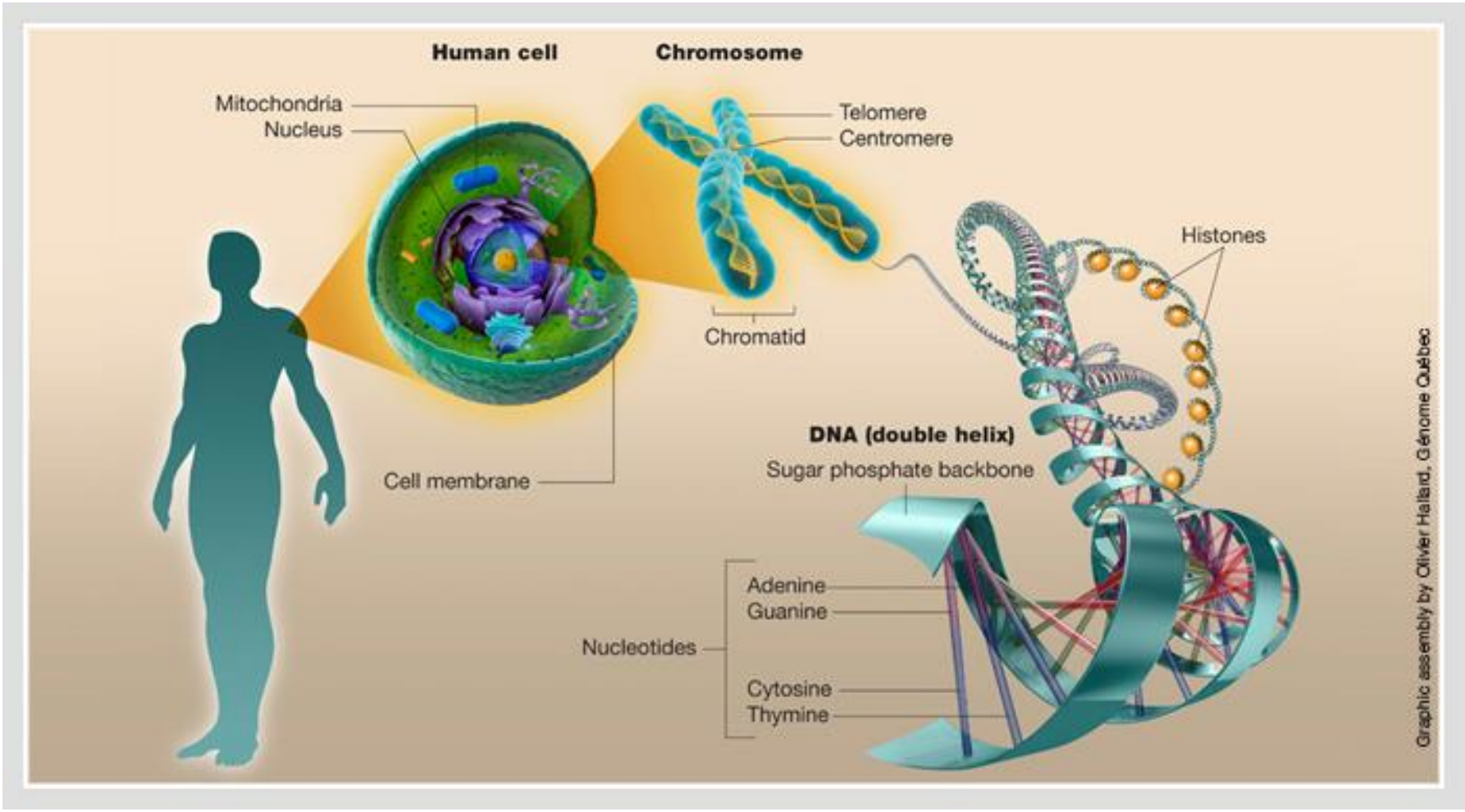
Cells and Genomes



Cells and Genomes

- The surface of our planet is populated by living things—curious, intricately organized chemical factories that take in matter from their surroundings and use these raw materials to generate copies of themselves.

- The living organisms appear extraordinarily diverse.

- Our ancestors, knowing nothing of cells or DNA, saw that all these things had something in common and they called that something “life”.

- We can now see that all living things are made of cells, and that these units of living matter all share the same machinery for their most basic functions.

- Living things, though infinitely varied when viewed from the outside, are fundamentally similar inside.

- It is estimated that there are more than 10 million—perhaps 100 million—living species on Earth today.

- Each species is different, and each reproduces itself faithfully, yielding progeny that belong to the same species: the parent organism hands down information specifying, in extraordinary detail, the characteristics that the offspring shall have.

- This phenomenon of heredity is central to the definition of life: it distinguishes life from other processes, such as the growth of a crystal, or the burning of a candle.

- Like the candle flame, the living organism consumes free energy to create and maintain its organization; but the free energy drives a hugely complex system of chemical processes that is specified by the hereditary information.

- All cells store their hereditary information in the same linear chemical code (DNA).

- Computers have made us familiar with the concept of information as a measurable quantity—a million bytes (to record a few hundred pages of text or an image from a digital camera), 600 million for the music on a CD, and so on.

- They have also made us well aware that the same information can be recorded in many different physical forms.

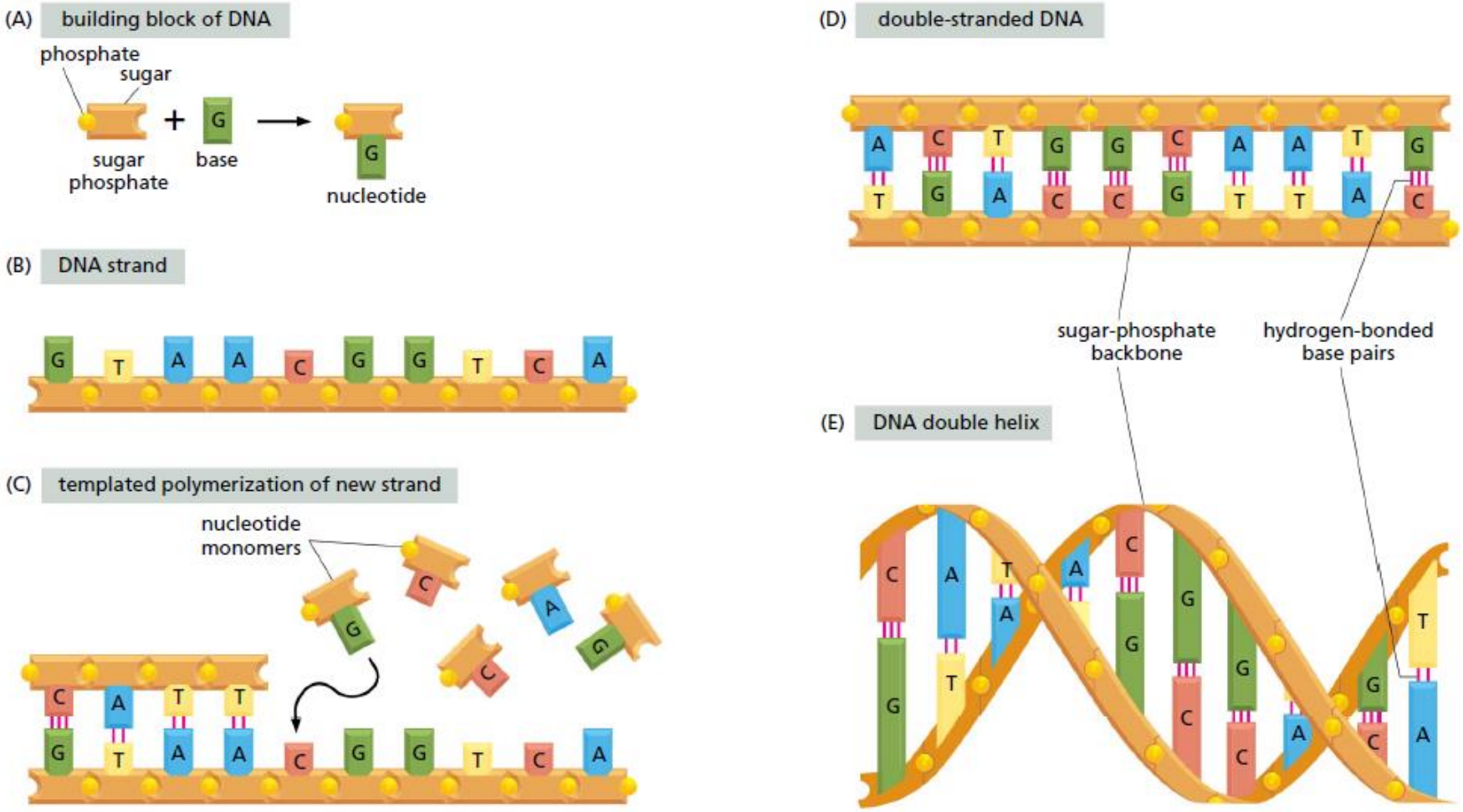
- Living cells, like computers, deal in information, and it is estimated that they have been evolving and diversifying for over 3.5 billion years.

- It is scarcely to be expected that they should all store their information in the same form, or that the archives of one type of cell should be readable by the information- handling machinery of another. And yet it is so.

- All living cells on Earth, without any known exception, store their hereditary information in the form of double-stranded molecules of DNA—long unbranched paired polymer chains, formed always of the same four types of monomers.

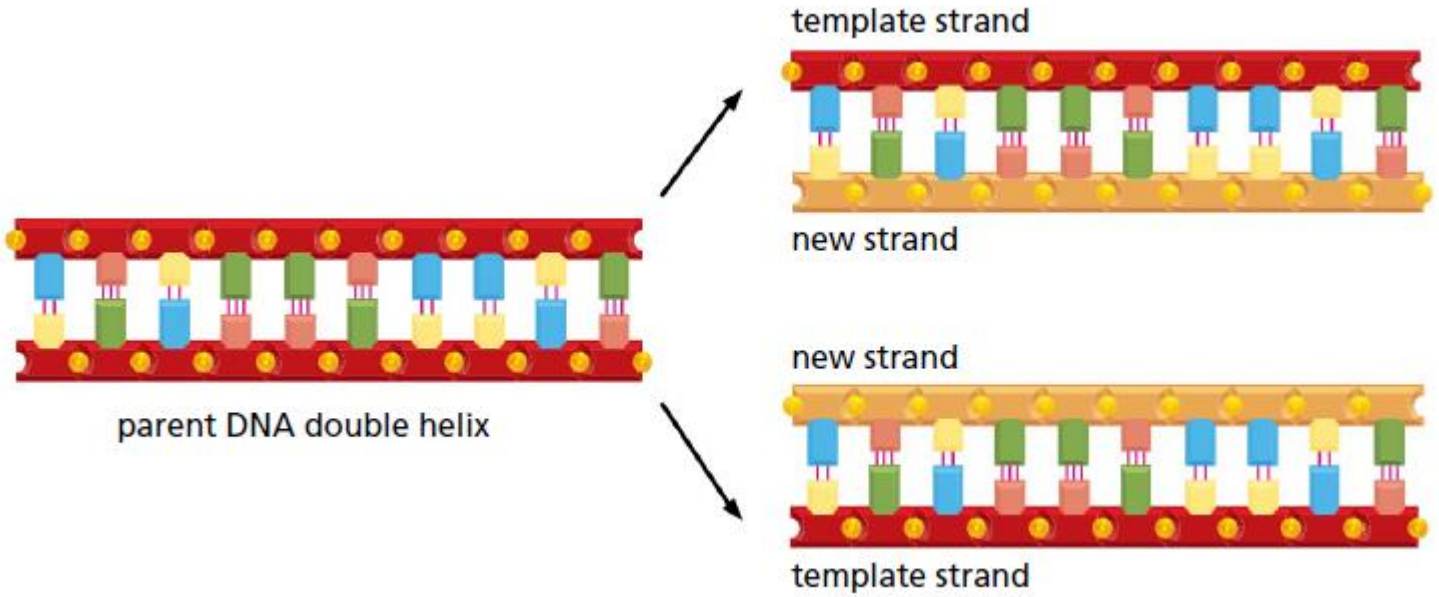
- These monomers have nicknames drawn from a four-letter alphabet—A, T, C, G—and they are strung together in a long linear sequence that encodes the genetic information, just as the sequence of 1s and 0s encodes the information in a computer file.

- We can take a piece of DNA from a human cell and insert it into a bacterium, or a piece of bacterial DNA and insert it into a human cell, and the information will be successfully read, interpreted, and copied.

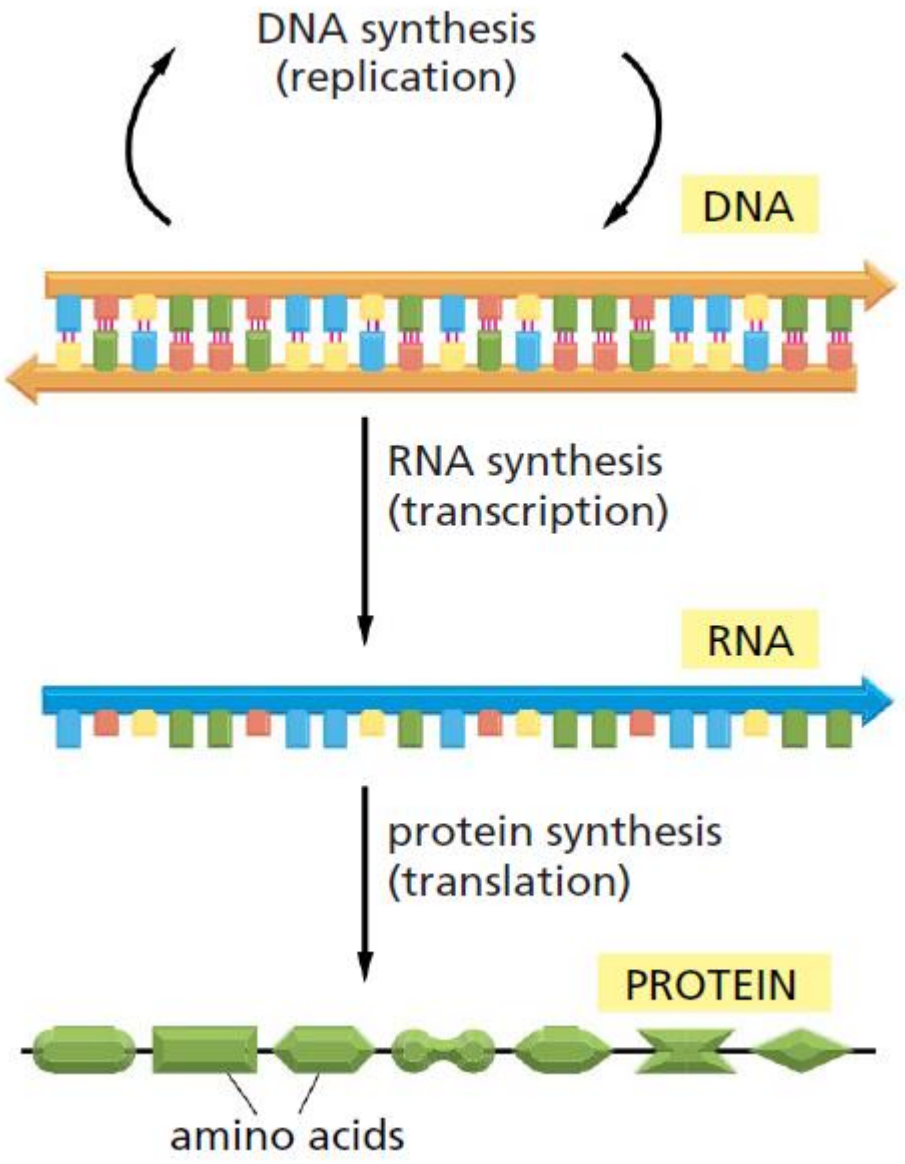

(A) building block of DNA
phosphate
sugar
sugar phosphate
base
nucleotide
(B) DNA strand
(C) templated polymerization of new strand
nucleotide monomers
(D) double-stranded DNA
sugar-phosphate backbone
hydrogen-bonded base pairs
(E) DNA double helix

- The bonds between the base pairs are weak compared with the sugar-phosphate links, and this allows the two DNA strands to be pulled apart without breakage of their backbones.

- Each strand then can serve as a template for the synthesis of a fresh DNA strand complementary to itself.



- To carry out its information-bearing function, DNA must do more than copy itself.
- It must also express its information, by letting it guide the synthesis of other molecules in the cell.
- This also occurs by a mechanism that is the same in all living organisms, leading first and foremost to the production of two other key classes of polymers: RNAs and proteins.



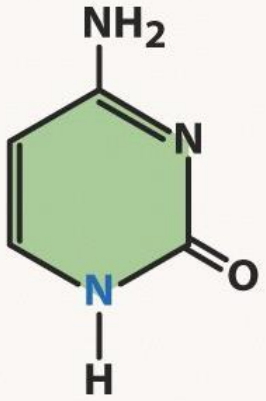
- In RNA, the backbone is formed of a slightly different sugar from that of DNA—ribose instead of deoxyribose—and one of the four bases is slightly different— uracil (U) in place of thymine (T); but the other three bases—A, C, and G—are the same, and all four bases pair with their complementary counterparts in DNA—the A, U, C, and G of RNA with the T, A, G, and C of DNA.

- The same segment of DNA can be used repeatedly to guide the synthesis of many identical RNA transcripts.

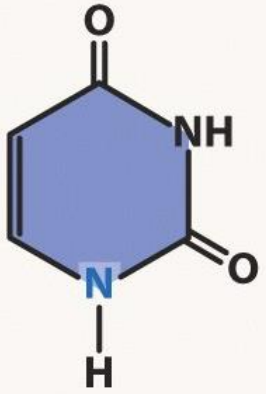
- Thus, whereas the cell's archive of genetic information in the form of DNA is fixed and sacred, the RNA transcripts are mass-produced and disposable.

- Transcripts function as intermediates in the transfer of genetic information: they mainly serve as messenger RNA (mRNA) to guide the synthesis of proteins according to the genetic instructions stored in the DNA.

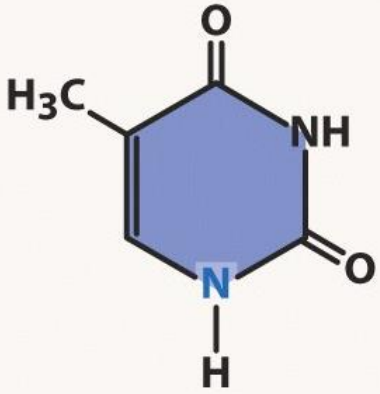
Nitrogen-containing bases



Cytosine (C)

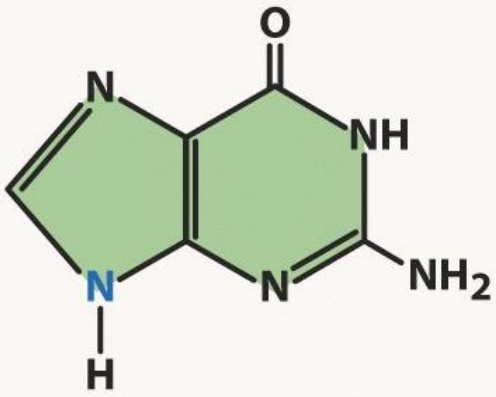


Uracil (U)

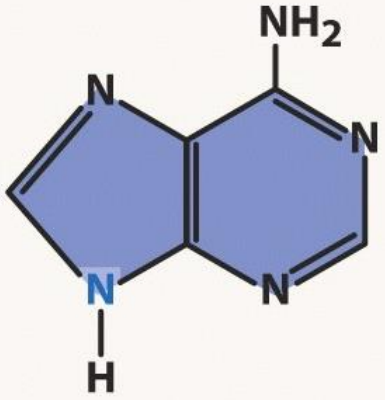


Thymine (T)

Pyrimidines



Guanine (G)



Adenine (A)

Purines

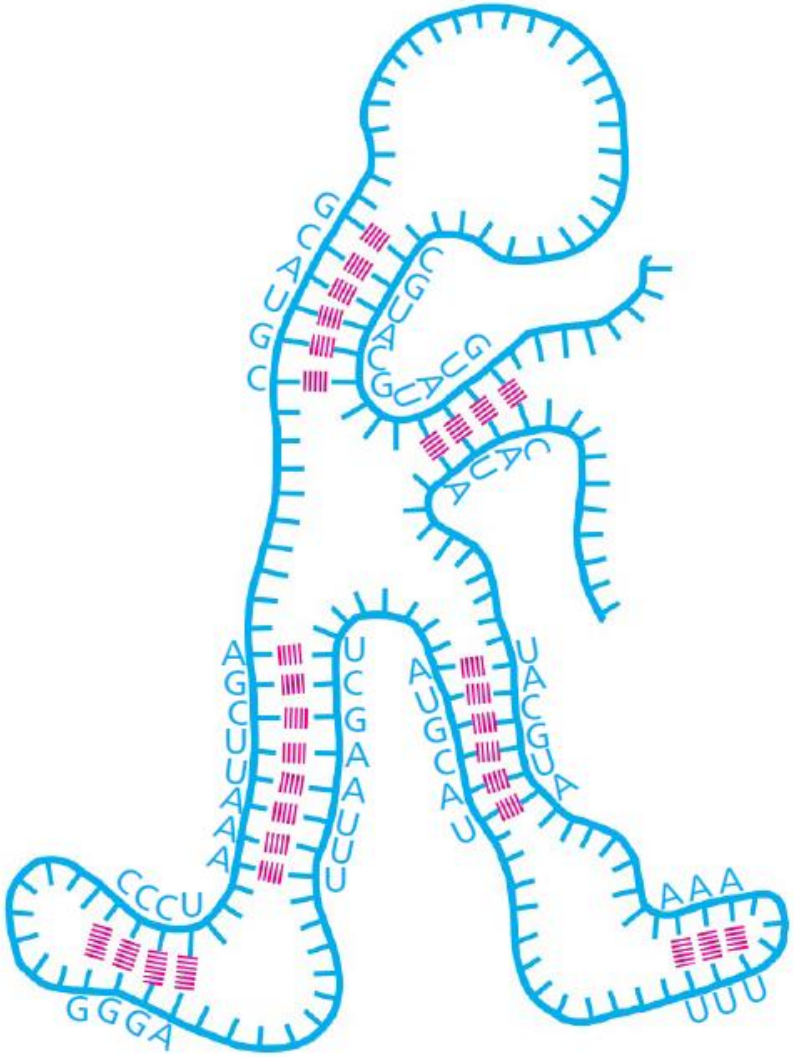
- RNA molecules have distinctive structures that can also give them other specialized chemical capabilities.

