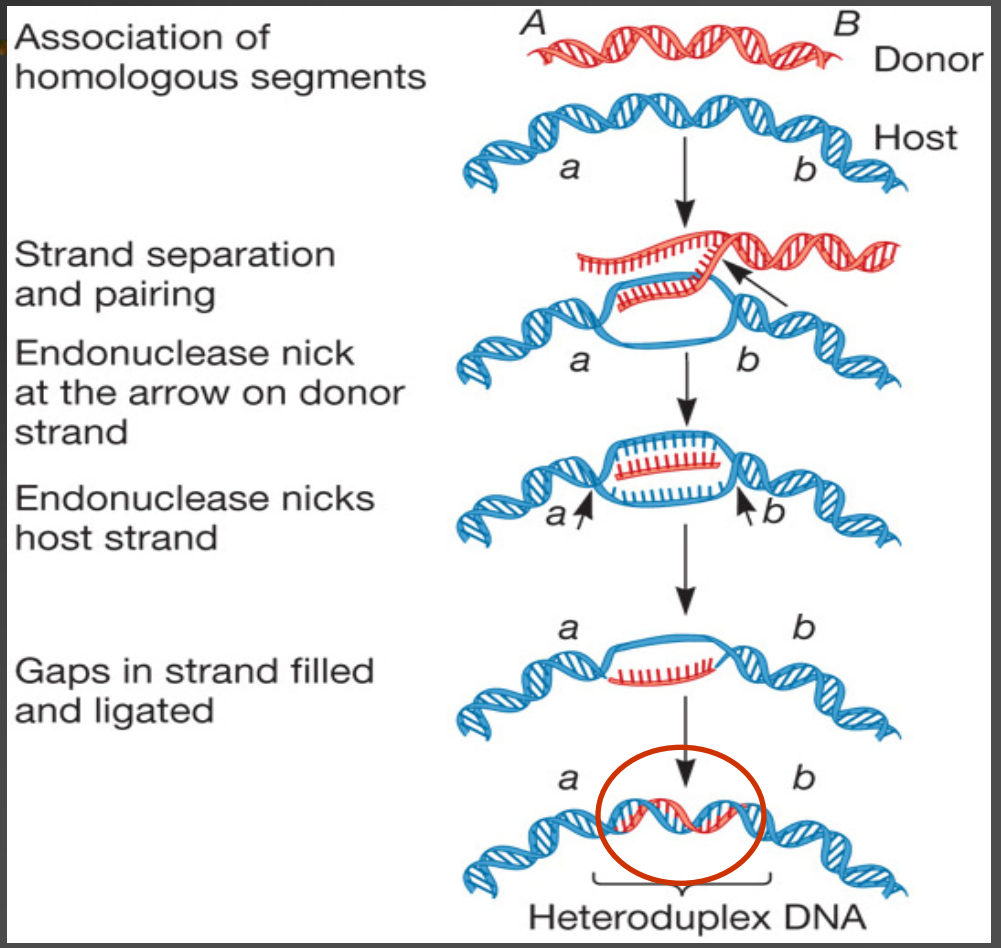


Figure 14.13

Site-specific recombination

- insertion of viral genome into host chromosomes
  - insertion of nonhomologous DNA
    - Only a small region of homology is required
- Transposition carried out by transposable elements
  - Simple transposition
  - Replicative transposition

Transposable Elements

- transposition
  - the movement of pieces of DNA around the genome
- transposable elements (transposons)
  - segments of DNA that carry genes for transposition
- widespread in bacteria, eukaryotes and archaea

Types of transposable elements

- IS element
  - 2 IR + transposase gene
- Transposons
  - 2 IS elements
  - additional genes
- Replicative transposons
  - Resolvase gene

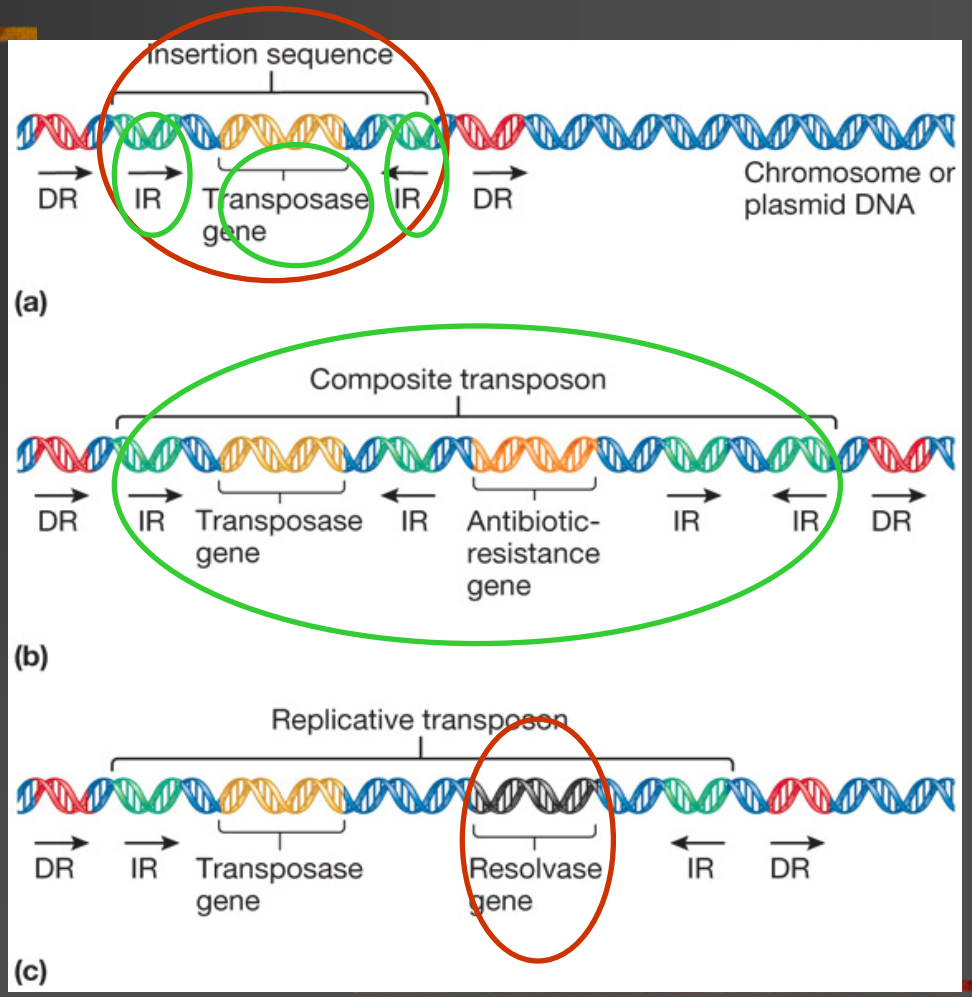


Figure 14.14

Simple Transposition- cut and paste

- Generation of **direct repeats** in host DNA flanking a transposon

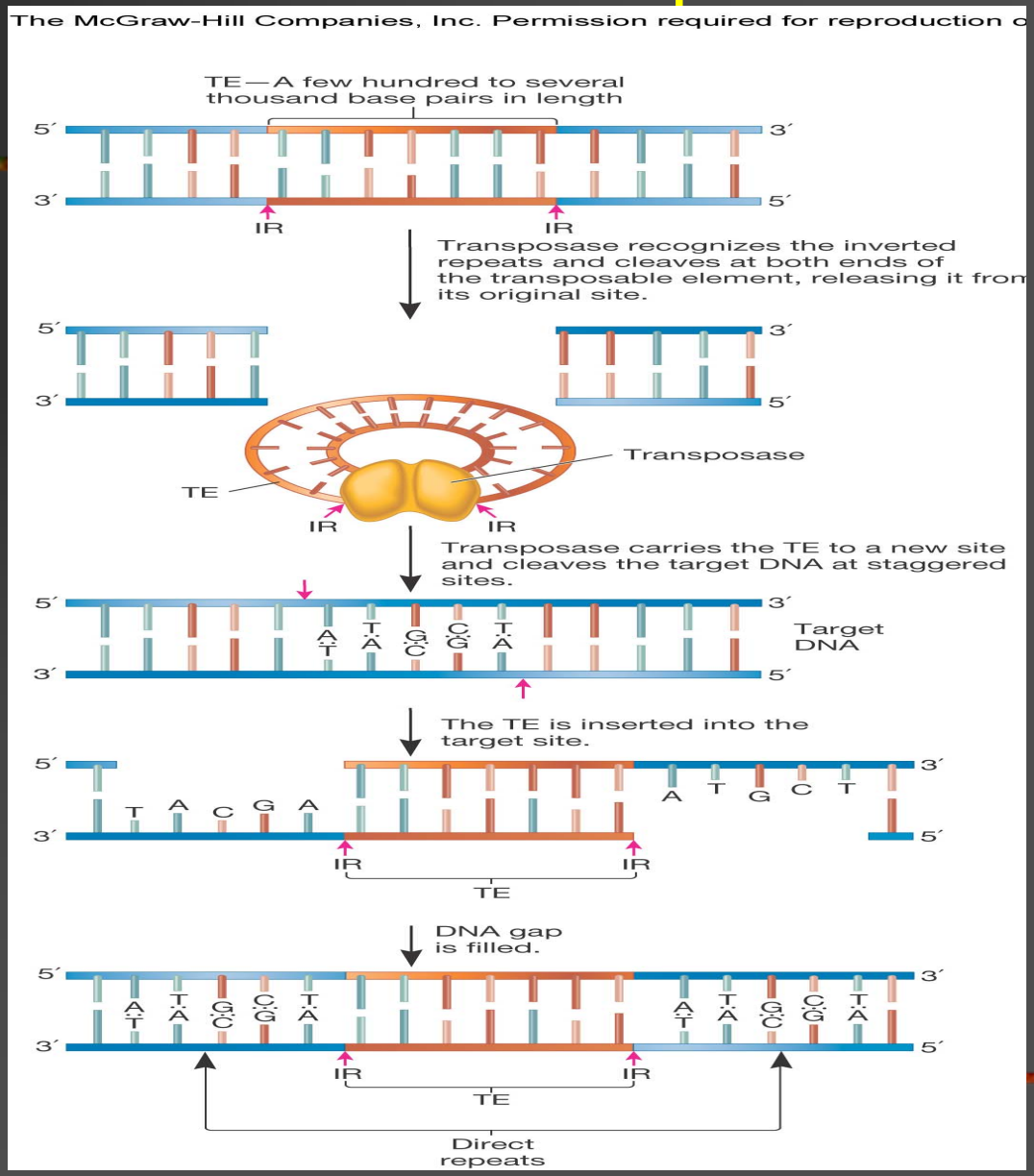


Figure 13.20

Replicative Transposition

- insertion generates direct repeats of flanking host DNA
- usually transposon replicated, remaining in original site, while duplicate inserts at another site
- Tn*3*

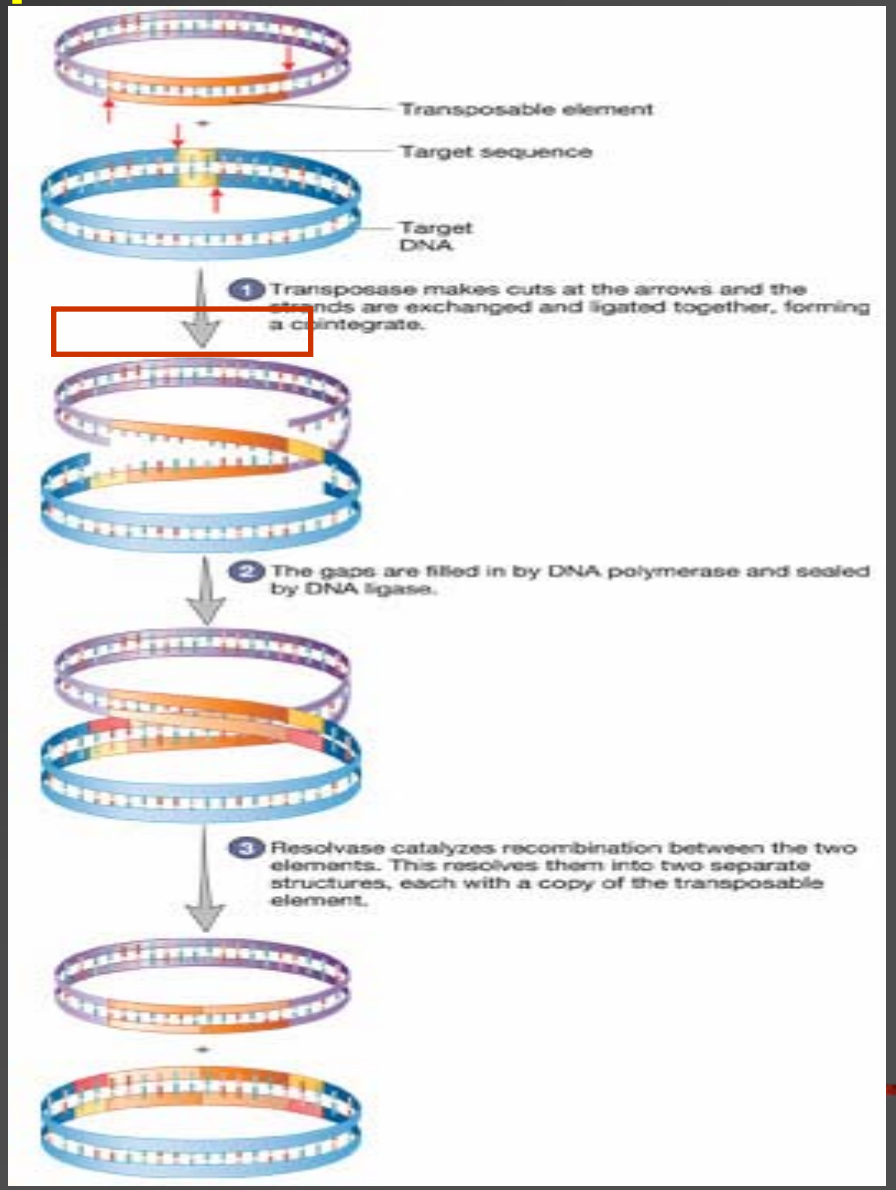


Figure 14.16

Effects of transposition

- mutation in coding region
- arrest of translation or transcription
  - interruption of promoter regions…
  - interruption of ribosomal binding sites…
- activation of genes
  - provide strong promoter sequences
- generation of new plasmids
  - Multi-drug resistance plasmids

R Plasmid results from Tn*3* Transposition

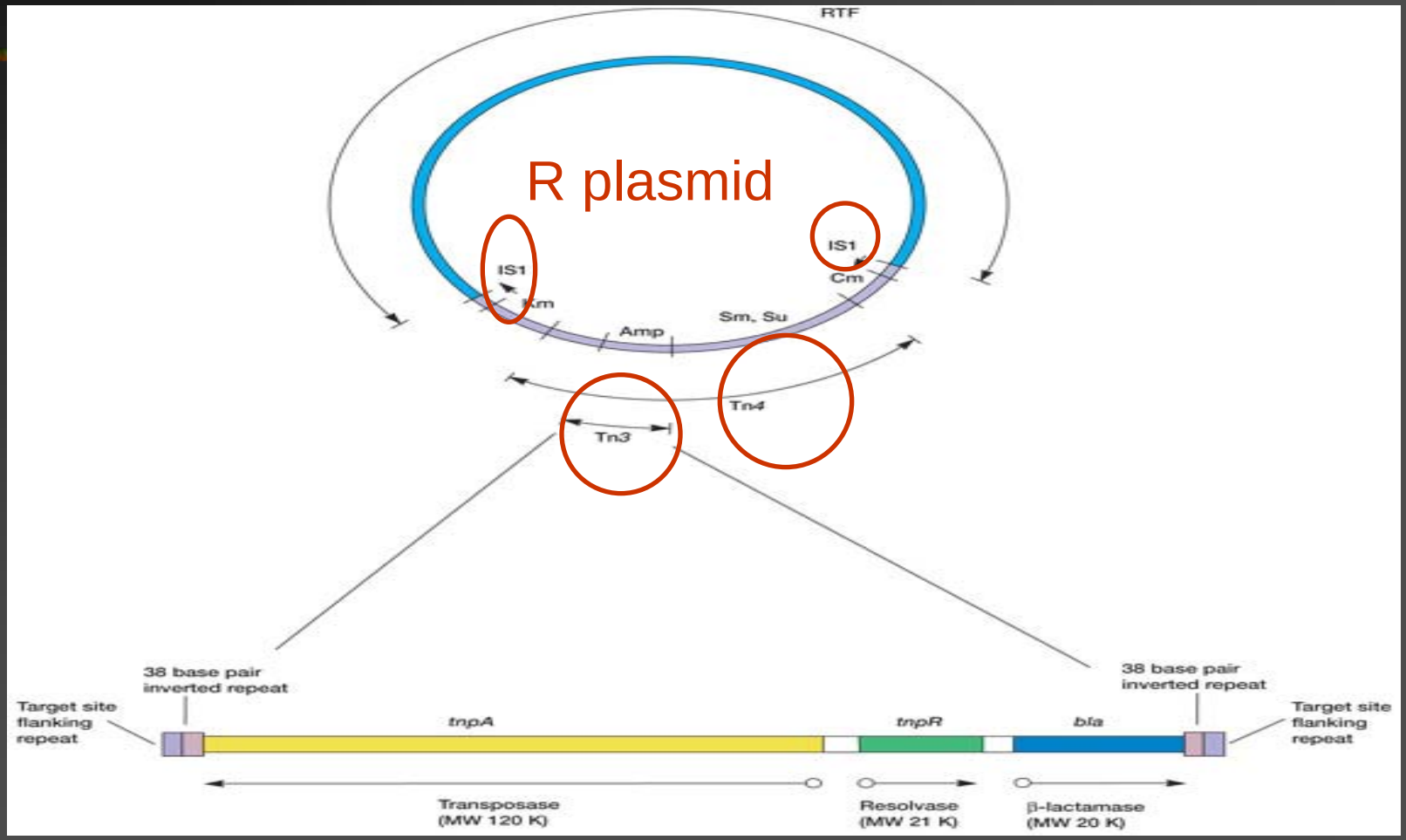


Figure 14.17

Plasmids

- small, autonomously replicating DNA molecules that can exist independently or, as episomes, integrate reversibly into the host chromosome
  - conjugative plasmids such as the F plasmid can transfer copies of themselves to other bacteria during conjugation
  - R plasmid, virulence plasmid, metabolic plasmid…

Bacterial conjugation

- The transfer of genes between bacteria that depends on direct cell to cell contact
  - mediated by the sex pilus encoded by F plasmid

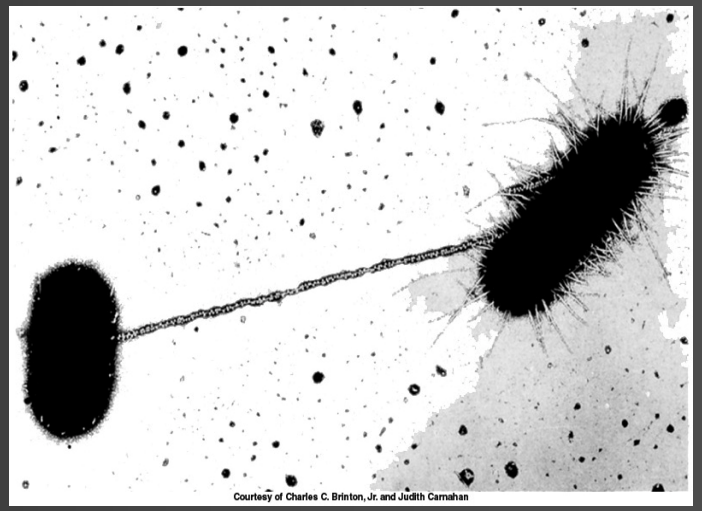


Fig. 14.19

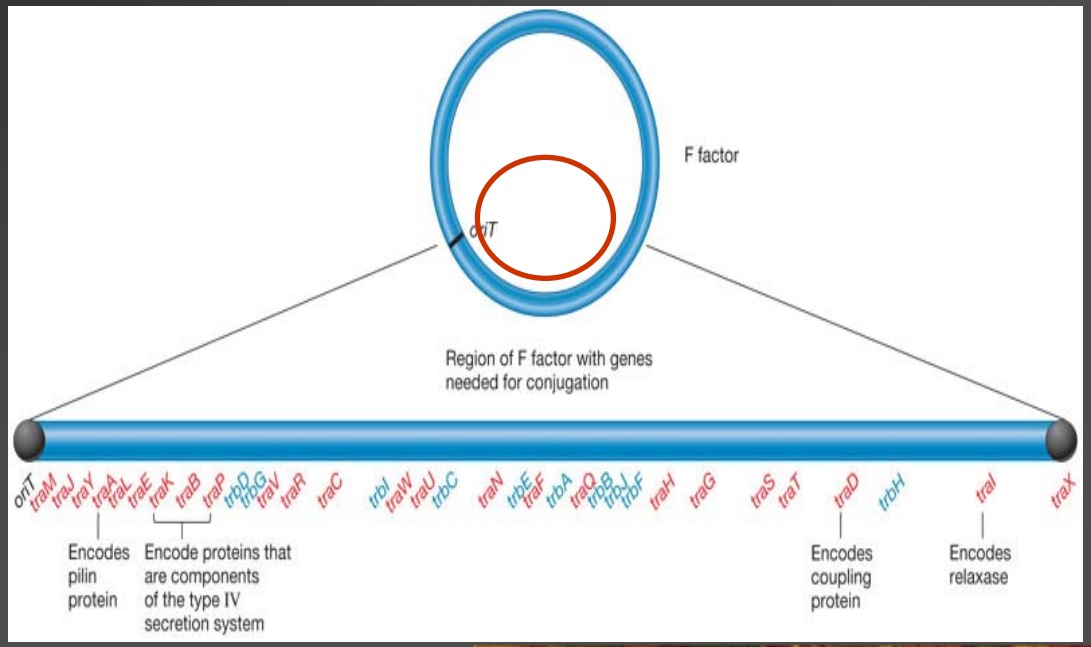


Fig. 14.18

F plasmid integration

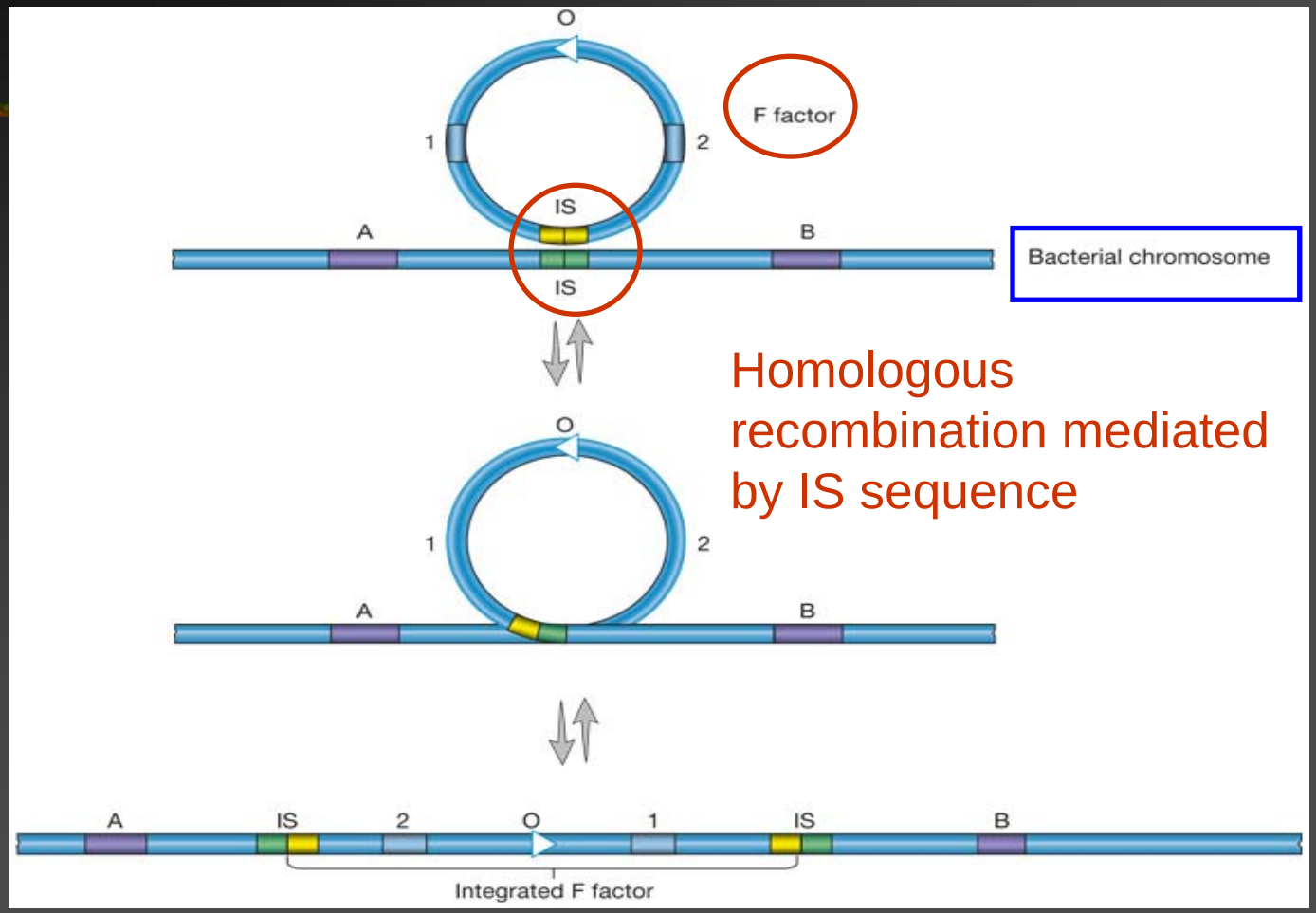
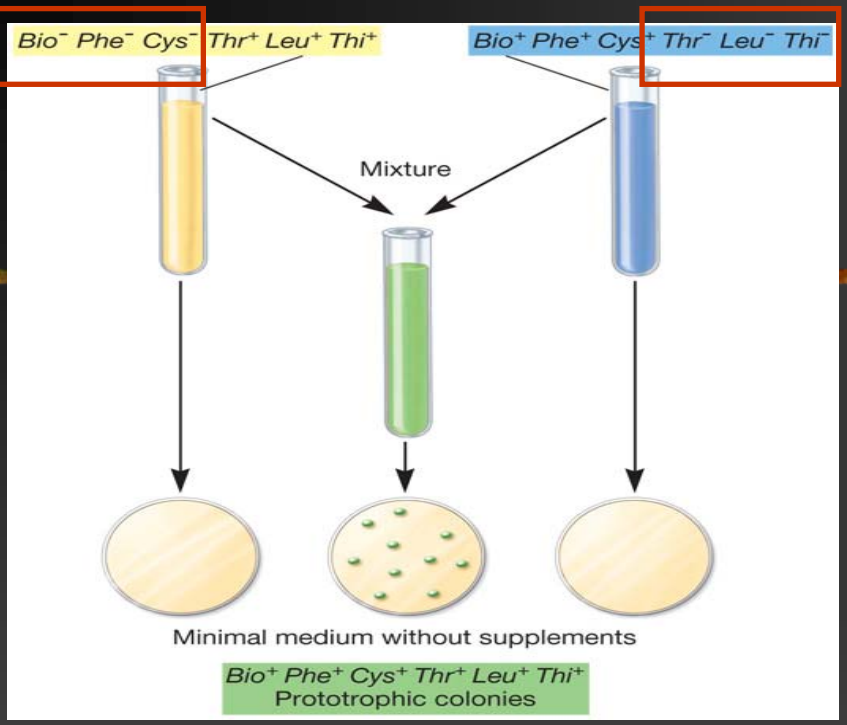


