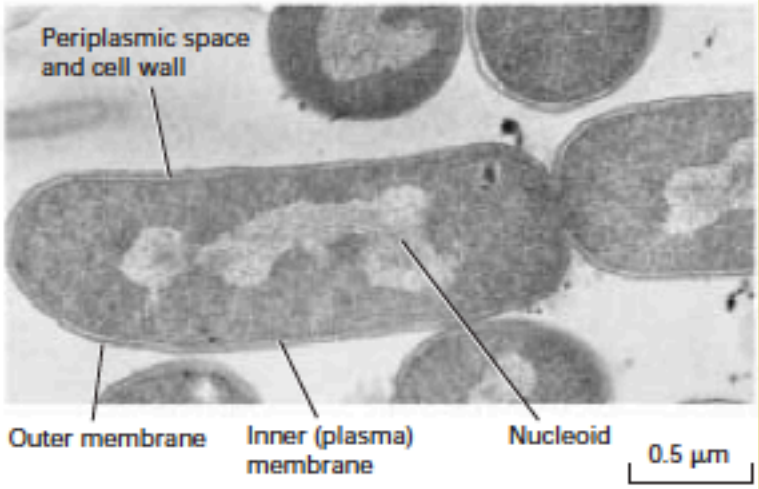
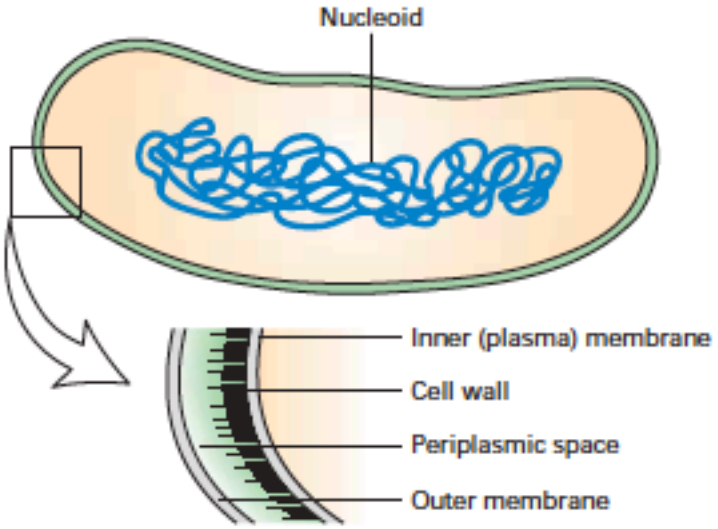
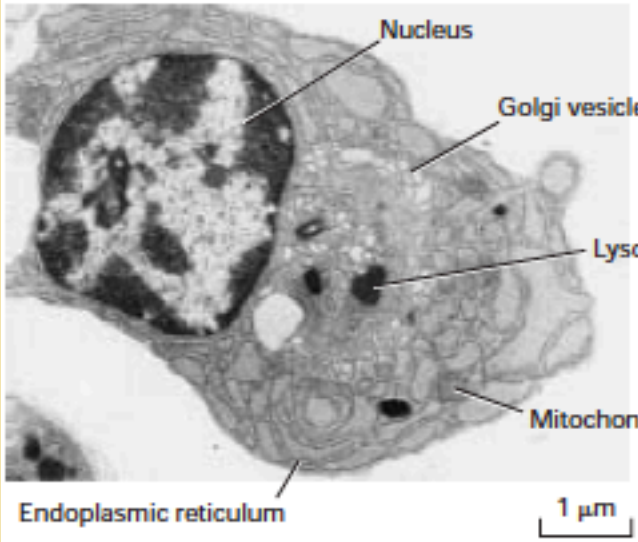
THE DIVERSITY OF GENOMES AND THE TREE OF LIFE

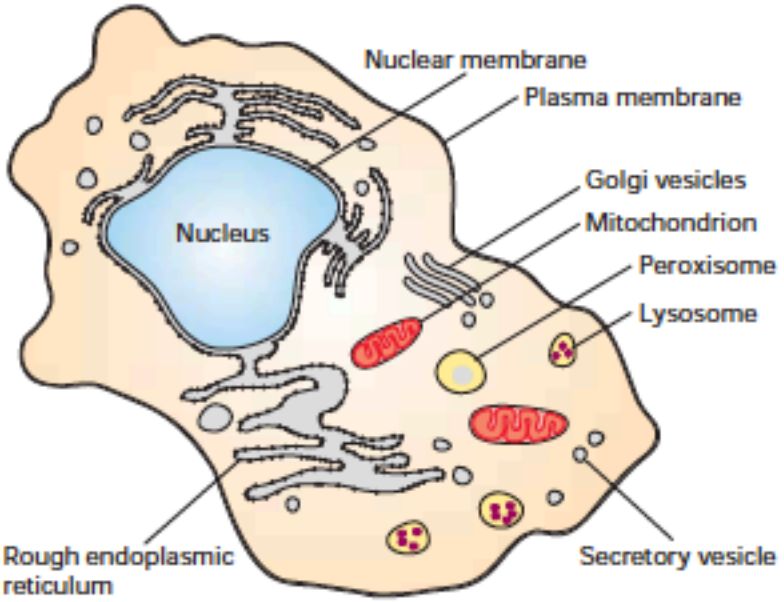
(a) Prokaryotic cell





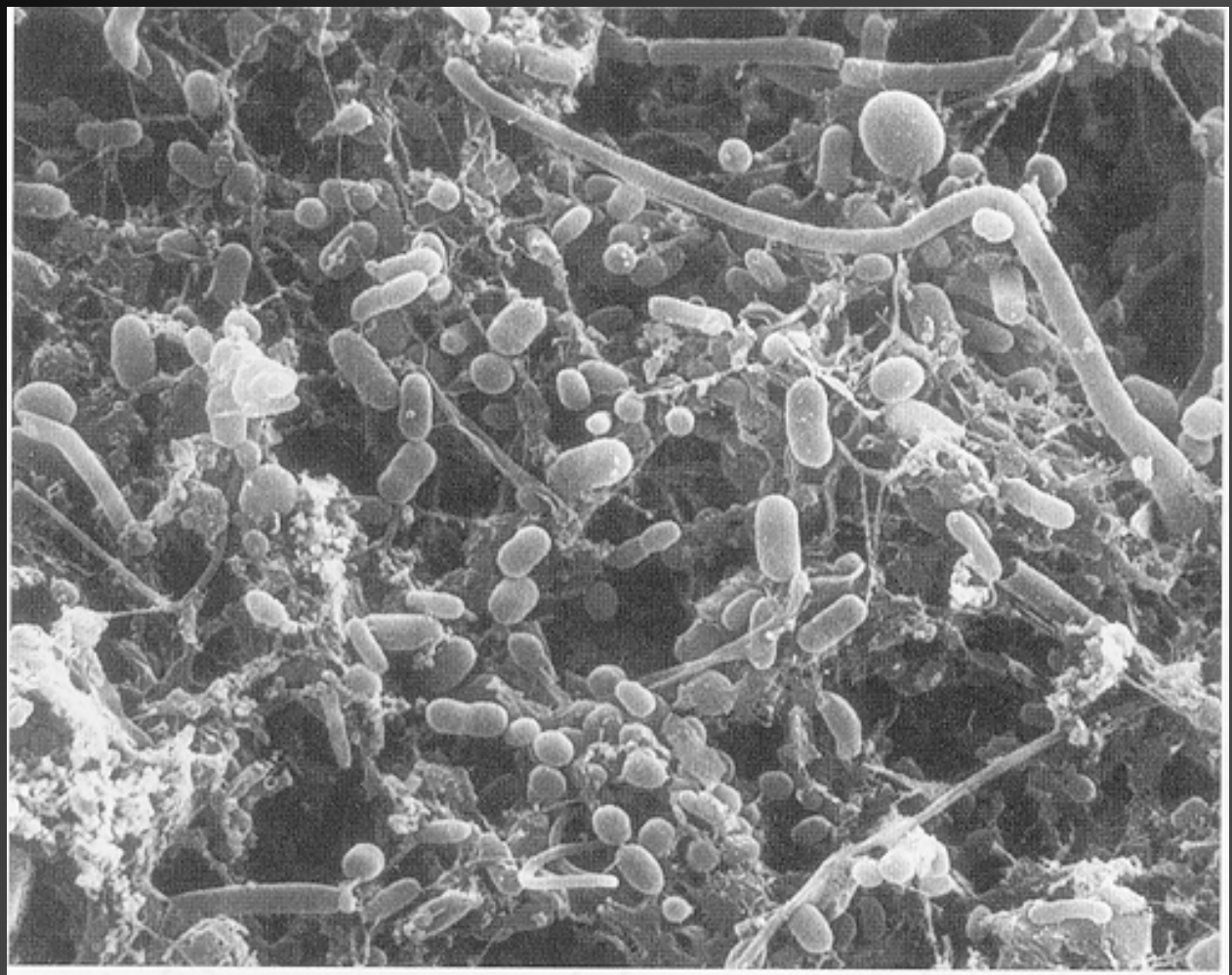
(b) Eukaryotic cell





Some Cells Fix Nitrogen and Carbon Dioxide for Others

Cells Can Be Powered by a Variety of Free-Energy Sources



25.9 Chemoautotrophs All of these archaea are chemoautotrophs that live near a hot-water vent along the Galapagos Rift in the eastern Pacific, where their fixing of carbon supports an entire community of organisms that thrives in total darkness.