- Being single-stranded, their backbone is flexible, so that the polymer chain can bend back on itself to allow one part of the molecule to form weak bonds with another part of the same molecule.

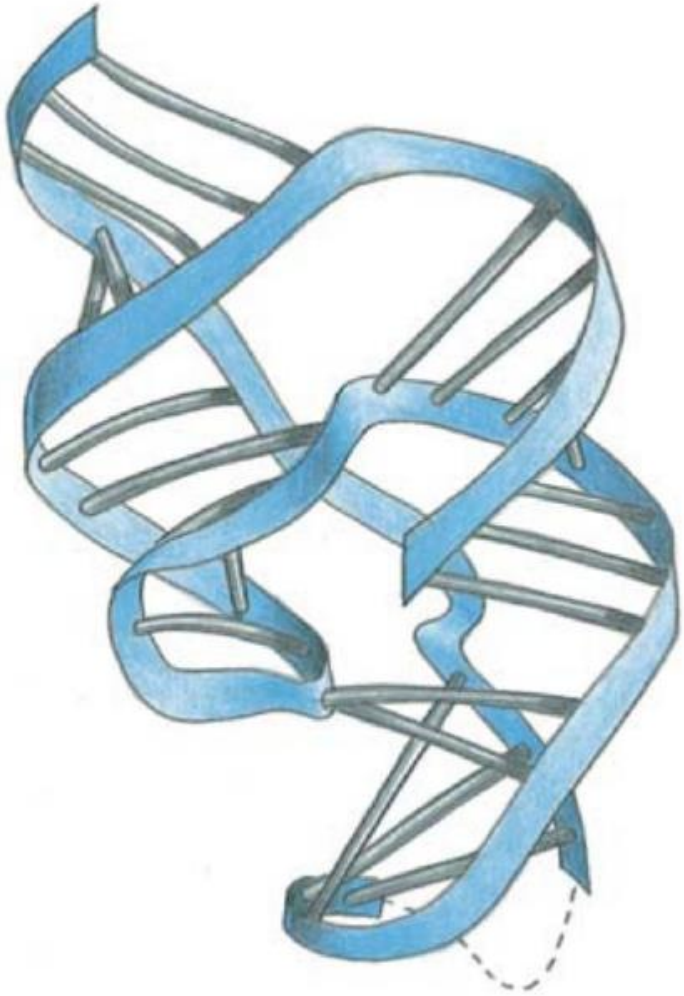
- This occurs when segments of the sequence are locally complementary.

- These types of internal associations can cause an RNA chain to fold up into a specific shape that is dictated by its sequence.

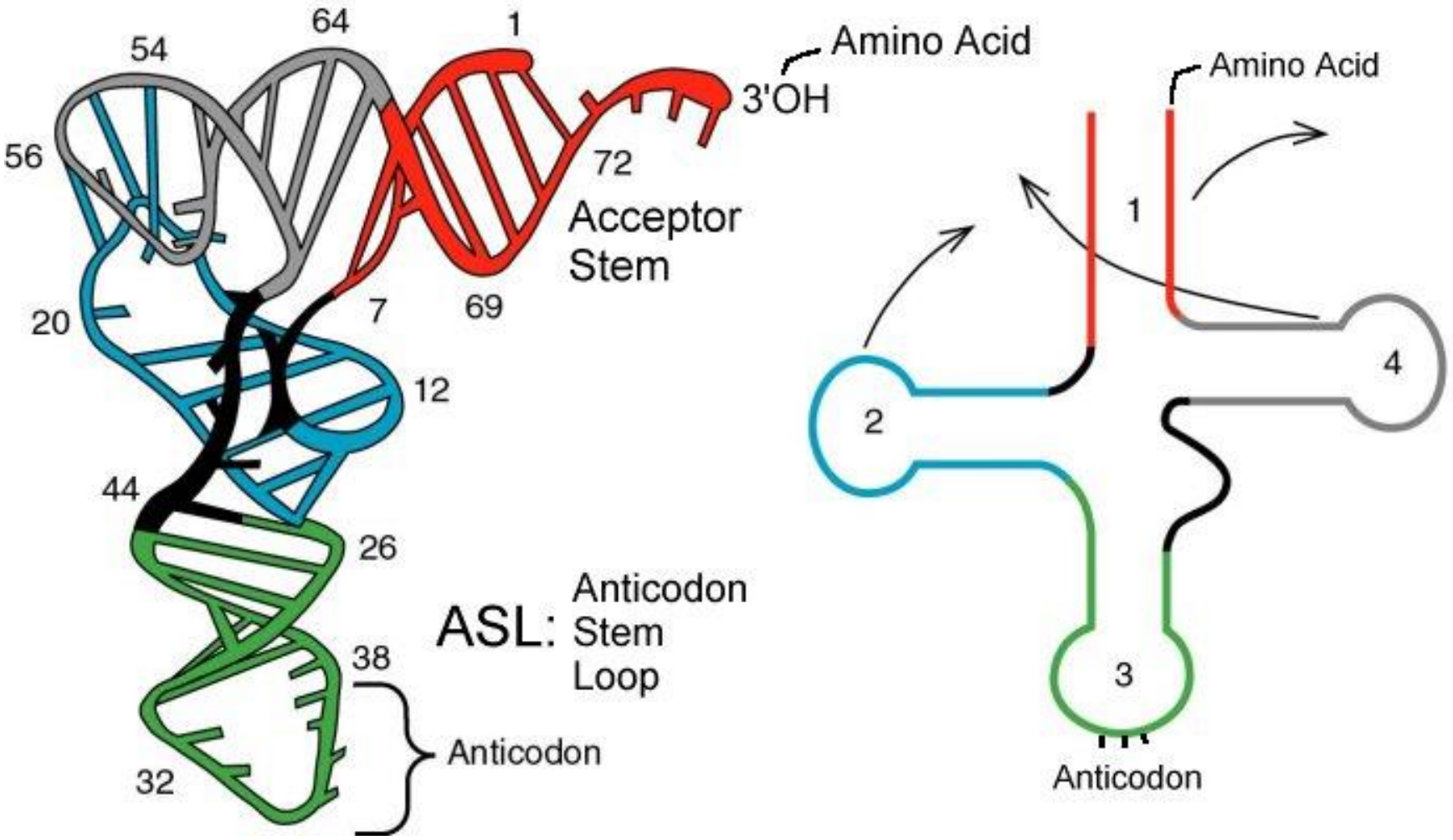
- The shape of the RNA molecule, in turn, may enable it to recognize other molecules by binding to them selectively—and even, in certain cases, to catalyze chemical changes in the molecules that are bound.



(A)



(B)

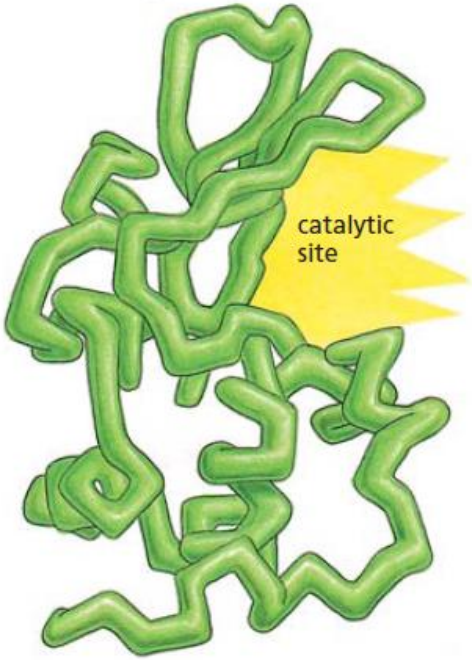

64
1
54
Amino Acid
3'OH
56
72
Acceptor
Stem
20
7
69
12
44
26
ASL:
Anticodon
Stem
Loop
38
Anticodon
32
Amino Acid
1
4
2
3
Anticodon

- Protein molecules, like DNA and RNA molecules, are long unbranched polymer chains, formed by stringing together monomeric building blocks drawn from a standard repertoire that is the same for all living cells.

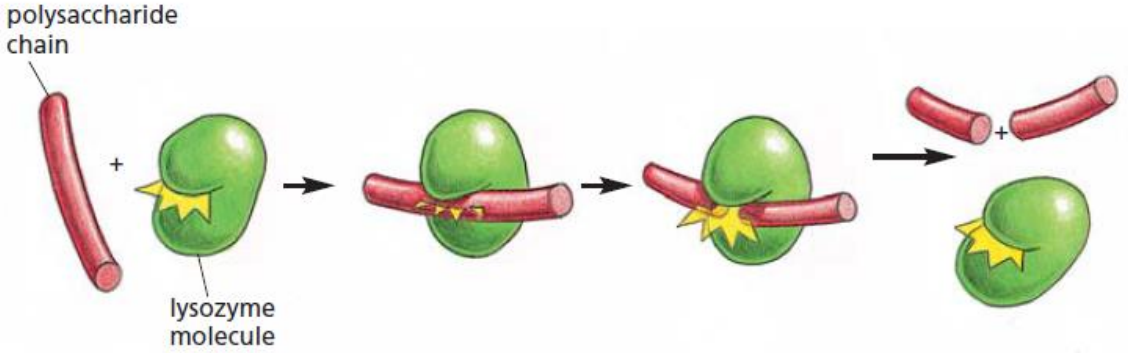
- Like DNA and RNA, they carry information in the form of a linear sequence of symbols, in the same way as a human message written in an alphabetic script.

- The monomers of protein, the amino acids, are quite different from those of DNA and RNA, and there are 20 types, instead of 4.

- Each of the protein molecules, or polypeptides, created by joining amino acids in a particular sequence folds into a precise three-dimensional form with reactive sites on its surface



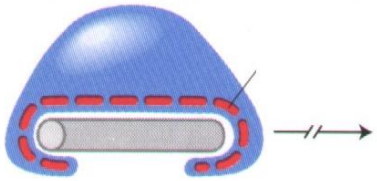
(A) lysozyme



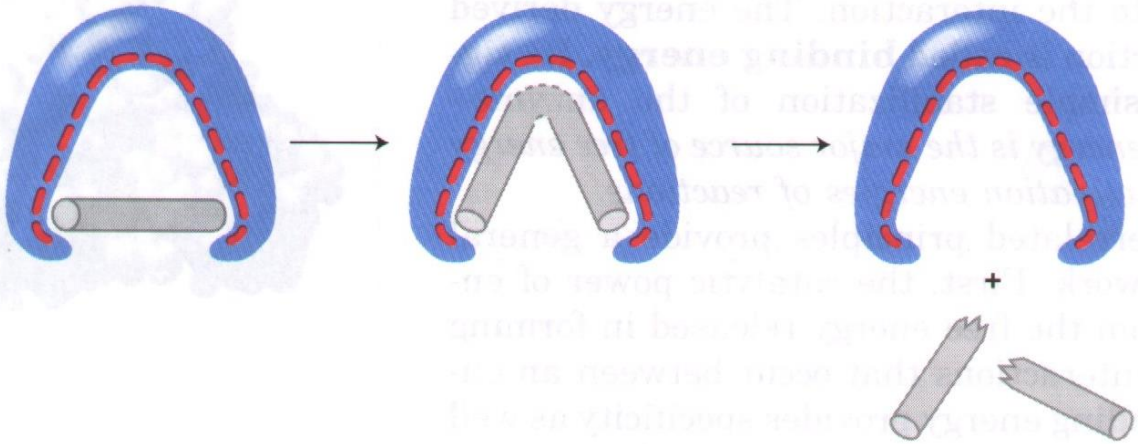
(B)

Figure 1–7 How a protein molecule acts as catalyst for a chemical reaction. (A) In a protein molecule the polymer chain folds up to into a specific shape defined by its amino acid sequence. A groove in the surface of this particular folded molecule, the enzyme lysozyme, forms a catalytic site. (B) A polysaccharide molecule *(red)*—a polymer chain of sugar monomers—binds to the catalytic site of lysozyme and is broken apart, as a result of a covalent bond-breaking reaction catalyzed by the amino acids lining the groove.

Enzyme complementary to substrate



Enzyme complementary to transition state



- The translation of genetic information from the 4-letter alphabet of polynucleotides into the 20-letter alphabet of proteins is a complex process.

- The rules of this translation seem in some respects neat and rational, in other respects strangely arbitrary, given that they are (with minor exceptions) identical in all living things.

- These arbitrary features, it is thought, reflect frozen accidents in the early history of life—chance properties of the earliest organisms that were passed on by heredity and have become so deeply embedded in the constitution of all living cells that they cannot be changed without disastrous effects.

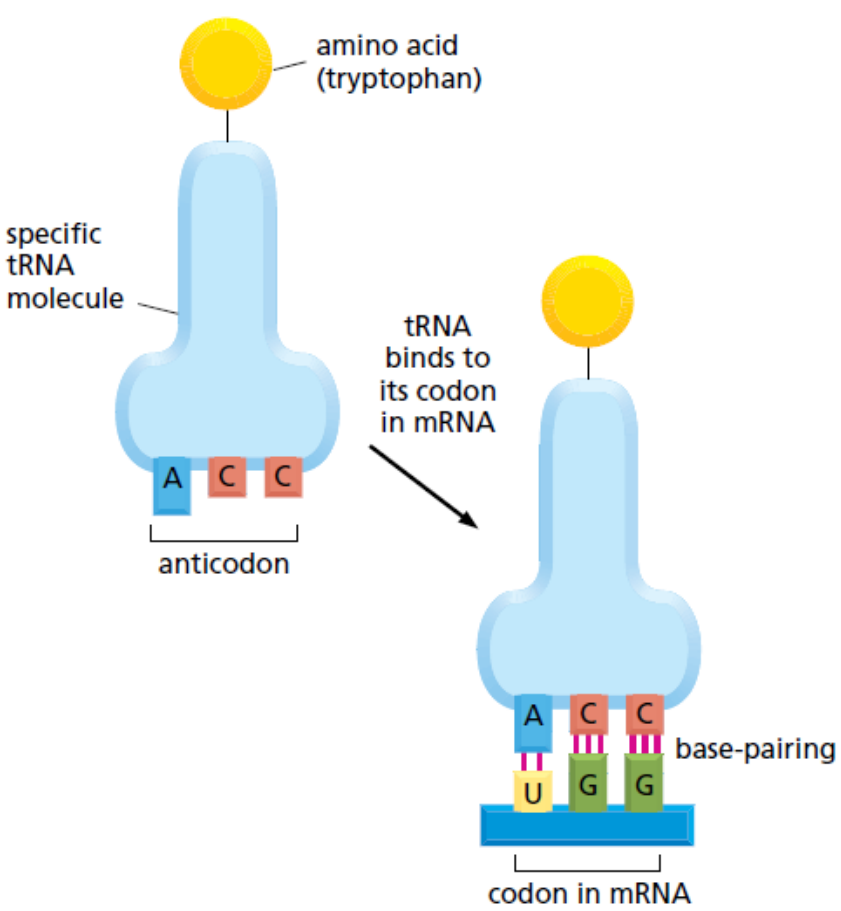
- The information in the sequence of a messenger RNA molecule is read out in groups of three nucleotides at a time: each triplet of nucleotides, or codon, specifies (codes for) a single amino acid in a corresponding protein.

- The code is read out by a special class of small RNA molecules, the transfer RNAs (tRNAs).

- Each type of tRNA becomes attached at one end to a specific amino acid, and displays at its other end a specific sequence of three nucleotides—an anticodon—that enables it to recognize, through base-pairing, a particular codon or subset of codons in mRNA.

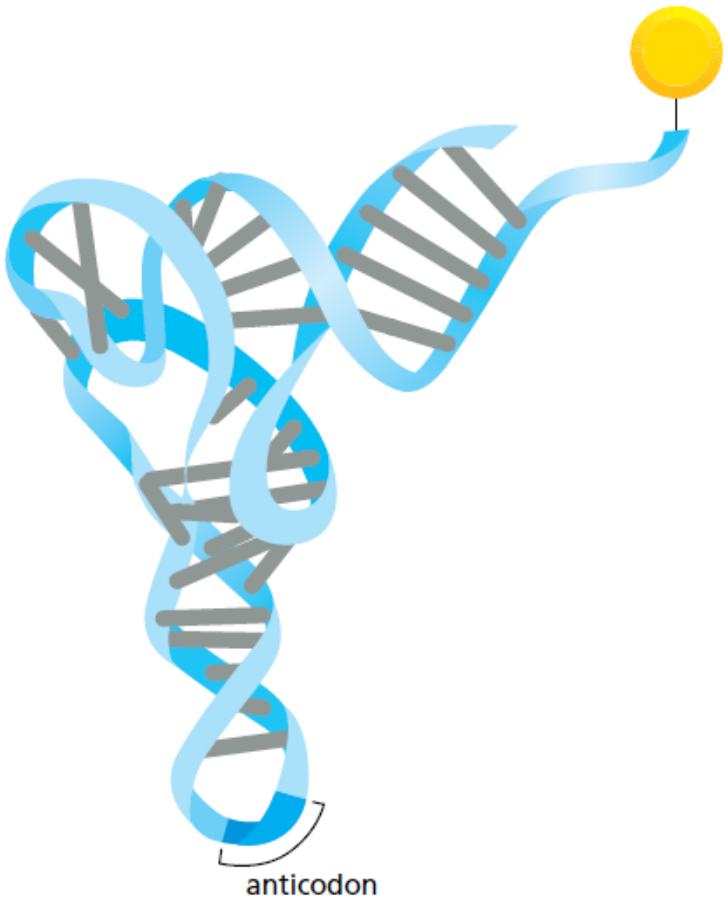
- For synthesis of protein, a succession of tRNA molecules charged with their appropriate amino acids have to be brought together with an mRNA molecule and matched up by base-pairing through their anticodons with each of its successive codons.

- The amino acids then have to be linked together to extend the growing protein chain, and the tRNAs, relieved of their burdens, have to be released.



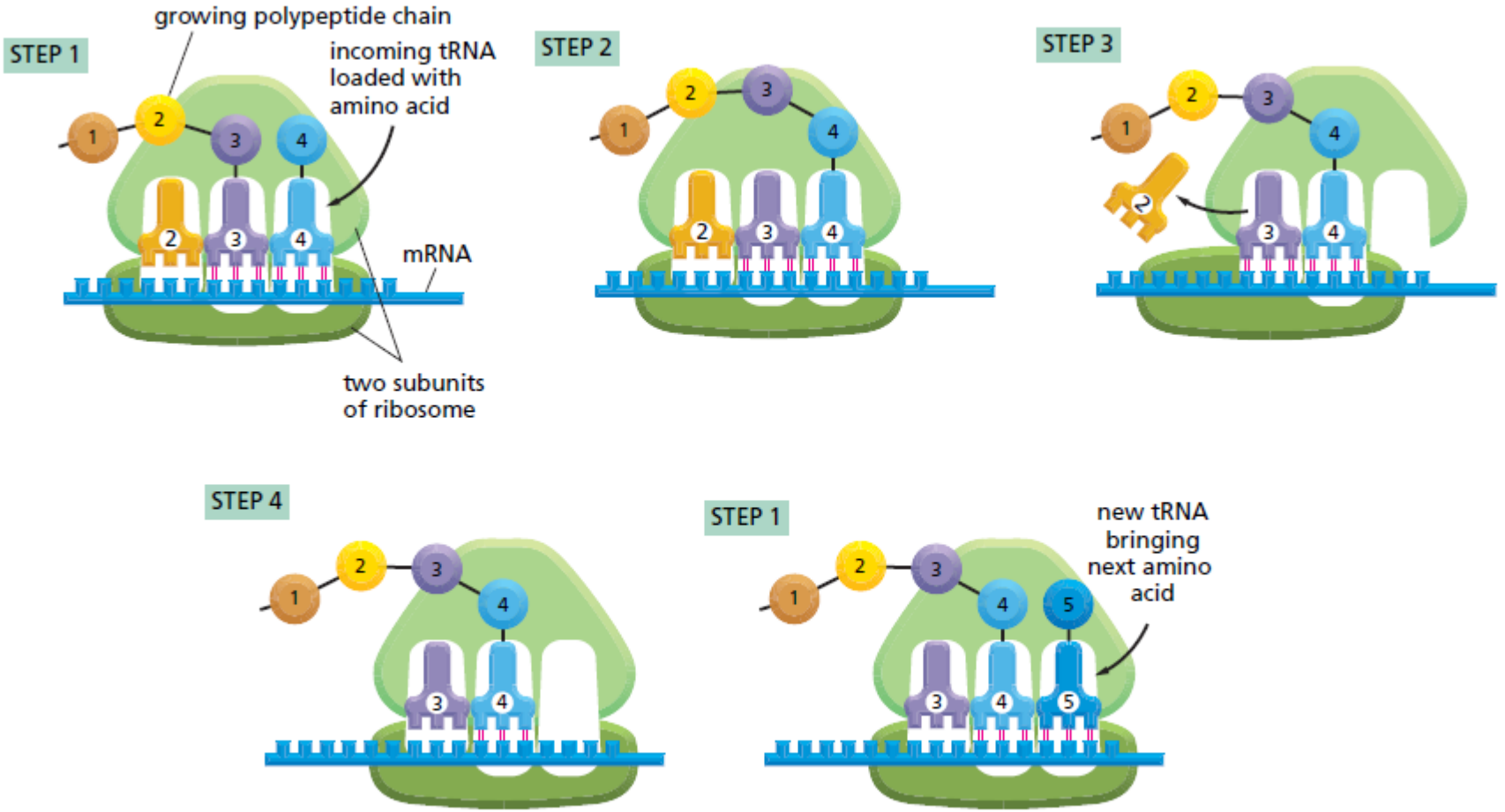
(A)

NET RESULT: AMINO ACID IS SELECTED BY ITS CODON



(B)

❖ This whole complex of processes is carried out by a giant multimolecular machine, the ribosome, formed of two main chains of RNA, called ribosomal RNAs (rRNAs), and more than 50 different proteins.