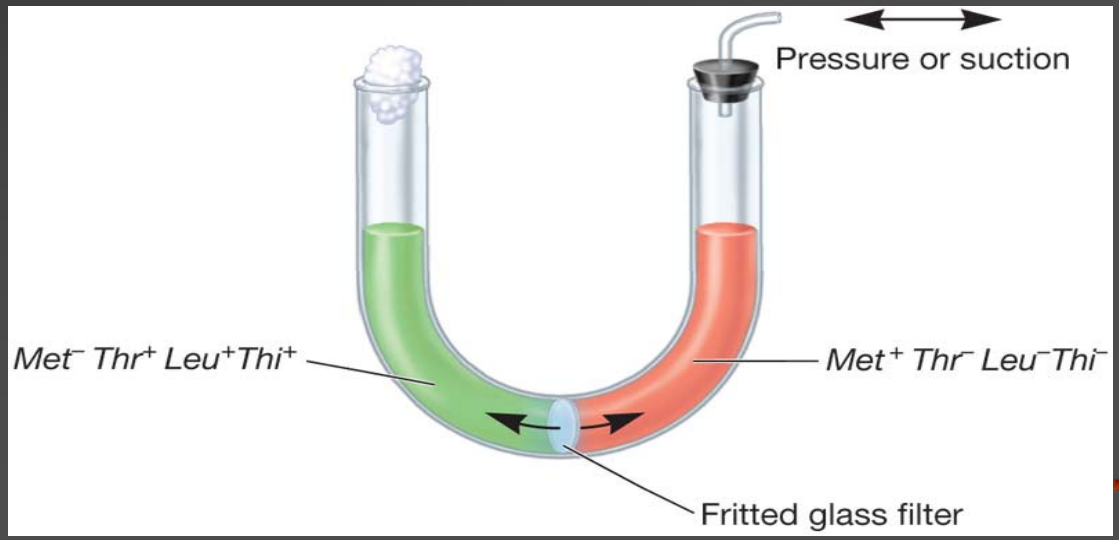
Figure 14.20



- transfer of DNA between cells discovered by Lederberg and Tatum (1946)

The U-tube experiment

- cell-to-cell contact is necessary for the DNA transfer, which demonstrated by Bernard Davis (1950)



F-mediated conjugation

- $F^+ \rightarrow F^-$
  - A polar gene transfer
  - replicated by rolling-circle and the duplicate is transferred
  - recipients usually become F^+
  - donor remains F^+

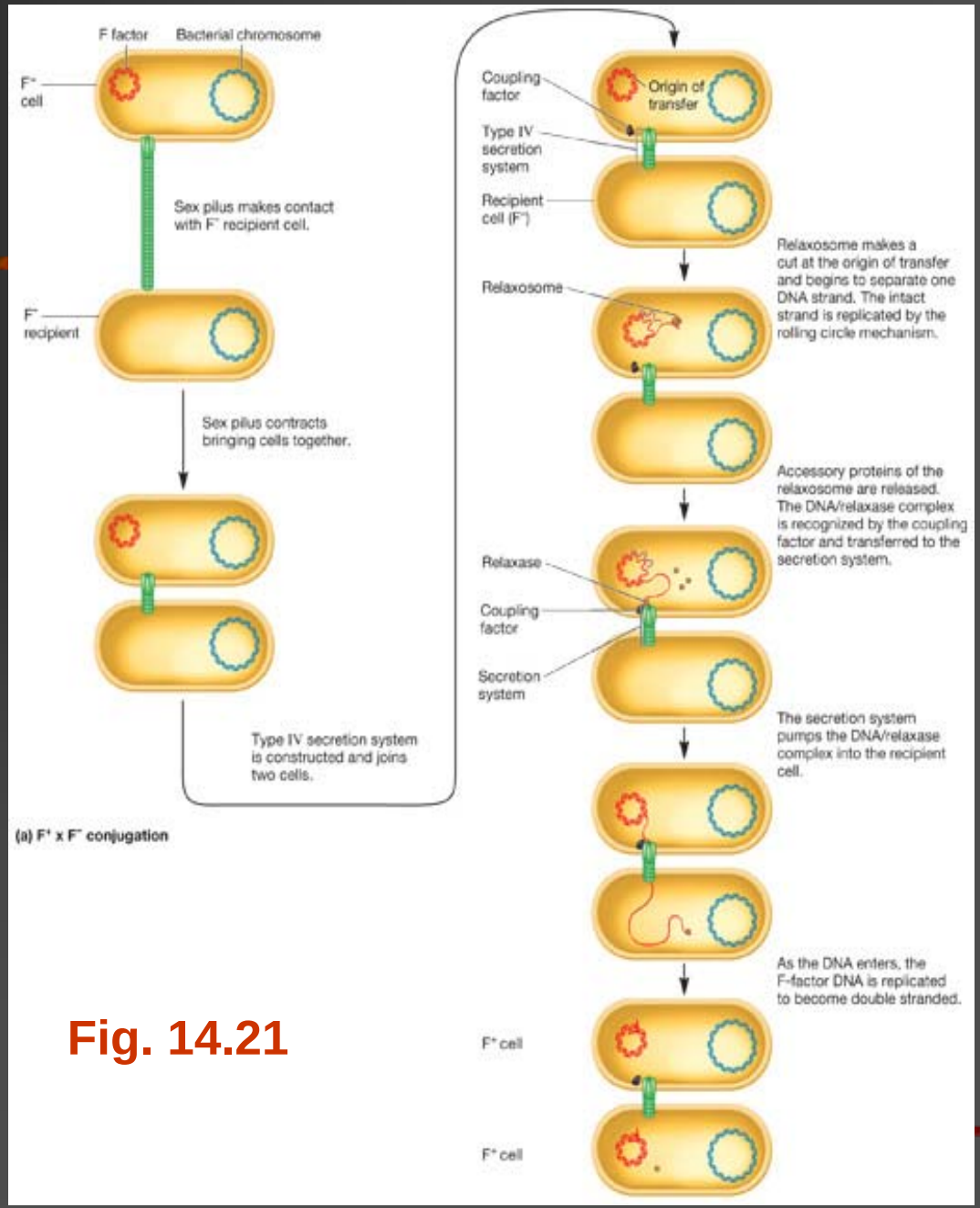


Fig. 14.21

The type IV secretion system

- Tra proteins
  - TraA pilin

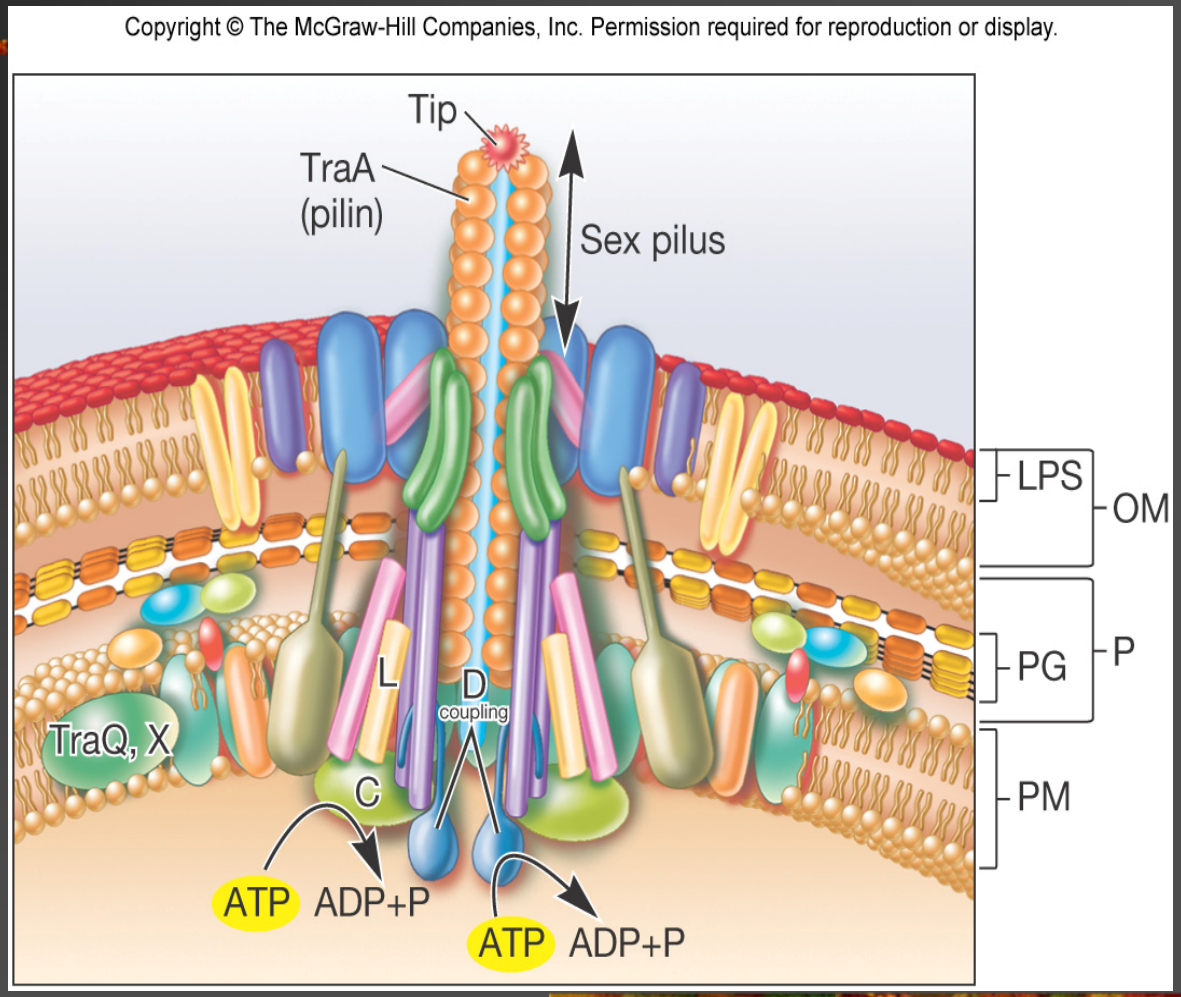
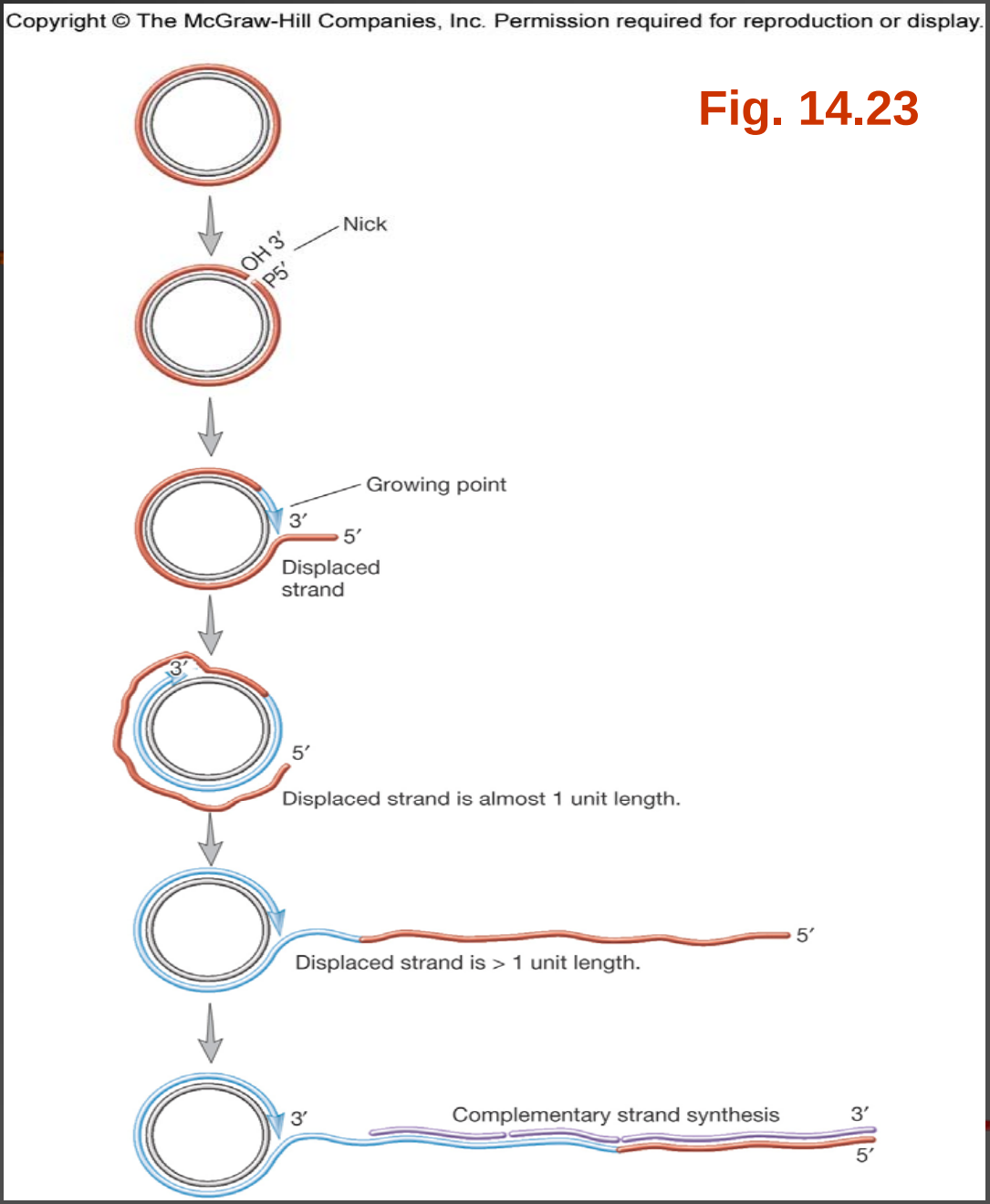


Fig. 14.22

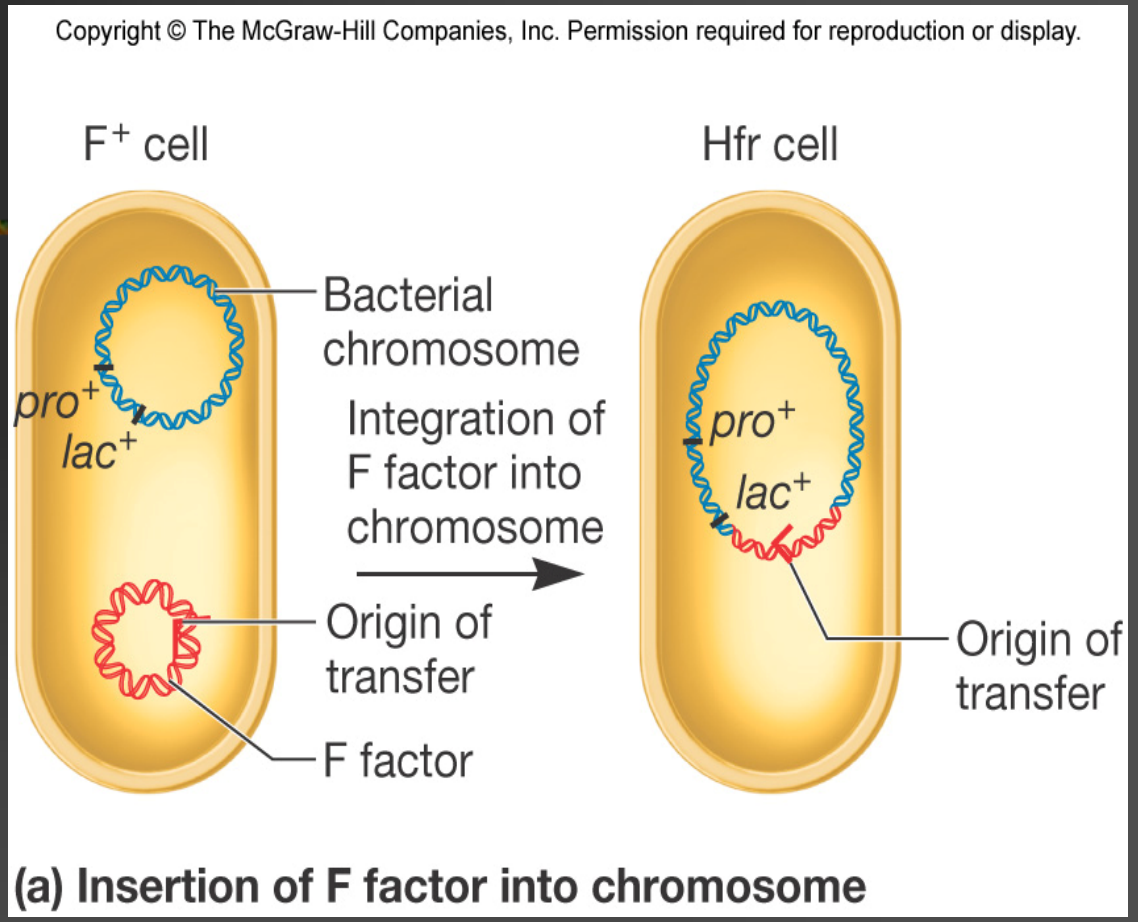
Rolling-circle replication

- A 1S tail is generated→ 2S by synthesis of a complimentary strand



Hfr cells

- Integration of F to chromosome to generate Hfr cell
- *h*igh *f*requency of *r*ecombinants



HFr → F⁻ conjugation

- a complete copy of the F factor is usually not transferred

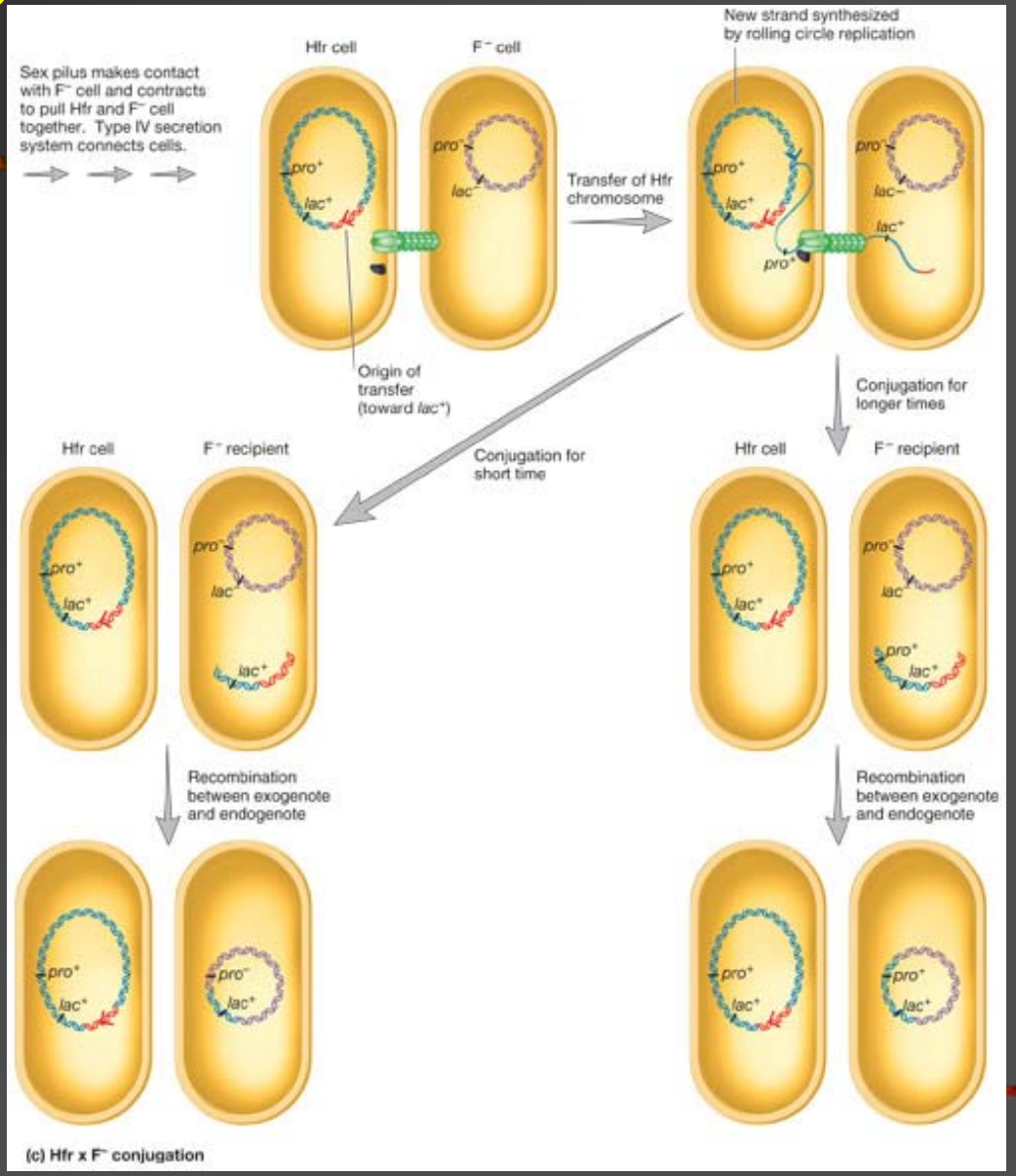


Fig. 14.24 (b)

F’ cells

- An error excision caused the *A* gene of Hfr cell is picked up by the F factor (red)