The geology of a hot hydrothermal vent in the ocean floor

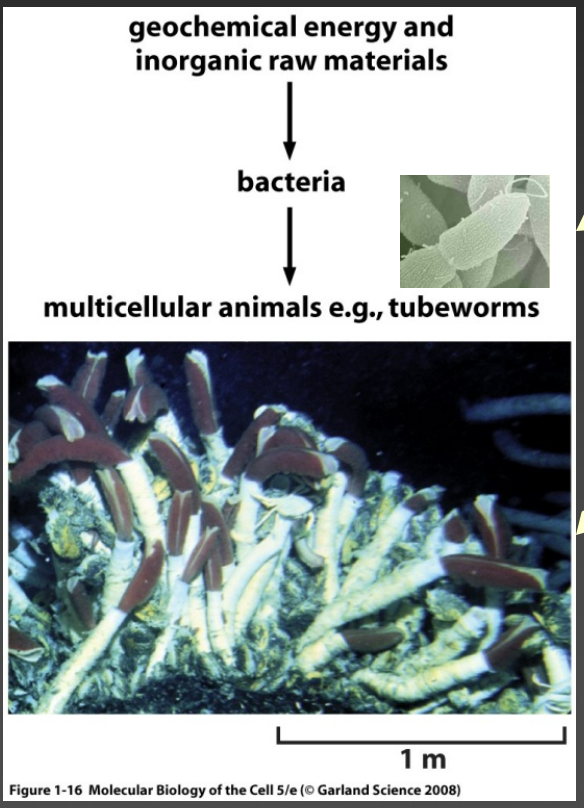


Figure 1-16 Molecular Biology of the Cell 5/e (© Garland Science 2008)

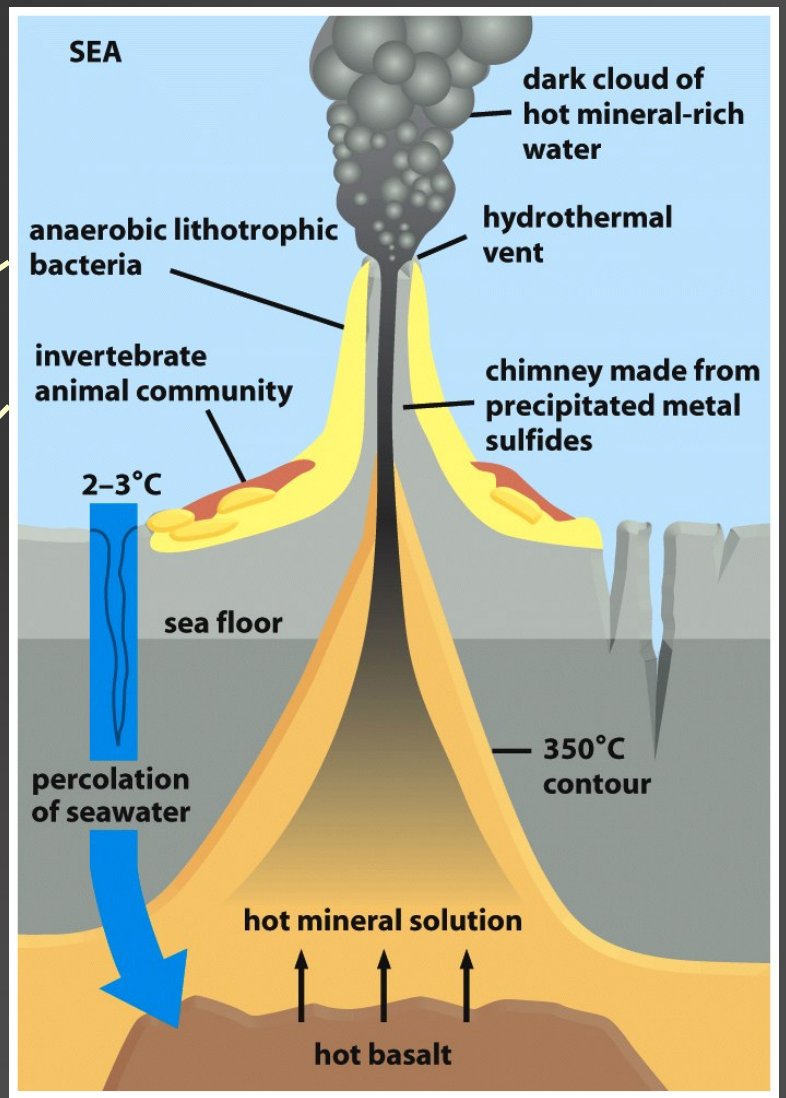


Figure 1-15 Molecular Biology of the Cell 5/e (© Garland Science 2008)

The phototrophic bacterium Anabaena cylindrica viewed in the light microscope.

The cells of this species form long, multicellular filaments.

Most of the cells (labeled V) perform photosynthesis, while others become specialized for nitrogen fixation (labeled H) or develop into resistant spores (labeled S).

A lithotrophic bacterium.

Beggiatoa, which lives in sulfurous environments, gets its energy by oxidizing H2S and can fix carbon even in the dark.

Note the yellow deposits of sulfur inside the cells.

The Greatest Biochemical Diversity Exists Among Prokaryotic Cells

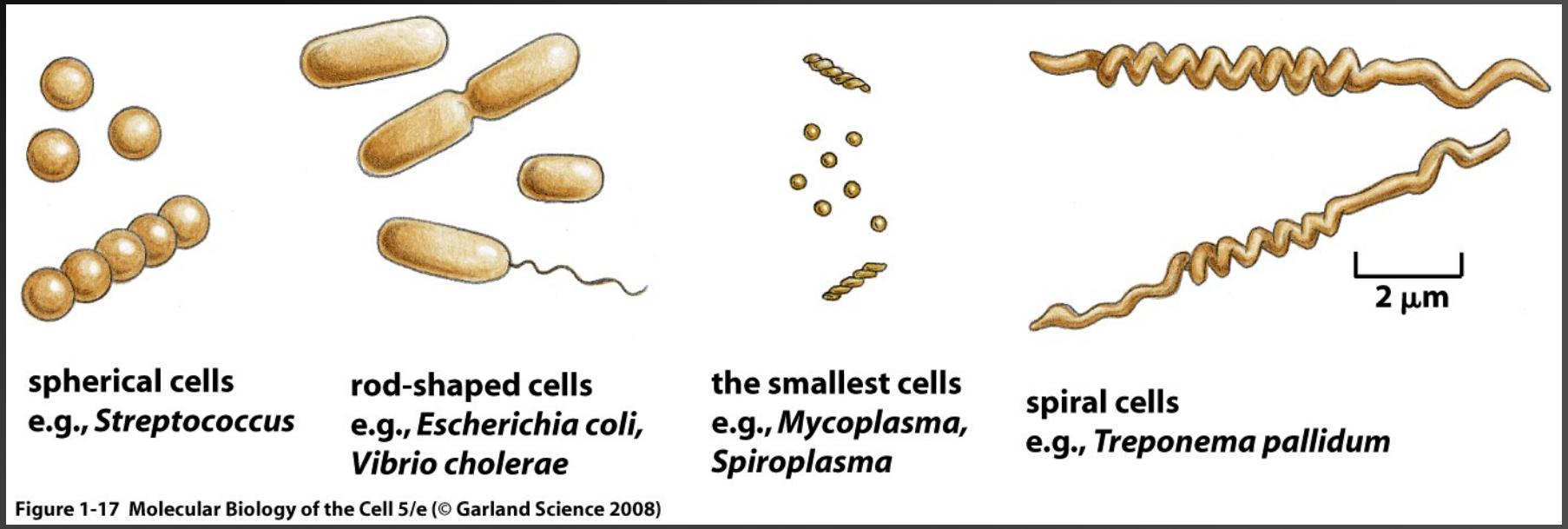


Figure 1-17 Molecular Biology of the Cell 5/e (© Garland Science 2008)

Most prokaryotic cells are small and simple in outward appearance. They live mostly as independent individuals or in loosely organized communities, rather than as multicellular organisms. They are typically spherical or rod-shaped and measure a few micrometers in linear dimension.