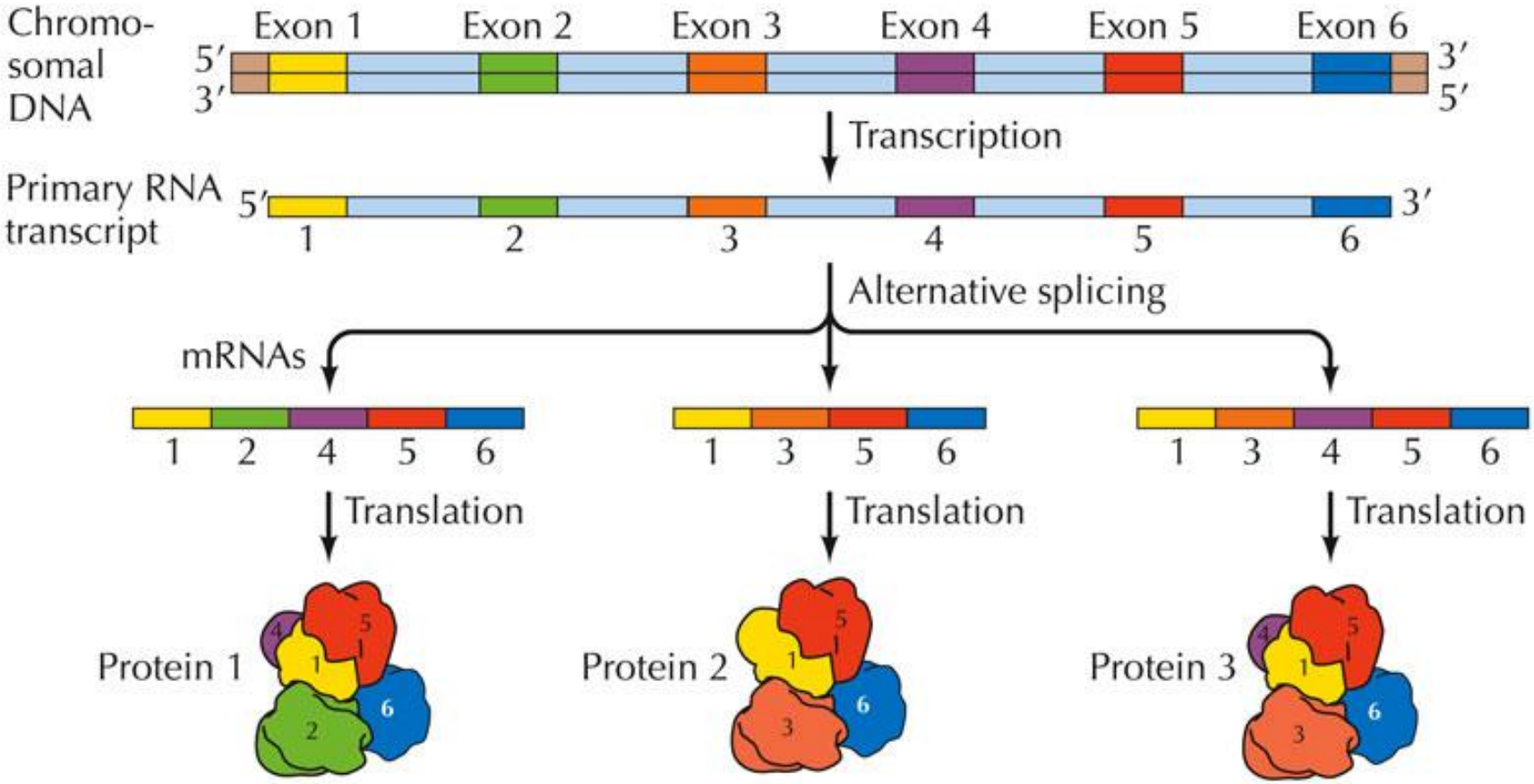
- DNA molecules as a rule are very large, containing the specifications for thousands of proteins.

- Individual segments of the entire DNA sequence are transcribed into separate mRNA molecules, with each segment coding for a different protein.

- Each such DNA segment represents one gene.

- A complication is that RNA molecules transcribed from the same DNA segment can often be processed in more than one way, so as to give rise to a set of alternative versions of a protein, especially in more complex cells such as those of plants and animals.

- A gene therefore is defined, more generally, as the segment of DNA sequence corresponding to a single protein or set of alternative protein variants.

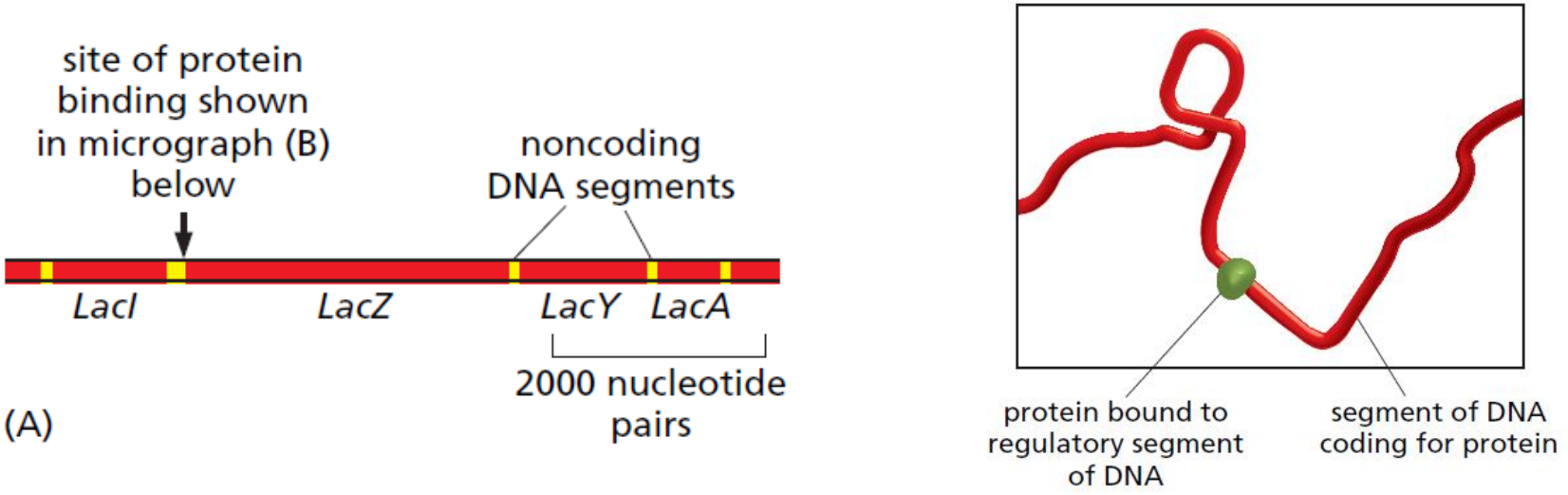

Chromo-
somal
DNA
Exon 1
Exon 2
Exon 3
Exon 4
Exon 5
Exon 6
5′
3′
3′
5′
Transcription
Primary RNA
transcript
5′
3′
1
2
3
4
5
6
Alternative splicing
mRNAs
1 2 4 5 6
1 3 5 6
1 3 4 5 6
Translation
Translation
Translation
Protein 1
Protein 2
Protein 3

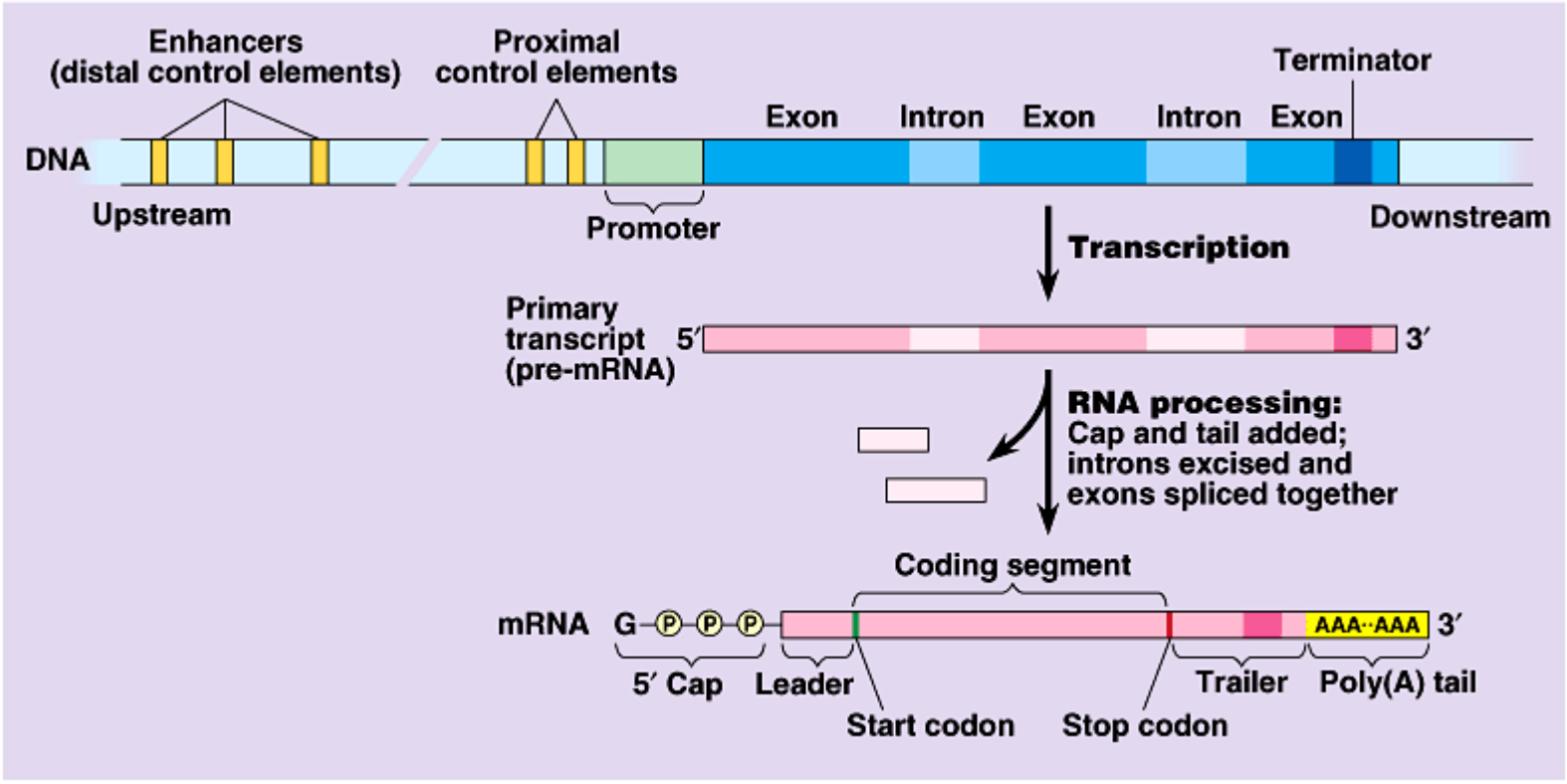
- In all cells, the expression of individual genes is regulated: instead of manufacturing its full repertoire of possible proteins at full tilt all the time, the cell adjusts the rate of transcription and translation of different genes independently, according to need.

- Stretches of regulatory DNA are interspersed among the segments that code for protein, and these noncoding regions bind to special protein molecules that control the local rate of transcription.

- Other noncoding DNA is also present, some of it serving, for example, as punctuation, defining where the information for an individual protein begins and ends.

- In this way, the genome of the cell—that is, the total of its genetic information as embodied in its complete DNA sequence—dictates not only the nature of the cell's proteins, but also when and where they are to be made.





- The basic principles of biological information transfer are simple enough, but how complex are real living cells? In particular, what are the minimum requirements?

- We can get a rough indication by considering a species that has one of the smallest known genomes—the bacterium Mycoplasma genitalium.

- This organism lives as a parasite in mammals, and its environment provides it with many of its small molecules ready-made.

- Nevertheless, it still has to make all the large molecules—DNA, RNAs, and proteins—required for the basic processes of heredity.

- It has only about 480 genes in its genome of 580,070 nucleotide pairs, representing 145,018 bytes of information

- The success of living organisms based on DNA, RNA, and protein, out of the infinitude of other chemical forms that we might conceive of, has been spectacular.

- They have populated the oceans, covered the land, infiltrated the Earth's crust, and molded the surface of our planet.

- Living things are not confined to the familiar temperate realm of land, water, and sunlight inhabited by plants and plant-eating animals.

- They can be found in the darkest depths of the ocean, in hot volcanic mud, in pools beneath the frozen surface of the Antarctic, and buried kilometers deep in the Earth's crust.

- The genetic information for every organism is written in the universal language of DNA sequences, and the DNA sequence of any given organism can be obtained by standard biochemical techniques.

- Living organisms obtain their free energy in different ways. Some, such as animals, fungi, and the bacteria that live in the human gut, get it by feeding on other living things or the organic chemicals they produce; such organisms are called organotrophic (from the Greek word trophe, meaning “food”).

- Others derive their energy directly from the nonliving world. These fall into two classes: those that harvest the energy of sunlight, and those that capture their energy from energy-rich systems of inorganic chemicals in the environment (chemical systems that are far from chemical equilibrium).

- Organisms of the former class are called phototrophic (feeding on sunlight); those of the latter are called lithotrophic (feeding on rock).

- Organotrophic organisms could not exist without these primary energy converters, which are the most plentiful form of life.

- To make a living cell requires matter, as well as free energy. DNA, RNA, and protein are composed of just six elements: hydrogen, carbon, nitrogen, oxygen, sulfur, and phosphorus.

- These are all plentiful in the nonliving environment, in the Earth's rocks, water, and atmosphere, but not in chemical forms that allow easy incorporation into biological molecules.

- Atmospheric N_2 and CO_2, in particular, are extremely unreactive, and a large amount of free energy is required to drive the reactions that use these inorganic molecules to make the organic compounds needed for further biosynthesis—that is, to fix nitrogen and carbon dioxide, so as to make N and C available to living organisms.

- Many types of living cells lack the biochemical machinery to achieve this fixation, and rely on other classes of cells to do the job for them.

- From simple microscopy, it has long been clear that living organisms can be classified on the basis of cell structure into two groups: the eucaryotes and the procaryotes.

- Eucaryotes keep their DNA in a distinct membrane-enclosed intracellular compartment called the nucleus.

- Procaryotes have no distinct nuclear compartment to house their DNA.

- Plants, fungi, and animals are eucaryotes; bacteria are procaryotes, as are archaea.

- Procaryotic cells live in an enormous variety of ecological niches, and they are astonishingly varied in their biochemical capabilities—far more so than eucaryotic cells.

- According to one estimate, at least 99% of procaryotic species remain to be characterized.

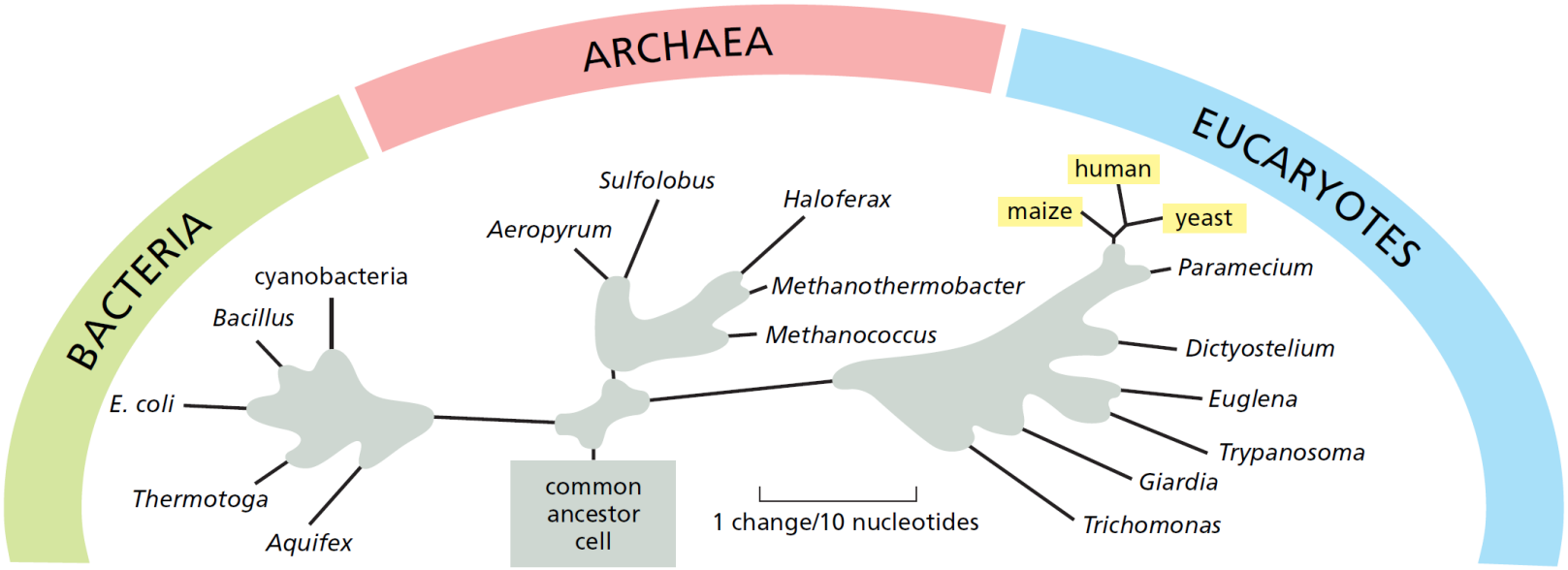


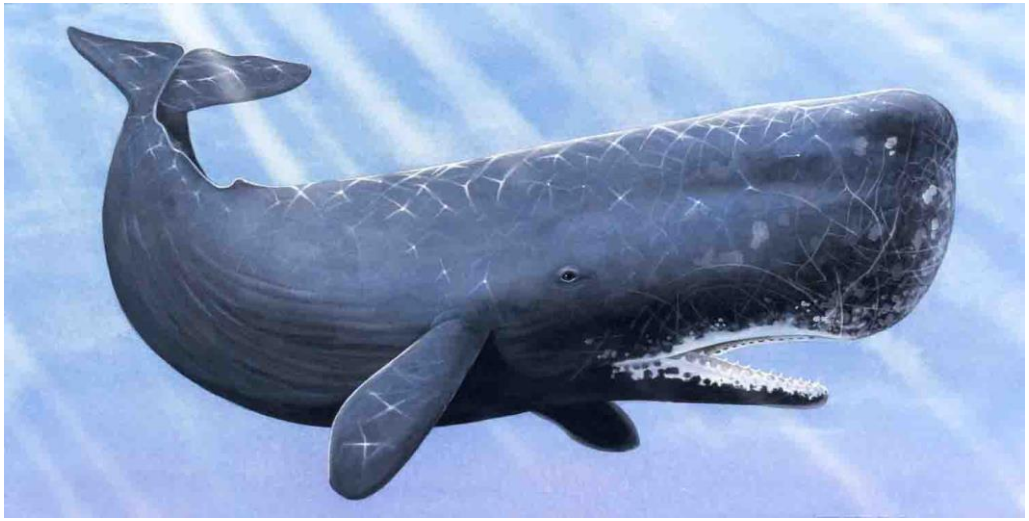
Figure 1–21 The three major divisions (domains) of the living world. Note that traditionally the word *bacteria* has been used to refer to procaryotes in general, but more recently has been redefined to refer to eubacteria specifically. The tree shown here is based on comparisons of the nucleotide sequence of a ribosomal RNA 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 (see Figure 1–22). The parts of the tree shrouded in *gray cloud* represent uncertainties about details of the true pattern of species divergence in the course of evolution: comparisons of nucleotide or amino acid sequences of molecules other than rRNA, as well as other arguments, lead to somewhat different trees. There is general agreement, however, as to the early divergence of the three most basic domains—the bacteria, the archaea, and the eucaryotes.