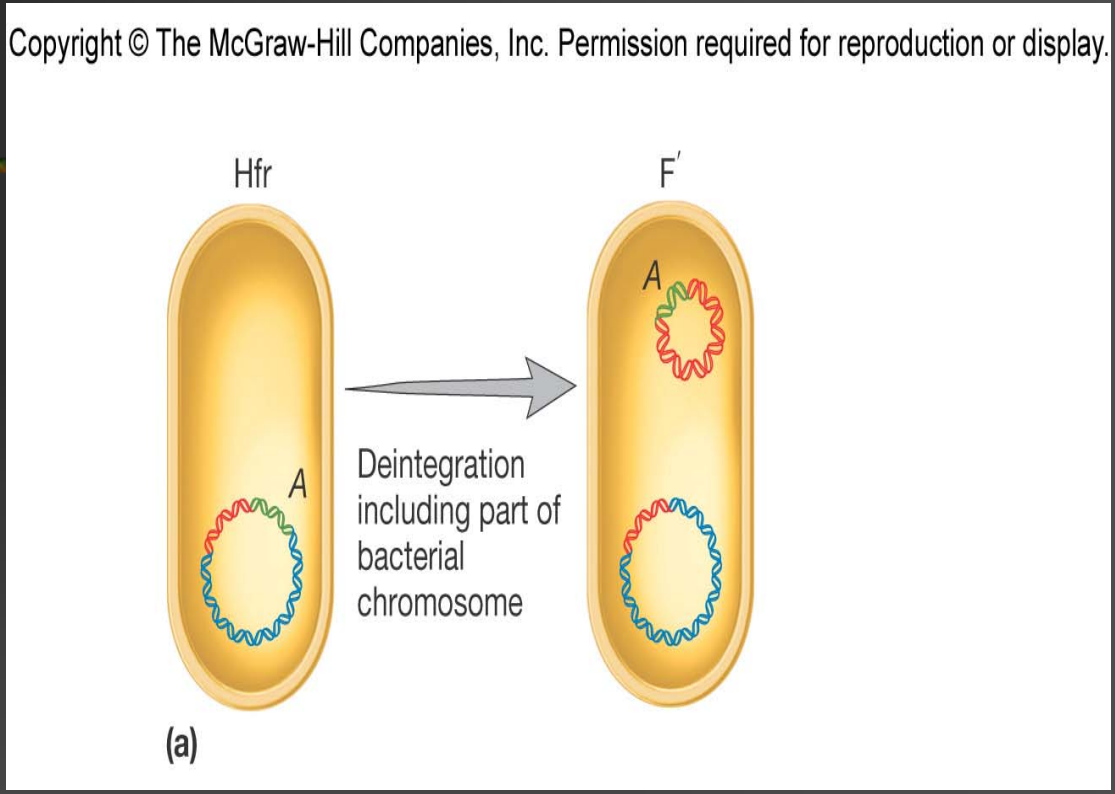


Fig. 14.25(a)

F’ conjugation

F’ x F- Mating

- formed by incorrect excision from chromosome
- some of the F factor is left behind in the host chromosome
- some host genes have been removed along with some of the F factor
  - these genes can be transferred to a second host cell by conjugation

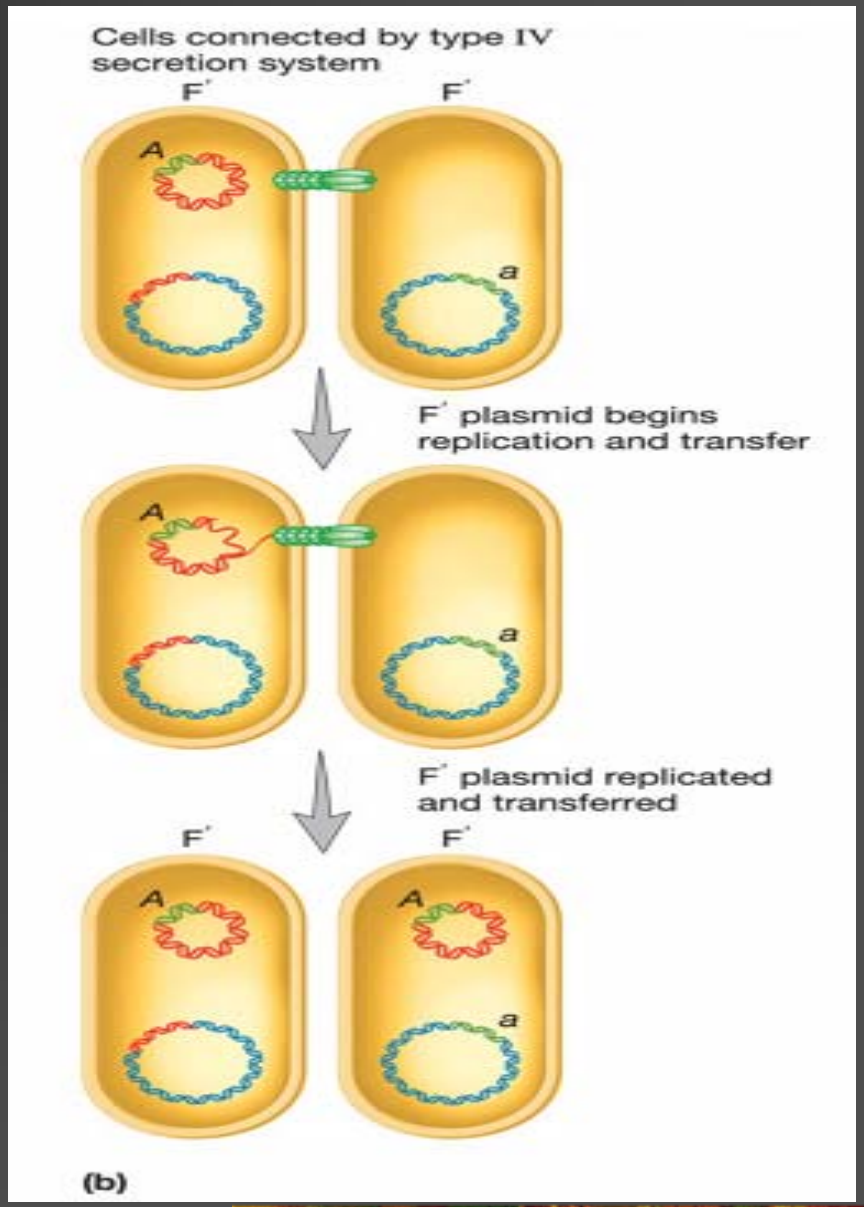


Fig 14.25 (b)

DNA transformation

- uptake of DNA fragments and incorporation into recipient
  - competent cell
- uptake of a plasmid

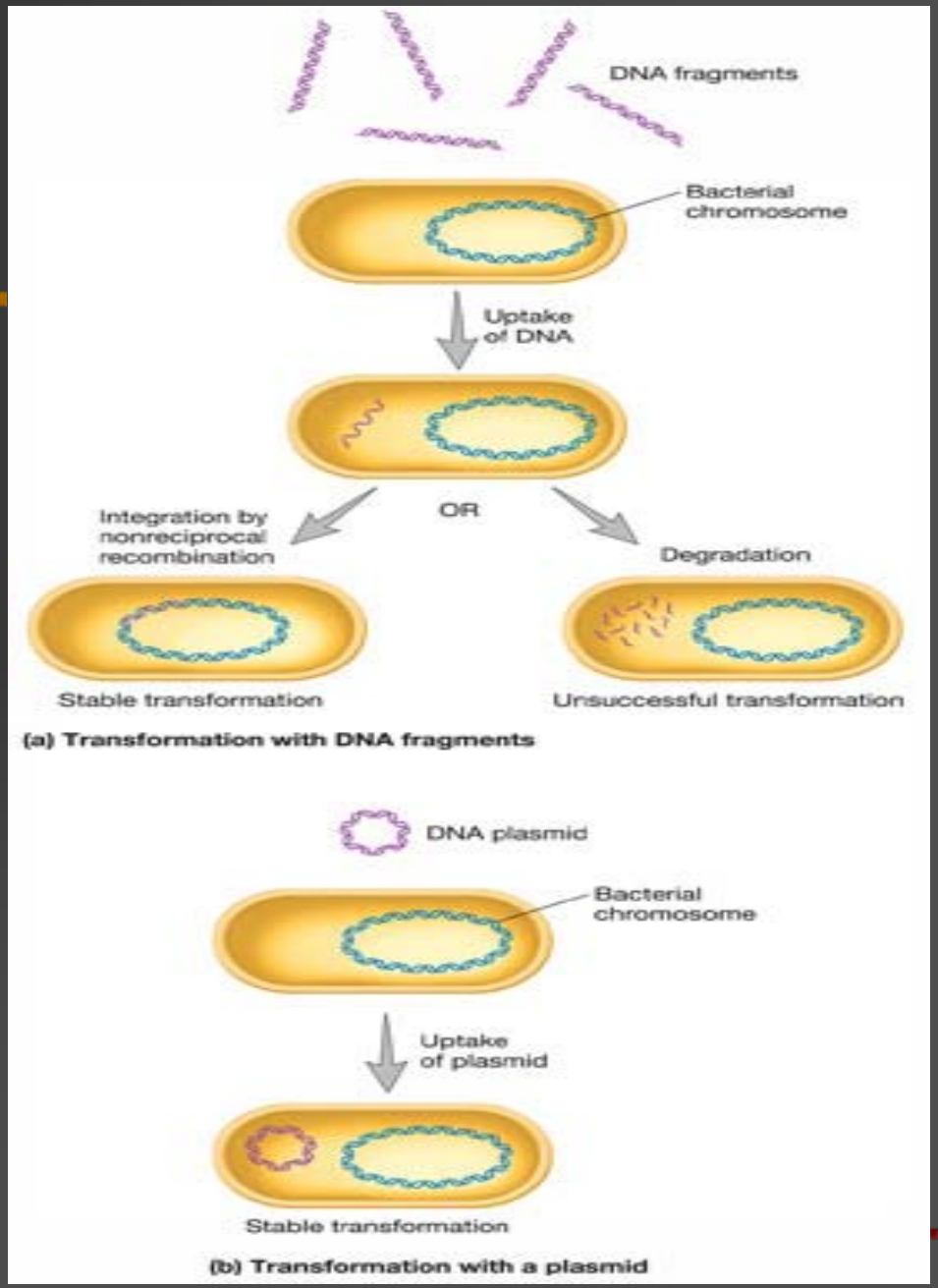


Figure 14.25

Transformation in *Streptococcus pneumoniae*

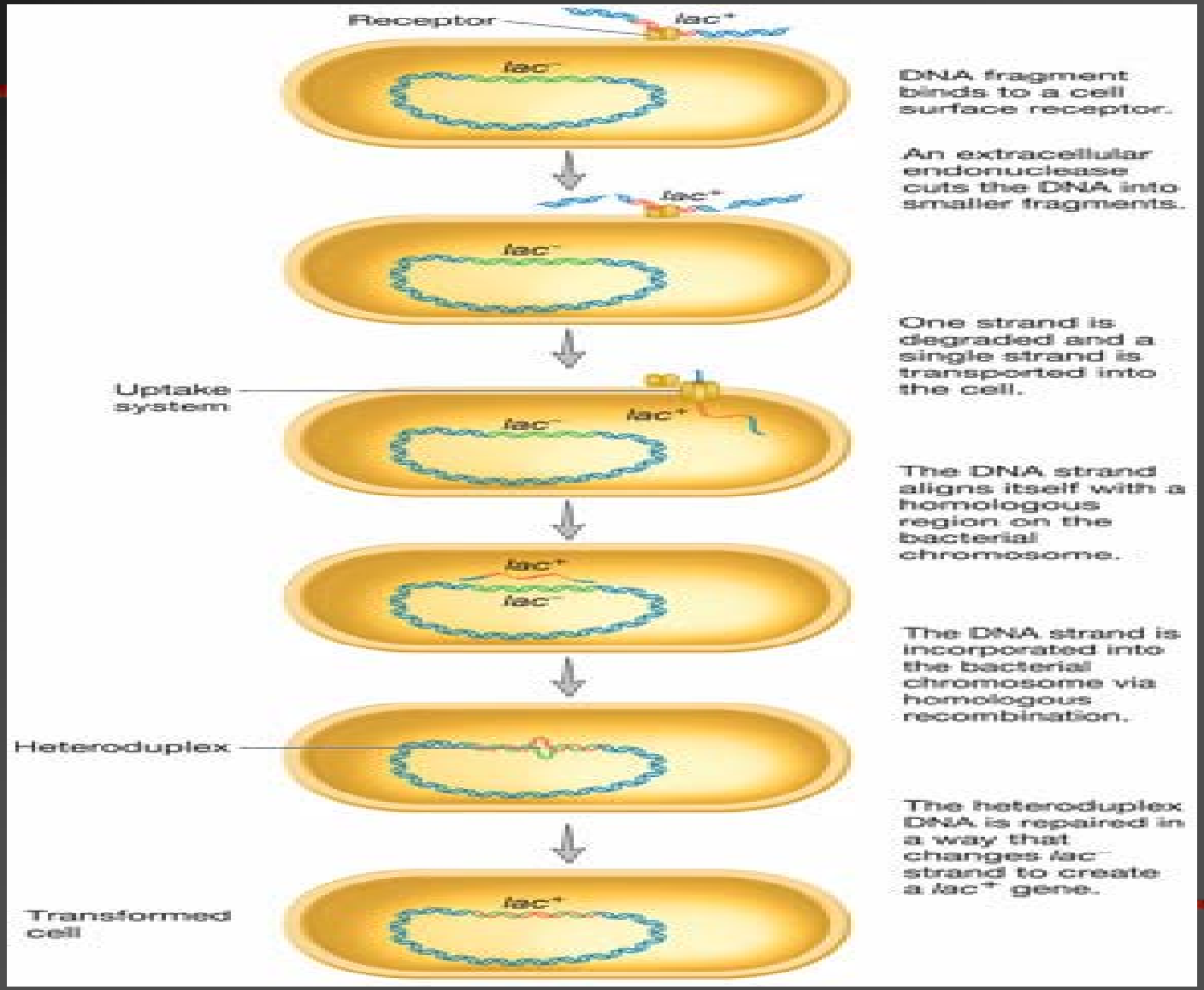


Fig 14.27

DNA Uptake Systems

Natural competence

G(-) *Neisseria sp.*

G(+) *Bacillus sp.*

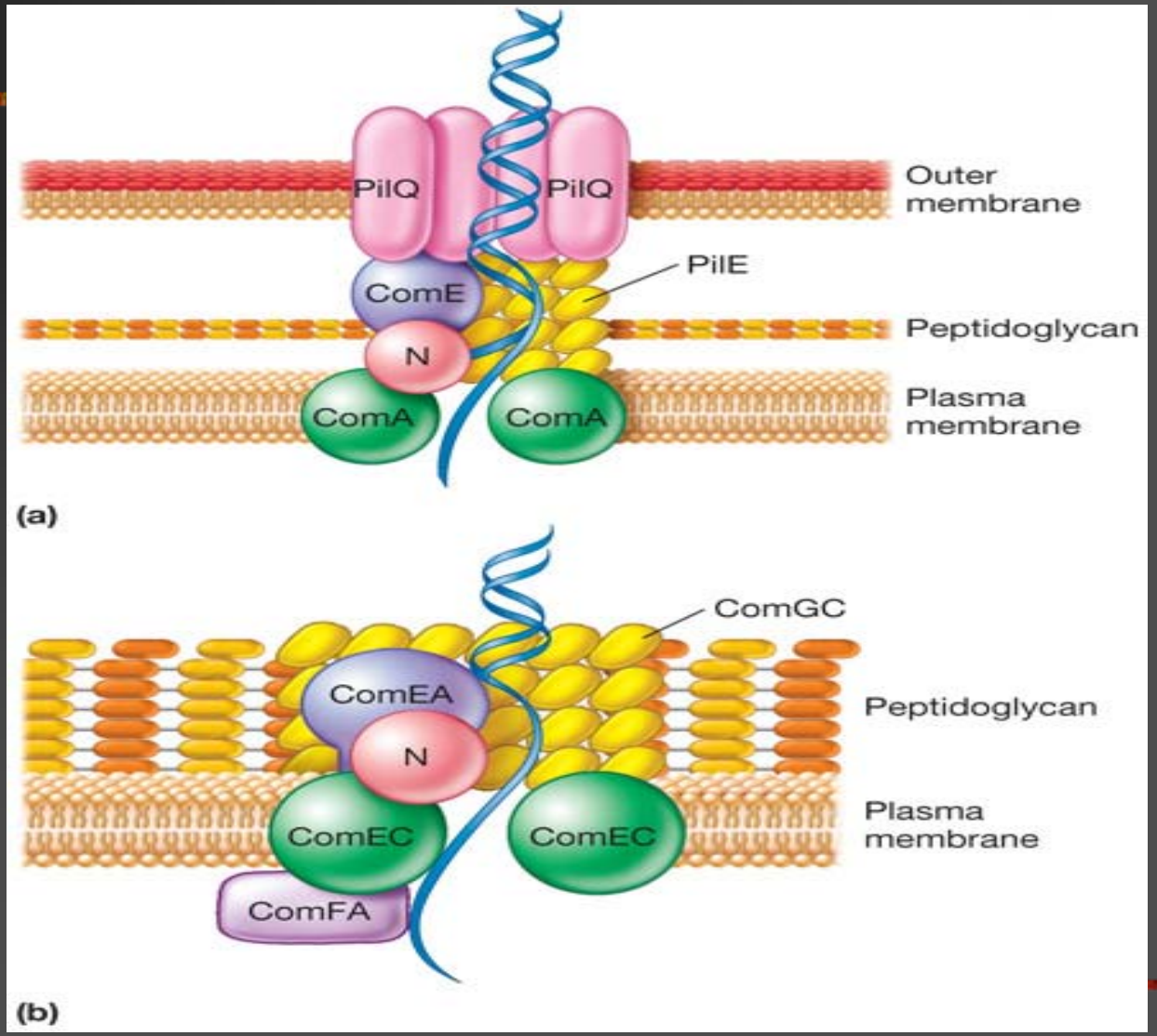


Fig 14.28

Transduction- transfer of genes by phages

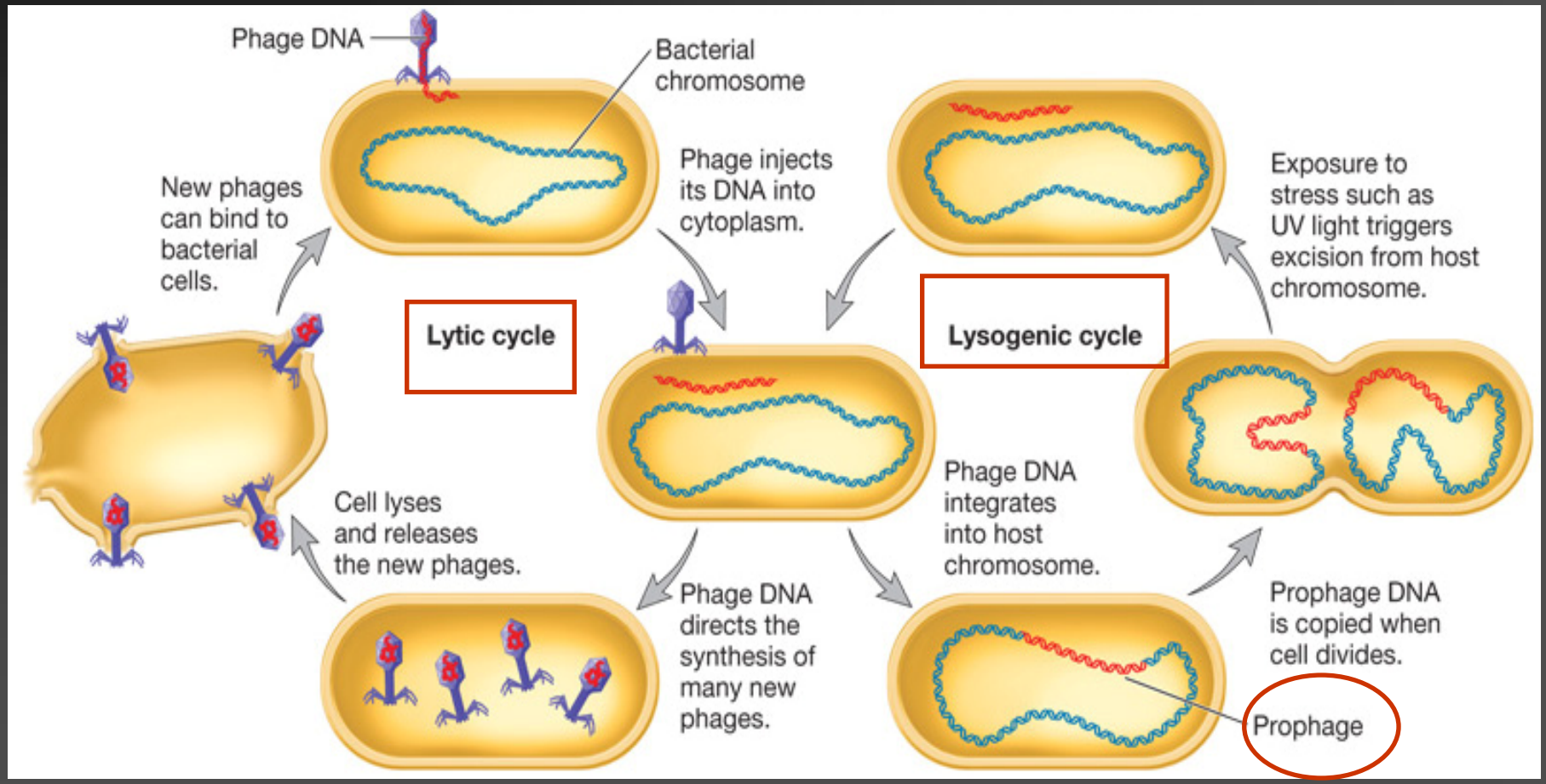


Fig 14.25

Generalized transduction

- any part of bacterial genome can be transferred
- occurs during lytic cycle
- during viral assembly, fragments of host DNA mistakenly packaged into phage head

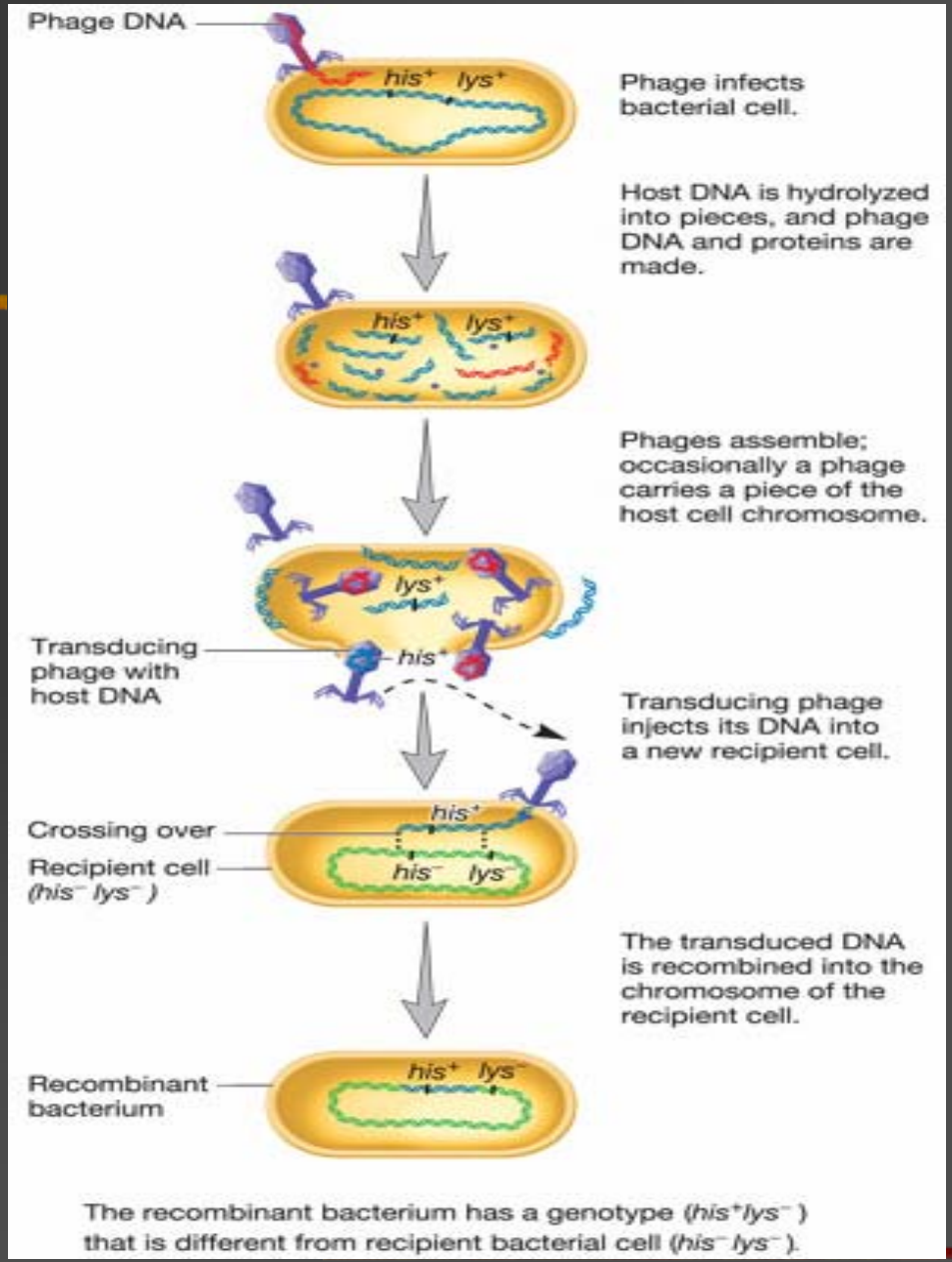


Figure 14.30

Specialized Transduction

- carried out only by temperate phages that have established lysogeny
- only specific portion of bacterial genome is transferred
- occurs when prophage is incorrectly excised