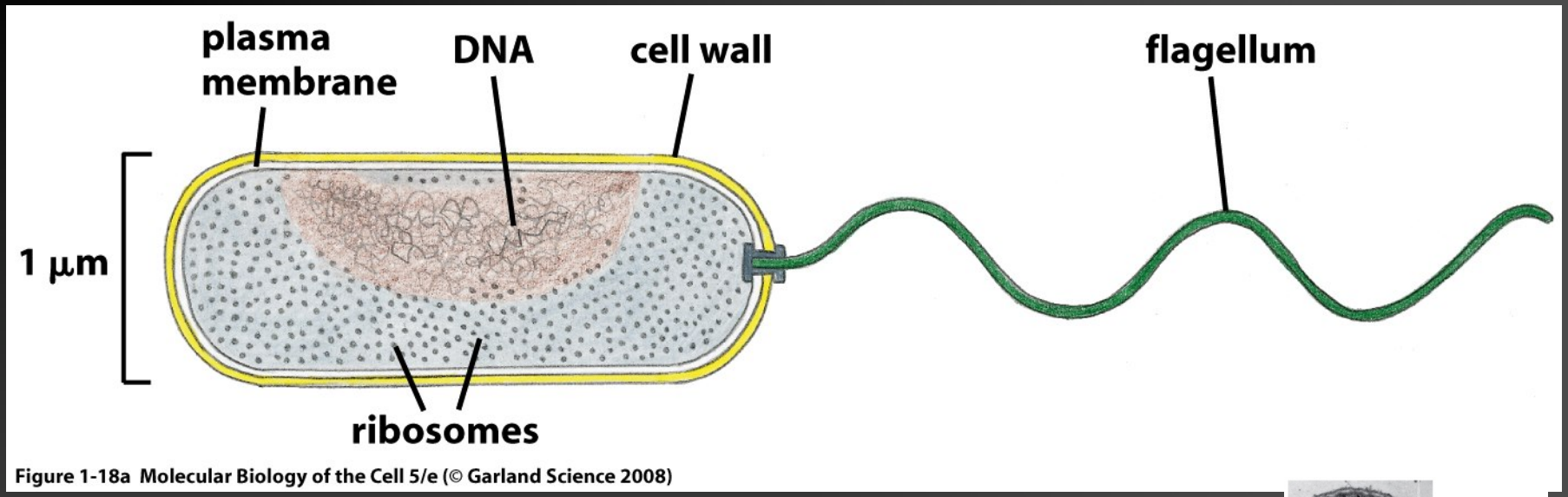


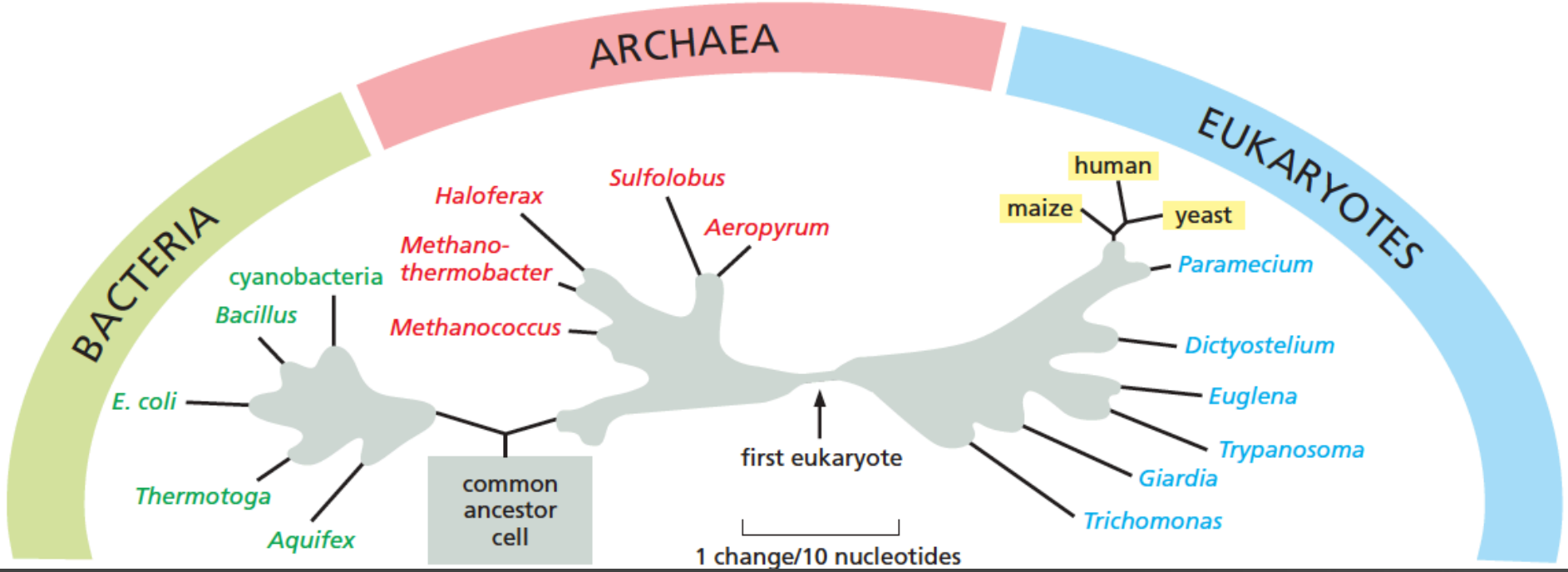
Figure 1-18a Molecular Biology of the Cell 5/e (© Garland Science 2008)

Vibrio cholerae

Figure 1-18b Molecular Biology of the Cell 5/e (© Garland Science 2008)

In the electron microscope, cell interior of bacteria appears as a matrix of varying texture without any discernible organized internal structure

The three major divisions (domains) of the living world



The tree shown here is based on comparisons of the nucleotide sequence of a ribosomal RNA (rRNA) subunit in the different species, and the distances in the diagram represent estimates of the numbers of evolutionary changes that have occurred in this molecule in each lineage

Some Genes Evolve Rapidly;

Others Are Highly Conserved

Small rRNA gene similarity comparison

Genetic information conserved since the days of the last common ancestor of all living things.

A part of the gene for the smaller of the two main rRNA components of the ribosome is shown. (The complete molecule is about 1500–1900 nucleotides long, depending on species.) Corresponding segments of nucleotide sequence from an archaean (Methanococcus jannaschii), a bacterium (Escherichia coli), and a eukaryote (Homo sapiens) are aligned.

Most Bacteria and Archaea Have 1000–6000 Genes

EUBACTERIA	info	habitat	X1000(kb) genome	genes
Mycoplasma genitalium	smallest genome of any known cell	human genital tract	580	468
Synechocystis sp.	photosynthetic, oxygen-generating (cyanobacterium)	lakes and streams	3573	3168
Escherichia coli	laboratory favorite	human gut	4639	4289
Helicobacter pylori	causes stomach ulcers and predisposes to stomach cancer	human stomach	1667	1590
Bacillus subtilis	bacterium	soil	4214	4099
Aquifex aeolicus	lithotrophic; lives at high temperatures	hydrothermal vents	1551	1544
Mycobacterium tuberculosis	causes tuberculosis	human tissues	4447	4402
Treponema pallidum	spirochaete; causes syphilis	human tissues	1138	1041
Rickettsia prowazekii	bacterium most closely related to mitochondria; causes typhus	lice and humans (intracellular parasite)	1111	834
Thermotoga maritima	organotrophic; lives at high temperatures	hydrothermal vents	1860	1877
ARCHAEA				
Methanococcus jannaschii	lithotrophic, anaerobic, methane-producing	hydrothermal vents	1664	1750
Archaeoglobus fulgidus	lithotrophic or organotrophic, anaerobic, sulfate-reducing	hydrothermal vents	2178	2493
Aeropyrum pernix	aerobic, organotrophic hot-steam vents	coastal volcanic	669	2620
EUCARYOTES				
Saccharomyces cerevisiae (budding yeast)	minimal model eucaryote	grape skins, beer	12,069	~6300
Arabidopsis thaliana (wall cress)	model organism for flowering plants	soil and air	~142,000	~26,000
Caenorhabditis elegans (nematode worm)	simple animal with perfectly predictable development	soil	~97,000	~19,000
Drosophila melanogaster (fruit fly)	key to the genetics of animal development	rotting fruit	~137,000	~14,000
Homo sapiens (human)	most intensively studied mammal	houses	~3,200,000	~30,000