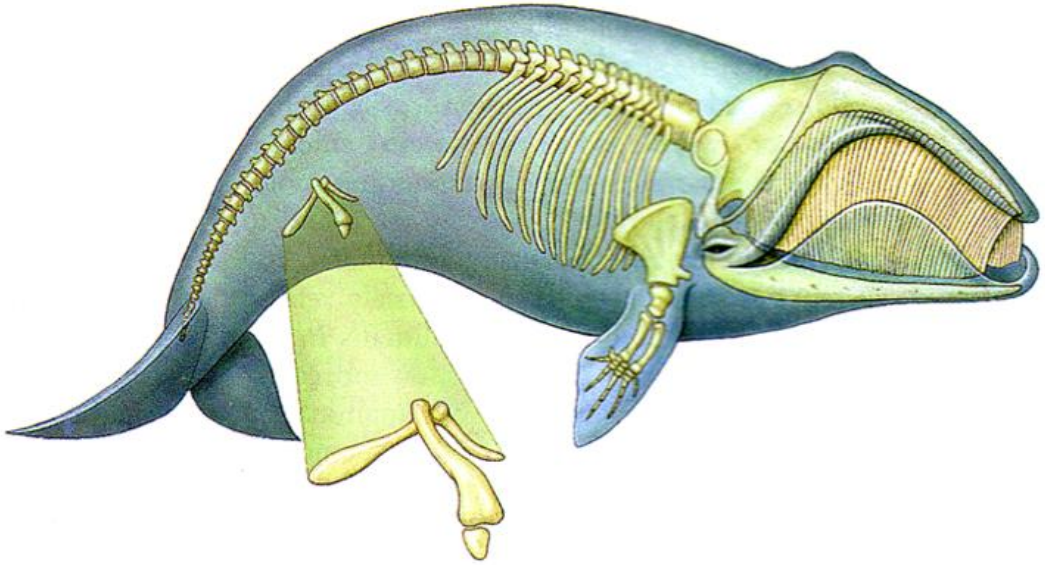
- The classification of living things has traditionally depended on comparisons of their outward appearances: we can see that a fish has eyes, jaws, backbone, brain, and so on, just as we do, and that a worm does not; that a rosebush is cousin to an apple tree, but less similar to a grass.

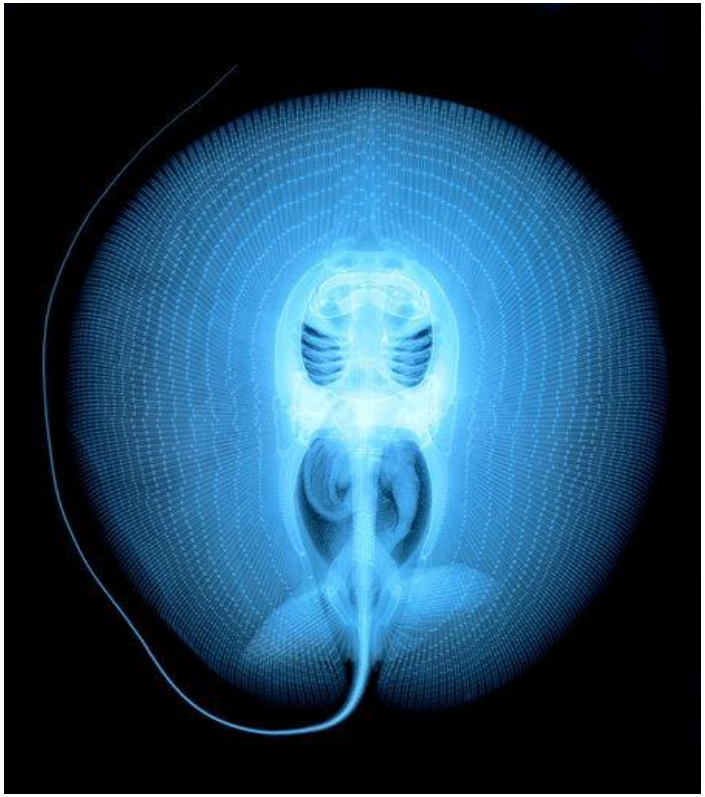
- As Darwin showed, we can readily interpret such close family resemblances in terms of evolution from common ancestors, and we can find the remains of many of these ancestors preserved in the fossil record.

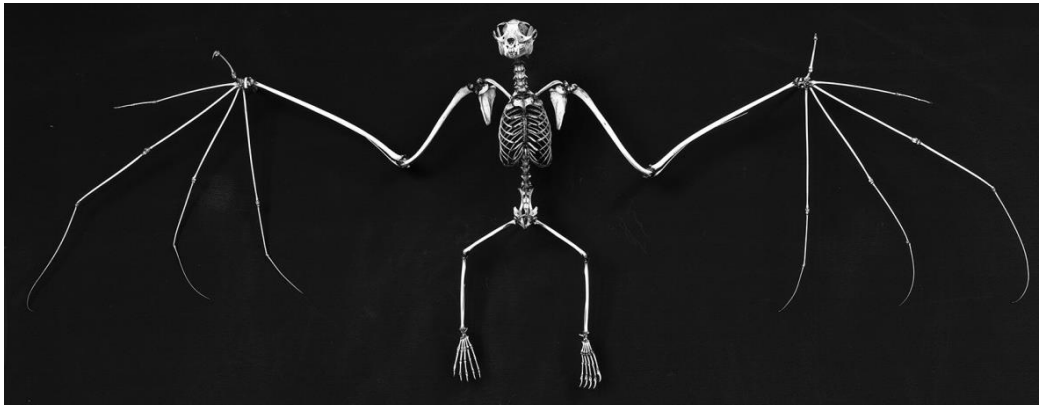
- In this way, it has been possible to begin to draw a family tree of living organisms, showing the various lines of descent, as well as branch points in the history, where the ancestors of one group of species became different from those of another.

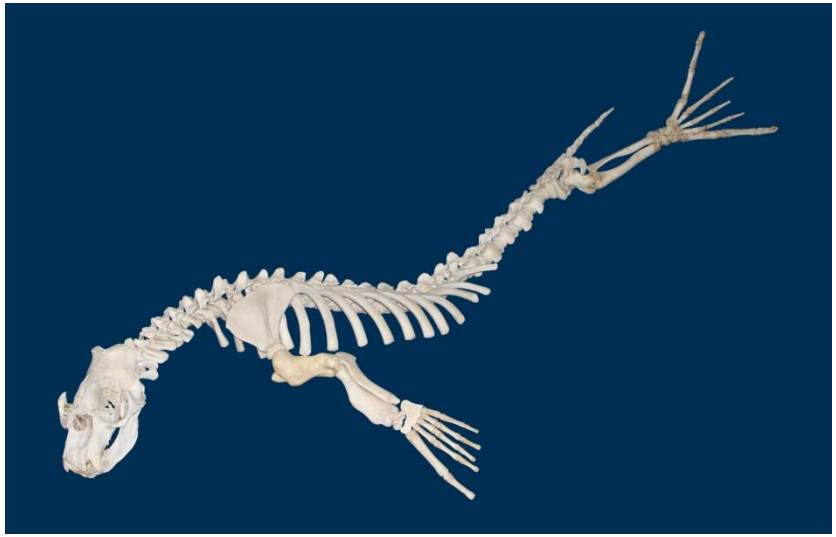
- When the disparities between organisms become very great, however, these methods begin to fail. How do we decide whether a fungus is closer kin to a plant or to an animal?

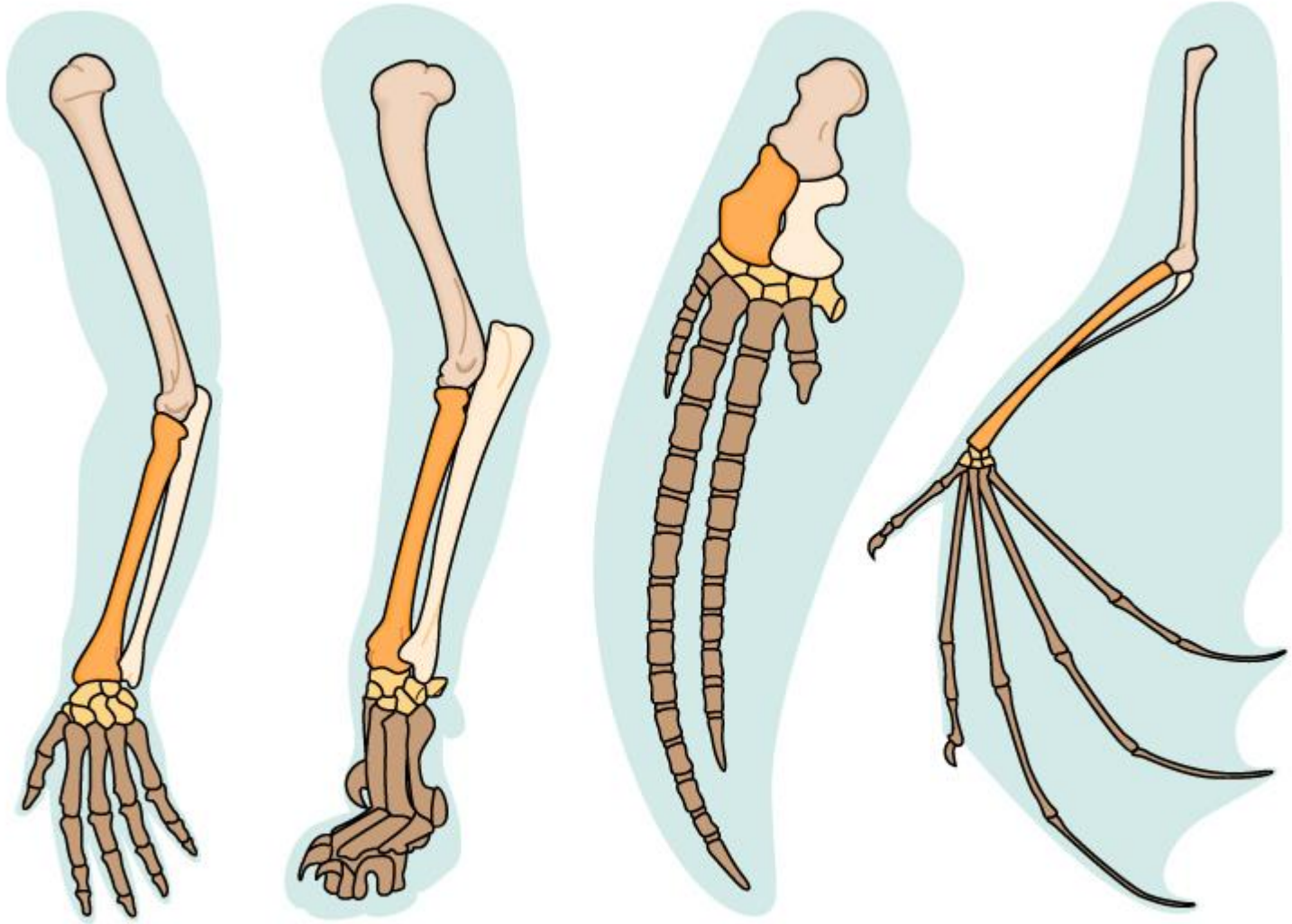


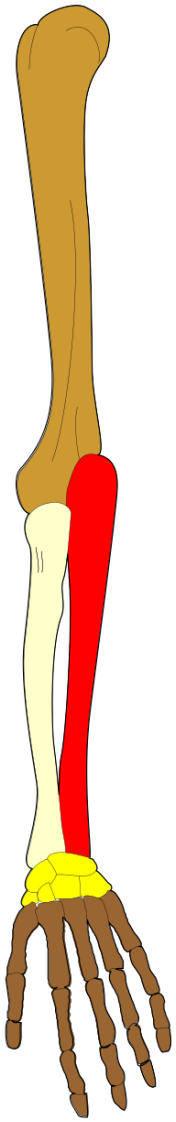




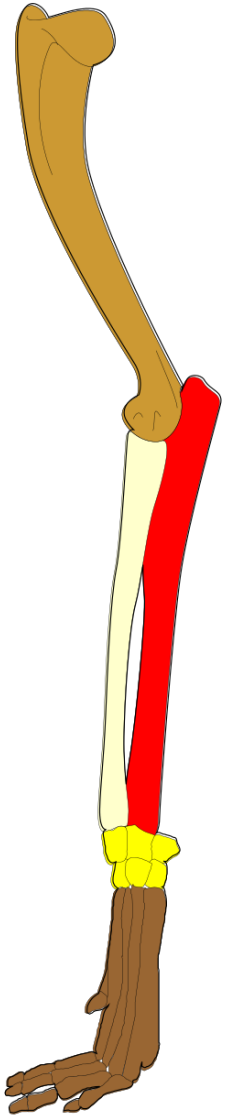




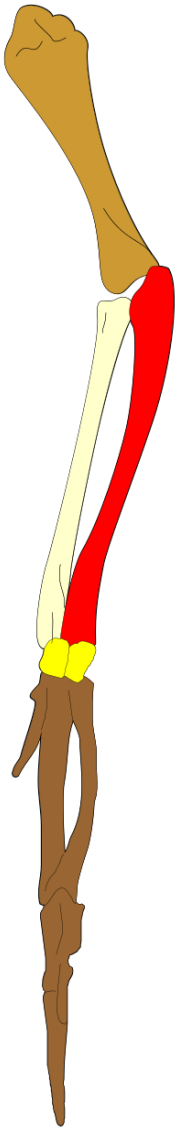




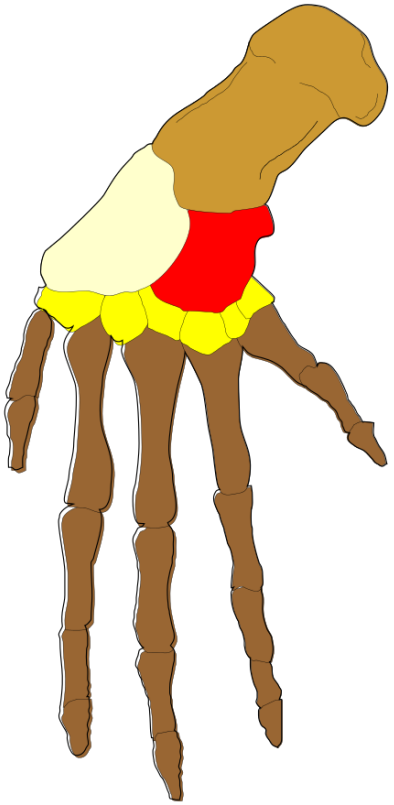
Human



Dog



Bird



Whale

- Genome analysis has given us a simpler, more direct, and more powerful way to determine evolutionary relationships.

- The complete DNA sequence of an organism defines its nature with almost perfect precision and in exhaustive detail.

- Moreover, this specification is in a digital form—a string of letters—that can be entered straightforwardly into a computer and compared with the corresponding information for any other living thing.

- Because DNA is subject to random changes that accumulate over long periods of time (as we shall see shortly), the number of differences between the DNA sequences of two organisms can provide a direct, objective, quantitative indication of the evolutionary distance between them.

- Both in the storage and in the copying of genetic information, random accidents and errors occur, altering the nucleotide sequence—that is, creating mutations.

- Therefore, when a cell divides, its two daughters are often not quite identical to one another or to their parent.

- On rare occasions, the error may represent a change for the better; more probably, it will cause no significant difference in the cell's prospects; and in many cases, the error will cause serious damage—for example, by disrupting the coding sequence for a key protein.

- Changes due to mistakes of the first type will tend to be perpetuated, because the altered cell has an increased likelihood of reproducing itself.

- Changes due to mistakes of the second type—selectively neutral changes—may be perpetuated or not: in the competition for limited resources, it is a matter of chance whether the altered cell or its cousins will succeed.

- But changes that cause serious damage lead nowhere: the cell that suffers them dies, leaving no progeny.

- Through endless repetition of this cycle of error and trial—of mutation and natural selection—organisms evolve: their genetic specifications change, giving them new ways to exploit the environment more effectively, to survive in competition with others, and to reproduce successfully.

- Clearly, some parts of the genome change more easily than others in the course of evolution.

- A segment of DNA that does not code for protein and has no significant regulatory role is free to change at a rate limited only by the frequency of random errors.

- In contrast, a gene that codes for a highly optimized essential protein or RNA molecule cannot alter so easily: when mistakes occur, the faulty cells are almost always eliminated.

- Genes of this latter sort are therefore highly conserved.

- Through 3.5 billion years or more of evolutionary history, many features of the genome have changed beyond all recognition; but the most highly conserved genes remain perfectly recognizable in all living species.

- The studies that led to the classification of the living world into the three domains of bacteria, archaea, and eucaryotes were based chiefly on analysis of one of the two main RNA components of the ribosome—the so-called smallsubunit ribosomal RNA.

- Because translation is fundamental to all living cells, this component of the ribosome has been well conserved since early in the history of life on Earth.

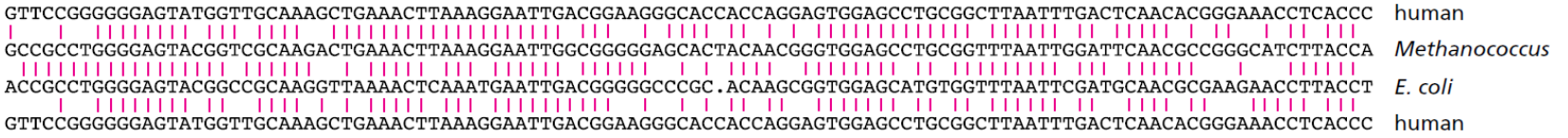


Figure 1–22 Genetic information conserved since the days of the last common ancestor of all living things. A part of the gene for the smaller of the two main RNA components of the ribosome is shown. (The complete molecule is about 1500–1900 nucleotides long, depending on species.) Corresponding segments of nucleotide sequence from an archaean *(Methanococcus jannaschii),* a bacterium *(Escherichia coli)* and a eucaryote *(Homo sapiens)* are aligned. Sites where the nucleotides are identical between species are indicated by a vertical line; the human sequence is repeated at the bottom of the alignment so that all three two-way comparisons can be seen. A dot halfway along the *E. coli* sequence denotes a site where a nucleotide has been either deleted from the bacterial lineage in the course of evolution, or inserted in the other two lineages. Note that the sequences from these three organisms, representative of the three domains of the living world, all differ from one another to a roughly similar degree, while still retaining unmistakable similarities.

- Most procaryotic cells carry very little superfluous baggage; their genomes are small, with genes packed closely together and minimal quantities of regulatory DNA between them.

- Most bacterial and archaeal genomes contain between 10^6 and 10^7 nucleotide pairs, encoding 1000–6000 genes.

- The raw material of evolution is the DNA sequence that already exists: there is no natural mechanism for making long stretches of new random sequence.

- In this sense, no gene is ever entirely new. Innovation can, however, occur in several ways

- 1. Intragenic mutation: an existing gene can be 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 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. Segment shuffling: two or more existing genes can be broken and rejoined 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 information from parent to progeny.

- Each of these types of change leaves a characteristic trace in the DNA sequence of the organism, providing clear evidence that all four processes have occurred.

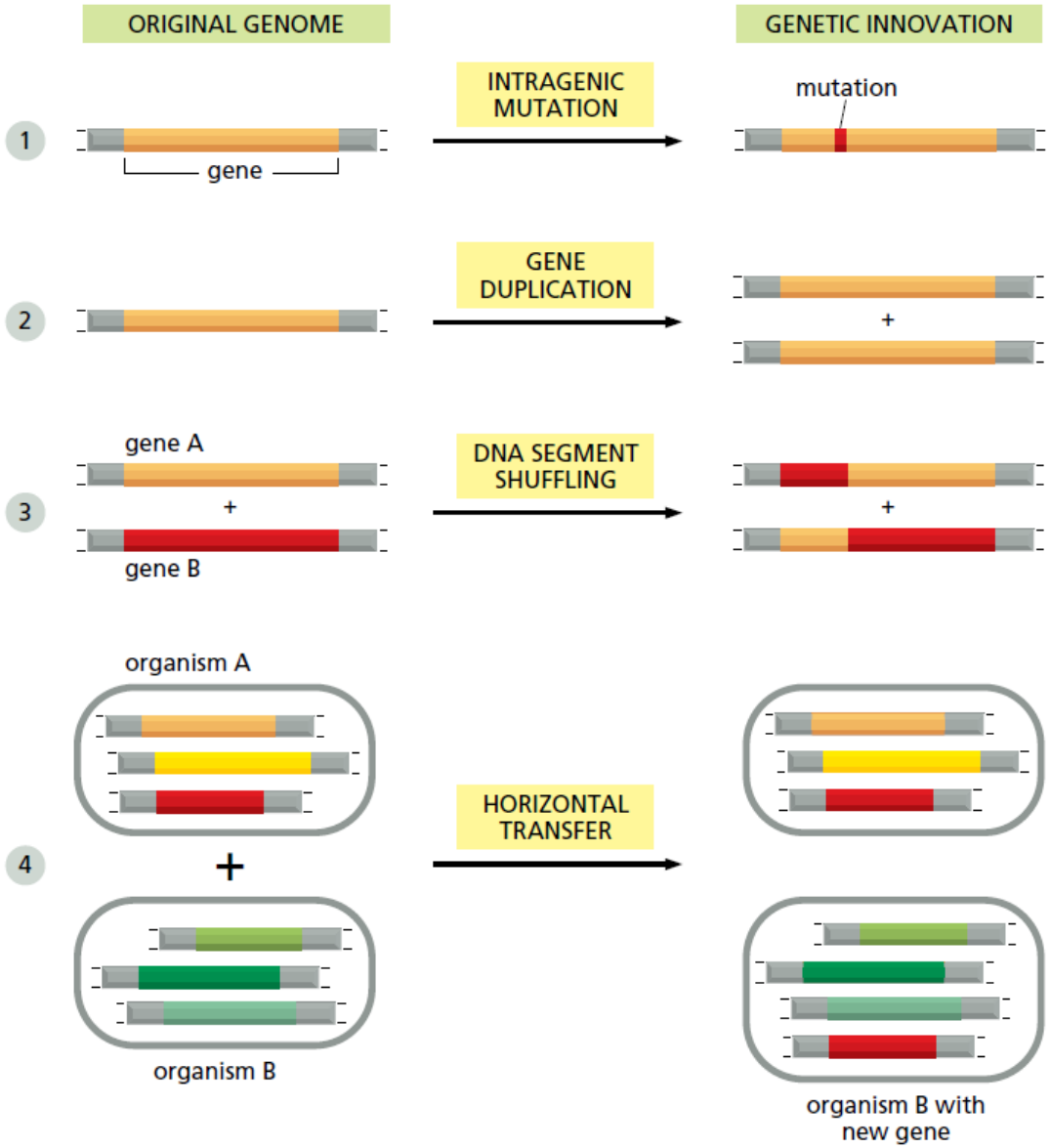


Figure 1–23 Four modes of genetic innovation and their effects on the DNA sequence of an organism. A special form of horizontal transfer occurs when two different types of cells enter into a permanent symbiotic association. Genes from one of the cells then may be transferred to the genome of the other, as we shall see below when we discuss mitochondria and chloroplasts.