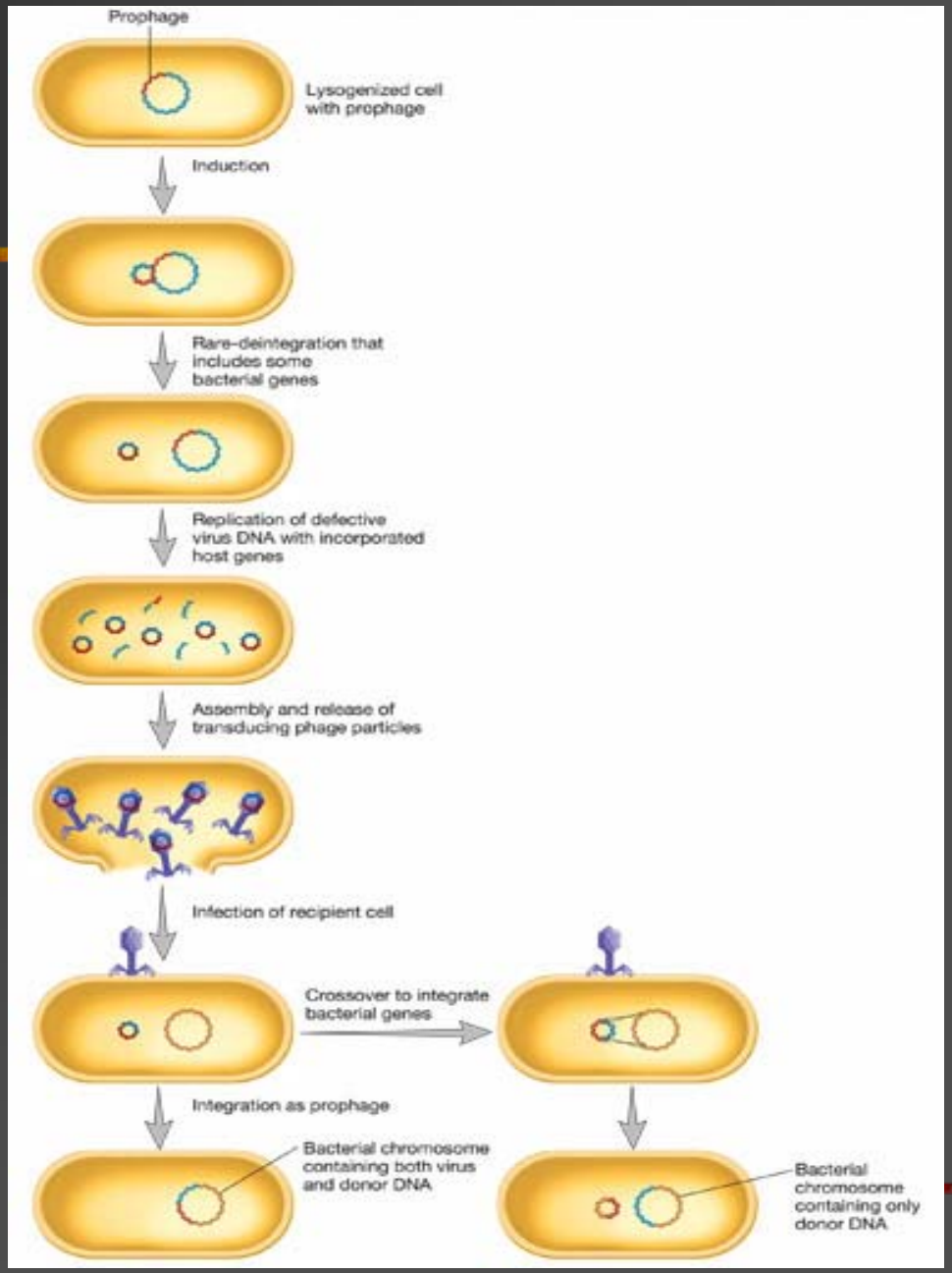
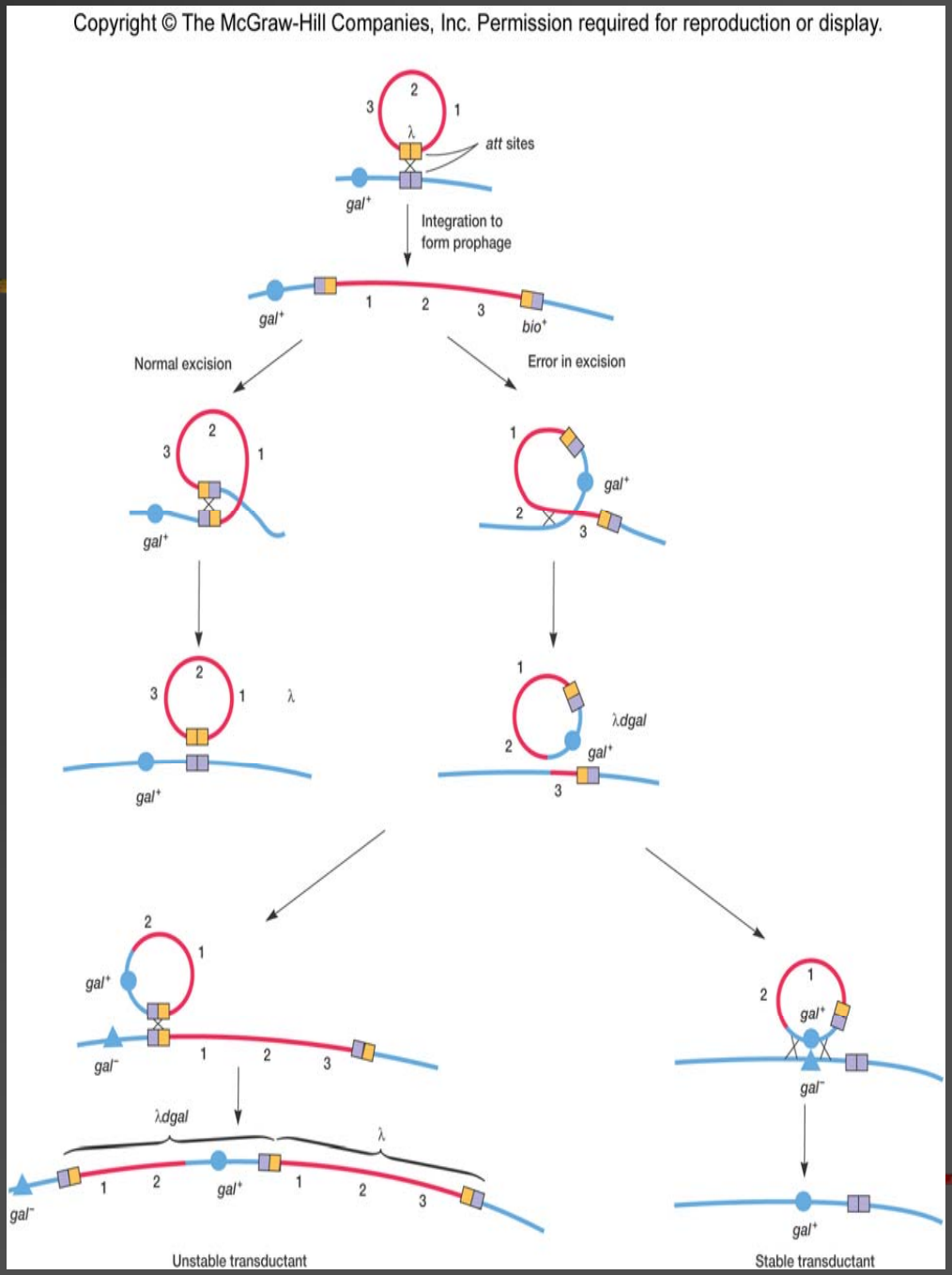


Fig 14.31

Transduction for phage Lamda



Transmission of drug resistance

- The origin of drug resistance
  - R plasmid
    - Tn5, Tn21, Tn551
  - Intergon
    - several resistance genes carried together as gene cassettes

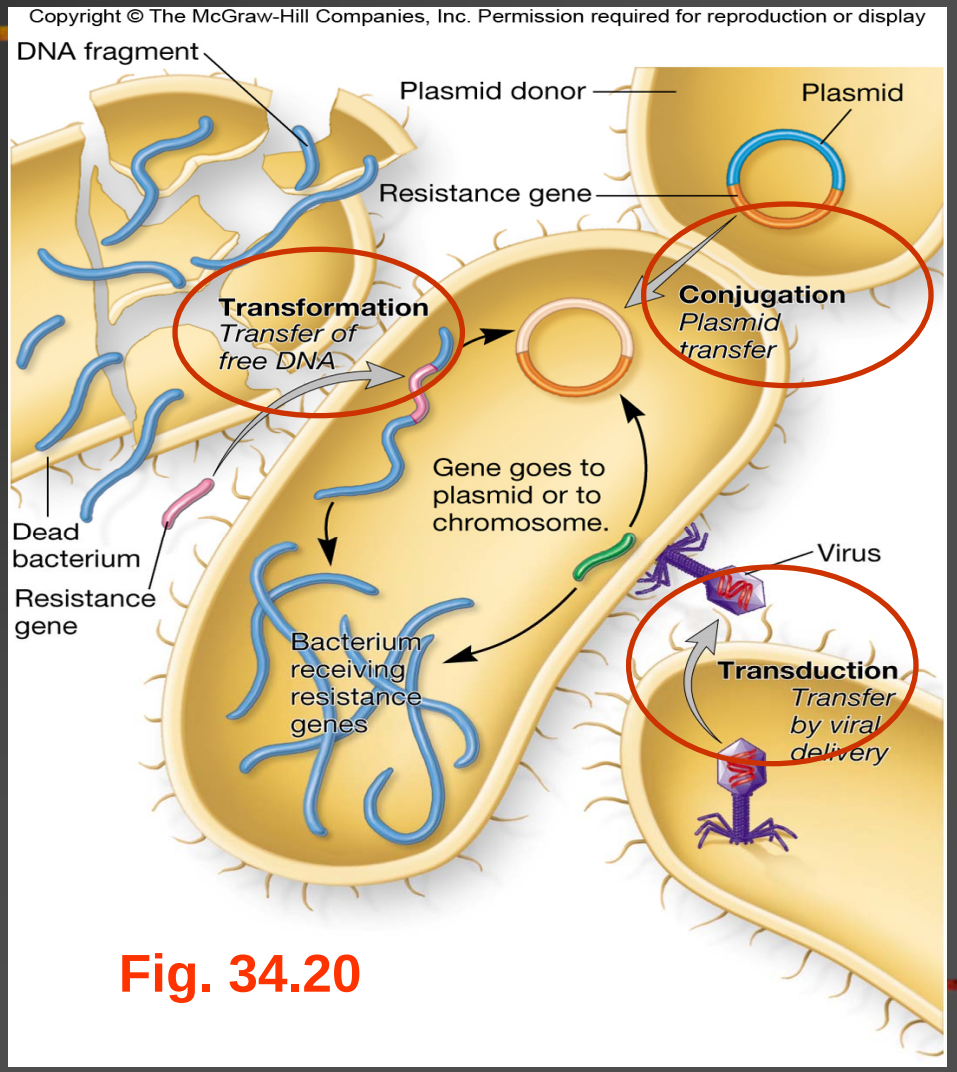
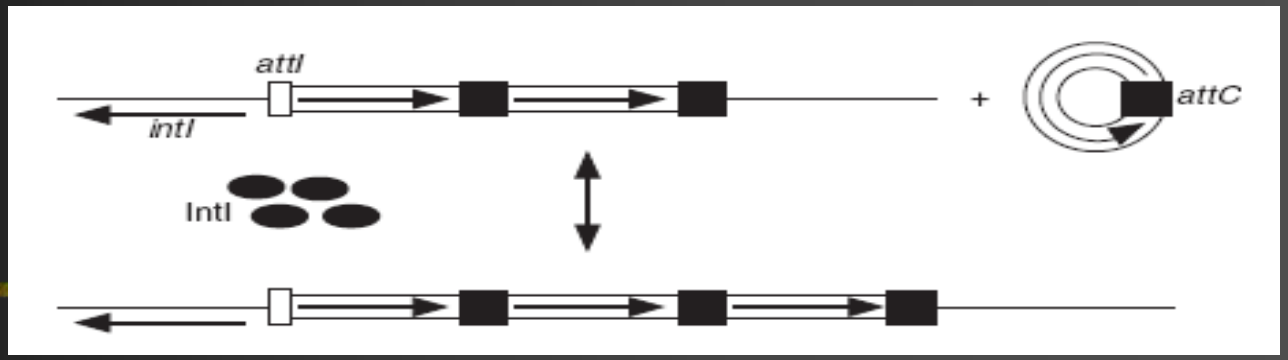


Fig. 34.20

Integron

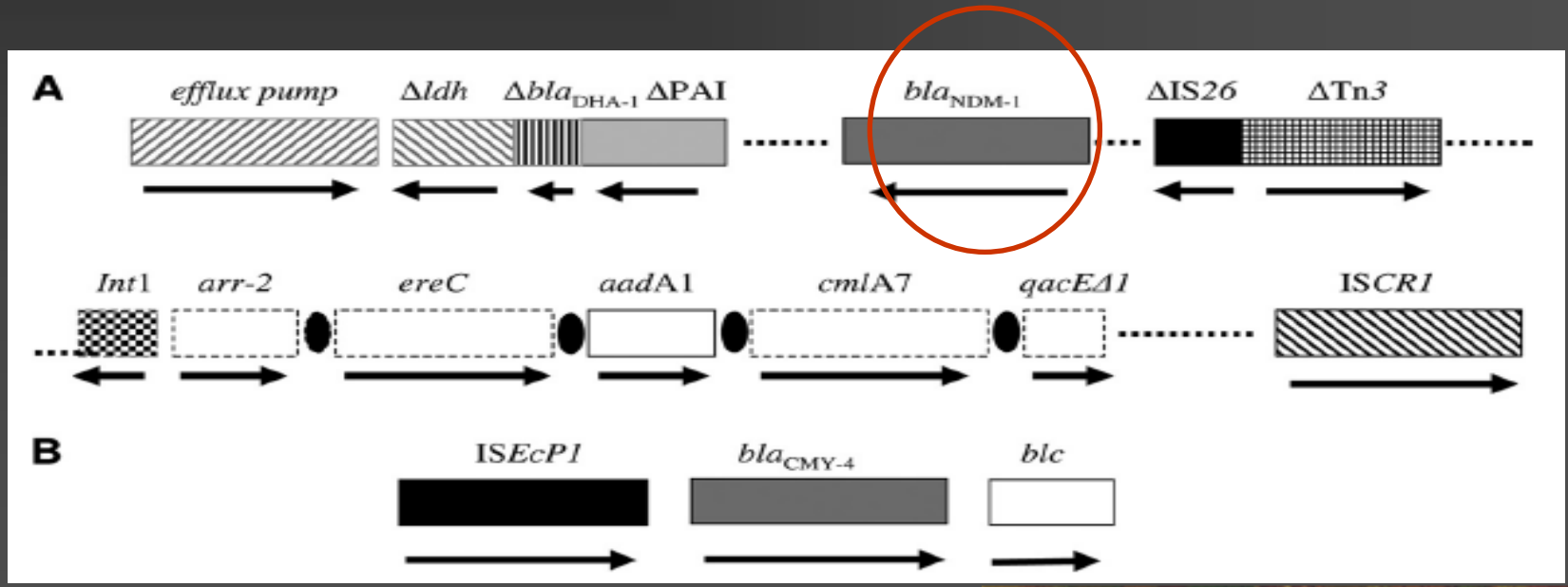


- A gene cassette typically consists of little more than a single promoter-less gene and a recombination site
  - *intI* encodes an integrase of tryosine recombinase family
  - *attC* and *attI*
  - At least 130 different cassettes that carry known or predicted antibiotic resistance genes, along with many cassettes of unknown function

FEMS Microbiol Rev 33 (2009) 757–784

Three characterized antibiotic resistance-conferring regions from *K. pneumoniae* 05-506

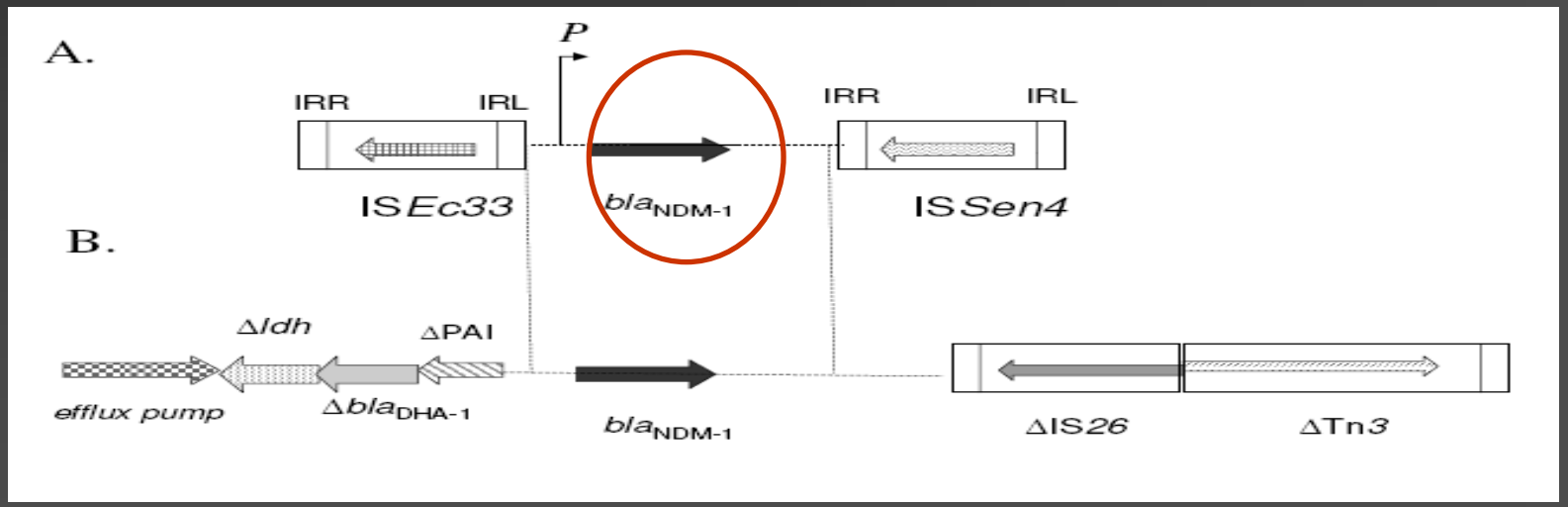
- A UTI (urinary tract infection) isolate of KPC strain 05-506, from a Swedish patient of Indian origin traveled to New Delhi, India, which carries a new subgroup of metallo-lactamase



Antimicrob. Agents Chemother., Dec. 2009, p. 5046–5054

NDM-1 producing *E. coli*

- Emergence of NDM-1-producing multidrug resistant *Escherichia coli* in Australia



Antimicrob. Agents Chemother. doi:10.1128

Chapter 16

Microbial Genomics-1

Genomics

- study of molecular organization of genomes, their information content, and gene products they encode

Determining DNA Sequences

- Sanger Method (1975)
  - uses dideoxynucleoside triphosphates (ddNTP) as chain terminator
- automated systems
  - use dideoxynucleotides labeled with fluorescent dyes

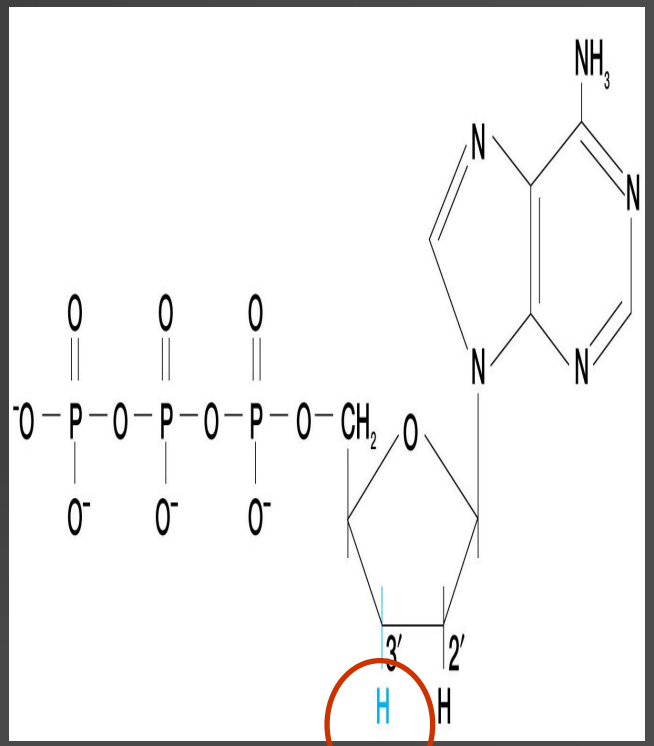


Fig. 16.1

Chain termination method

- Four reactions
  - **1S DNA template**
  - **primer**
  - **DNA polymerase**
  - **4 dNTP**
  - **one ddNTP**
- random insertion of ddNTP generates different lengths of DNA fragments
- fragments separated electrophoretically
- sequence read