New Genes Are Generated from Preexisting Genes

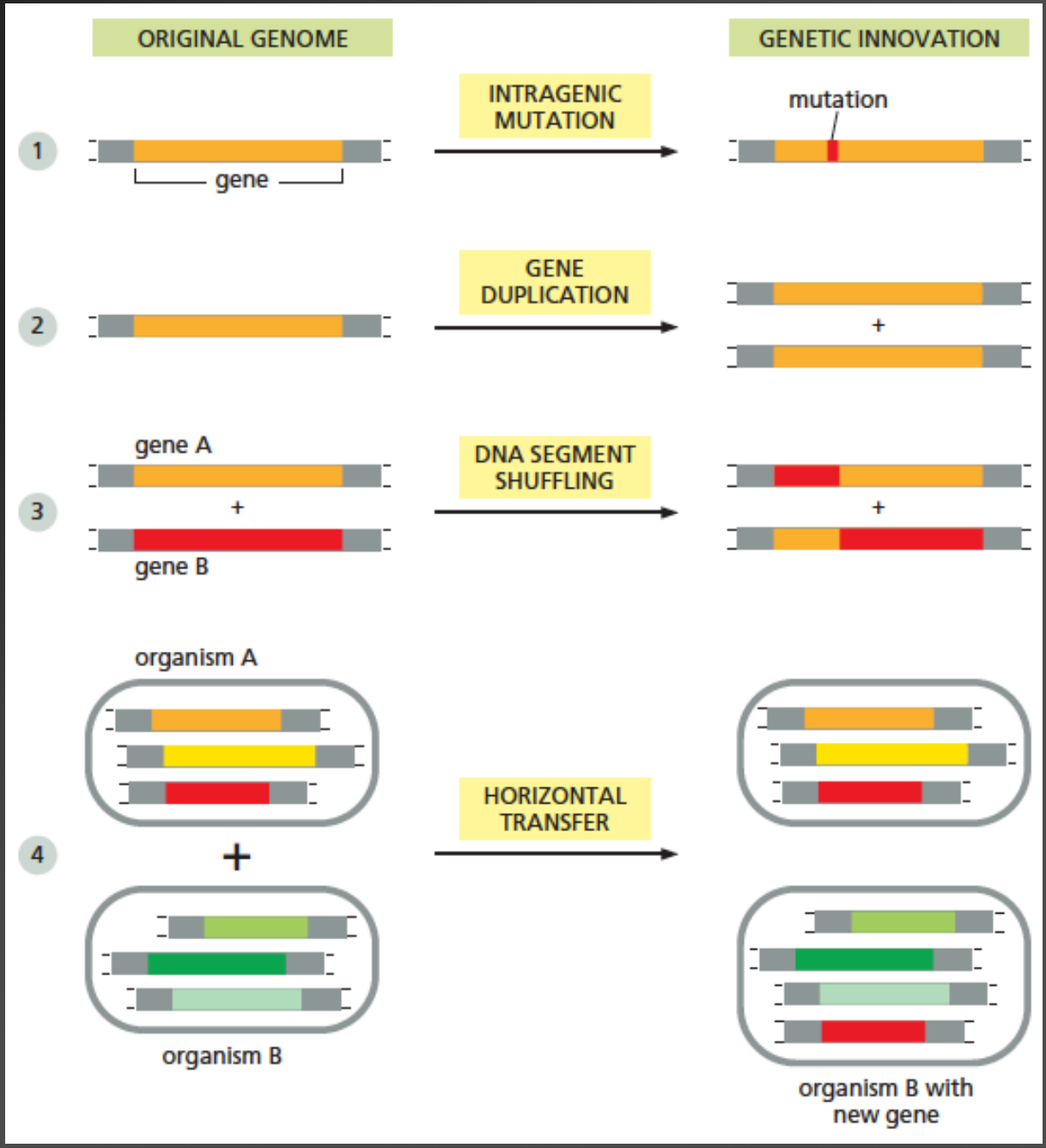
1. **Intragenic mutation:** an existing gene can be randomly modified by changes in its DNA sequence, through various types of error that occur mainly in the process of DNA replication.

2. **Gene duplication:** an existing gene can be accidentally duplicated so as to create a pair of initially identical genes within a single cell; these two genes may then diverge in the course of evolution.

3. **DNA segment shuffling:** two or more existing genes can break and rejoin to make a hybrid gene consisting of DNA segments that originally belonged to separate genes.

4. **Horizontal (intercellular) transfer:** a piece of DNA can be transferred from the genome of one cell to that of another—even to that of another species. This process is in contrast with the usual vertical transfer of genetic infor-mation from parent to progeny.

Four modes of genetic innovation and their effects on the DNA sequence of an organism.



Gene Duplications Give Rise to Families of Related Genes Within a Single Cell

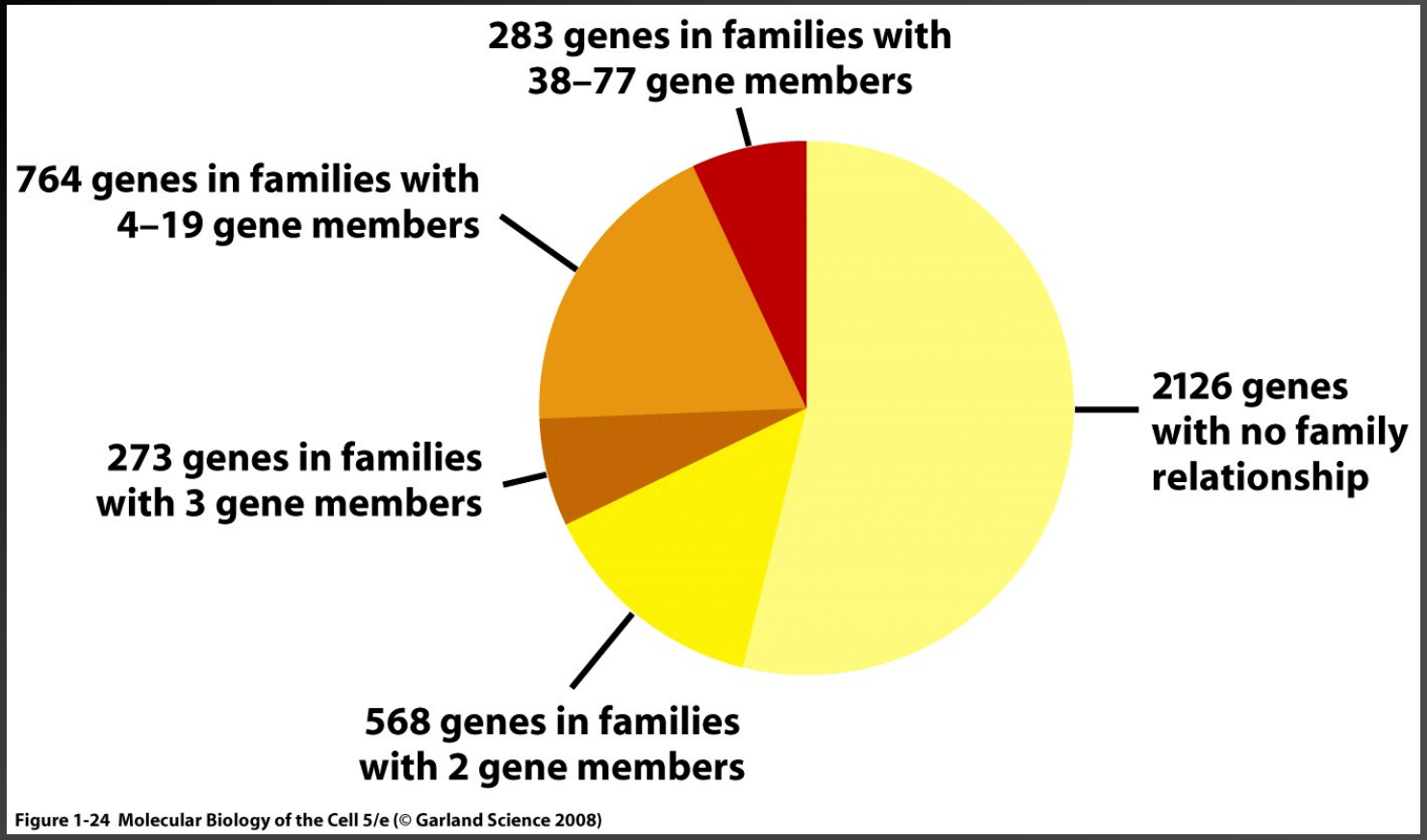
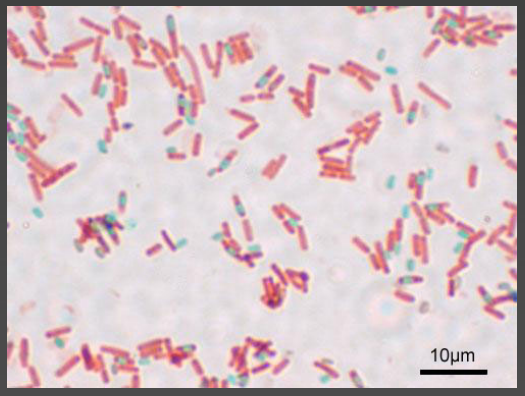
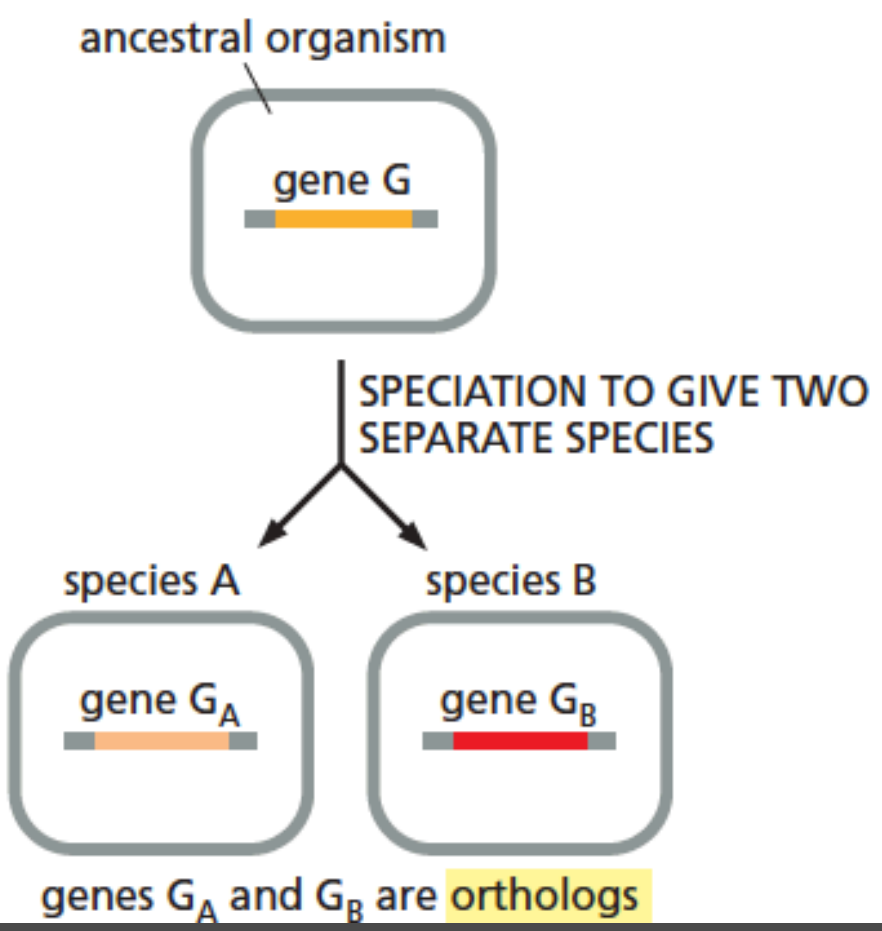