- A cell duplicates its entire genome each time it divides into two daughter cells.

- However, accidents occasionally result in the inappropriate duplication of just part of the genome, with retention of original and duplicate segments in a single cell.

- Once a gene has been duplicated in this way, one of the two gene copies is free to mutate and become specialized to perform a different function within the same cell.

- Repeated rounds of this process of duplication and divergence, over many millions of years, have enabled one gene to give rise to a family of genes that may all be found within a single genome.

- Genes that are related by descent in this way—that is, genes in two separate species that derive from the same ancestral gene in the last common ancestor of those two species—are called orthologs.

- Related genes that have resulted from a gene duplication event within a single genome—and are likely to have diverged in their function—are called paralogs.

- Genes that are related by descent in either way are called homologs, a general term used to cover both types of relationship

- From the DNA sequences, it is usually easy to recognize that two genes in different species are homologous; it is much more difficult to decide, without other information, whether they stand in the precise evolutionary relationship of orthologs.

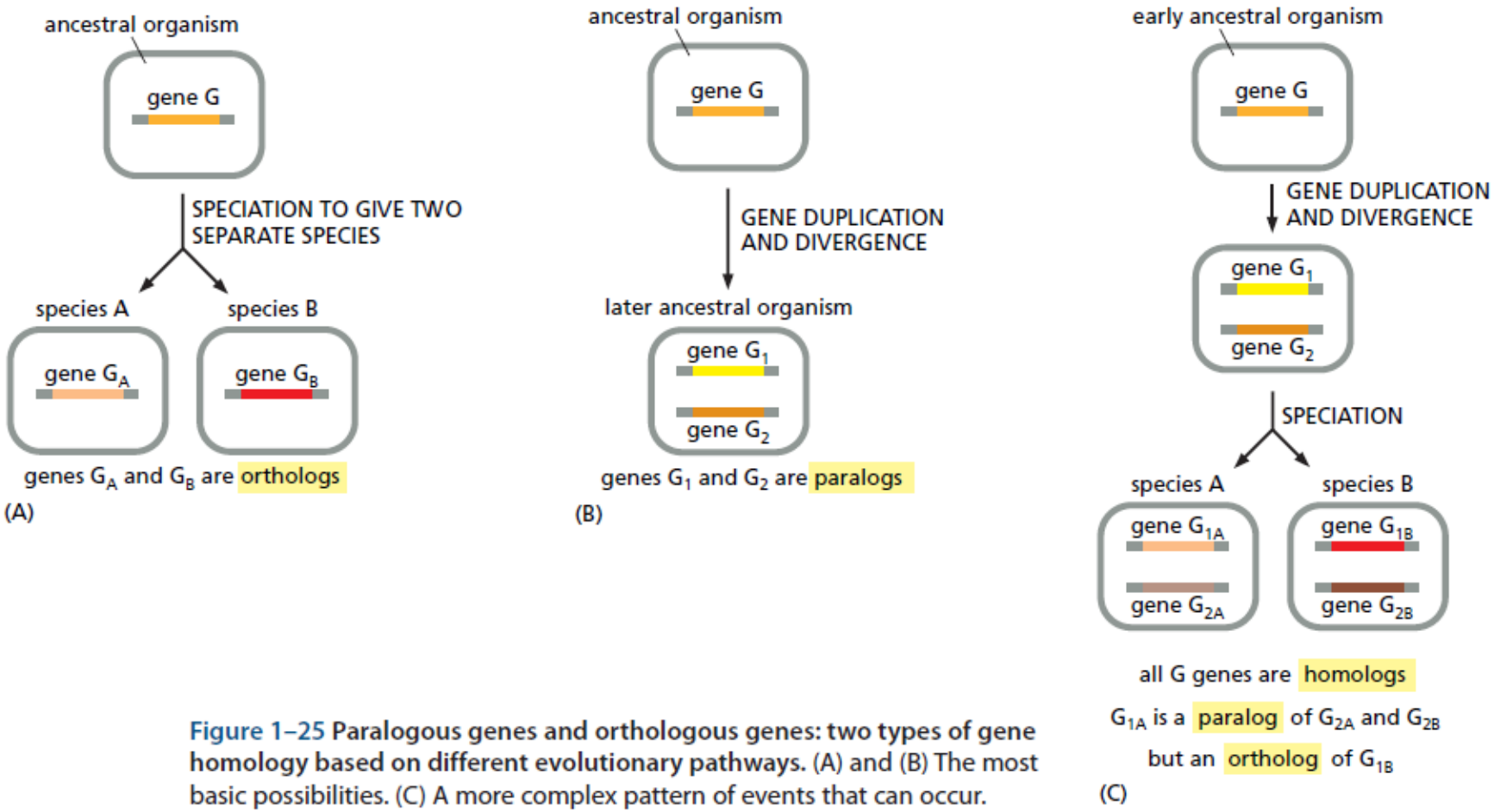
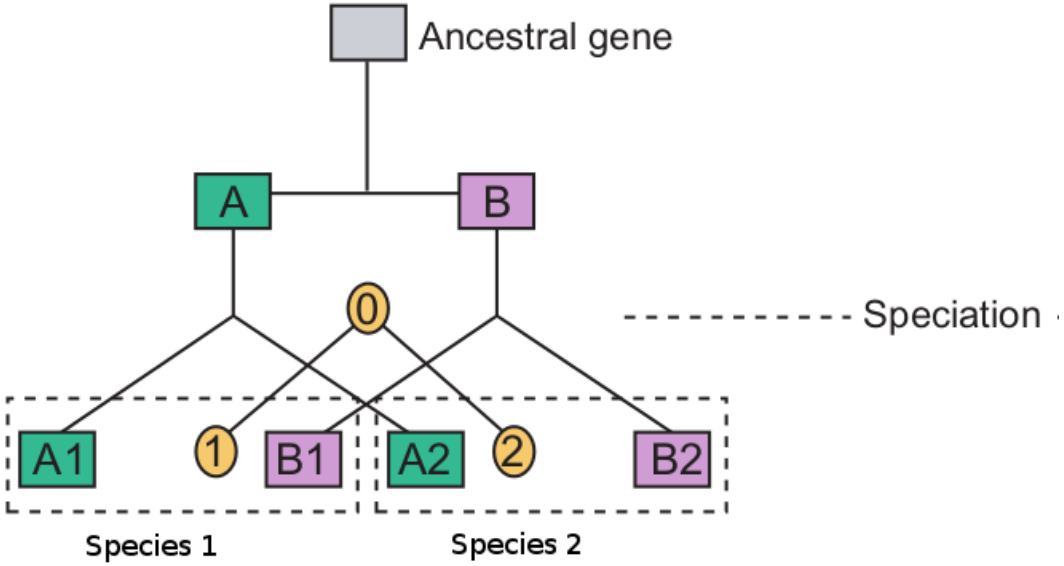


Figure 1–25 Paralogous genes and orthologous genes: two types of gene homology based on different evolutionary pathways. (A) and (B) The most basic possibilities. (C) A more complex pattern of events that can occur.

Orthologs are homologous genes that are the result of a **speciation event**.
Paralogs are homologous genes that are the result of a **duplication event**.

Part **(a)** of the diagram above shows a hypothetical evolutionary history of a gene. The ancestral genome had two copies of this gene (A and B) which were **paralogs**. At some point, the ancestral species split into two daughter species, each of whose genomes contain two copies of the ancestral duplicated gene (A1,A2 and B1,B2).



These genes are all homologous to one another but are they paralogs or orthologs?

A1 and B1 are **paralogs** A1 and B2 are **paralogs** A2 and B1 are **paralogs**
A2 and B2 are **paralogs**

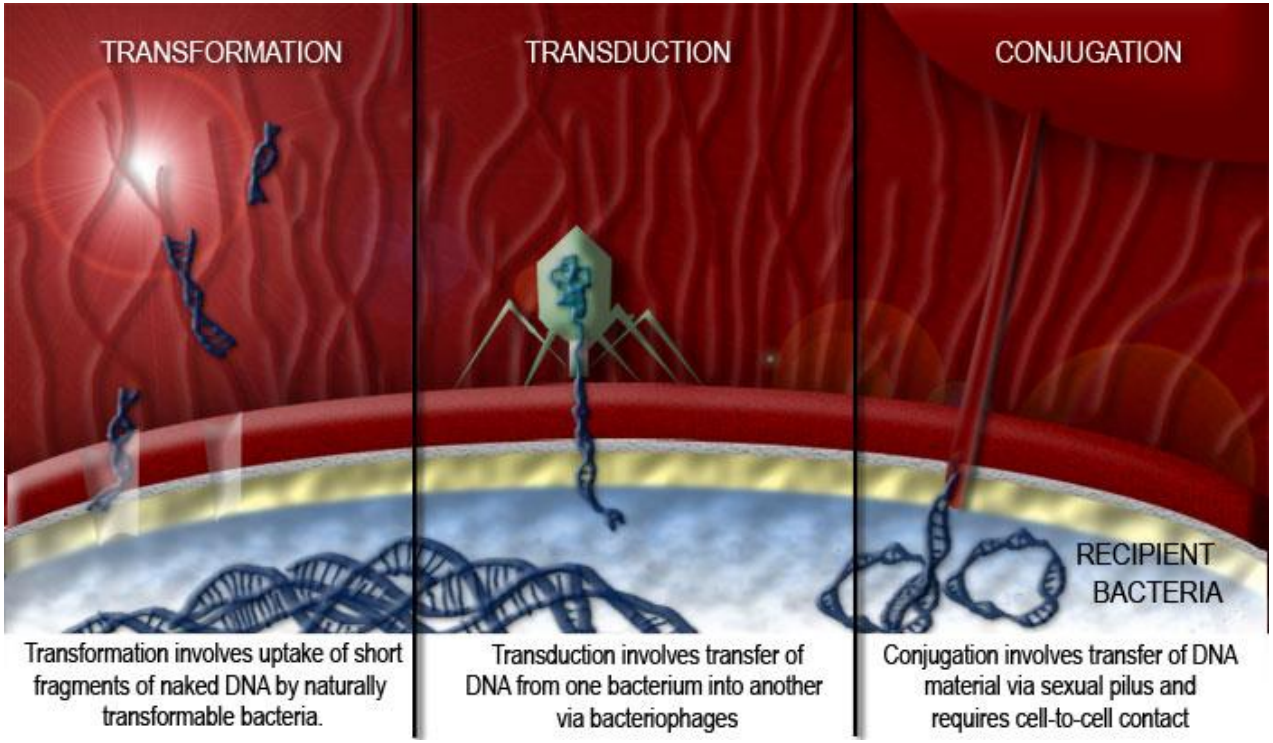
A1 and A2 are **orthologs**. B1 and B2 are **orthologs**

- Procaryotes also provide examples of the horizontal transfer of genes from one species of cell to another.

- Viruses are not themselves living cells but can act as vectors for gene transfer: they are small packets of genetic material that have evolved as parasites on the reproductive and biosynthetic machinery of host cells.

- They replicate in one cell, emerge from it with a protective wrapping, and then enter and infect another cell, which may be of the same or a different species.

- Often, the infected cell will be killed by the massive proliferation of virus particles inside it; but sometimes, the viral DNA, instead of directly generating these particles, may persist in its host for many cell generations as a relatively innocuous passenger, either as a separate intracellular fragment of DNA, known as a plasmid, or as a sequence inserted into the cell's regular genome.

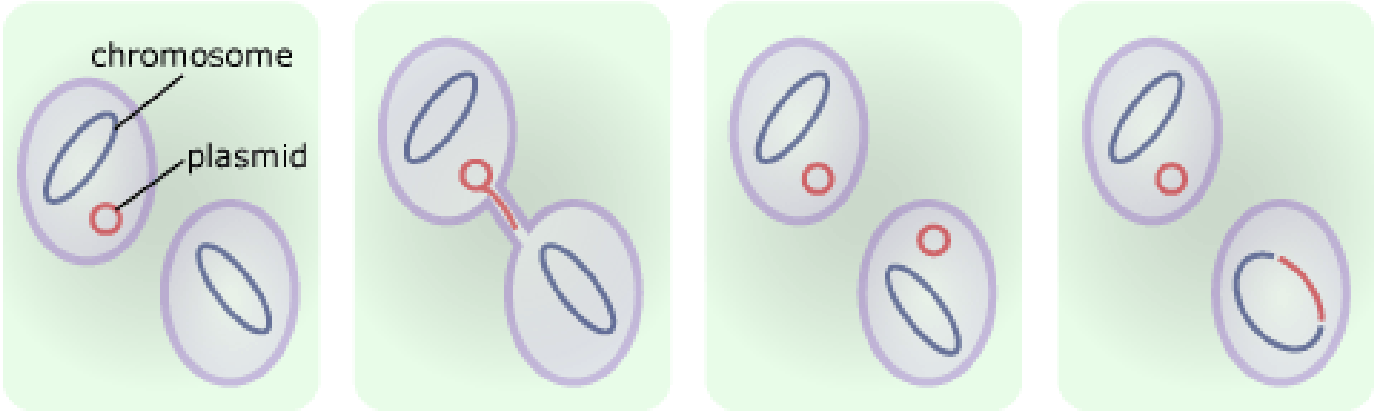


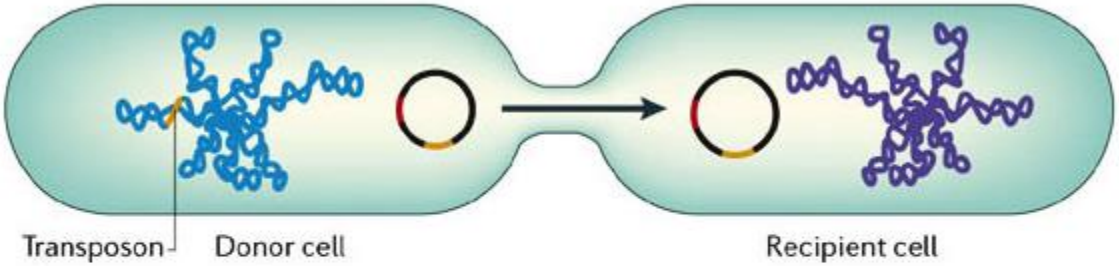
1. One bacterium contains a plasmid to be transferred.

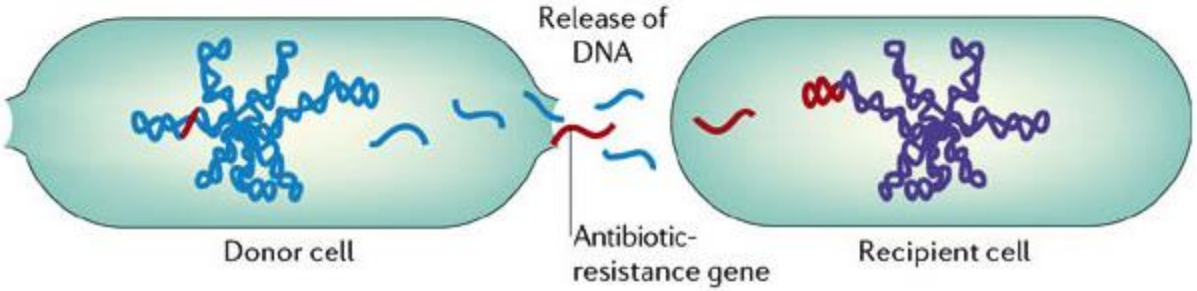
2. A connection forms and the plasmid is copied.

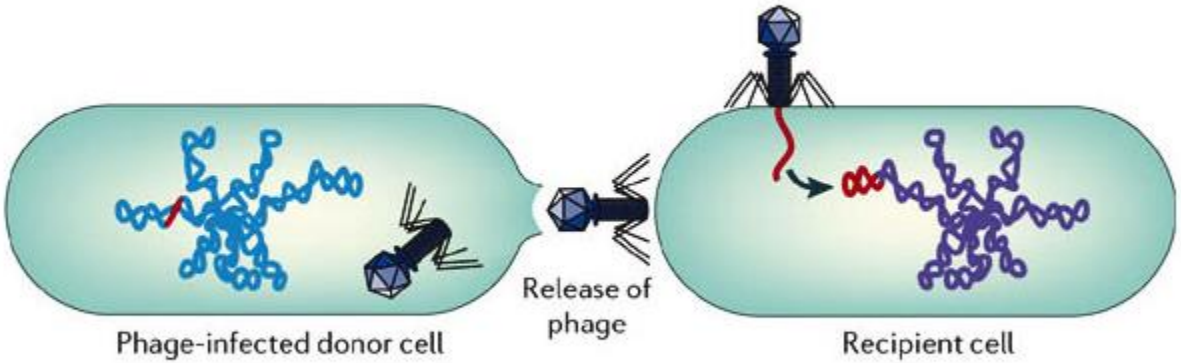
3. Both bacteria now contain the plasmid.

4. The recipient may even integrate the plasmid into its chromosome.




Transposon
Donor cell
Recipient cell


Release of DNA
Donor cell
Antibiotic-resistance gene
Recipient cell


Phage-infected donor cell
Release of phage
Recipient cell

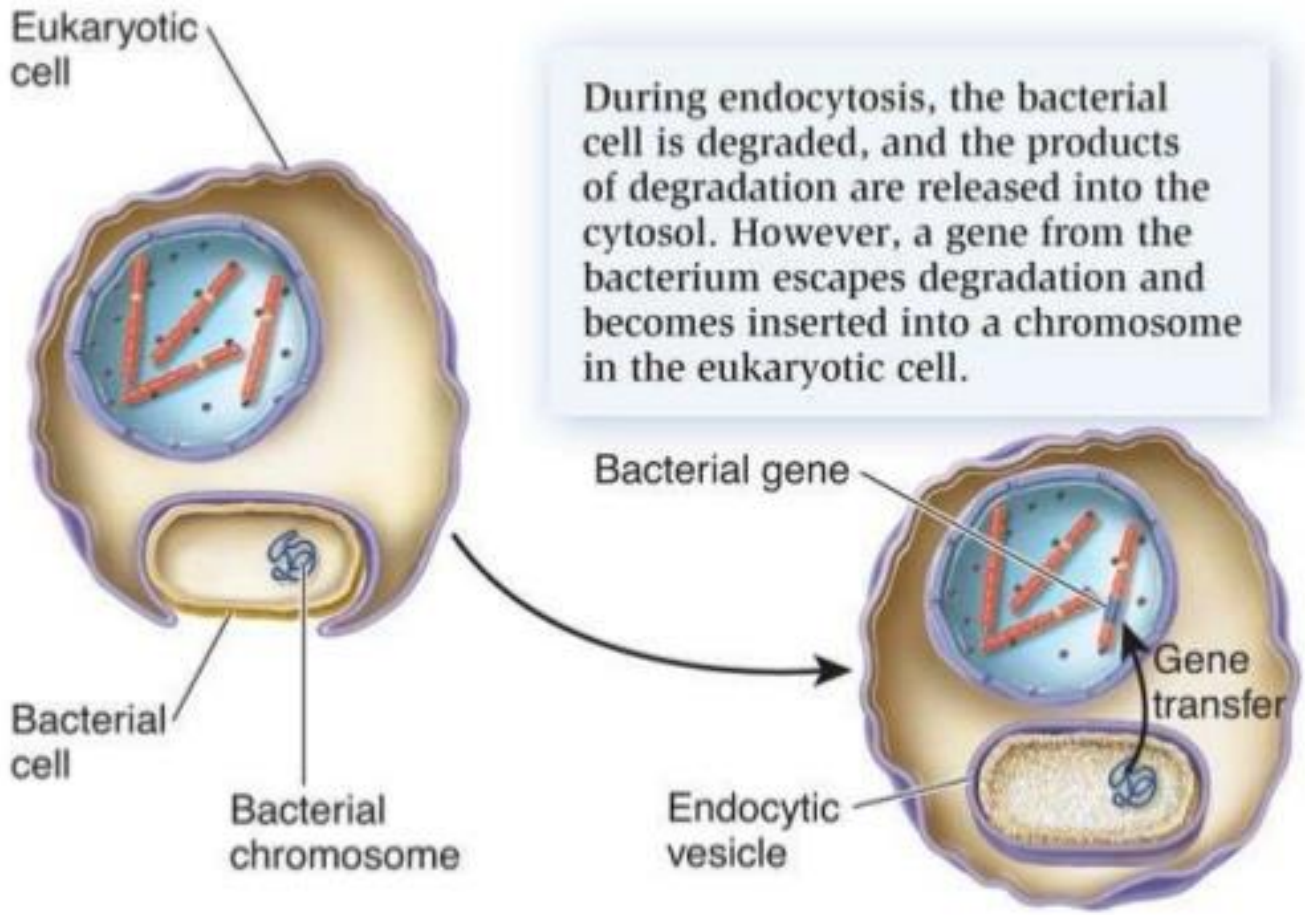
- In their travels, viruses can accidentally pick up fragments of DNA from the genome of one host cell and ferry them into another cell.