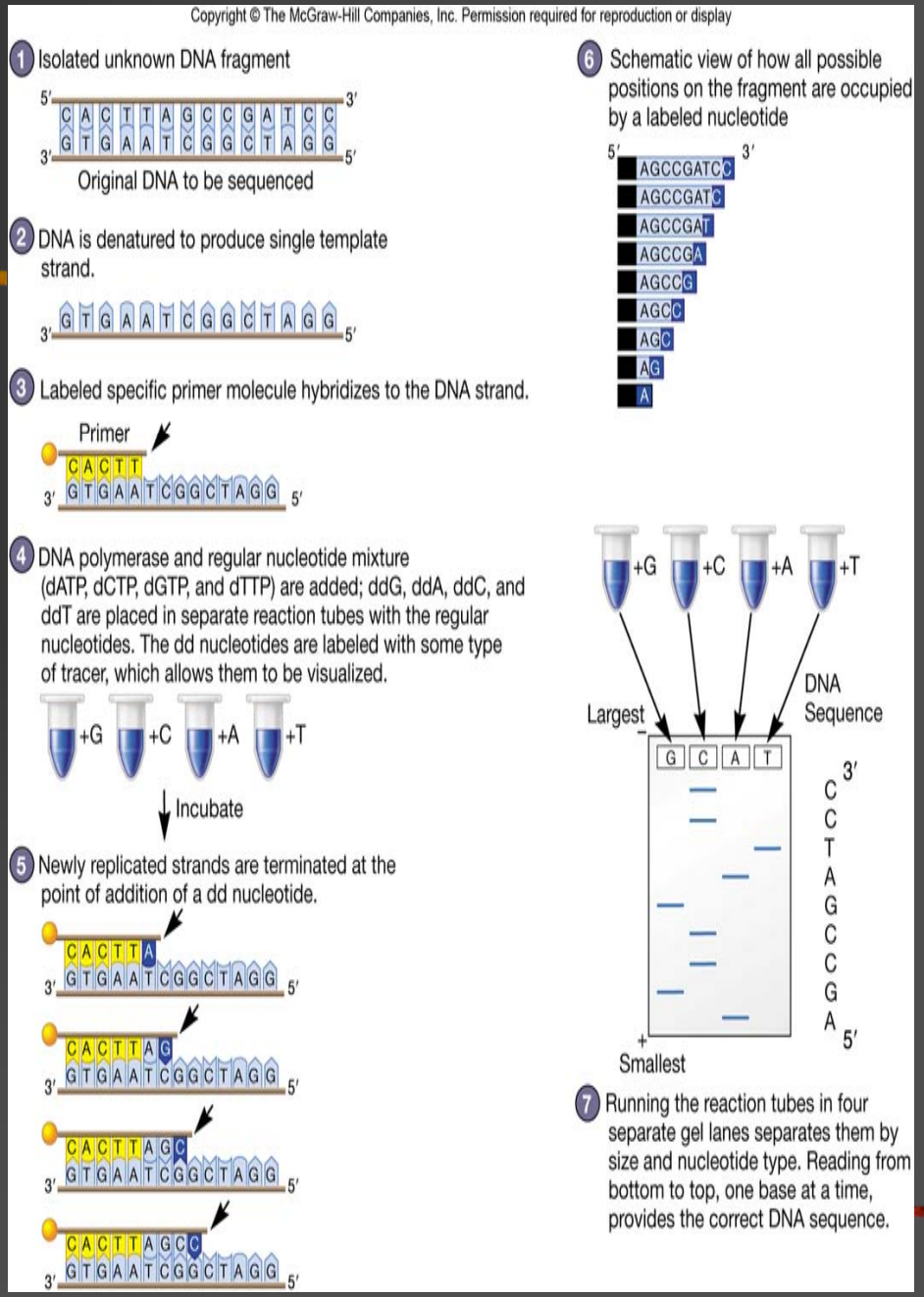


Fig. 16.2

Automated Sanger DNA sequencing

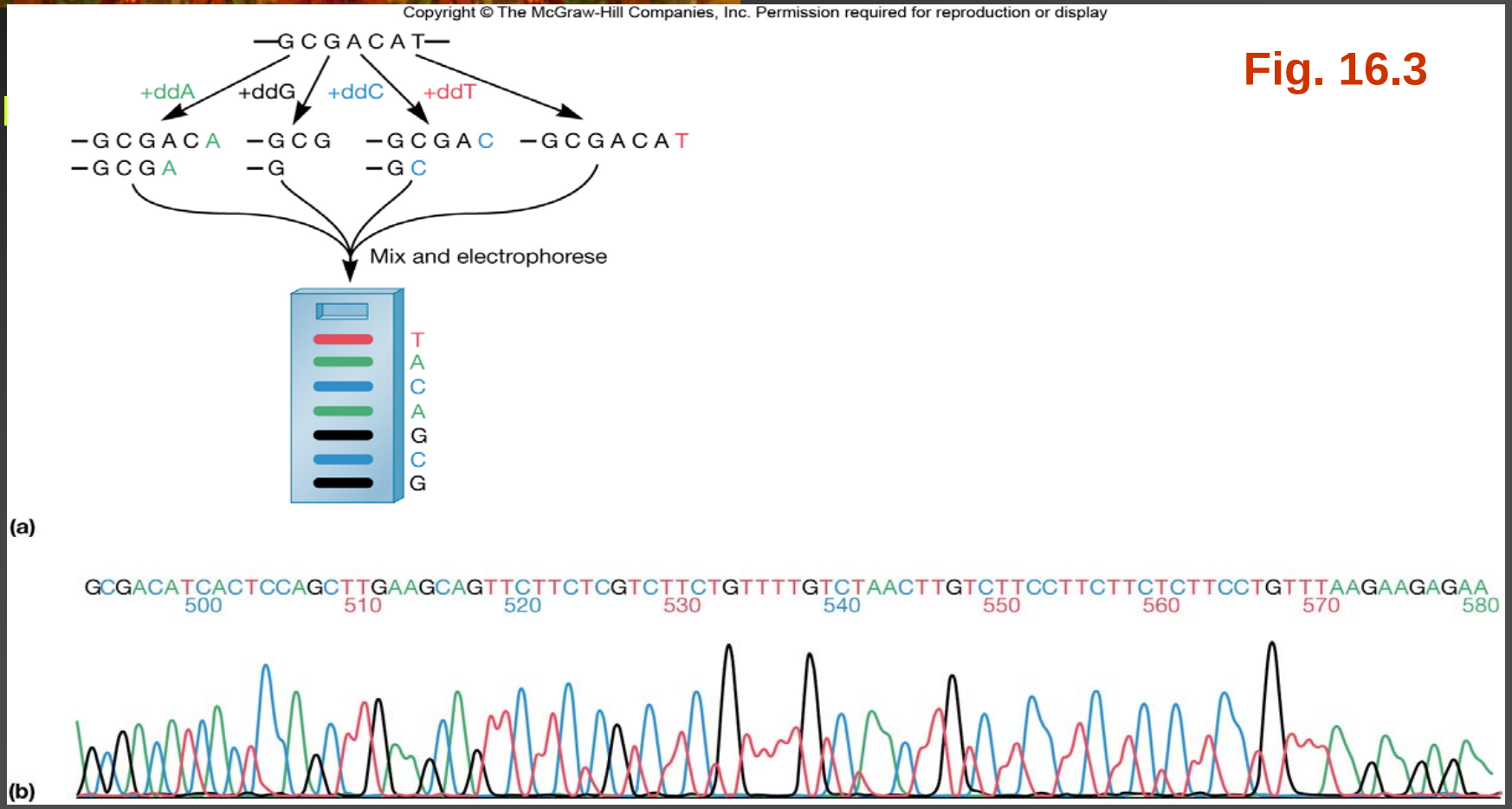
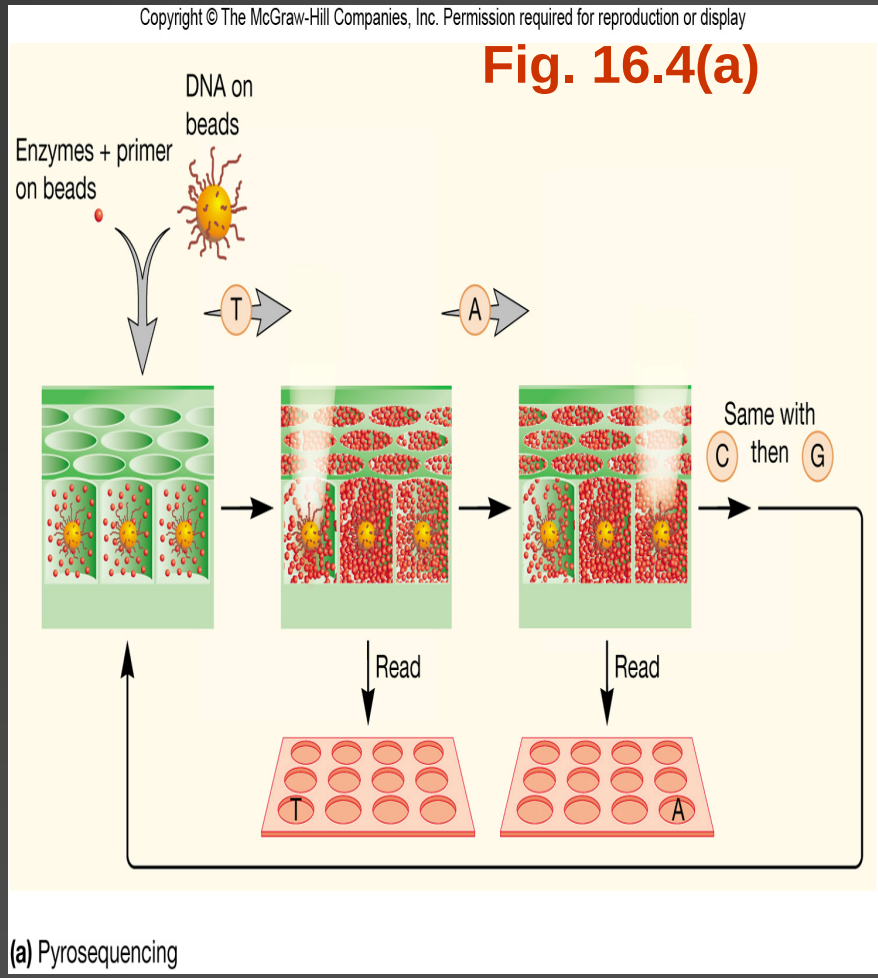


Fig. 16.3

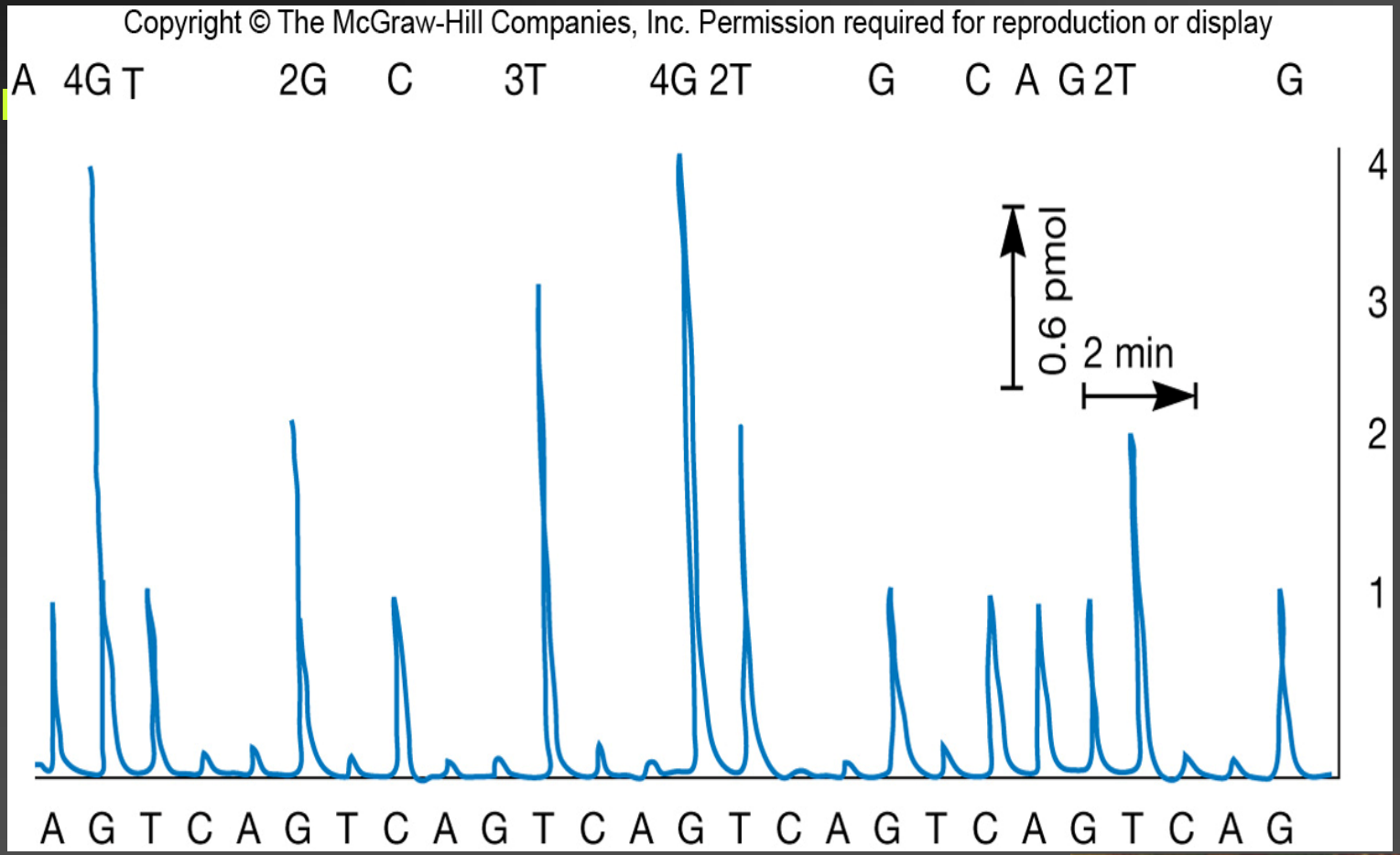
Post-Sanger DNA sequencing- pyrosequencing

- Each bead coated with PCR-amplified chromosomal fragments (300-500 bp)
- pyrophosphatase sequencing
  - $(DNA)_n$+ dNTP→ $(DNA)_{n+1}$ +PPi by DNA polymerase
  - PPi + APS (adenosine phophosulfate)→ ATP + SO_4^{2-} by ATP sulfurylase
  - ATP+luciferin+O2→ AMP +ppi +oxylucerin +CO2+ light by luciferase



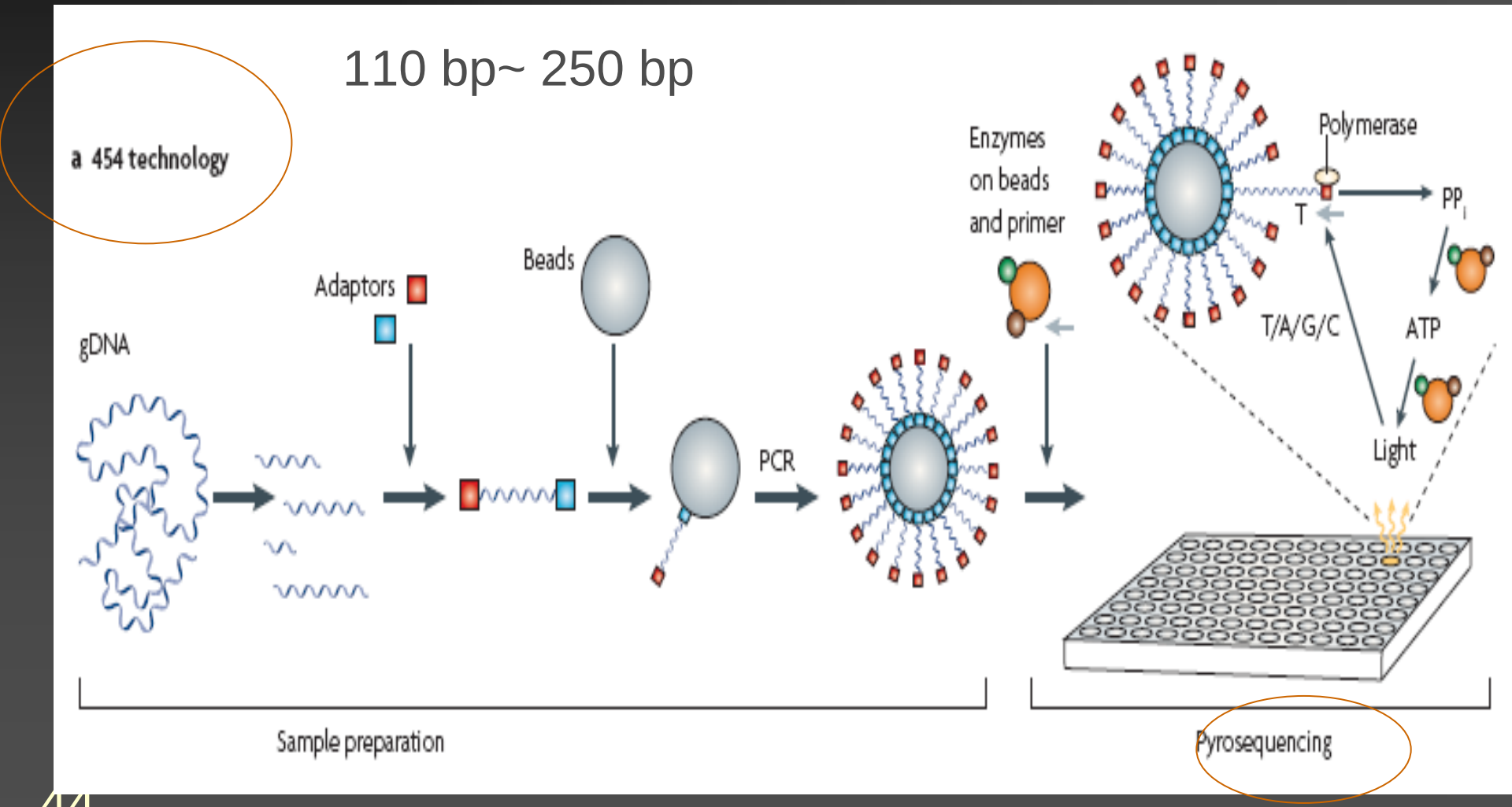
A Pyrogram

Fig. 16.5



Roche 454 Genome Sequencing

Nature Rev Microbiology (2008) May



SOLEXA sequencing

- A glass slide attached with the PCR amplification of immobilized DNA fragnents
- Sanger-type sequencing
  - 4 fluorophore terminators

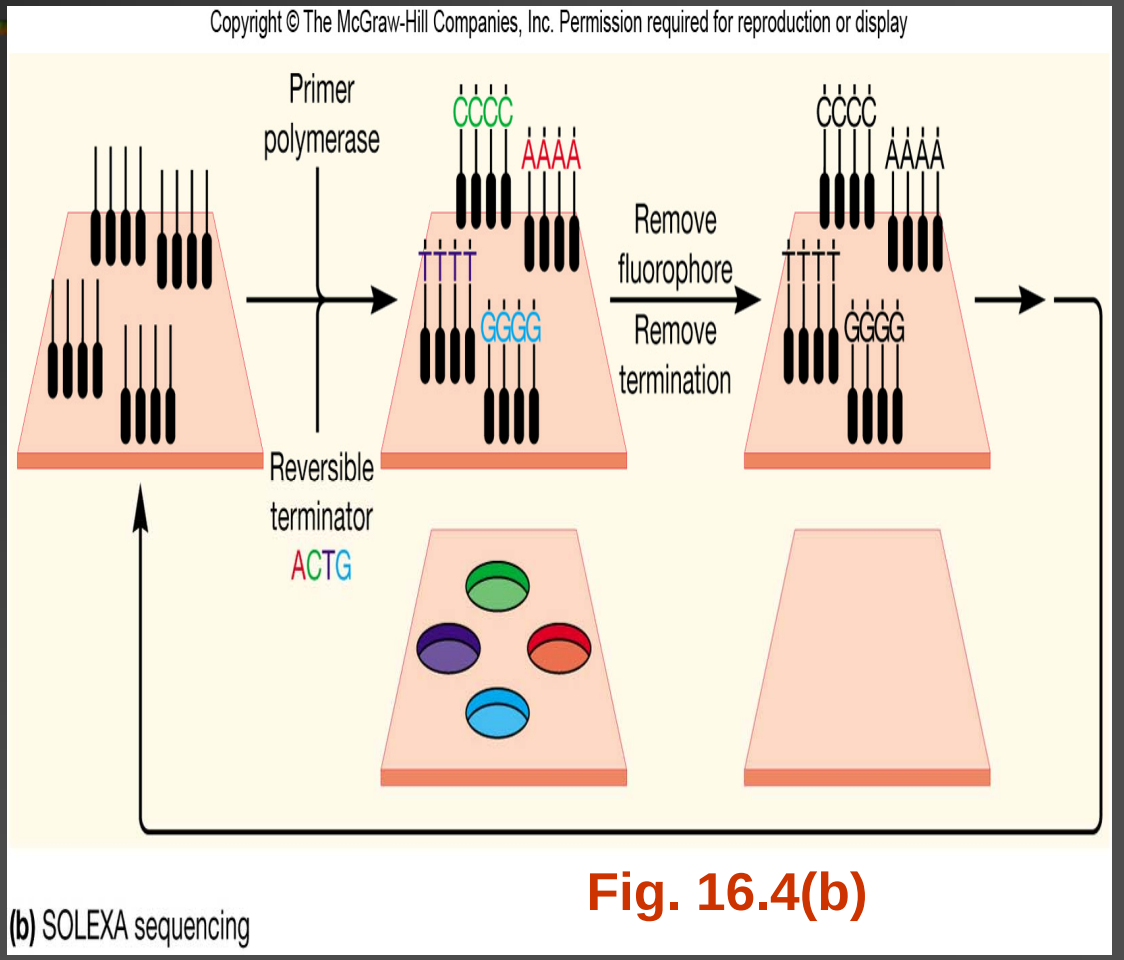


Fig. 16.4(b)

SOLiD sequencing

- *S*upported *o*ligo *li*gation *d*etection
- The DNA fragments are lengthen by
  - **the addition of a short piece of DNA (the anchor primer)**
  - **eight-base oligonucleotides (by every possible combination of A-T-C-G) and each has a fluorescently labeled A, T, C, or G in the 4th and 5th positions**
  - **Ligase**
  - **Exciting with laser light**

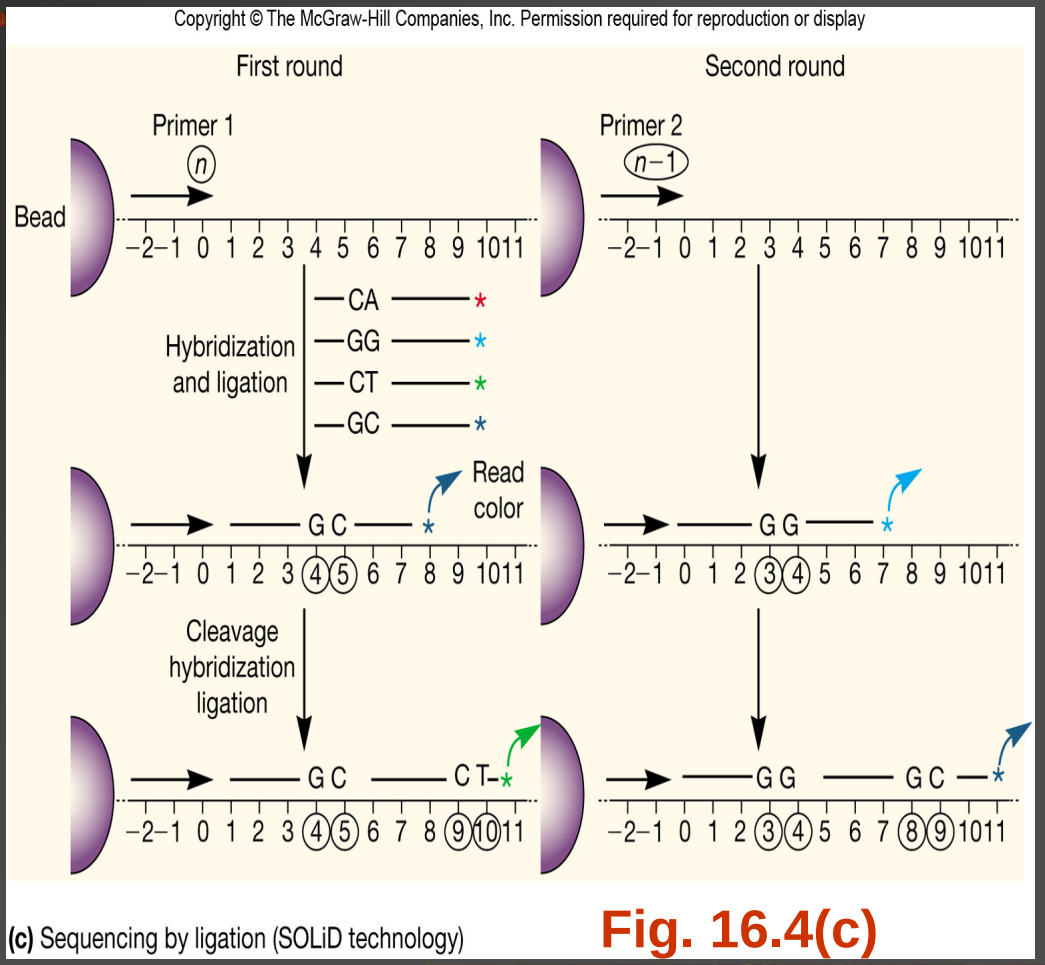


Fig. 16.4(c)

Whole-genome shotgun sequencing

- Library construction to provide the templates
  - genomic clones
- random sequencing
- fragments alignment and gap closure
  - Contig- a set of larger, contiguous nucleotide sequences
- Editing (proofread)
  - 1000 kb/ 24 h

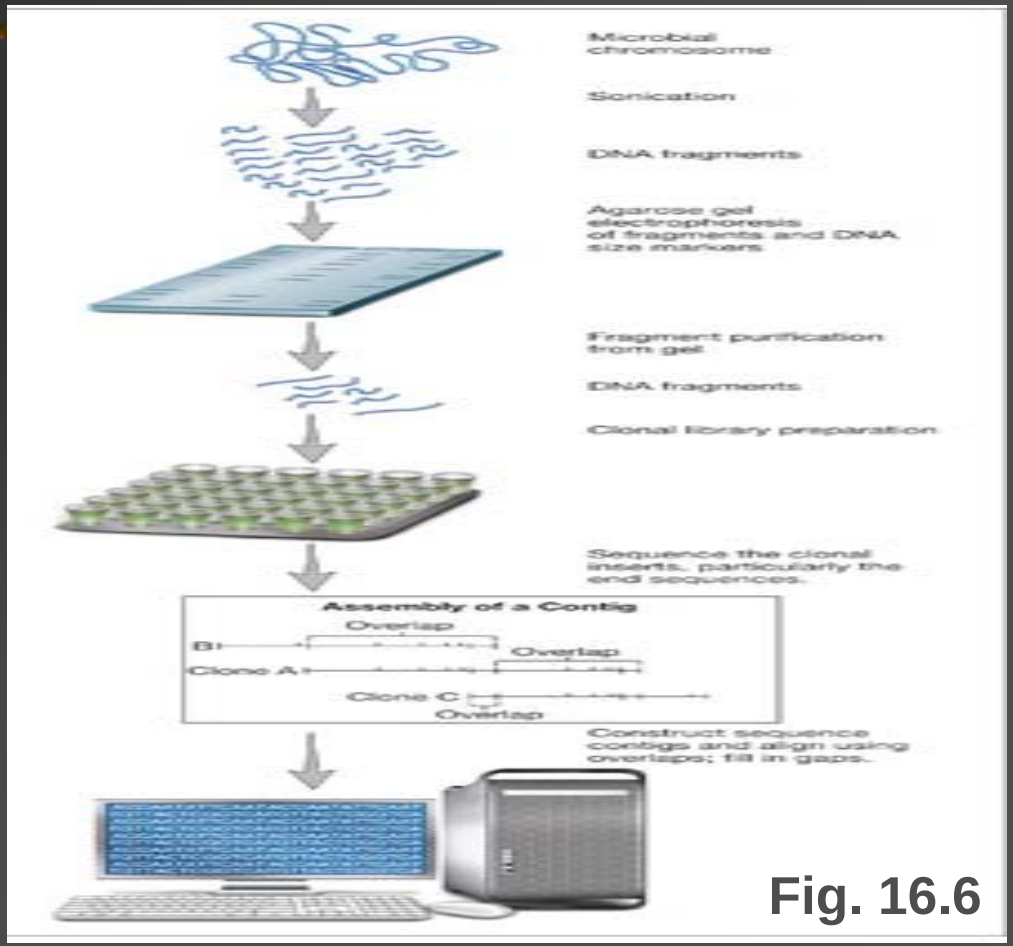


Fig. 16.6

1995, J. Craig Venter and Hamilton Smith (TIGR: the institute of genomic research)