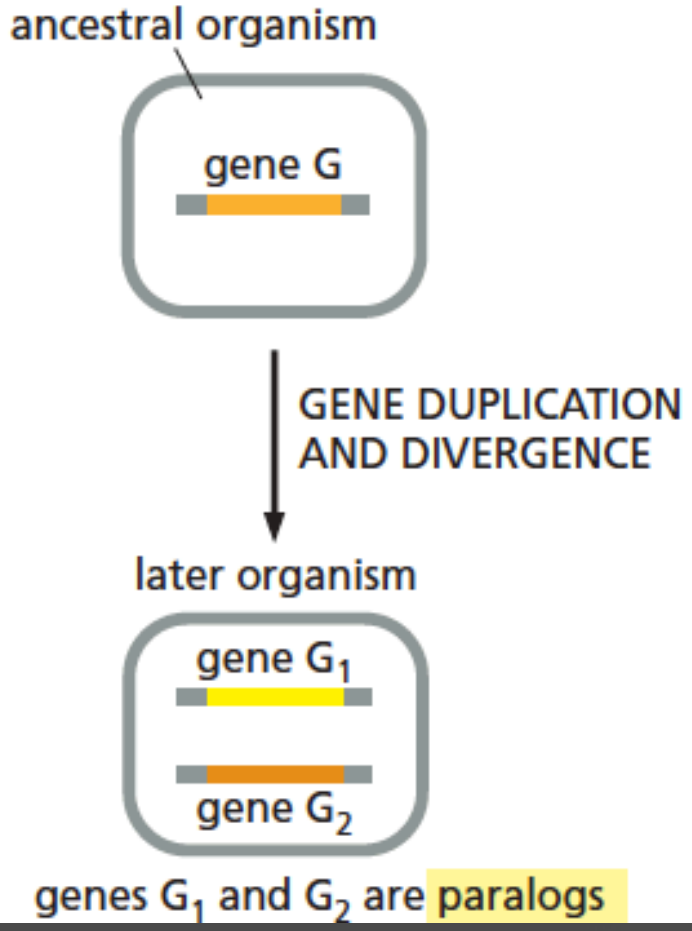
Figure 1-24 Molecular Biology of the Cell 5/e (© Garland Science 2008)



Families of evolutionarily related genes in the genome of Bacillus subtilis. The largest gene family in this bacterium consists of 77 genes coding for varieties of ABC transporters—a class of membrane transport proteins found in all three domains of the living world.

Ortholog vs Paralog Homology
Formation of new genes and gene families

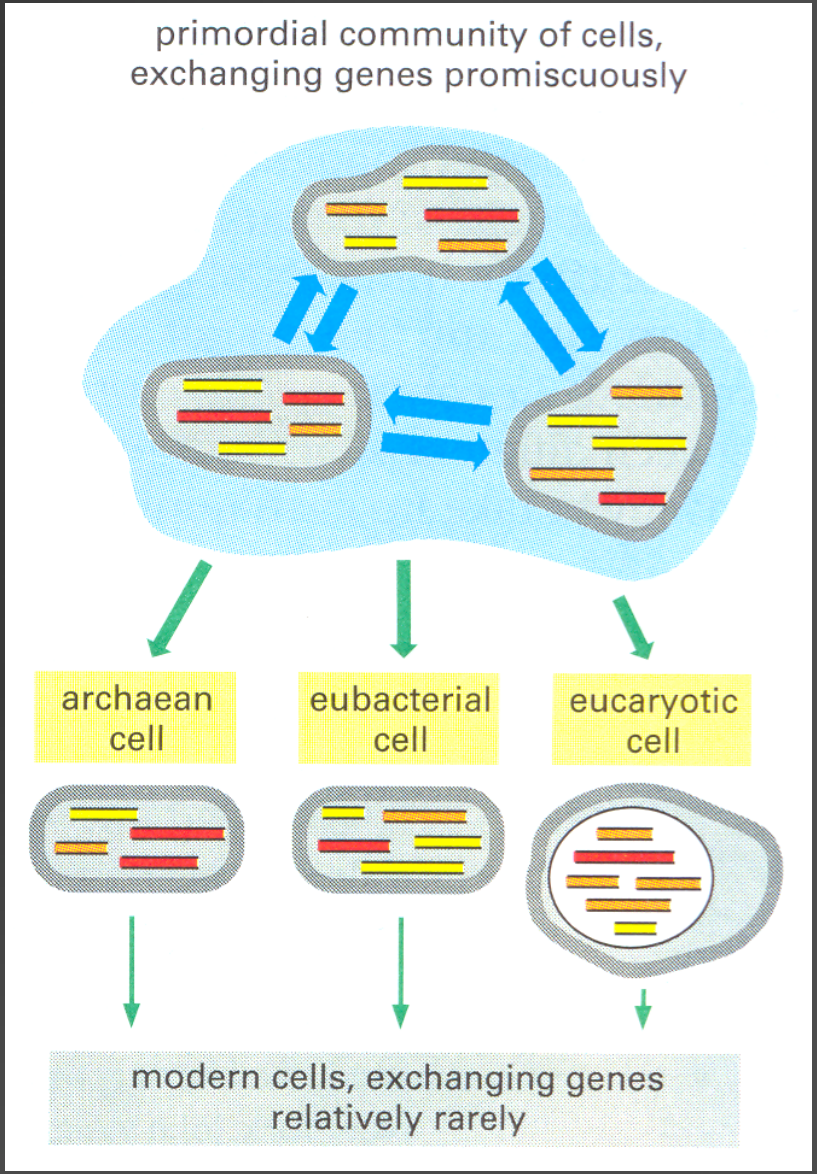




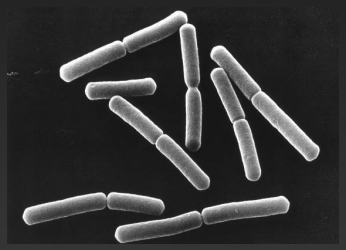
Horizontal gene transfers in early evolution.