- Such transfers of genetic material frequently occur in procaryotes, and they can also occur between eucaryotic cells of the same species.

- Horizontal transfers of genes between eucaryotic cells of different species are very rare, and they do not seem to have played a significant part in eucaryote evolution (although massive transfers from bacterial to eucaryotic genomes have occurred in the evolution of mitochondria and chloroplasts).

- Horizontal gene transfer among procaryotes may seem a surprising process, but it has a parallel in a phenomenon familiar to us all: sex.

Horizontal Gene Transfer in Eukaryote

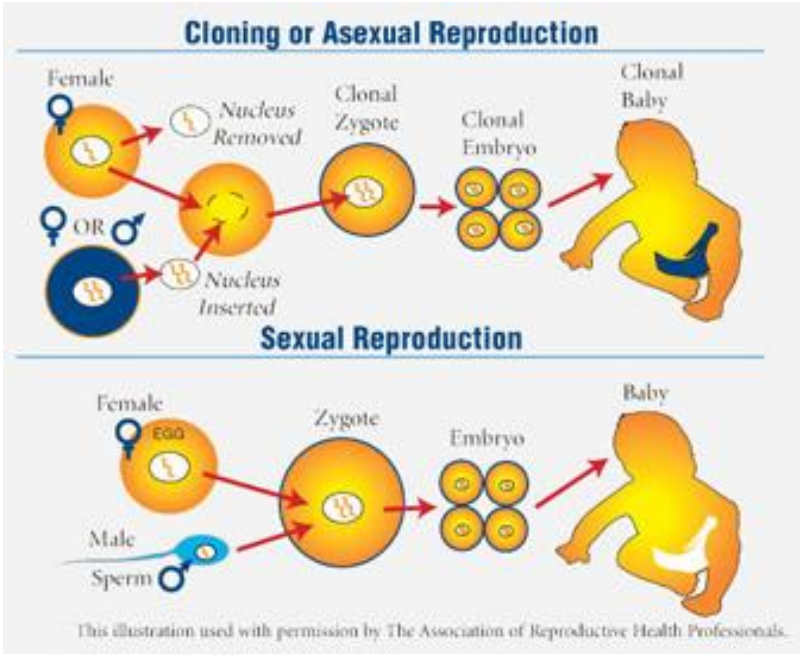


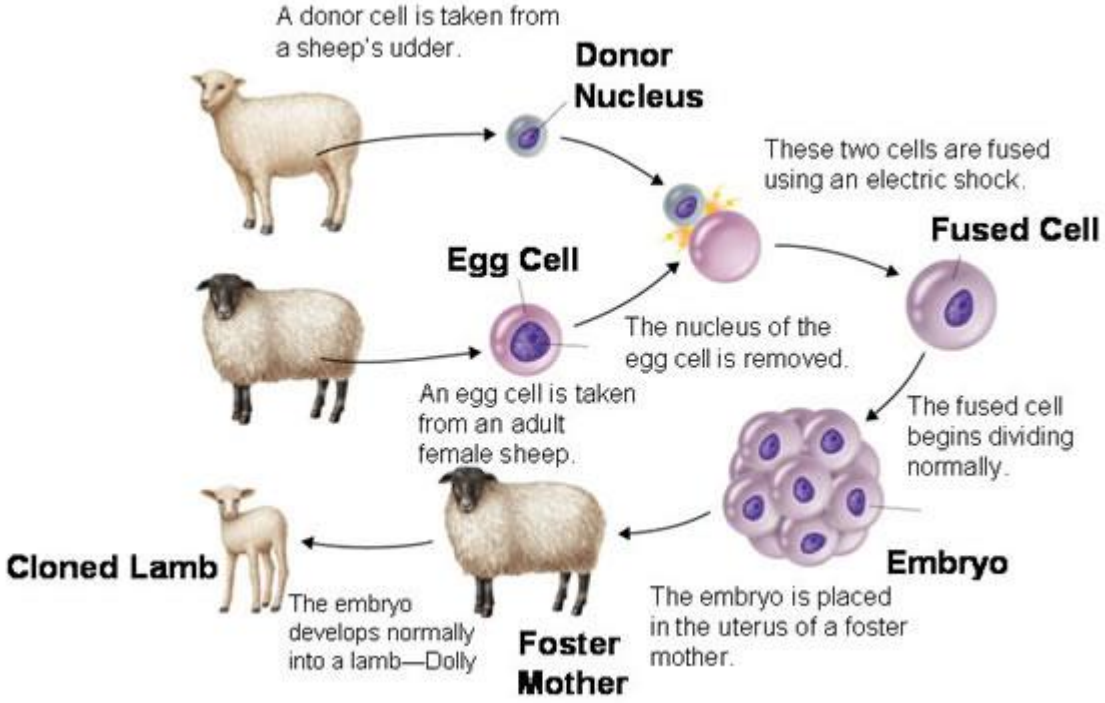
- In addition to the usual vertical transfer of genetic material from parent to offspring, sexual reproduction causes a large-scale horizontal transfer of genetic information between two initially separate cell lineages—those of the father and the mother.

- A key feature of sex, of course, is that the genetic exchange normally occurs only between individuals of the same species.

- Sexual reproduction is widespread (although not universal), especially among eucaryotes.

- Even bacteria indulge from time to time in controlled sexual exchanges of DNA with other members of their own species.


Cloning or Asexual Reproduction
Female
Nucleus Removed
Clonal Zygote
Clonal Embryo
Clonal Baby
OR
Nucleus Inserted
Sexual Reproduction
Female
EGG
Zygote
Embryo
Baby
Male
Sperm
This illustration used with permission by The Association of Reproductive Health Professionals.


A donor cell is taken from a sheep's udder.
Donor Nucleus
These two cells are fused using an electric shock.
Fused Cell
Egg Cell
The nucleus of the egg cell is removed.
An egg cell is taken from an adult female sheep.
The fused cell begins dividing normally.
Embryo
Cloned Lamb
The embryo develops normally into a lamb—Dolly
Foster Mother
The embryo is placed in the uterus of a foster mother.

- Without additional information, no amount of gazing at genome sequences will reveal the functions of genes.

- We may recognize that gene B is like gene A, but how do we discover the function of gene A in the first place?

- And even if we know the function of gene A, how do we test whether the function of gene B is truly the same as the sequence similarity suggests?

- How do we connect the world of abstract genetic information with the world of real living organisms?

- The analysis of gene functions depends on two complementary approaches: genetics and biochemistry.

- Genetics starts with the study of mutants: we either find or make an organism in which a gene is altered, and examine the effects on the organism's structure and performance.

- Biochemistry examines the functions of molecules: we extract molecules from an organism and then study their chemical activities.

- By combining genetics and biochemistry and examining the chemical abnormalities in a mutant organism, it is possible to find those molecules whose production depends on a given gene.

- At the same time, studies of the performance of the mutant organism show us what role those molecules have in the operation of the organism as a whole.

- Thus, genetics and biochemistry together provide a way to relate genes, molecules, and the structure and function of the organism.

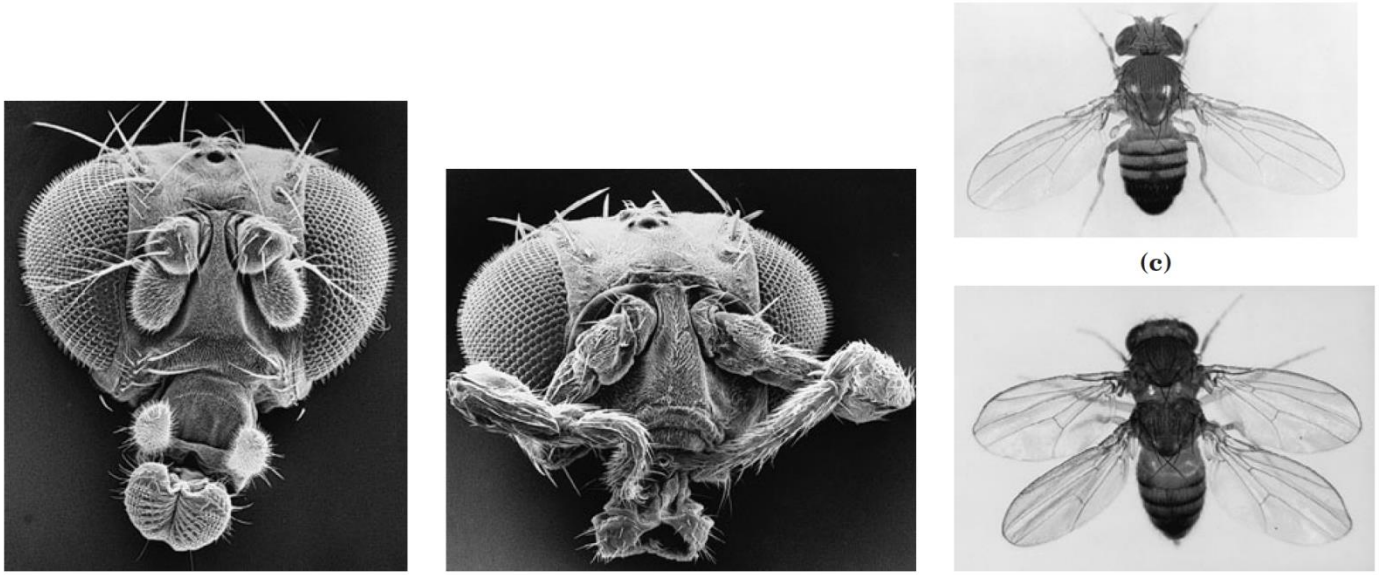


FIGURE 28–43 **Effects of mutations in *Hox* genes in *Drosophila*.** **(a)** Normal head. **(b)** Homeotic mutant (*antennapedia*) in which antennae are replaced by legs. **(c)** Normal body structure. **(d)** Homeotic mutant (*bithorax*) in which a segment has developed incorrectly to produce an extra set of wings.

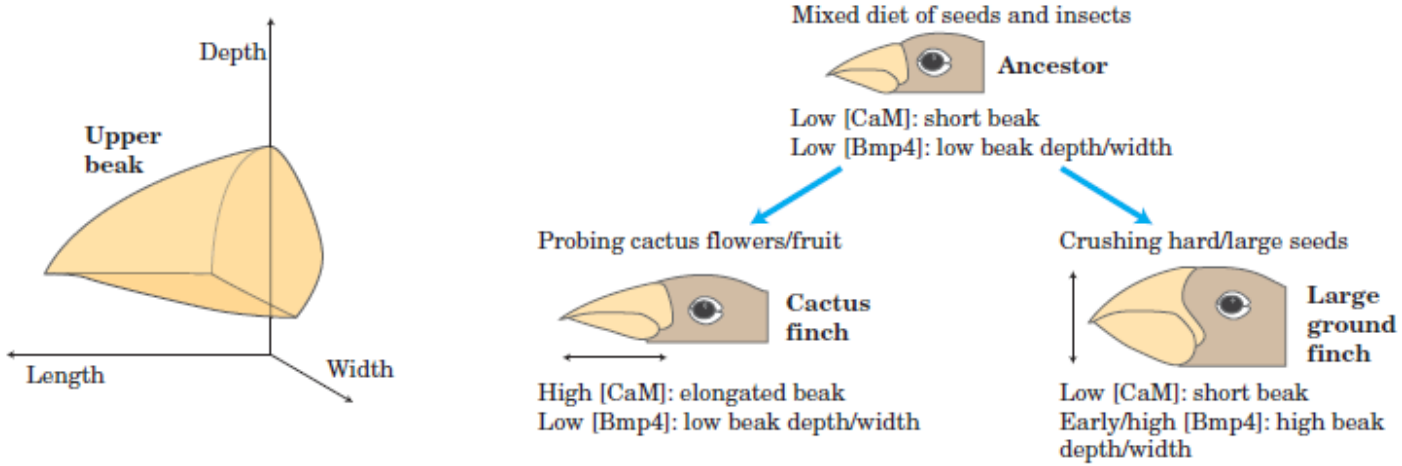
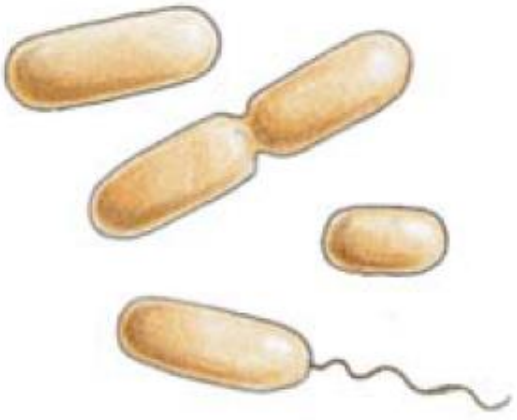


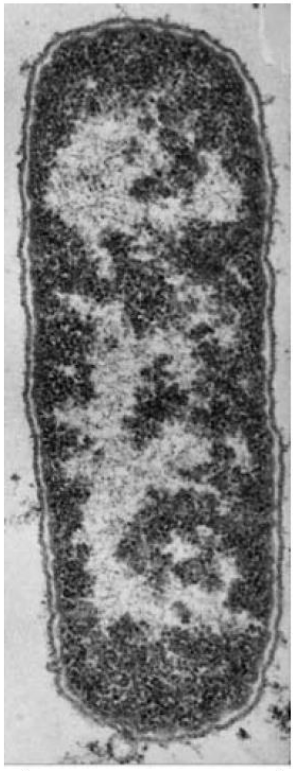
FIGURE 1 Evolution of new beak structures to exploit new food sources. In the Galapagos finches, the different beak structures of the cactus finch and the large ground finch, which feed on different, specialized food sources, were produced to a large extent by a few mutations that altered the timing and level of expression of just two genes: those encoding calmodulin (CaM) and Bmp4.

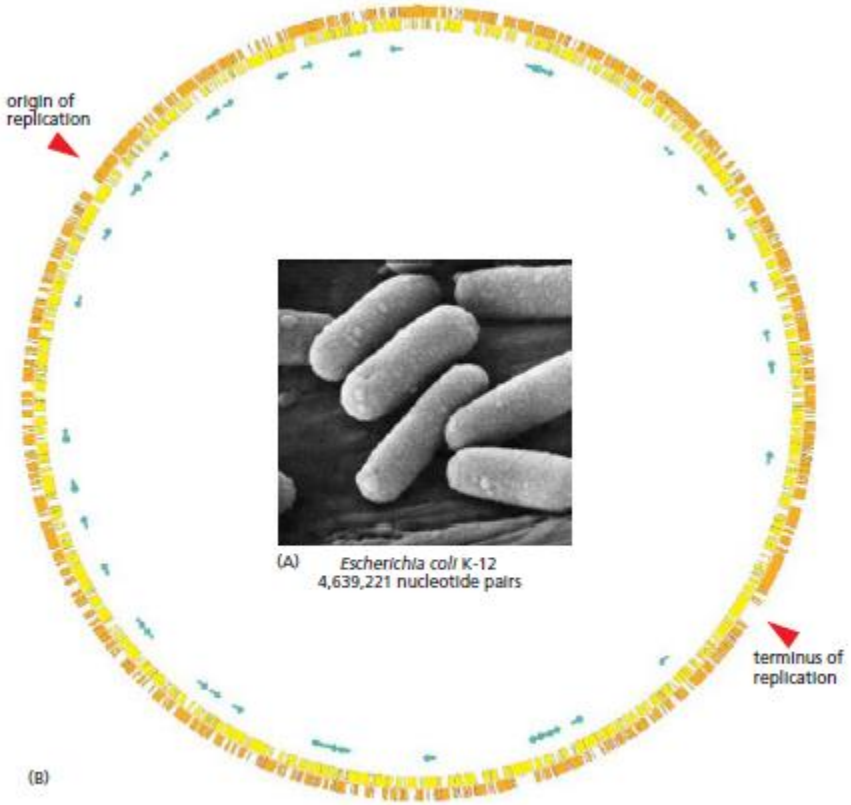
- Because living organisms are so complex, large communities of biologists have become dedicated to studying different aspects of the same model organism.

- In the enormously varied world of bacteria, the spotlight of molecular biology has for a long time focused intensely on just one species: Escherichia coli, or *E. Coli.*



rod-shaped cells
e.g., *Escherichia coli*,
Vibrio cholerae





- This small, rod-shaped bacterial cell normally lives in the gut of humans and other vertebrates, but it can be grown easily in a simple nutrient broth in a culture bottle.

- It adapts to variable chemical conditions and reproduces rapidly, and it can evolve by mutation and selection at a remarkable speed.

- As with other bacteria, different strains of *E. coli*, though classified as members of a single species, differ genetically to a much greater degree than do different varieties of a sexually reproducing organism such as a plant or animal.

- One E. coli strain may possess many hundreds of genes that are absent from another, and the two strains could have as little as 50% of their genes in common.

- In molecular terms, we know more about *E. coli* than about any other living organism. Most of our understanding of the fundamental mechanisms of life has come from studies of *E. Coli*.

- Eucaryotic cells, in general, are bigger and more elaborate than procaryotic cells, and their genomes are bigger and more elaborate, too.

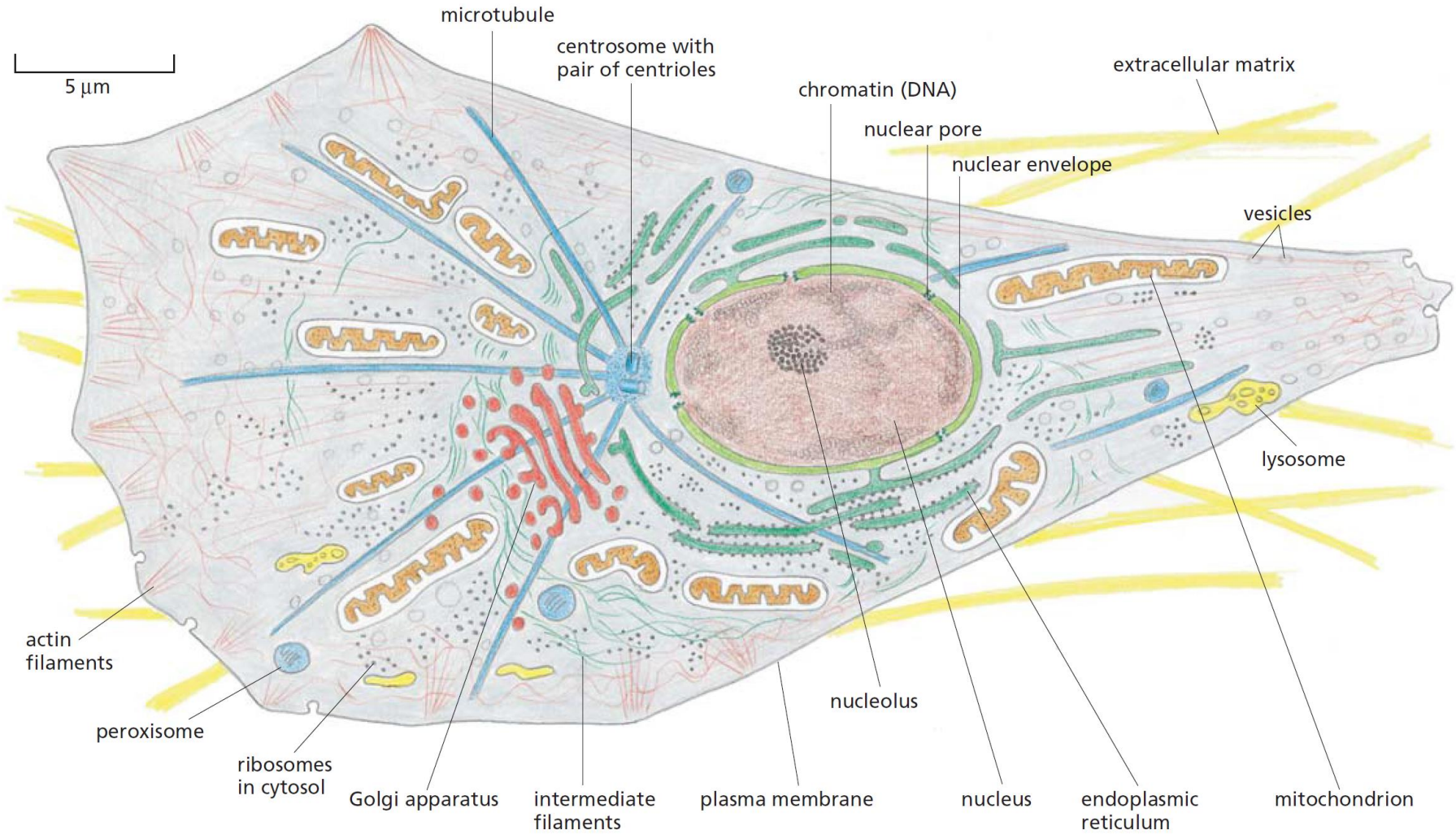
- The greater size is accompanied by radical differences in cell structure and function.