Single-Cell genomic sequencing

- Multiple strand displacement (NDA)
  - **DNA from single cell (down to 10^{-15} g)**
  - **random primers (hexamers)**
  - **phage phi29 DNA polymerase**
  - **dNTPs**
- Many new strands (12 kb~100 kb) are synthesized→ cloning→ sequencing

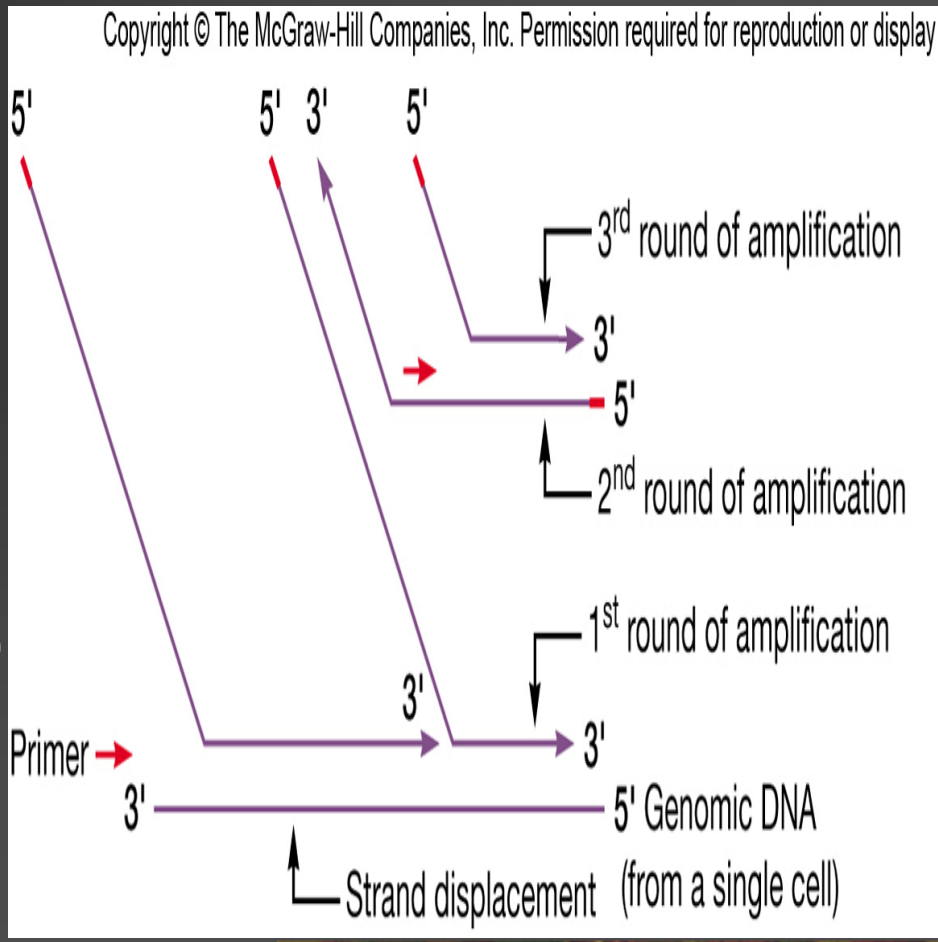


Fig. 16.7

Bioinformatics

- Combine Biology, mathematics, computer science, and Statistics to generate data on genome content, structure, and arrangement
  - also provides data on protein structure and function
  - DNA sequence data stored in large databases
    - International Nucleic Acid Sequence Data Library (GenBank)
  - analysis of genome data using computers
    - *in silico* analysis

Genome annotation

- To determine location of genes and nature of genes or presumed genes
- identifies open reading frame (ORF) in genome
  - a reading frame > 100 codons that is not interrupted by a stop codon
  - there is an apparent ribosomal binding site at the 5’ end and terminator sequences at the 3’ end
- To assign tentative function of gene

Finding potential protein coding genes
- CDS (coding sequences)

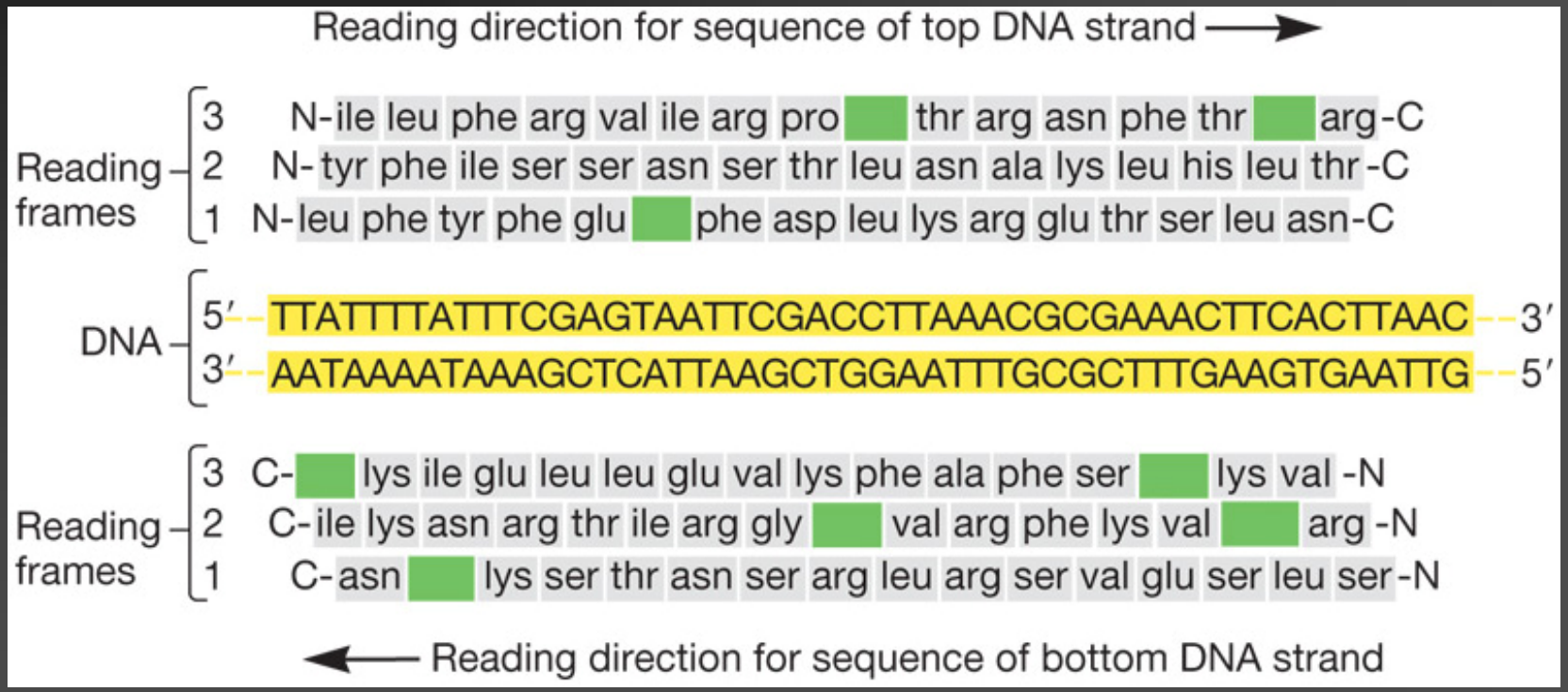


Fig 16.8

Alignment of the conserved regions

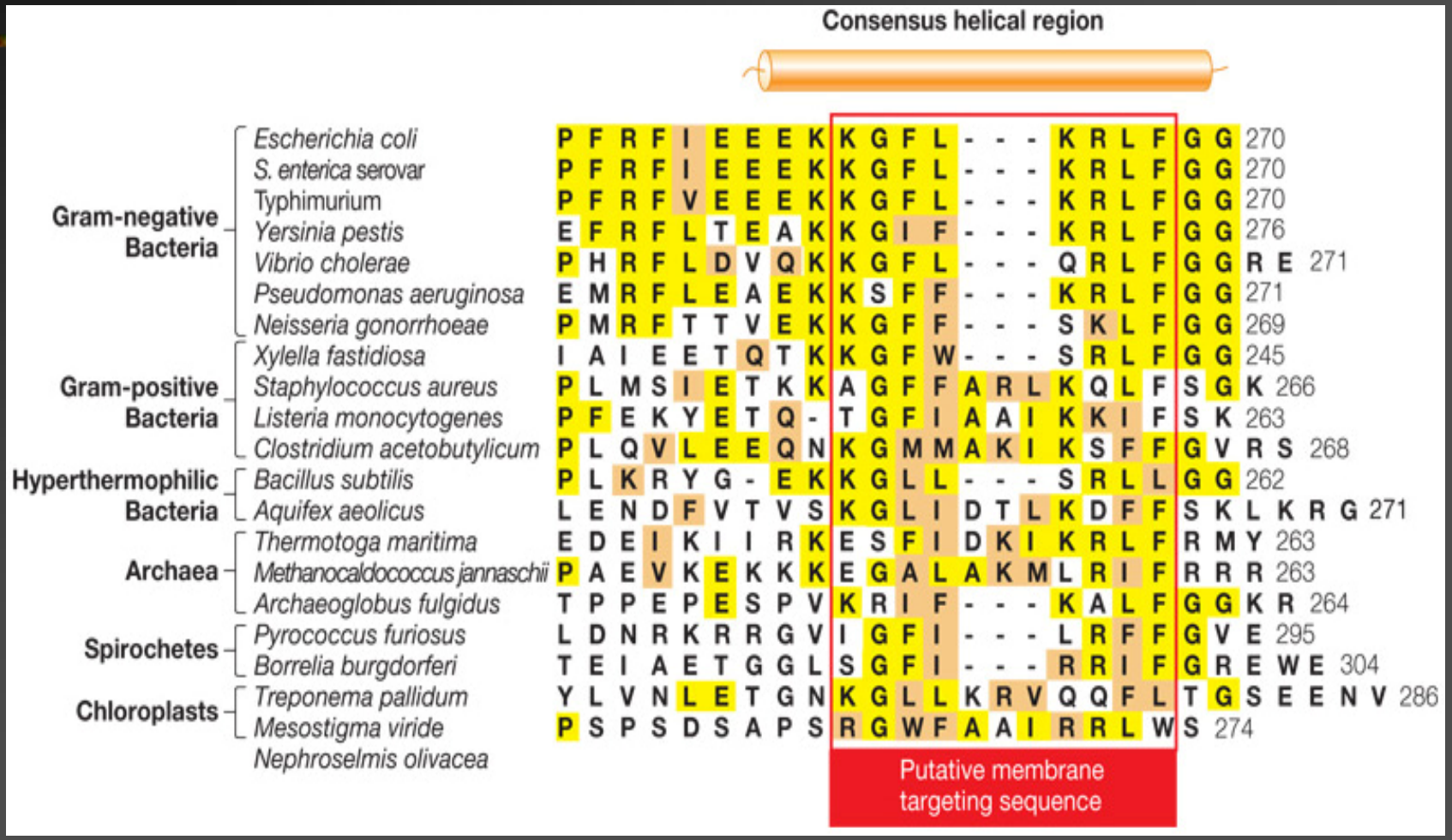


Fig 16.9 Pfam database

Chapter 16

16. Microbial Genomics- 2

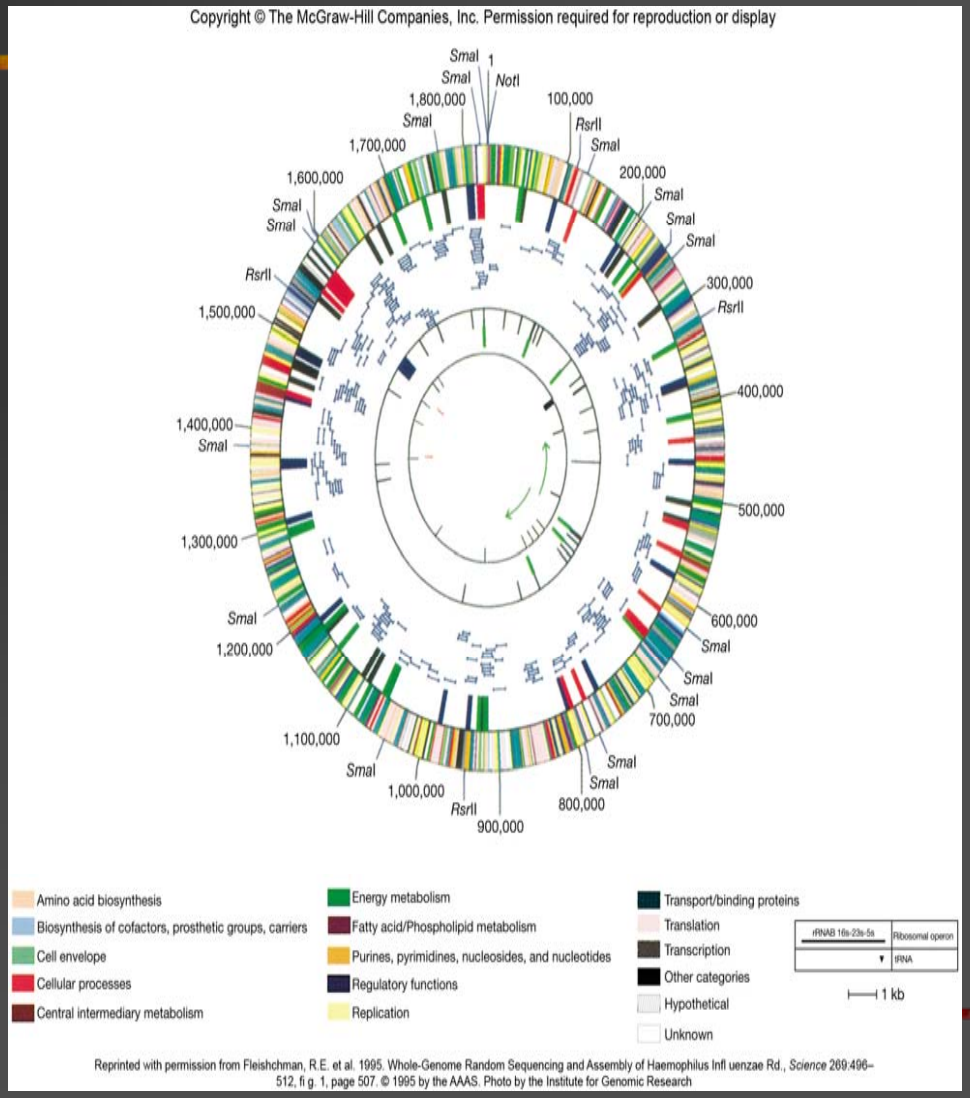
11-17-2010

Functional genomics

- Determination of how genome works
  - from alignment of gene sequences
    - paralogs – genes arose from gene duplication
    - orthologs – genes very similar and are predicted to have same function
  - involves analysis of translated amino acid sequence of presumed genes to understand protein structure and function
    - motif, a short pattern of amino acids, may represent a functional unit within the protein, such as the active site

Physical map of the *Haemophilus influenza* genome

- 1st annotated genome
  - Color-code genes
    - 1/3 of unknown function (white regions)
    - 65 regulatory genes (dark blue)
- > 700 complete bacterial and archaeal sequences now available



Treponema pallidum- syphilis 梅毒

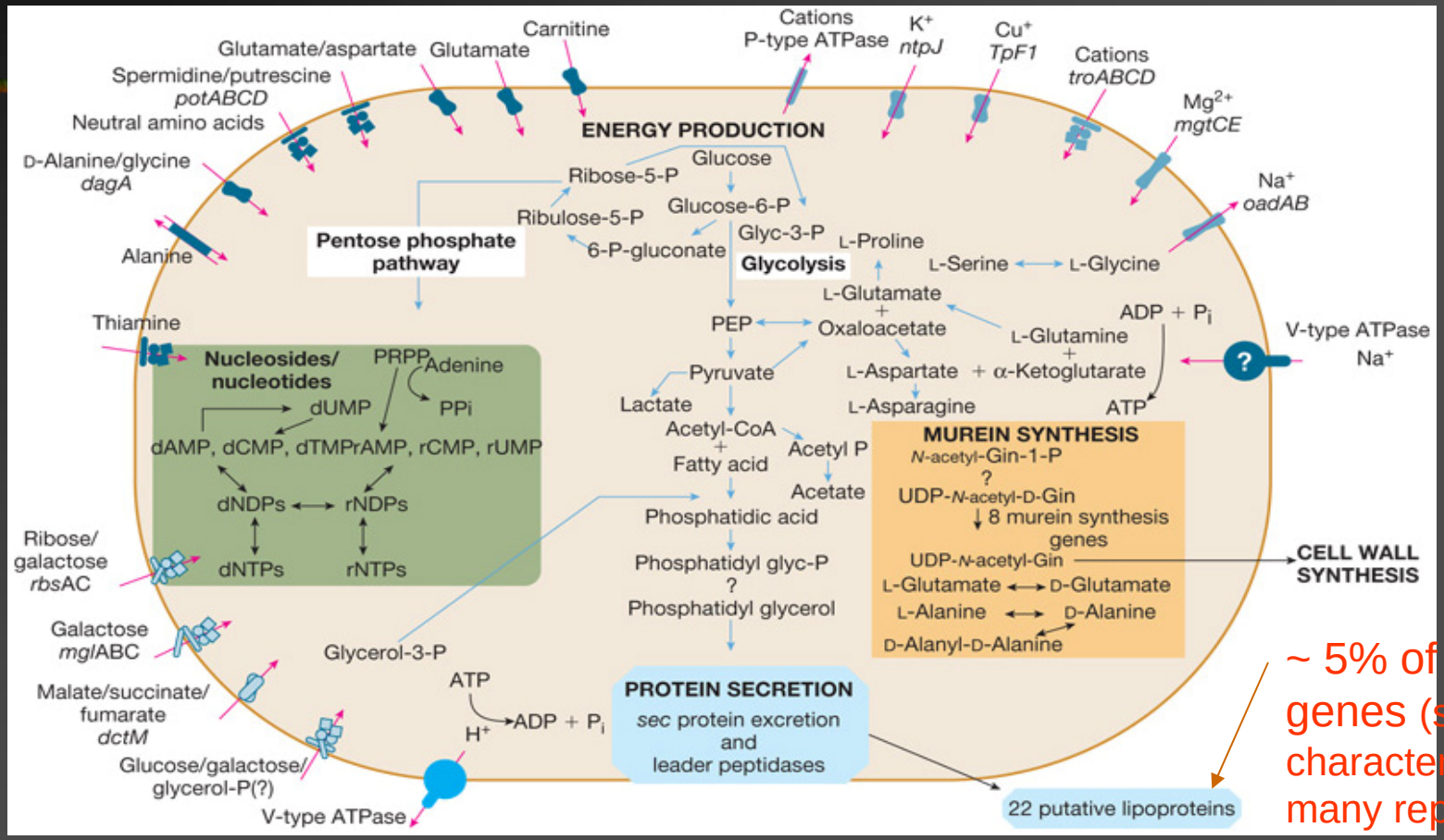
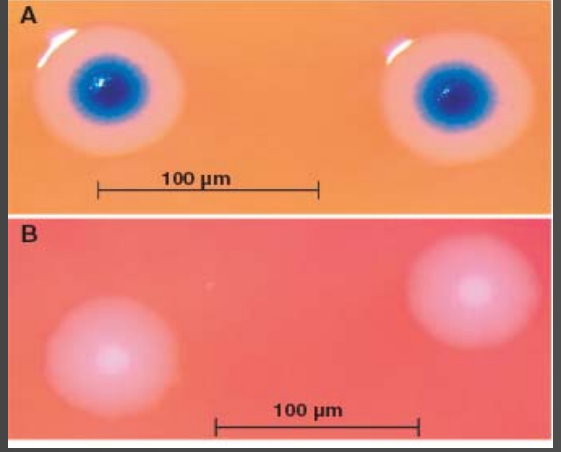


Figure 16.11

metabolically crippled- No TCA cycle and respiratory electron transport

~ 5% of the genes (some characterized by many repetitive sequences→ recombination→ generate new surface proteins)

First synthetic genome



- *Mycoplasma mycoides* JCVI-syn1.0
  - 1.08 million bp synthesized chromosome of a modified *M. mycoides* genome at the J. C. Venter Institute **(San Diego, CA, and Rockville, MD, USA)**
  - ~ 1 kb→ 10 kb in *E. coli* →100 kb in yeast →assembly of 10 X ~100 kb fragments in yeast→ mycoplasma
  - The new cells have expected phenotypic properties and are capable of continuous self-replication (>30 divisions) and making new set of proteins.

July 2010 Science 329:52-56

DNA microarray analysis (Gene Chips)

- Determine gene expression at a specific time
  - spotted arrays
    - prepared by robotic application of DNA probe
    - each DNA probe represents a single gene or ORF
      - **PCR product, cDNA or oligonucleotide**
      - **oligonucleotide probes from eukaryotes-**
        - **expressed sequence tags (EST)**
  - photolithography