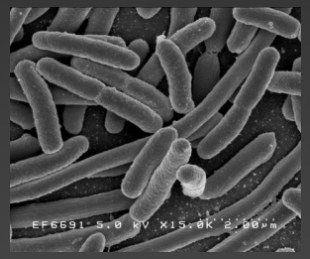
Early cells may have been less capable of maintaining their separate identities and exchanging genes more readily than now.

Three main lineages may have inherited different but overlapping subsets of gene from primordial community of cells.

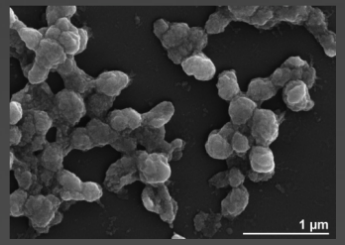


More Than 200 Gene Families Are Common to All Three Primary Branches of the Tree of Life

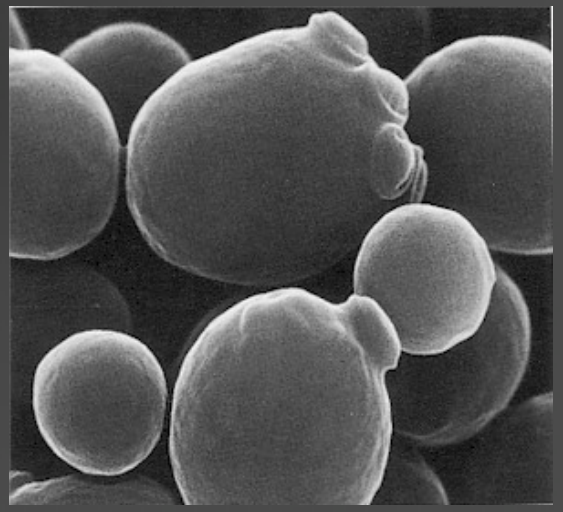




bakteria



archae



eukaryotes

The Number of Gene Families, Classified by Function, Common to All Three Domains of the Living World			
Information processing		**Metabolism**	
Translation	63	Energy production and conversion	19
Transcription	7	Carbohydrate transport and metabolism	16
Replication, recombination, and repair	13	Amino acid transport and metabolism	43
Cellular processes and signaling		Nucleotide transport and metabolism	15
Cell-cycle control, mitosis, and meiosis	2	Coenzyme transport and metabolism	22
Defense mechanisms	3	Lipid transport and metabolism	9
Signal transduction mechanisms	1	Inorganic ion transport and metabolism	8
Cell wall/membrane biogenesis	2	Secondary metabolite biosynthesis, transport, and catabolism	5
Intracellular trafficking and secretion	4	**Poorly characterized**	
Post-translational modification, protein turnover, chaperones	8	General biochemical function predicted; specific biological role unknown	24

For the purpose of this analysis, gene families are defined as “universal” if they are represented in the genomes of at least two diverse archaea (Archaeoglobus fulgidus and Aeropyrum pernix), two evolutionarily distant bacteria (Escherichia coli and Bacillus subtilis), and one eukaryote (yeast, Saccharomyces cerevisiae).

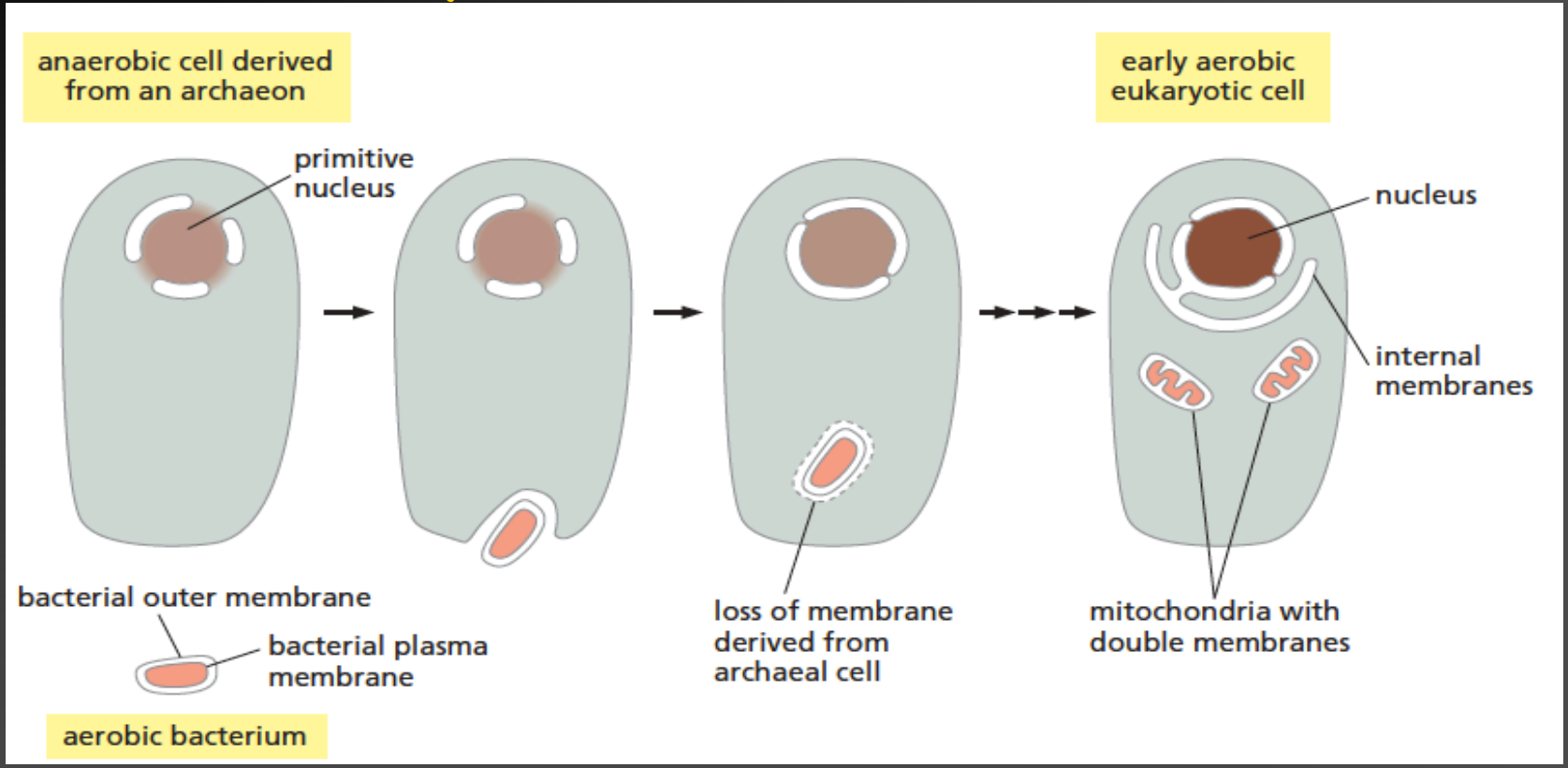
The Function of a Gene Can Often Be Deduced from Its Sequence

No.	Exon / Intron	Start	End	Start Phase	End Phase	Length	Sequence
	5' upstream sequence						..........gcaggagccagggctgggcataaaagtcagggcagagccatctattgctt
1	ENSPTRE00000032829	4,978,050	4,977,909	-	2	142	ACATTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAAACAGACACCATGGTGCACC TGACTCCTGAGGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTGAACGTGGATGAAG TTGGTGGTGAGGCCCTGGGCAG
	Intron 1-2	4,977,908	4,977,779			130	gttggtatcaaggttacaagacagg.....tattgggctattttcccacccttag
2	ENSPTRE00000032833	4,977,778	4,977,556	2	0	223	GCTGCTGGTGGTCTACCCTTGGACCCAGAGGTTCTTTGAGTCCTTTGGGGATCTGTCCAC TCCTGATGCTGTTATGGGCAACCCTAAGGTGAAGGCTCATGGCAAGAAAGTGCTCGGTGC CTTTAGTGATGGCCTGGCTCACCTGGACAACCTCAAGGGCACCTTTGCCACACTGAGTGA GCTGCACTGTGACAAGCTGCACGTGGATCCTGAGAACTTCAGG
	Intron 2-3	4,977,555	4,976,706			850	gtgagtctatgggacccttgatgtt.....catacctcttgtcttcctcccacag
3	ENSPTRE00000405718	4,976,705	4,976,444	0	-	262	CTCCTGGGCAACGTGCTGGTCTGTGTGCTGGCCCATCACTTTGGCAAAGAATTCACCCCA CCAGTGCAGGCTGCCTATCAGAAAGTGGTGGCTGGTGTGGCTAATGCCCTGGCCCACAAG TATCACTAAGCTCGCTTTCTTGCTGTCCAATTTCTATTAAAGGTTCCTTTGTTCCCTAAG TCCAACTACTAAACTGGGGGATATTATGAAGGGCCTTGAGCATCTGGATTCTGCCTAATA AAAAACATTTATTTTCATTGCA
	3' downstream sequence						atggtgtatttaaattatttctgaatattttactaaaaagggaatgtggg.....