- By definition, eucaryotic cells keep their DNA in an internal compartment called the nucleus.

- The nuclear envelope, a double layer of membrane, surrounds the nucleus and separates the DNA from the cytoplasm.

- Eucaryotes also have other features that set them apart from procaryotes; they have a cytoskeleton and the nuclear envelope is only one part of a set of internal membranes.

- Lacking the tough cell wall of most bacteria, animal cells and the free-living eucaryotic cells called protozoa can change their shape rapidly and engulf other cells and small objects by phagocytosis.


5 µm
microtubule
centrosome with pair of centrioles
chromatin (DNA)
nuclear pore
nuclear envelope
extracellular matrix
vesicles
lysosome
actin filaments
peroxisome
ribosomes in cytosol
Golgi apparatus
intermediate filaments
plasma membrane
nucleolus
nucleus
endoplasmic reticulum
mitochondrion

- It is still a mystery how all these properties evolved, and in what sequence.

- One plausible view, however, is that they are all reflections of the way of life of a primordial eucaryotic cell that was a predator, living by capturing other cells and eating them.

- Such a way of life requires a large cell with a flexible plasma membrane, as well as an elaborate cytoskeleton to support and move this membrane.

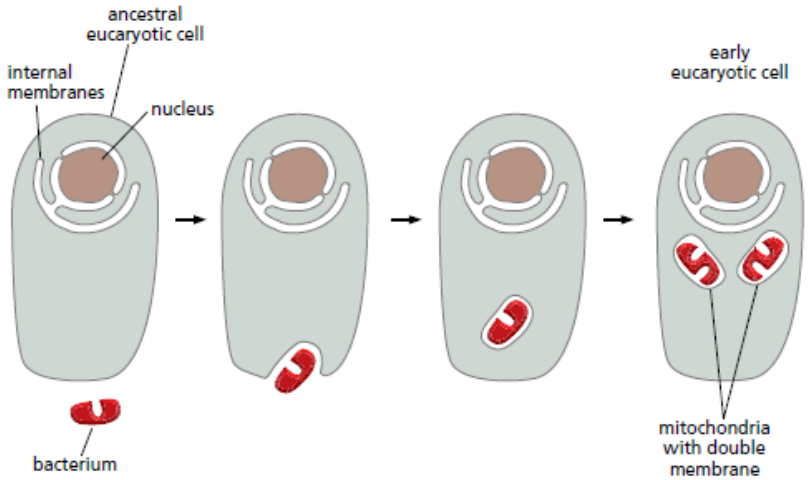
- It may also require that the cell's long, fragile DNA molecules be sequestered in a separate nuclear compartment, to protect the genome from damage by the movements of the cytoskeleton.

- A predatory way of life helps to explain another feature of eucaryotic cells. Almost all such cells contain mitochondria

- These small bodies in the cytoplasm, enclosed by a double layer of membrane, take up oxygen and harness energy from the oxidation of food molecules—such as sugars—to produce most of the ATP that powers the cell's activities.

- It is now generally accepted that mitochondria originated from free-living oxygen-metabolizing (aerobic) bacteria that were engulfed by an ancestral eucaryotic cell that could otherwise make no such use of oxygen (that is, was anaerobic).

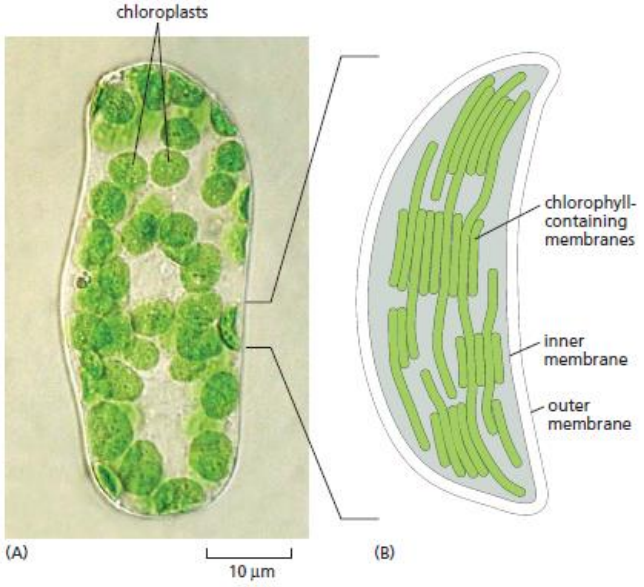
- This partnership between a primitive anaerobic eucaryotic predator cell and an aerobic bacterial cell is thought to have been established about 1.5 billion years ago, when the Earth's atmosphere first became rich in oxygen.

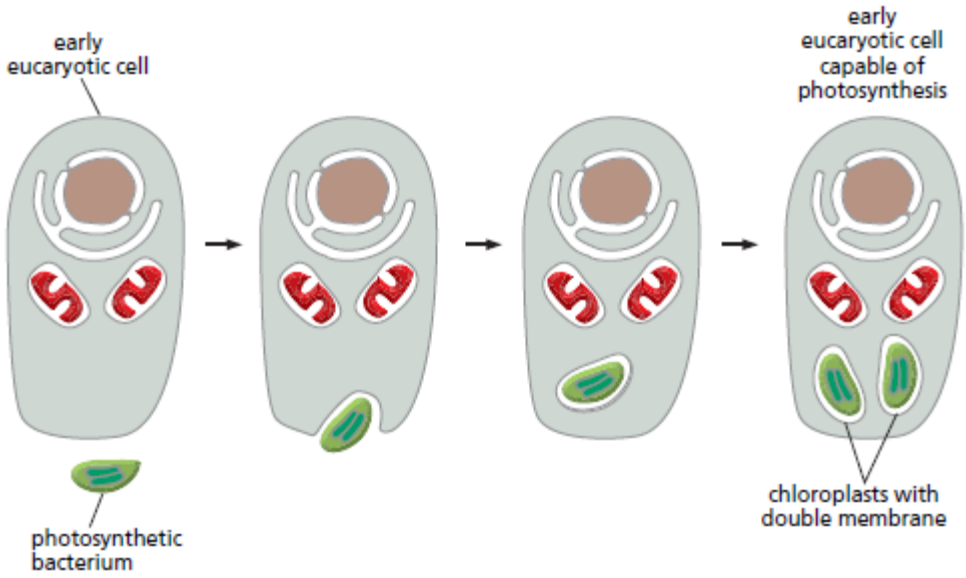


- Many eucaryotic cells—specifically, those of plants and algae—also contain another class of small membrane-enclosed organelles somewhat similar to mitochondria— the chloroplasts.

- Chloroplasts perform photosynthesis, using the energy of sunlight to synthesize carbohydrates from atmospheric carbon dioxide and water, and deliver the products to the host cell as food.

- Like mitochondria, chloroplasts have their own genome and almost certainly originated as symbiotic photosynthetic bacteria, acquired by cells that already possessed mitochondria .



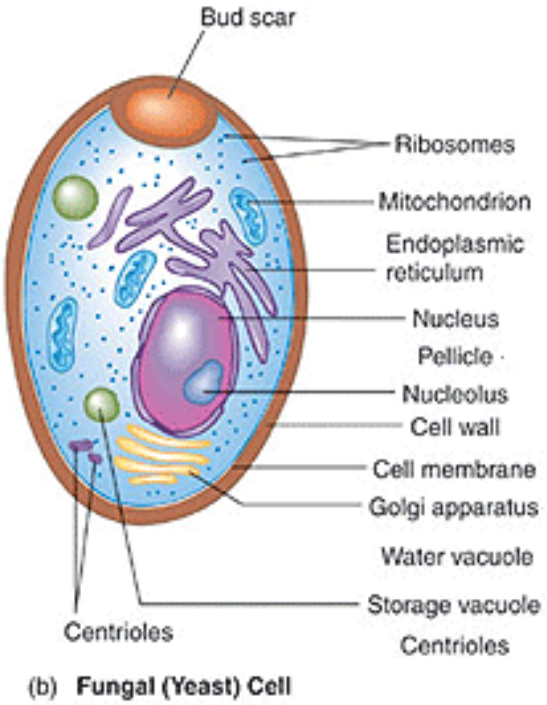


- If the ancestral eucaryote was indeed a predator on other organisms, we can view plant cells as eucaryotes that have made the transition from hunting to farming.

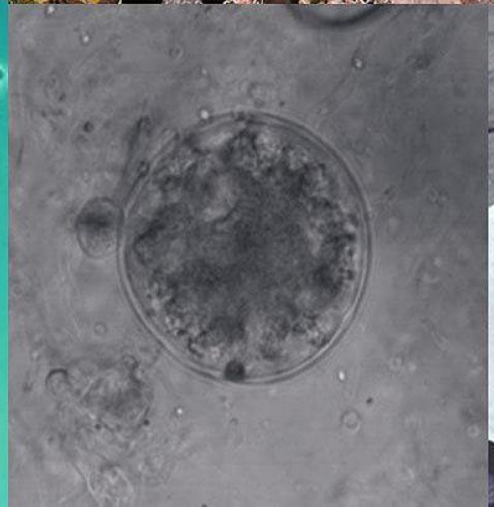
- Fungi represent yet another eucaryotic way of life.

- Fungal cells, like animal cells, possess mitochondria but not chloroplasts; but in contrast with animal cells and protozoa, they have a tough outer wall that limits their ability to move rapidly or to swallow up other cells.

- Fungi, it seems, have turned from hunters into scavengers: other cells secrete nutrient molecules or release them upon death, and fungi feed on these leavings—performing whatever digestion is necessary extracellularly, by secreting digestive enzymes to the exterior.



(b) **Fungal (Yeast) Cell**



- Natural selection has evidently favored mitochondria with small genomes, just as it has favored bacteria with small genomes.

- By contrast, the nuclear genomes of most eucaryotes seem to have been free to enlarge.

- Perhaps the eucaryotic way of life has made large size an advantage: predators typically need to be bigger than their prey, and cell size generally increases in proportion to genome size.

- Perhaps enlargement of the genome has been driven by the accumulation of parasitic transposable elements –segments of DNA that can insert copies of themselves at multiple sites in the genome.

- Whatever the explanation, the genomes of most eucaryotes are orders of magnitude larger than those of bacteria and archaea.

- Eucaryotes not only have more genes than procaryotes; they also have vastly more DNA that does not code for protein or for any other functional product molecule.

- Much of our noncoding DNA is almost certainly dispensable junk, there is more noncoding DNA than coding DNA, however at least some of the noncoding DNA certainly has important functions.

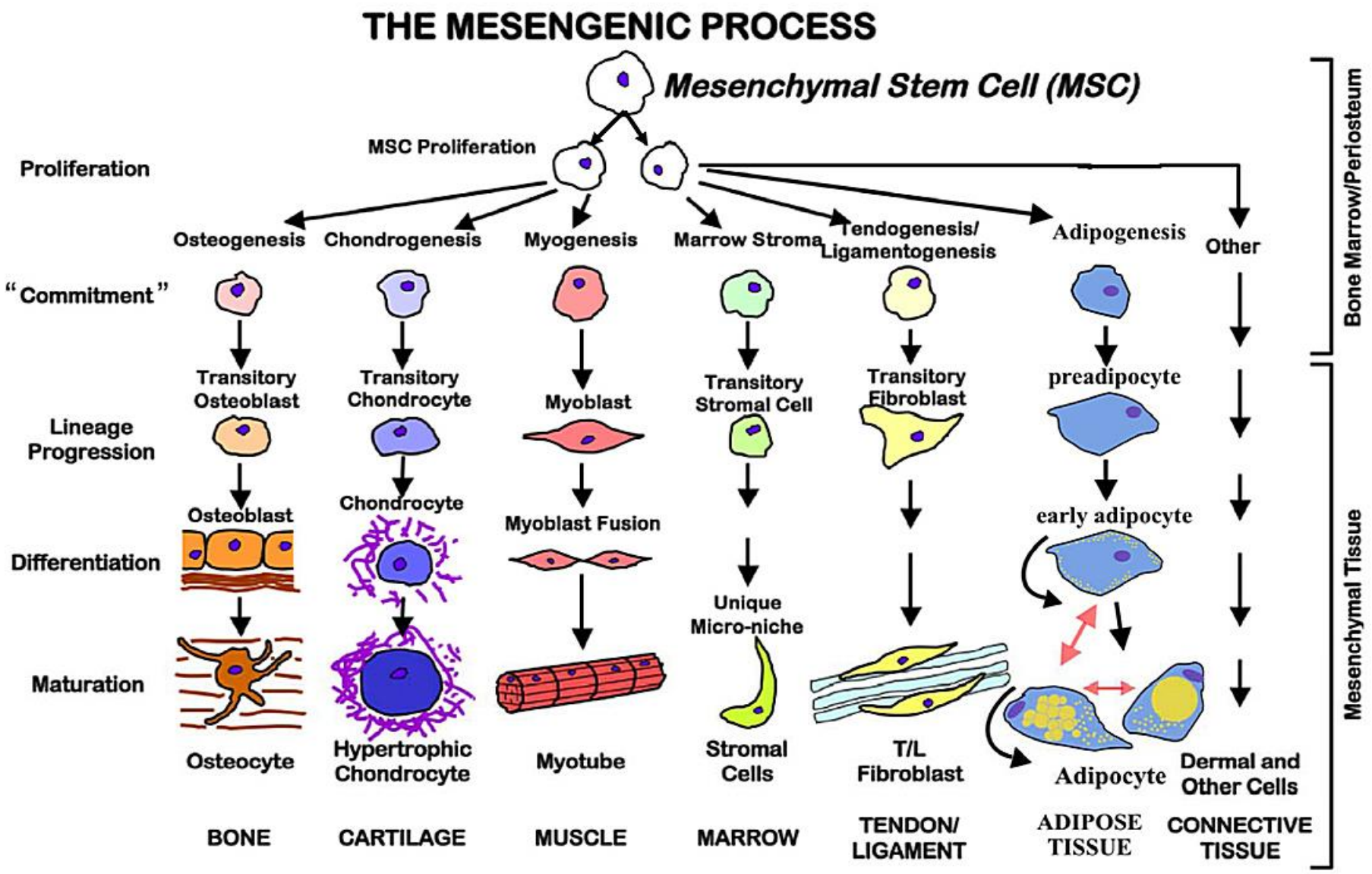
- In particular, it regulates the expression of adjacent genes.

- With this regulatory DNA, eucaryotes have evolved distinctive ways of controlling when and where a gene is brought into play.

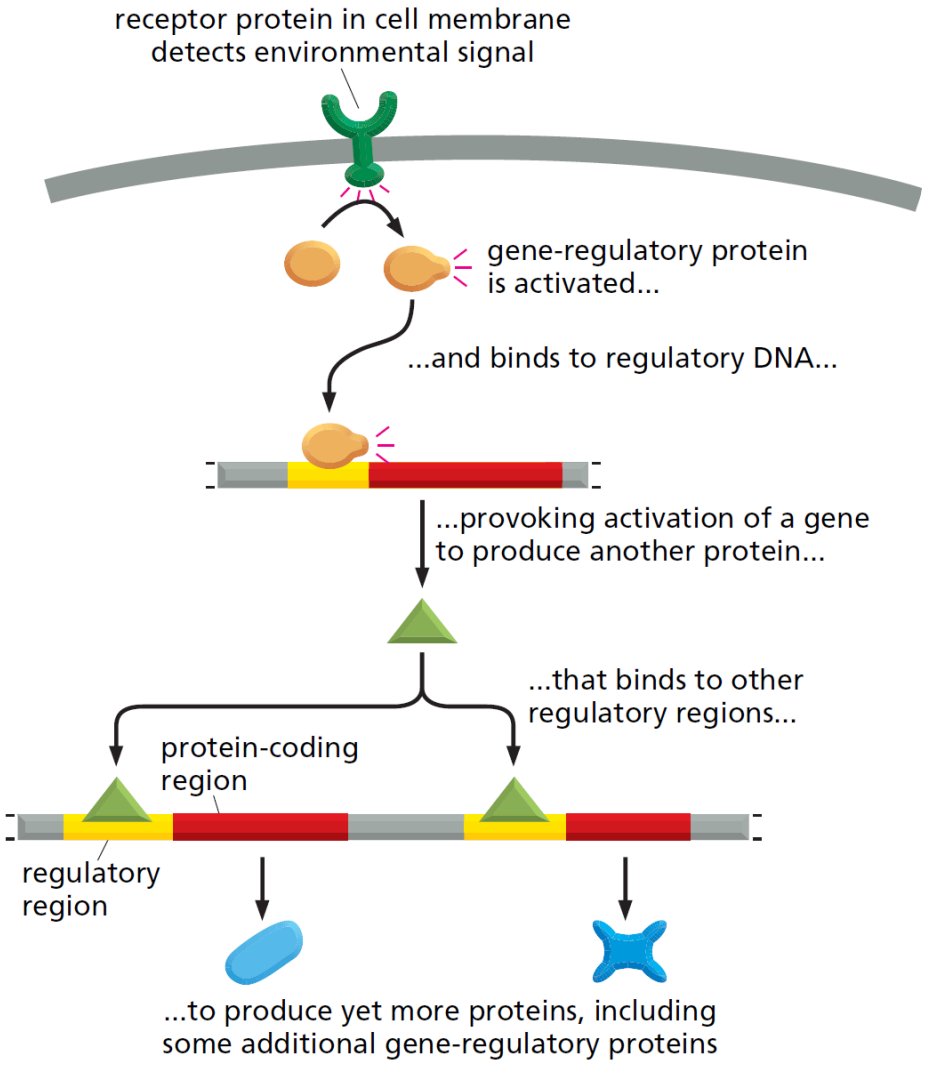
- This sophisticated gene regulation is crucial for the formation of complex multicellular organisms.

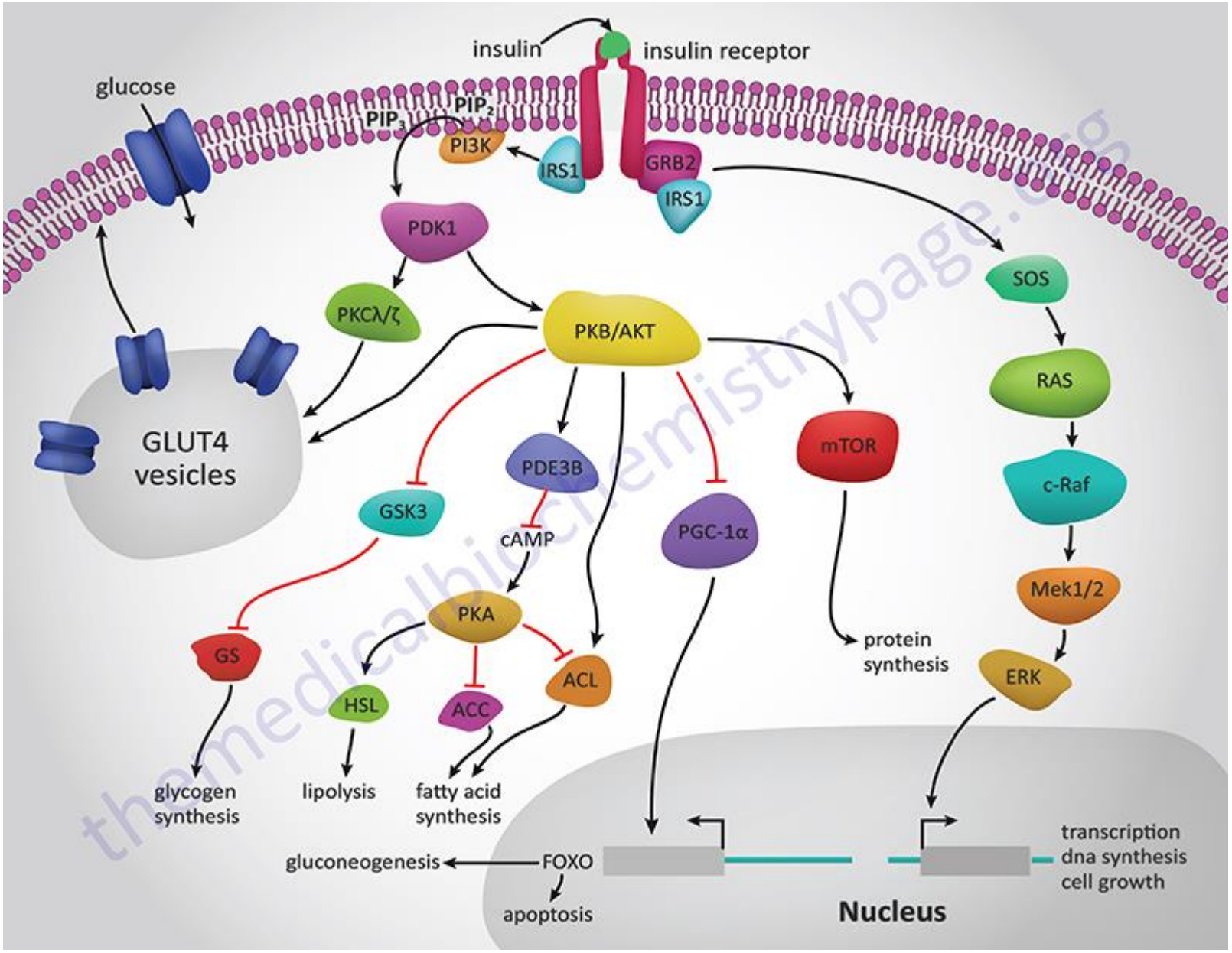
- The cells in an individual animal or plant are extraordinarily varied. Fat cells, skin cells, bone cells, nerve cells—they seem as dissimilar as any cells could be.