Construction of DNA chip- photolithography

Monitoring changes of gene expression

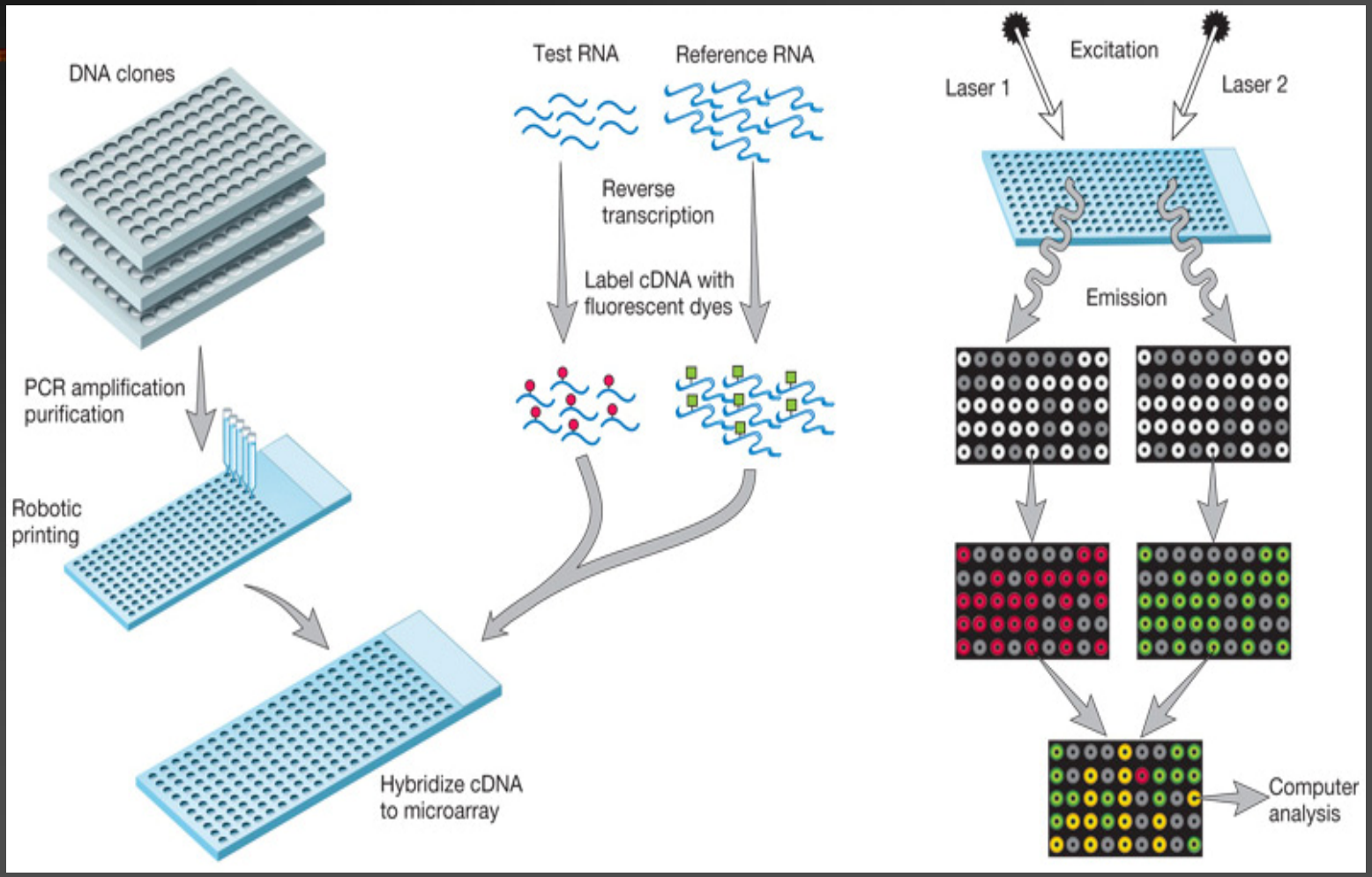
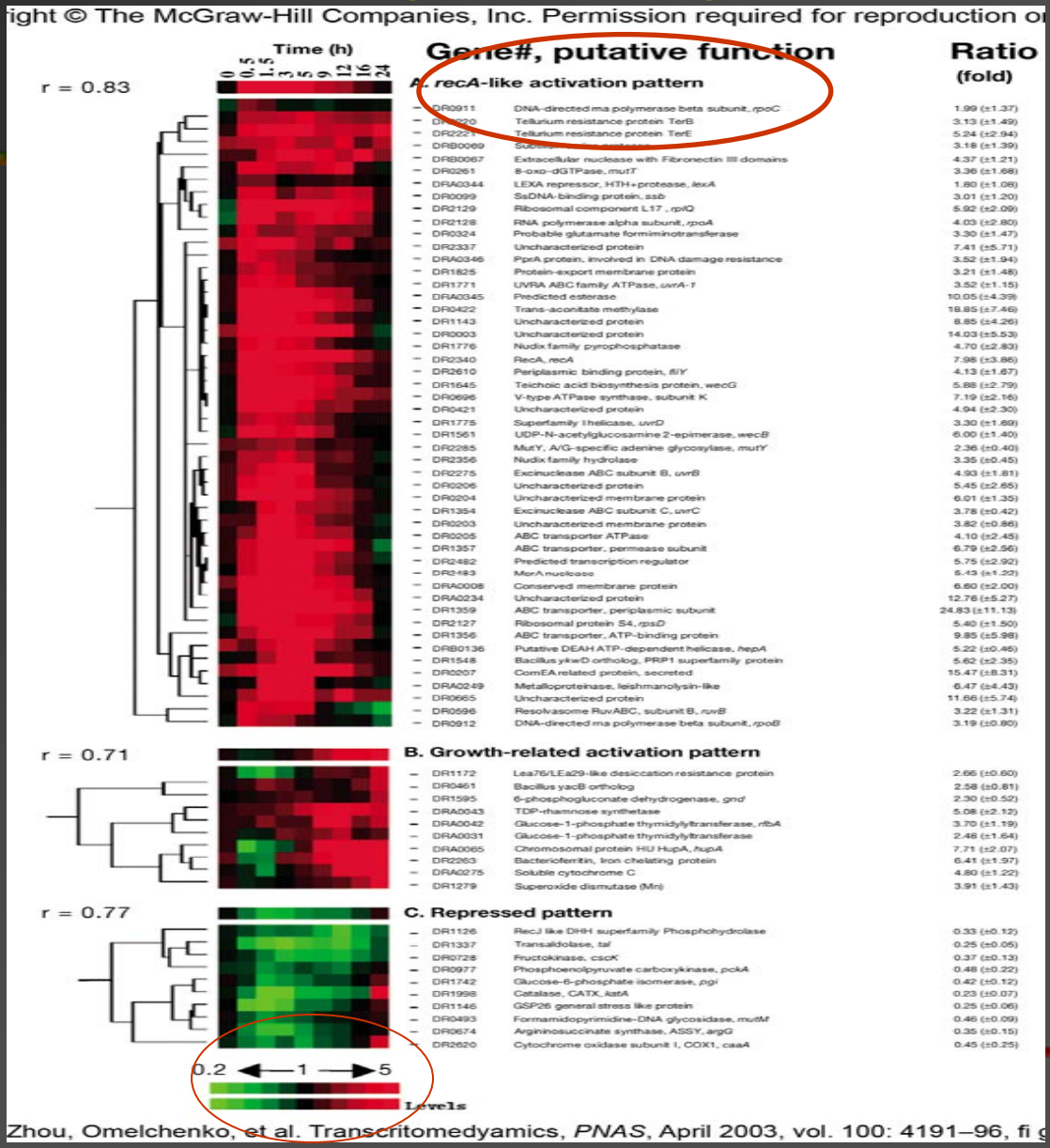


Figure 15.9

Hierarchical cluster analysis of gene expression

- Analysis of gene expression of *D. radiodurans* following exposure to γ-radiation
  - Each group of genes has been scored for relatedness
  - A tree has been generated with indication of **correlation coefficient** (r value)
  - Genes **induced** (**red**) or **repressed** (**green**) after radiation



Proteomics

- The study of the proteome (the entire collection of proteins)
  - often analyzed by two-dimensional gel electrophoresis
  - Functional proteomics to determine what is actually happening in cell proteome
  - Structural proteomics to resolve 3 dimensional structure of the proteins
    - protein modeling (on the basis of the assumption that proteins folds into a limited number of shapes)

Two-dimensional gel electrophoresis

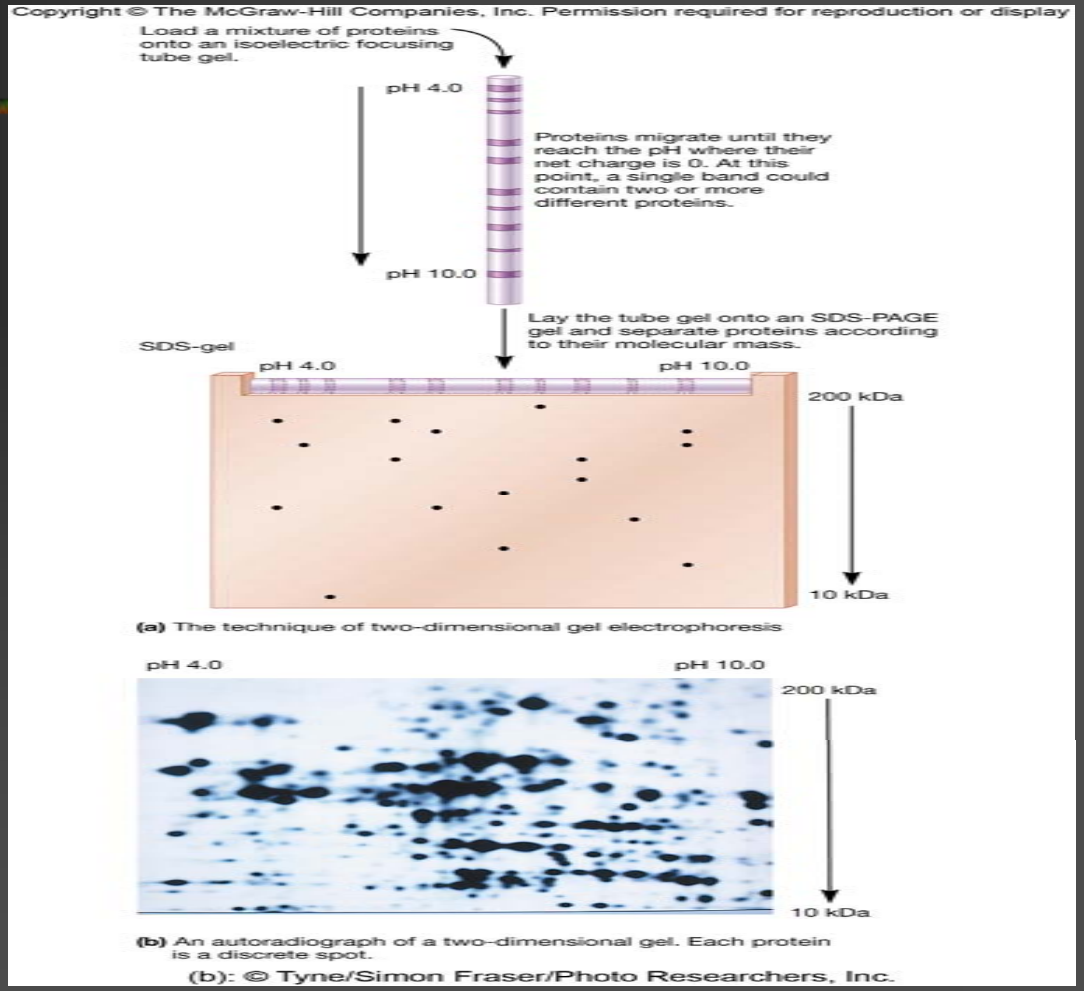


Fig 16.14

Tandem Mass Spectrometry

- unknown spot from 2-D gel is cut and cleaved
- fragments are analyzed by mass spectrometer
- mass of fragments is plotted
- protein tentatively identified from probable amino acid composition

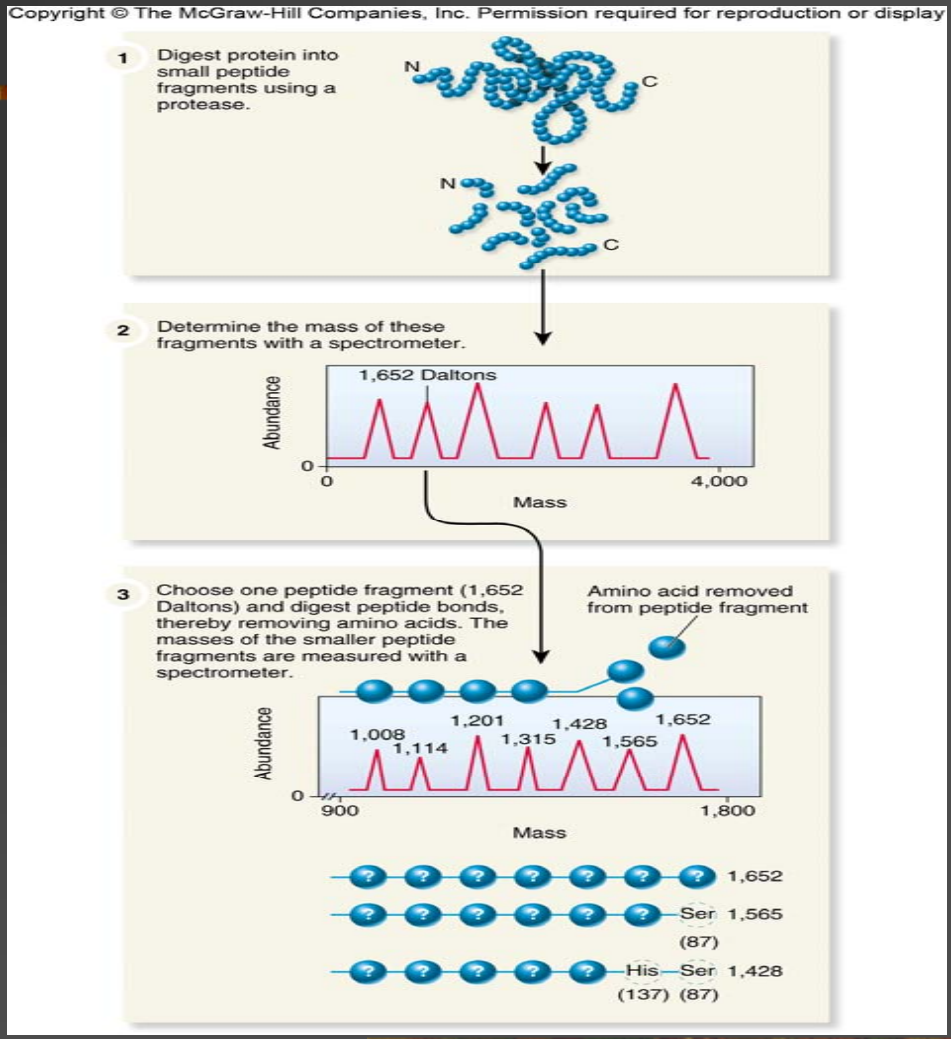
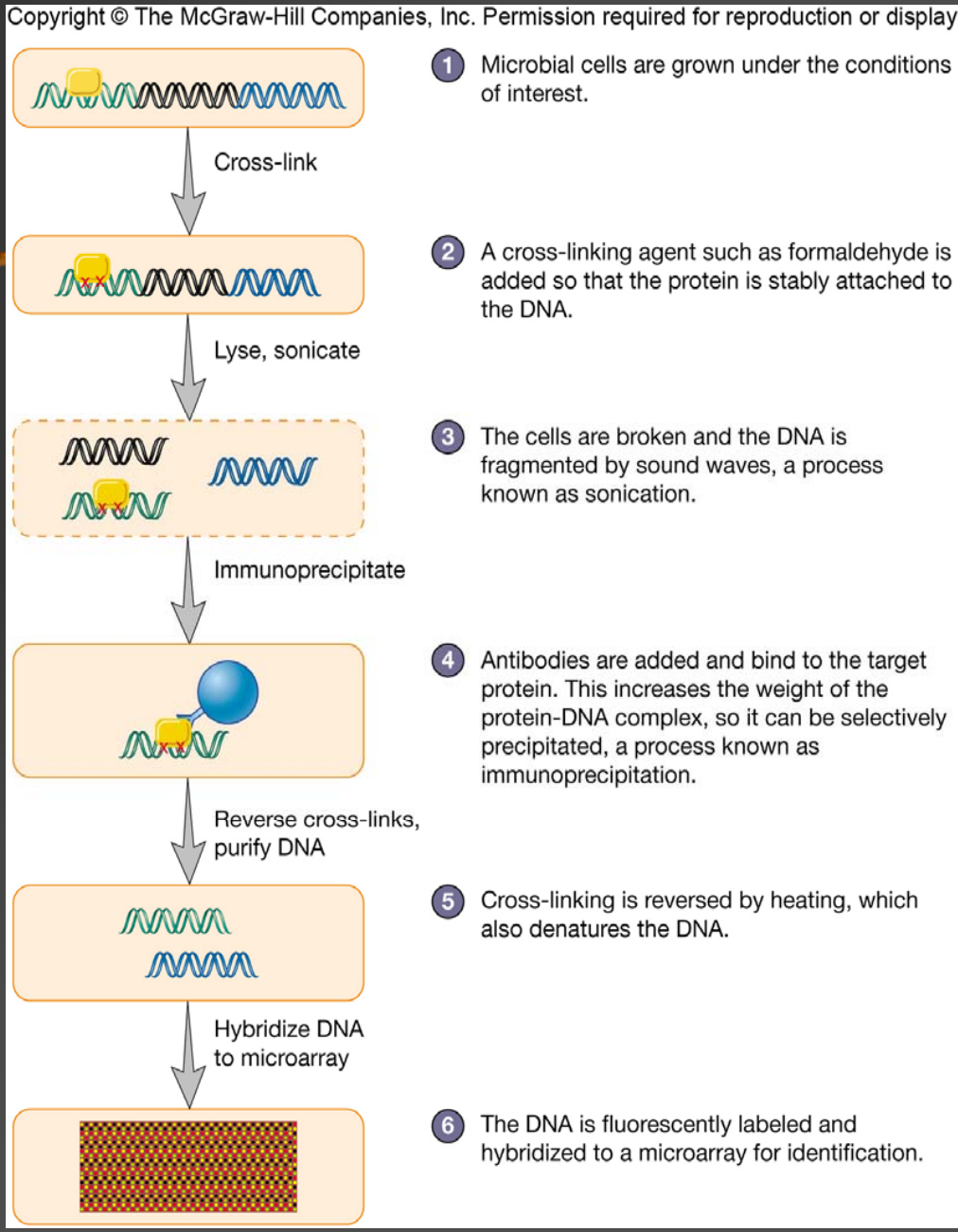


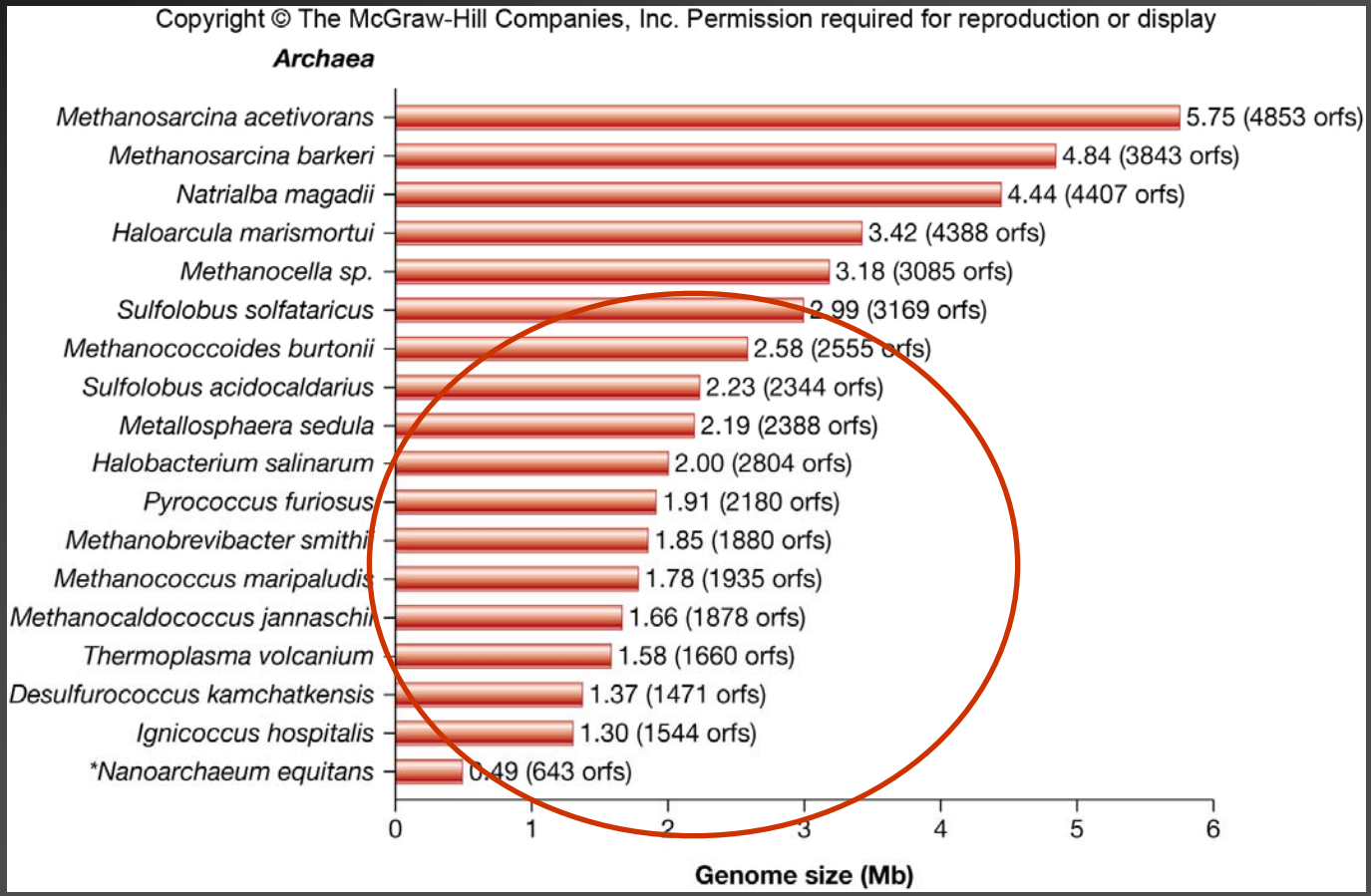
Fig 16.15

ChIP-Chip analysis

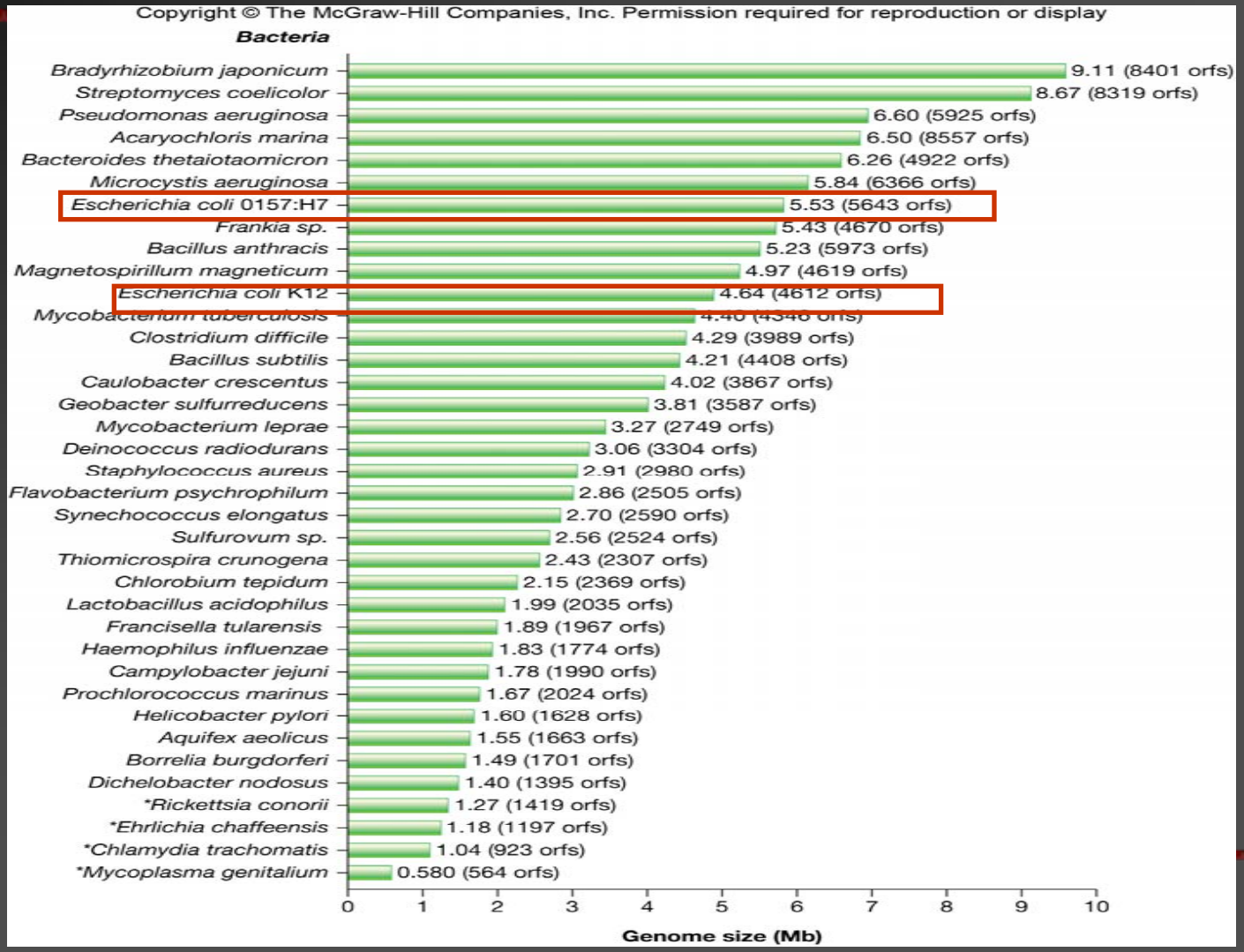
- probing DNA-protein interaction
  - EMSA (electrophoretic mobility shift assay) or called gel mobility shift assay
  - ChIP- chromatin immunoprecipitation
    - **search for protein binding DNA**