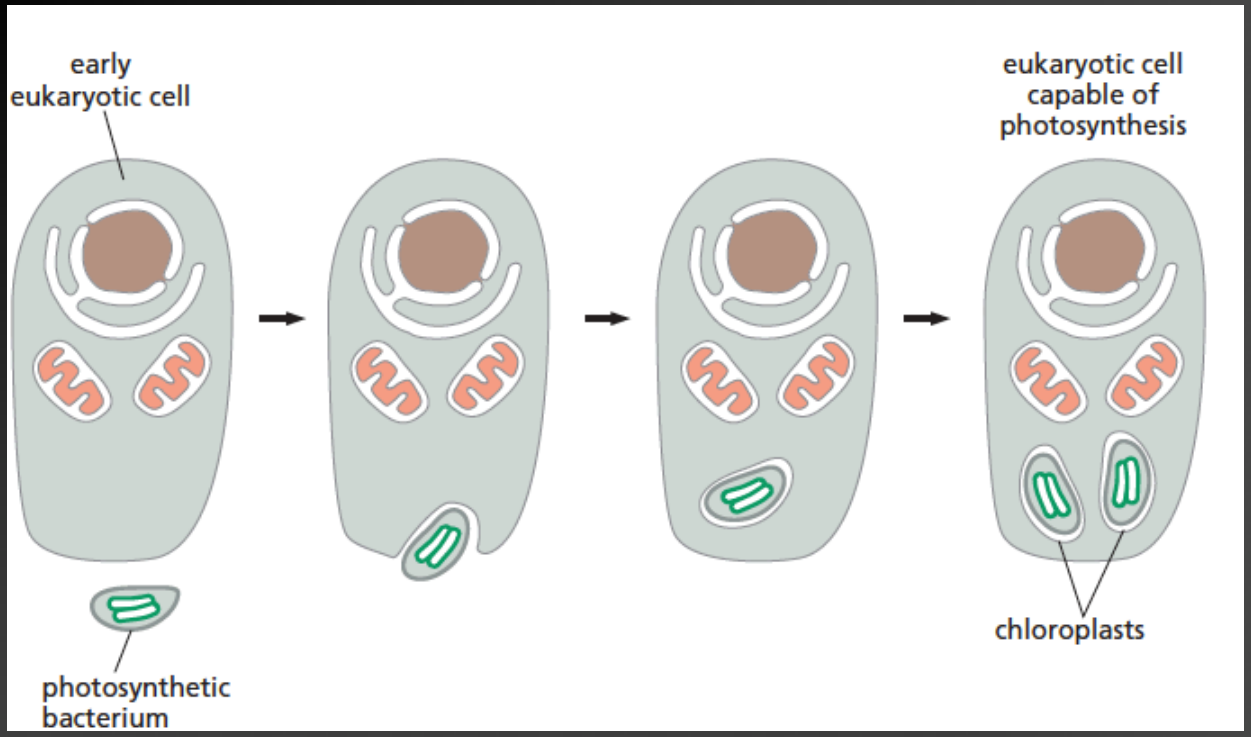
Hb beta
Gene cluster

?

Eukaryotes Have Hybrid Genomes



The origin of mitochondria.
An ancestral anaerobic predator cell (an archaeon) is thought to have engulfed the bacterial ancestor of mitochondria, initiating a symbiotic relationship. Clear evidence of a dual bacterial and archaeal inheritance can be discerned today in the genomes of all eukaryotes.



The origin of chloroplasts.
An early eukaryotic cell, already possessing mitochondria, engulfed a photosynthetic bacterium (a cyanobacterium) and retained it in symbiosis. Present-day chloroplasts are thought to trace their ancestry back to a single species of cyanobacterium that was adopted as an internal symbiont (an endosymbiont) over a billion years ago.

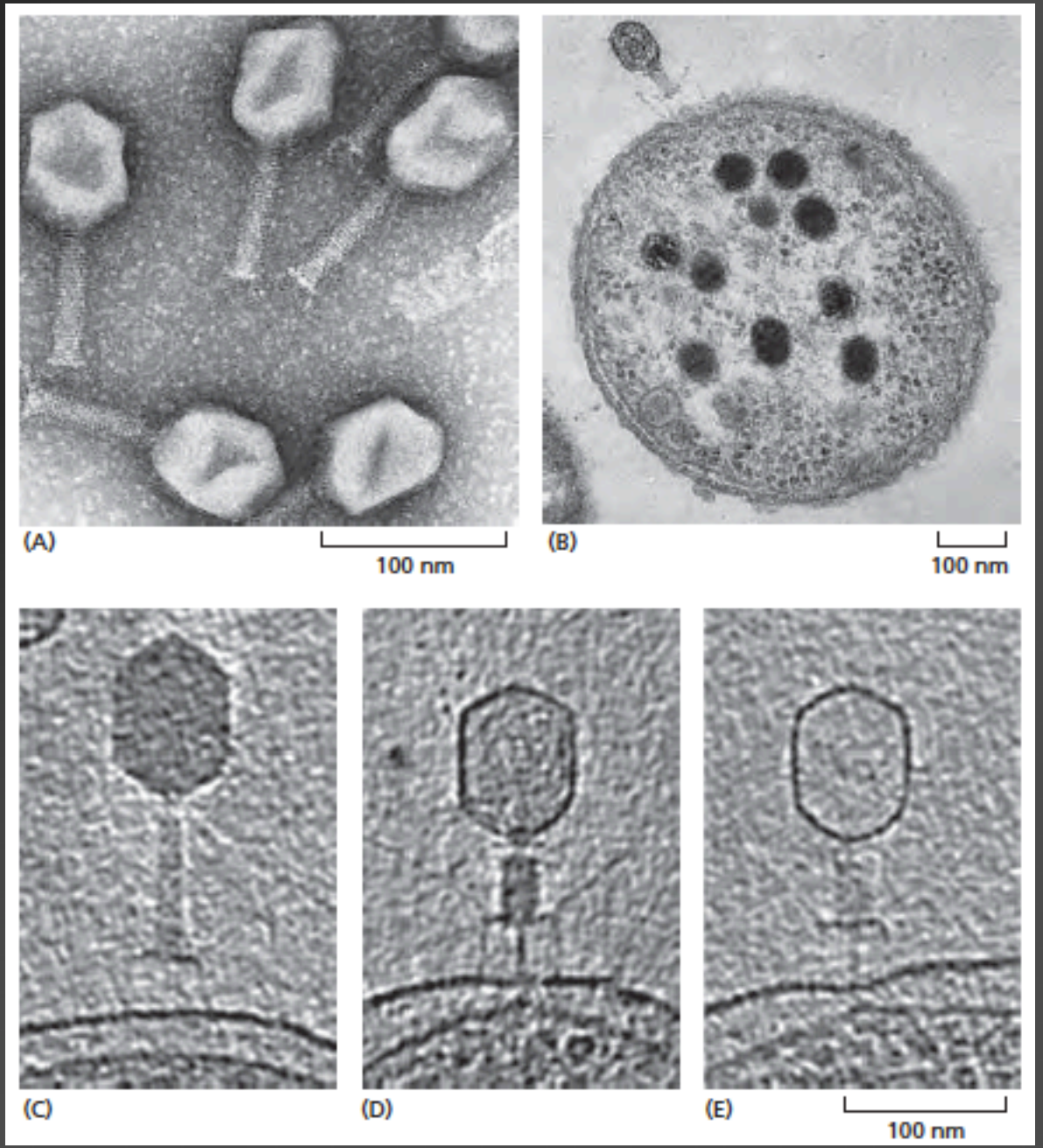
Genes Can Be Transferred Between Organisms, Both in the Laboratory and in Nature

The viral transfer of DNA into a cell.