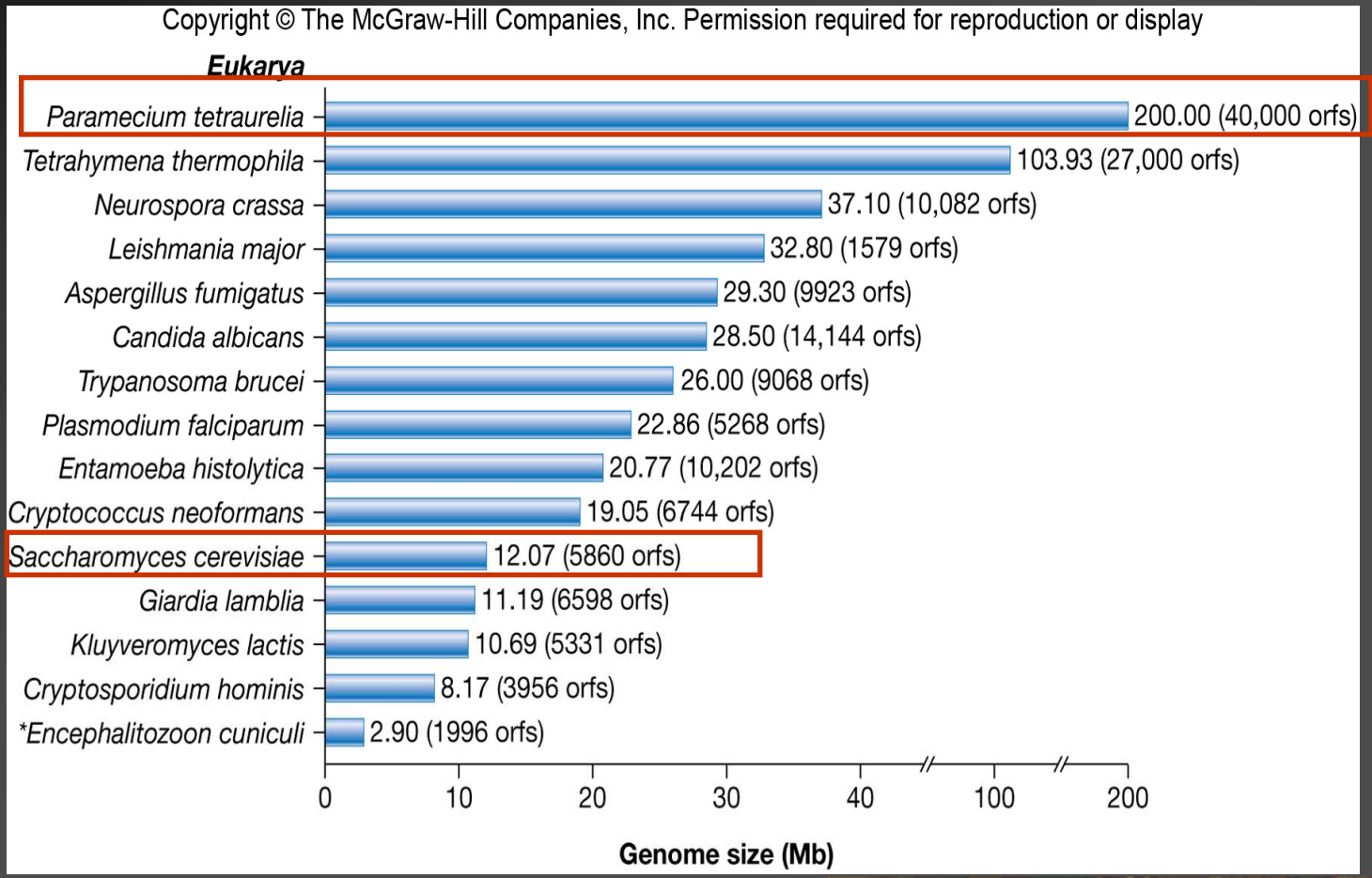
Comparative Genomics- *Archaea* genome size



Comparative Genomics- *Bacteria* genome size

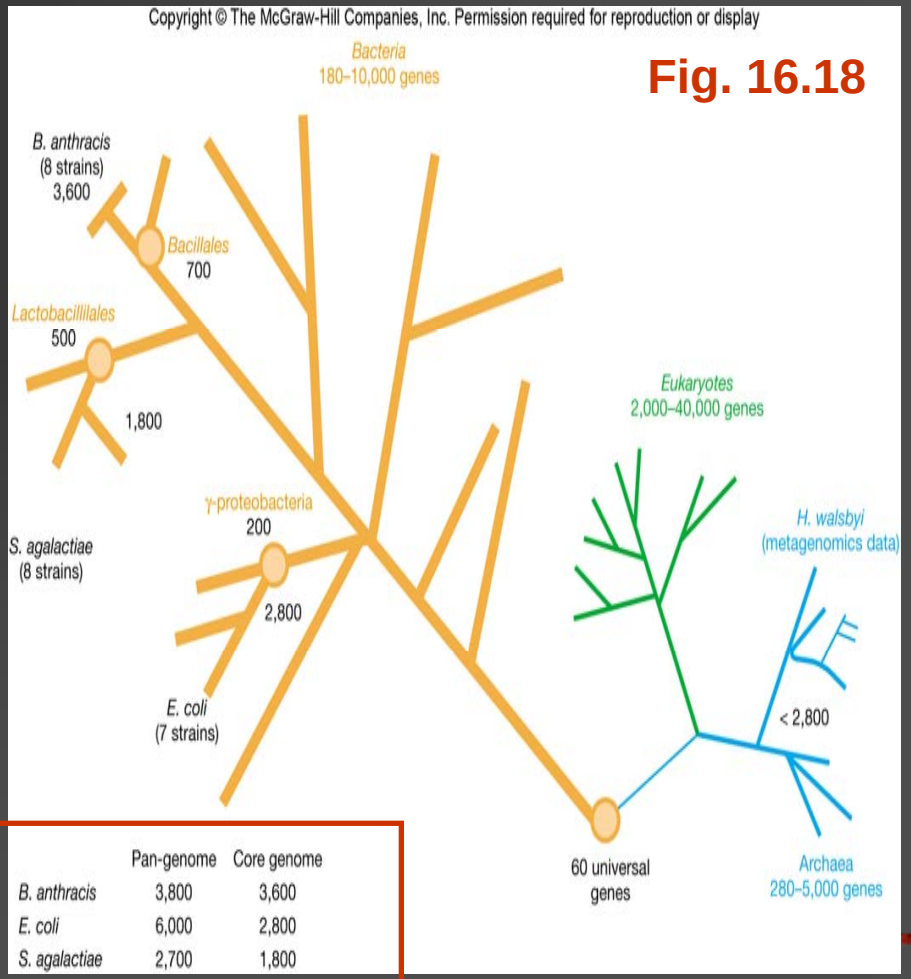


Comparative Genomics- *Eukarya* genome size



Core and Pan-genomes

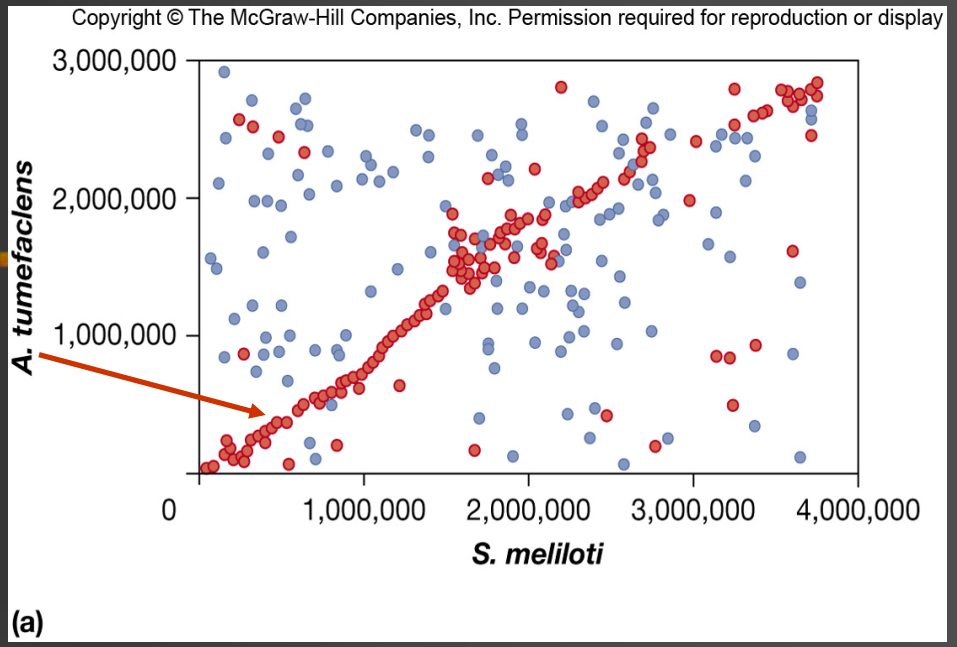
- Core genome- set of genes found in all members of a species (or monophyletic group)
- Pan genome- combination of all different genes found in all the different strains in a given species
  - **HGT** (horizontal gene transfer)
  - **Genomic island**
  - **Pathogenicity island**
    - **different GC contents**
    - **codon bias**



	Pan-genome	Core genome
B. anthracis	3,800	3,600
E. coli	6,000	2,800
S. agalactiae	2,700	1,800

Synteny

- Phylogenetic relationship analysis
- To examine how similar the organization of orthologous genes are in the compared genomes



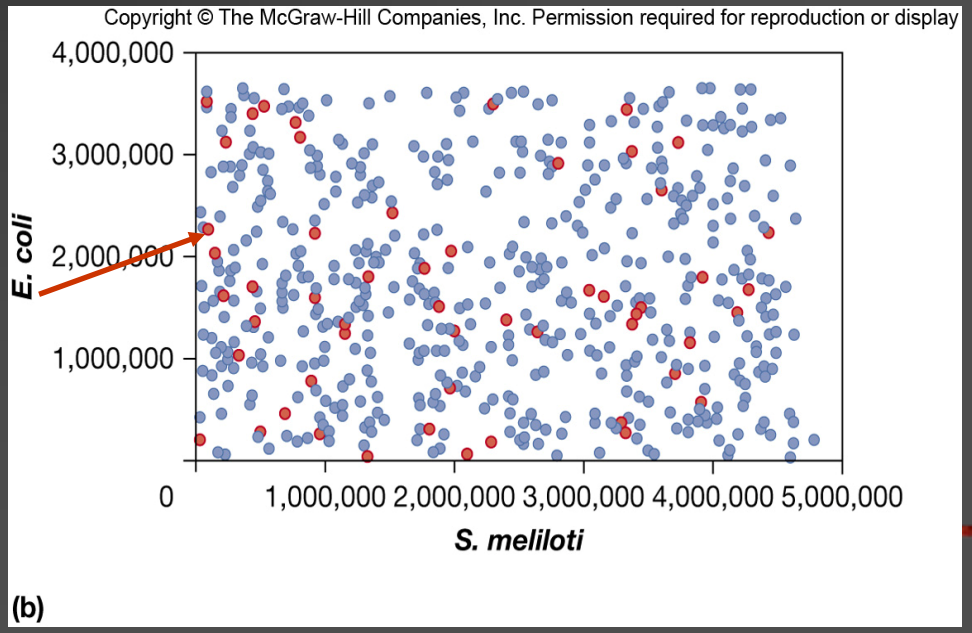
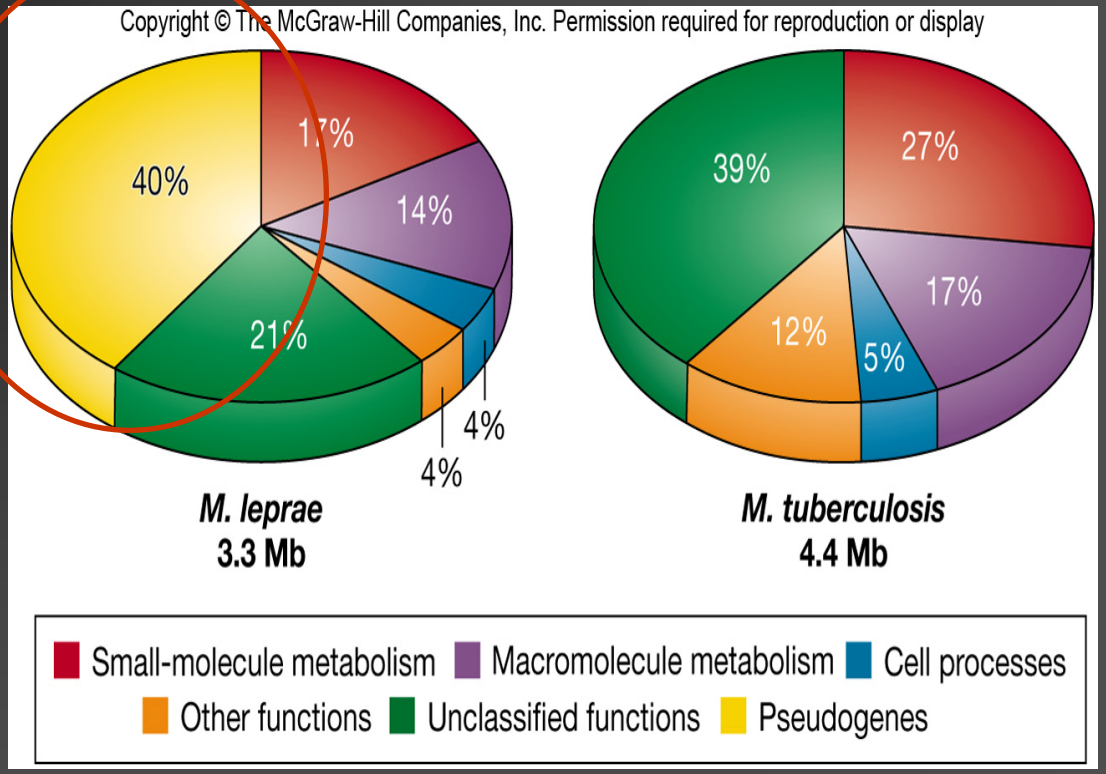


Fig. 16.19

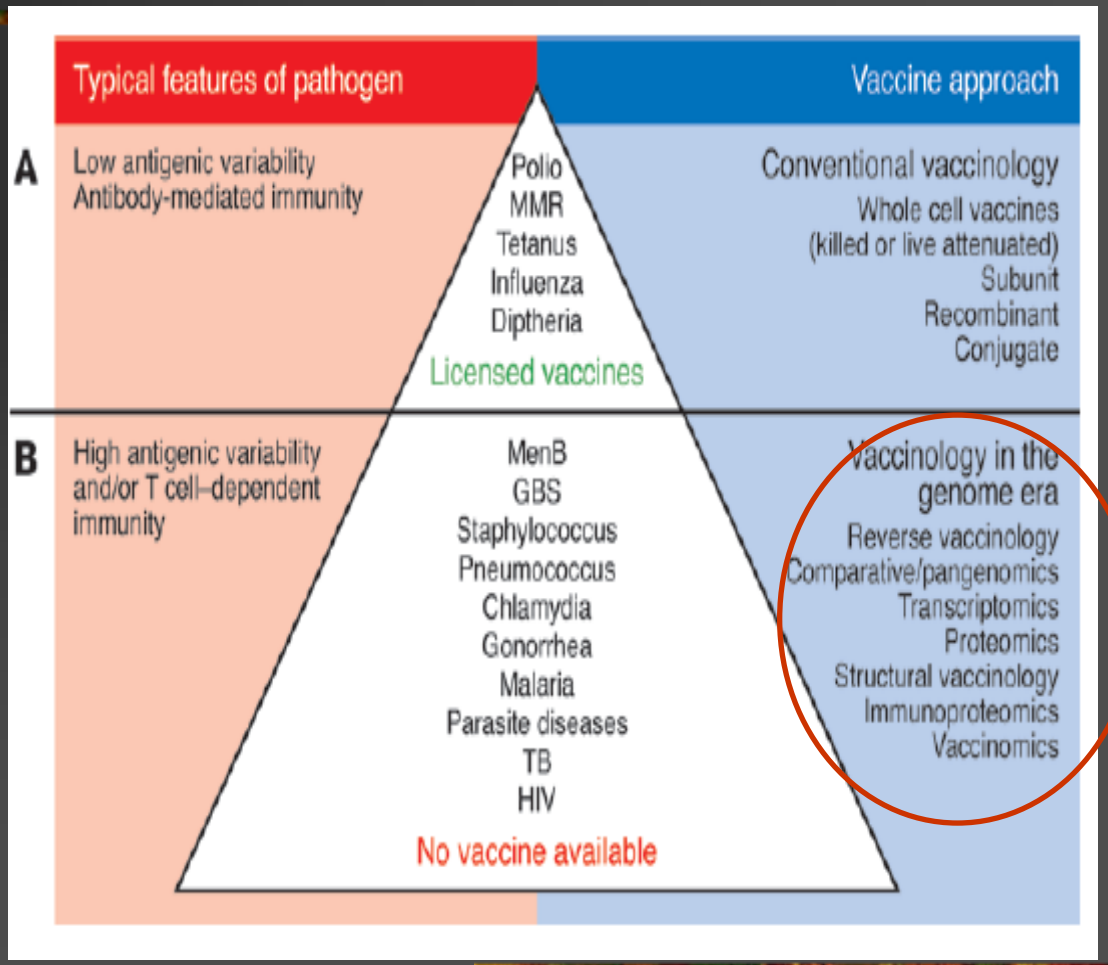
Comparative genomics analysis

- *M. leprae*- Leprosy
  - pseudogenes-degraded and nonfunctional genes
- *M. tuberculosis* (4.4 Mb) and *M. bovis* (4.3 Mb)- Tuberculosis
  - 99.5% sequence identity



Reverse Vaccinology

- Two types of reverse vaccinology
  - Single species genome is examined for vaccine targets
  - a pan-genomic approach



The Journal of Clinical Investigation (2009) 119: 2515-25.

Metagenomics- an environmental genomics

- used to learn more about the diversity and metabolic potential of microbial communities
  - shotgun sequencing DNA obtained directly from an environmental sample or series of related samples→ determine the presence and level of classes of genes
  - may serve to establish hypotheses concerning interactions between community members
  - being viewed as a baseline technology for understanding the ecology and evolution of microbial ecosystems
    - Phylochip

Construction and screening of genomic libraries directly from the environment

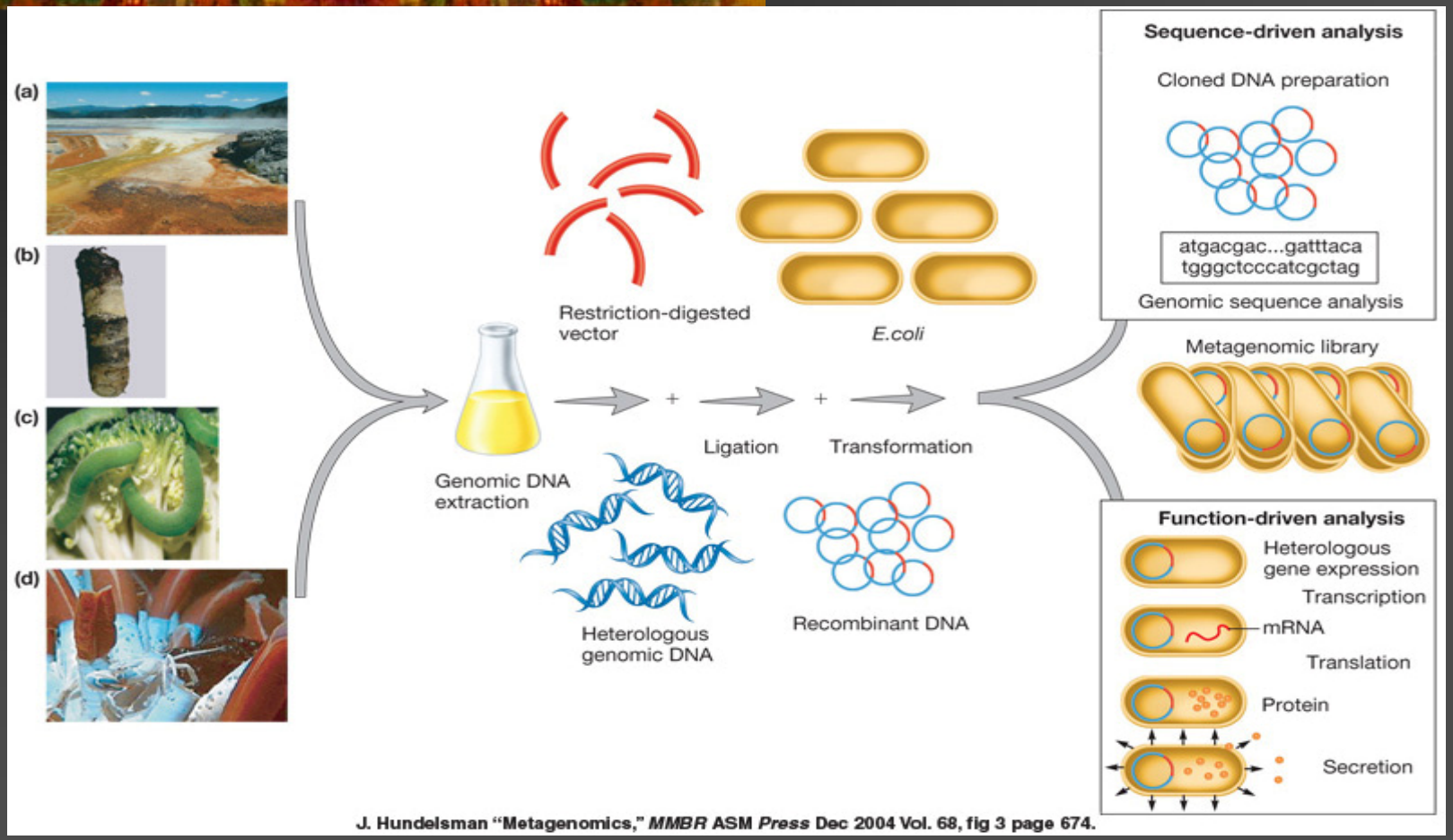
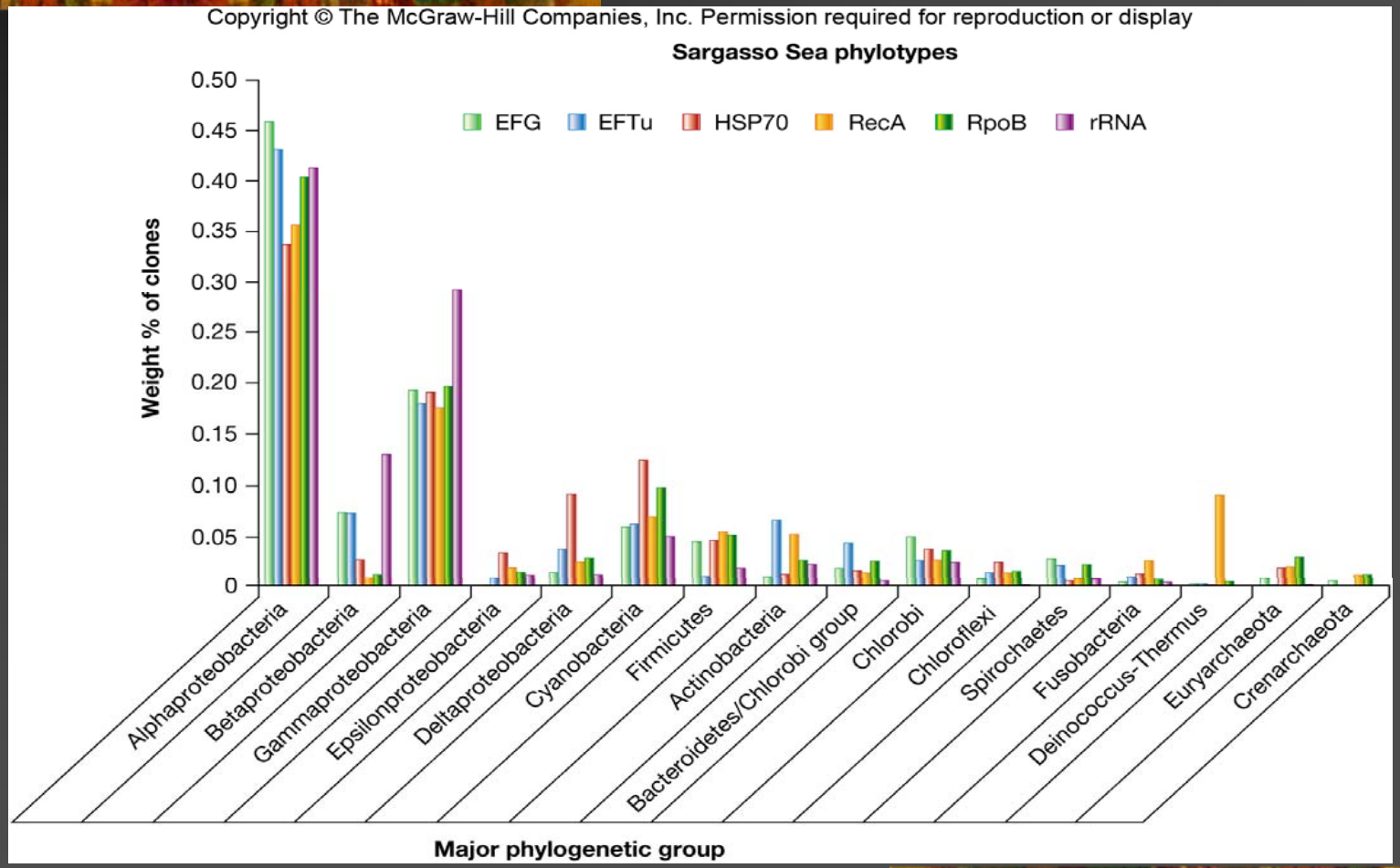


Figure 16.21

Microbiome- phylogenetic diversity of Sargasso sea microbes



Sargasso sea

- portion of the Atlantic ocean surrounds Bermuda
  - The Sargasso Sea is well known for the floating mats of Sargassum weed

The human microbiome project

- Launched by the US National Institutes of Health (NIH) as one of its major roadmap initiatives (2007, earmarking ~US$140 million)
  - to determine whether individuals share a core human microbiome
  - to understand whether changes in the human microbiome can be correlated with changes in human health
  - to develop the technological tools to support these goals
  - to address the ethical, legal and social implications of human microbiome research

Nature (2007) October 449:803-

Gut microbiome

- The human intestine is home to ~100 trillion microorganisms of at least 400 species
  - 10^{11}~10^{12} /ml bacteria in colon
  - Many gastrointestinal (GI) diseases are expected to be associated with disruption of host-bacterial interactions

PLOS PATHOGENS (2008)