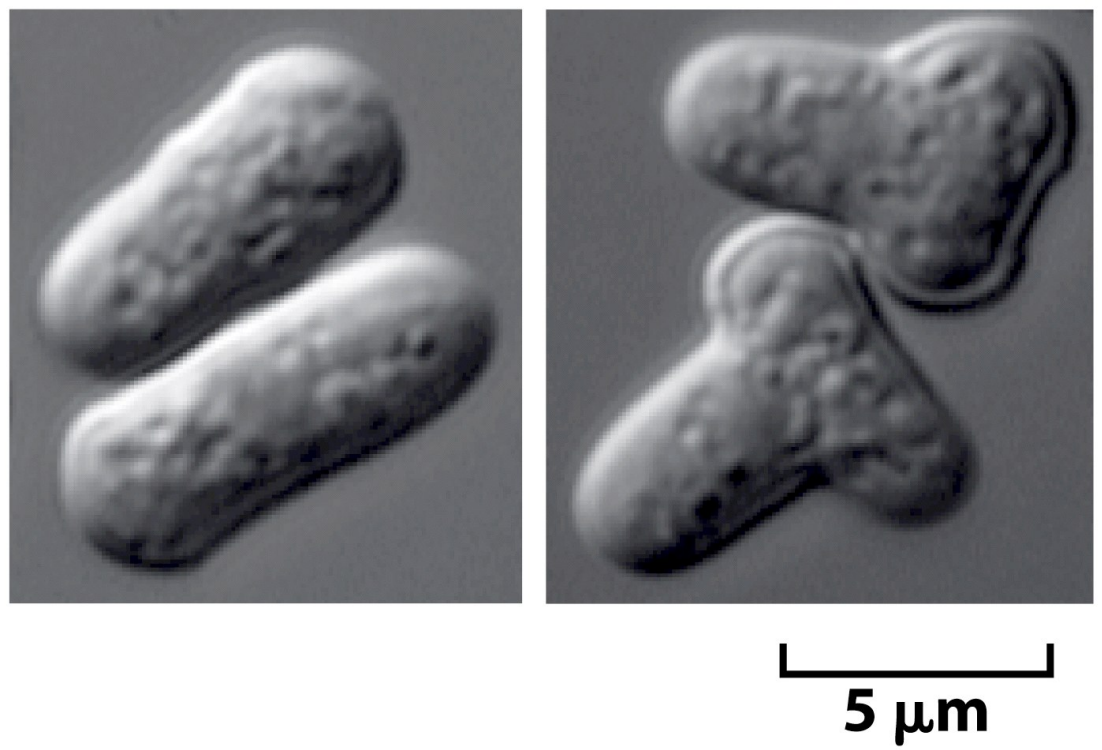
(A) An electron micrograph of particles of a bacterial virus, the T4 bacteriophage.

(B) A cross section of an E. coli bacterium with a T4 bacteriophage latched onto its surface.

(C-E) The process of DNA injection into the bacterium

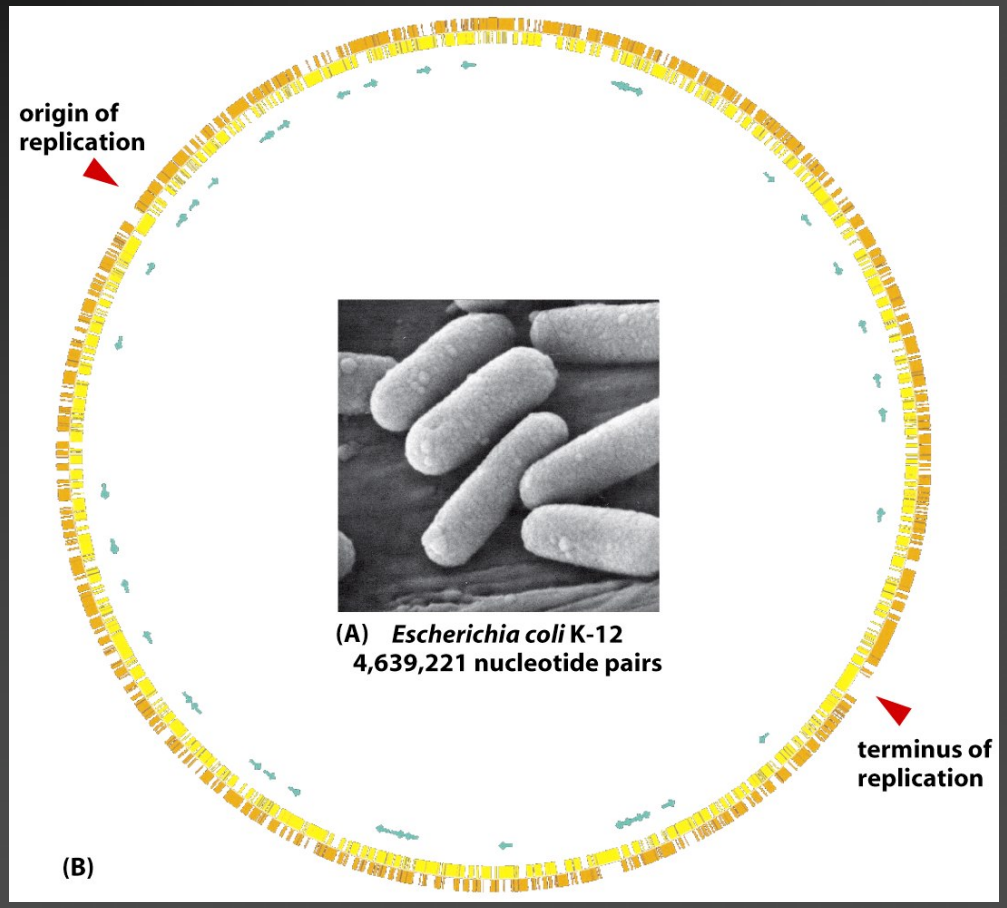


Mutations Reveal the Functions of Genes



A mutant phenotype reflecting the function of a gene. A normal yeast (of the species Schizosaccharomyces pombe) is compared with a mutant in which a change in a single gene has converted the cell from a cigar shape (left) to a T shape(right).

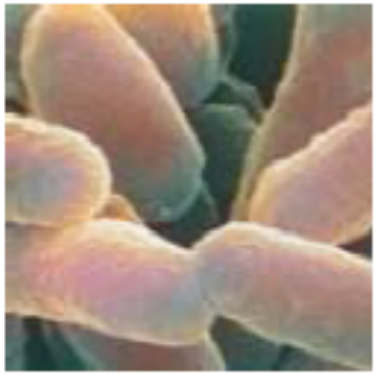
Molecular Biology Began with a Spotlight on E. coli



MODEL ORGANISMS

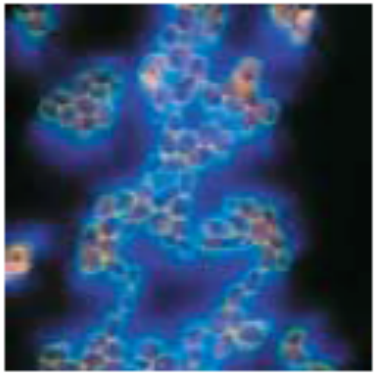
Viruses

Proteins involved in DNA, RNA, protein synthesis
Gene regulation
Cancer and control of cell proliferation
Transport of proteins and organelles inside cells
Infection and immunity
Possible gene therapy approaches



Bacteria

Proteins involved in DNA, RNA, protein synthesis, metabolism
Gene regulation
Targets for new antibiotics
Cell cycle
Signaling



Yeast (*Saccharomyces cerevisiae*)

Control of cell cycle and cell division
Protein secretion and membrane biogenesis
Function of the cytoskeleton
Cell differentiation
Aging
Gene regulation and chromosome structure

Roundworm (*Caenorhabditis elegans*)

Development of the body plan
Cell lineage
Formation and function of the nervous system
Control of programmed cell death
Cell proliferation and cancer genes
Aging
Behavior
Gene regulation and chromosome structure

Fruit fly (*Drosophila melanogaster*)

Development of the body plan
Generation of differentiated cell lineages
Formation of the nervous system, heart, and musculature
Programmed cell death
Genetic control of behavior
Cancer genes and control of cell proliferation
Control of cell polarization
Effects of drugs, alcohol, pesticides

Zebrafish

Development of vertebrate body tissues
Formation and function of brain and nervous system
Birth defects
Cancer

Mice, including cultured cells

Development of body tissues
Function of mammalian immune system
Formation and function of brain and nervous system
Models of cancers and other human diseases
Gene regulation and inheritance
Infectious disease

Plant (*Arabidopsis thaliana*)

Development and patterning of tissues
Genetics of cell biology
Agricultural applications
Physiology
Gene regulation
Immunity
Infectious disease

TABLE 1–2 Some Model Organisms and Their Genomes

Organism	Genome size* (nucleotide pairs)	Approximate number of genes
Escherichia coli (bacterium)	4.6×10^6	4,300
Saccharomyces cerevisiae (yeast)	12.1×10^6	6,600
Caenorhabditis elegans (roundworm)	100×10^6	20,222
Arabidopsis thaliana (plant)	119×10^6	25,000
Drosophila melanogaster (fruit fly)	137×10^6	17,660
Danio rerio (zebrafish)	1.4×10^9	25,592
Mus musculus (mouse)	2.7×10^9	22,619
Homo sapiens (human)	3.1×10^9	20,418

*Genome size includes an estimate for the amount of highly repeated DNA sequence not in genome databases.