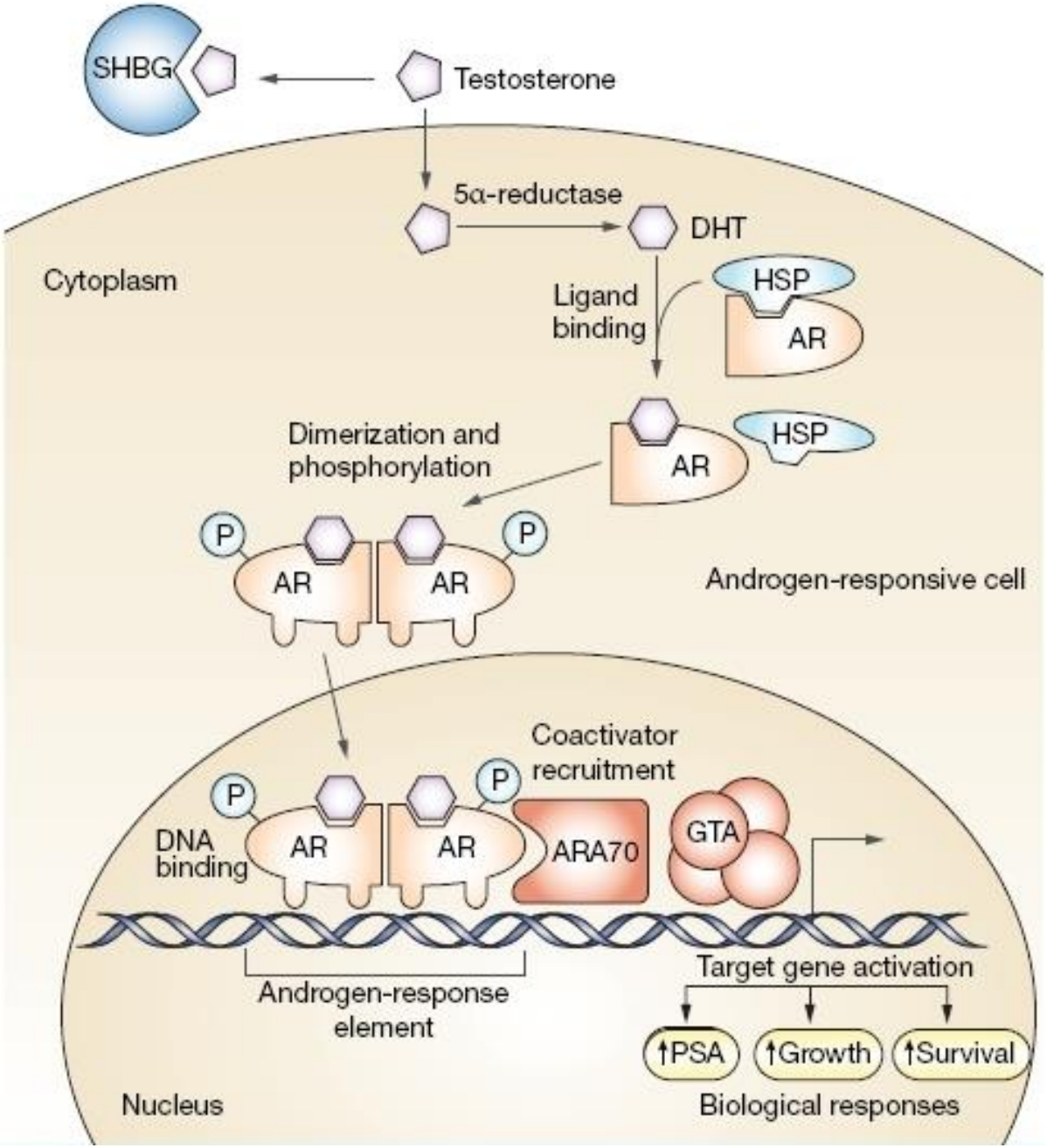
- Yet all these cell types are the descendants of a single fertilized egg cell, and all (with minor exceptions) contain identical copies of the genome of the species.


THE MESENGENIC PROCESS
Mesenchymal Stem Cell (MSC)
MSC Proliferation
Proliferation
Osteogenesis
Chondrogenesis
Myogenesis
Marrow Stroma
Tendogenesis/ Ligamentogenesis
Adipogenesis
Other
"Commitment"
Transitory Osteoblast
Transitory Chondrocyte
Myoblast
Transitory Stromal Cell
Transitory Fibroblast
preadipocyte
Lineage Progression
Osteoblast
Chondrocyte
Myoblast Fusion
early adipocyte
Differentiation
Unique Micro-niche
Maturation
Osteocyte
Hypertrophic Chondrocyte
Myotube
Stromal Cells
T/L Fibroblast
Adipocyte
Dermal and Other Cells
BONE
CARTILAGE
MUSCLE
MARROW
TENDON/ LIGAMENT
ADIPOSE TISSUE
CONNECTIVE TISSUE
Bone Marrow/Periosteum
Mesenchymal Tissue

- The differences result from the way in which the cells make selective use of their genetic instructions according to the cues they get from their surroundings in the developing embryo.




insulin
insulin receptor
glucose
PIP3
PIP2
PI3K
IRS1
GRB2
IRS1
PDK1
PKCλ/ζ
PKB/AKT
SOS
RAS
c-Raf
Mek1/2
ERK
GLUT4
vesicles
GSK3
PDE3B
cAMP
PGC-1α
mTOR
protein
synthesis
PKA
GS
HSL
ACC
ACL
glycogen
synthesis
lipolysis
fatty acid
synthesis
gluconeogenesis
FOXO
apoptosis
transcription
dna synthesis
cell growth
Nucleus


SHBG
Testosterone
5α-reductase
DHT
Cytoplasm
Ligand binding
HSP
AR
Dimerization and phosphorylation
AR
HSP
P
AR
AR
P
Androgen-responsive cell
Coactivator recruitment
P
P
DNA binding
AR
AR
ARA70
GTA
Androgen-response element
Target gene activation
↑PSA
↑Growth
↑Survival
Nucleus
Biological responses

- Many species of eucaryotic cells lead a solitary life—some as hunters (the protozoa), some as photosynthesizers (the unicellular algae), some as scavengers (the unicellular fungi, or yeasts).

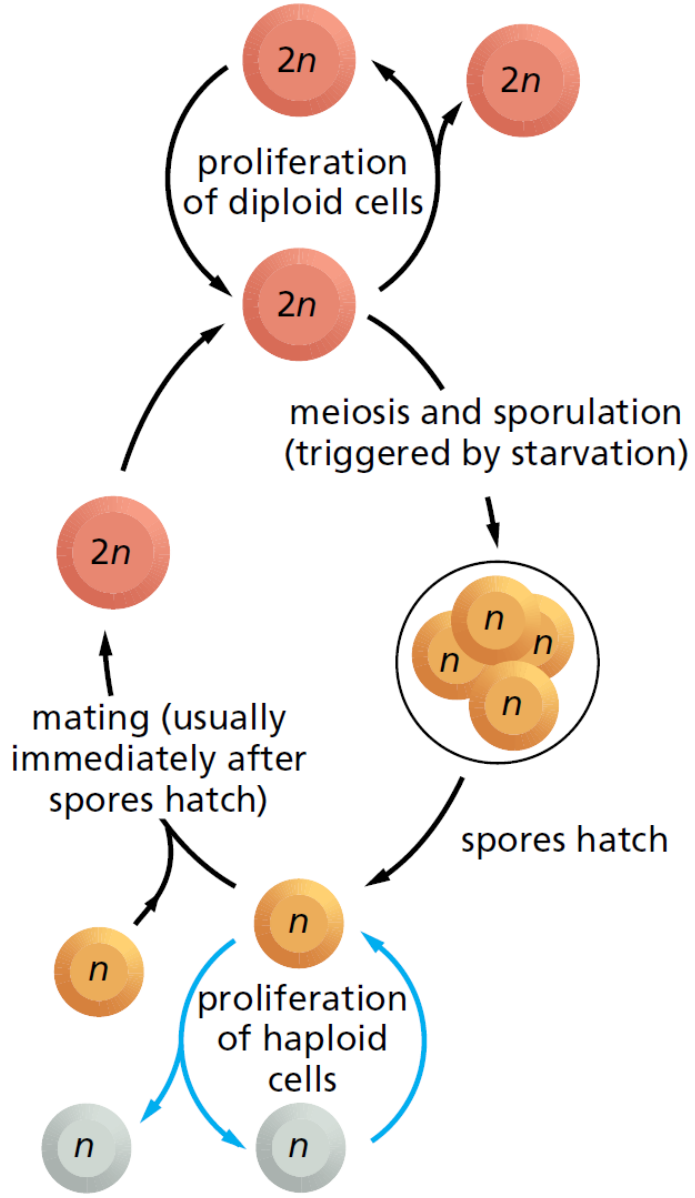
- In terms of their ancestry and DNA sequences, protists are far more diverse than the multicellular animals, plants, and fungi.

- The molecular and genetic complexity of eucaryotes is daunting. Even more than for procaryotes, biologists need to concentrate their limited resources on a few selected model organisms to fathom this complexity.

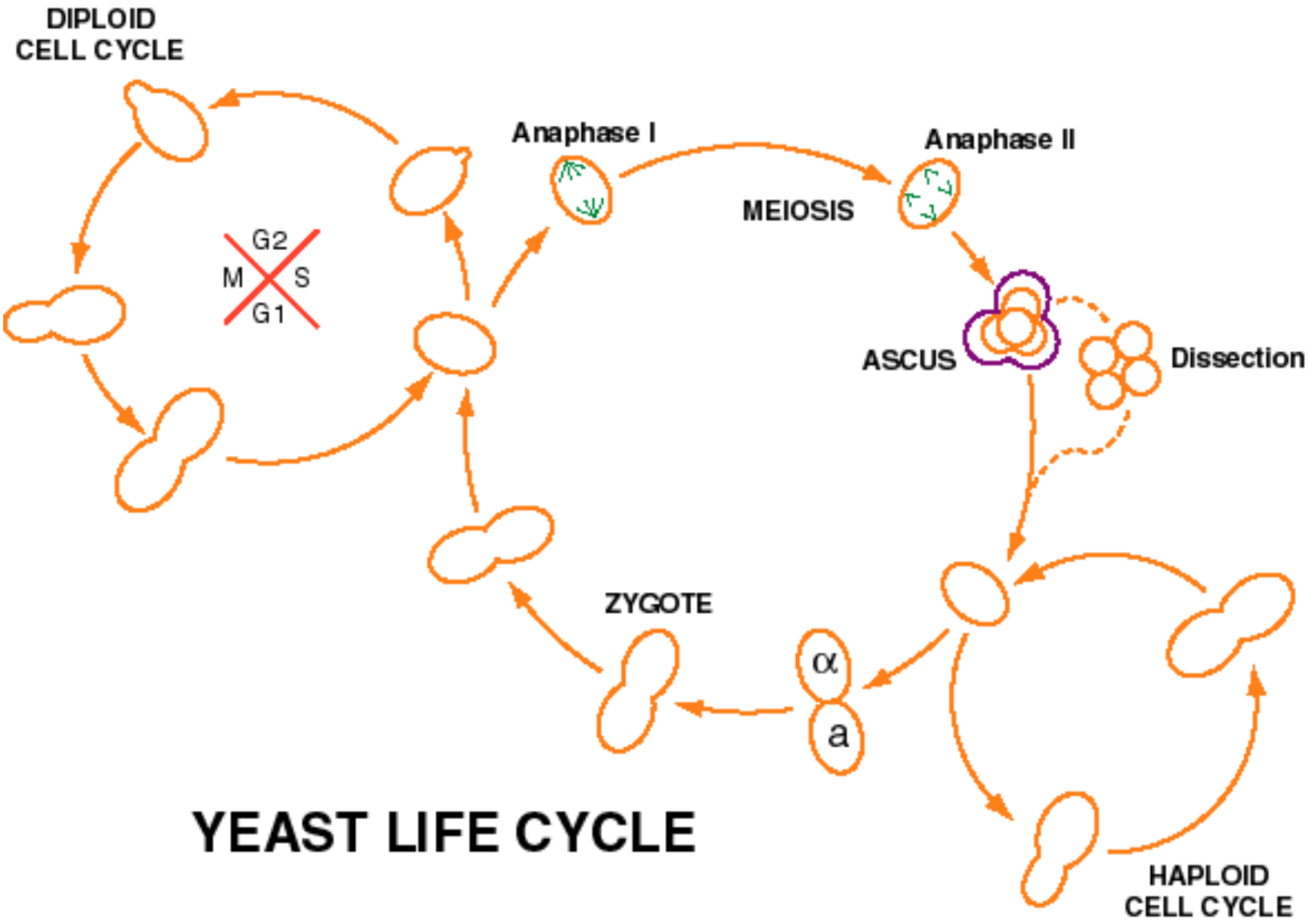
- The popular choice for this role of minimal model eucaryote has been the yeast *Saccharomyces cerevisiae* —the same species that is used by brewers of beer and bakers of bread.

- S. cerevisiae is a small, single-celled member of the kingdom of fungi and thus, it is closely related to animals as it is to plants.

- When nutrients are plentiful, it grows and divides almost as rapidly as a bacterium.
- It can reproduce either vegetatively (that is, by simple cell division), or sexually: two yeast cells that are haploid (possessing a single copy of the genome) can fuse to create a cell that is diploid (containing a double genome).
- The diploid cell can undergo meiosis (a reduction division) to produce cells that are once again haploid.



BUDDING YEAST LIFE CYCLE



YEAST LIFE CYCLE

- Because of the close evolutionary relationship between all flowering plants, we can, once again, get insight into the cell and molecular biology of this whole class of organisms by focusing on just one or a few species for detailed analysis.

- Out of the several hundred thousand species of flowering plants on Earth today, molecular biologists have chosen to concentrate their efforts on a small weed, the common Thale cress *Arabidopsis thaliana.*

- It can be grown indoors in large numbers, and produces thousands of offspring per plant after 8–10 weeks.

- Multicellular animals account for the majority of all named species of living organisms, and for the largest part of the biological research effort.

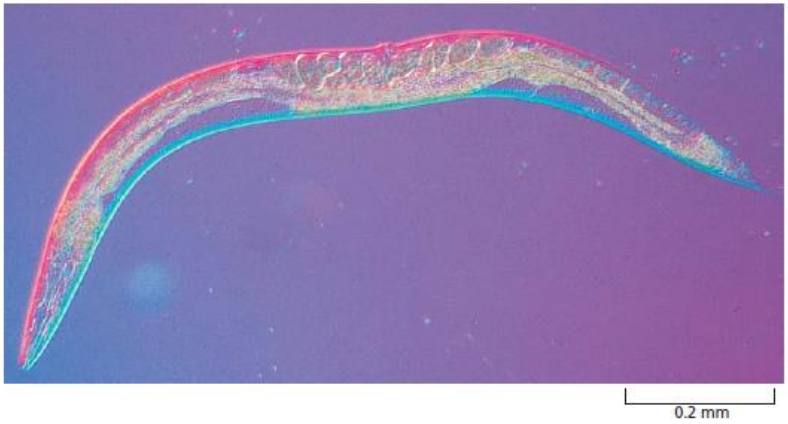
- Four species have emerged as the foremost model organisms for molecular genetic studies.

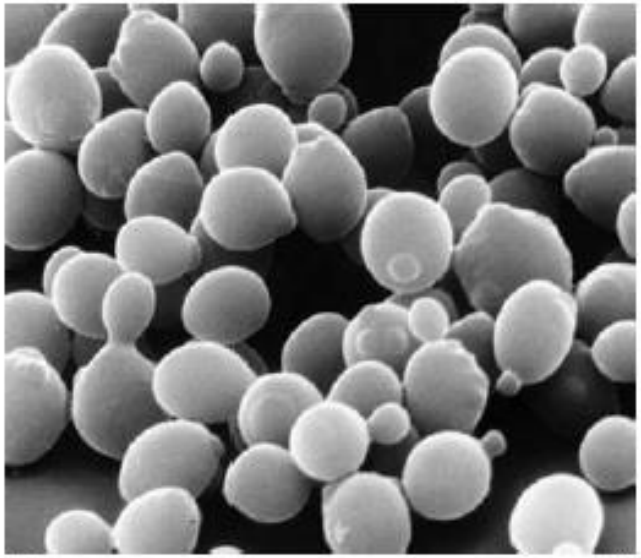
- In order of increasing size, they are the nematode worm *Caenorhabditis elegans*, the fly *Drosophila melanogaster*, the mouse *Mus musculus*, and the human, *Homo sapiens*.

- Each of these has had its genome sequenced.

- *Caenorhabditis elegans* is a small, harmless relative of the eelworm that attacks crops.

- With a life cycle of only a few days, an ability to survive in a freezer indefinitely in a state of suspended animation, a simple body plan, and an unusual life cycle that is well suited for genetic studies, it is an ideal model organism.





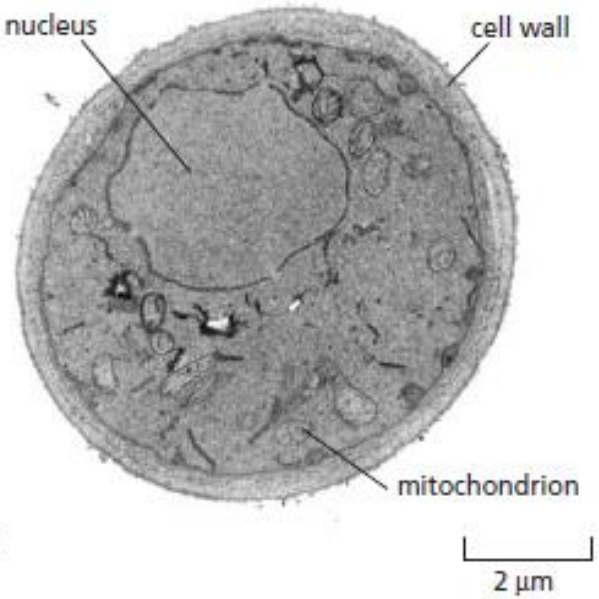
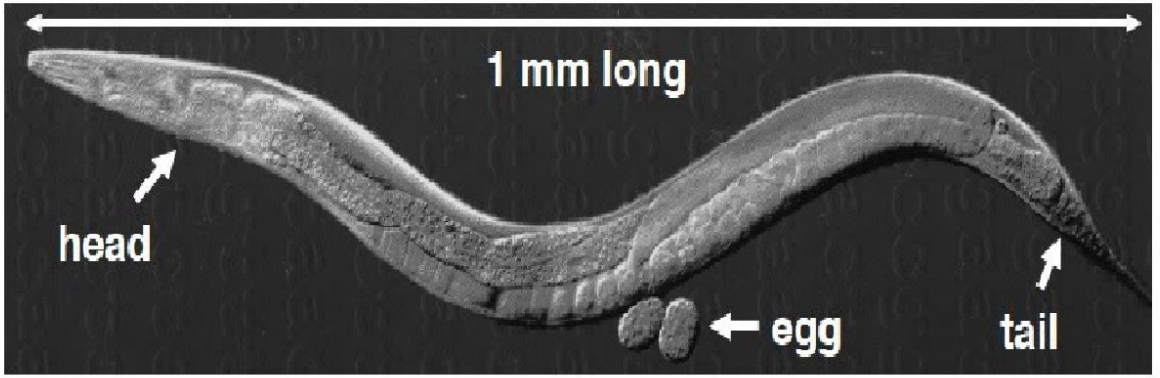
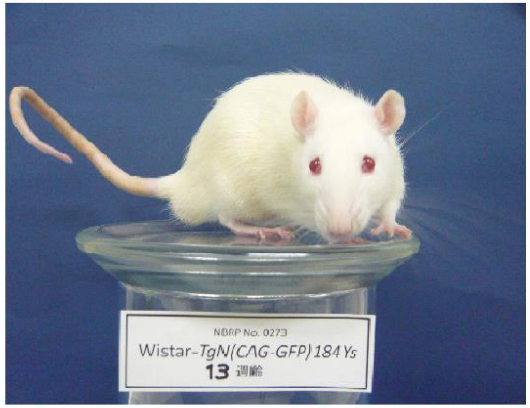


Figure 1–46 ***Arabidopsis thaliana*****, the plant chosen as the primary model for studying plant molecular genetics.** (Courtesy of Toni Hayden and the John Innes Foundation.)

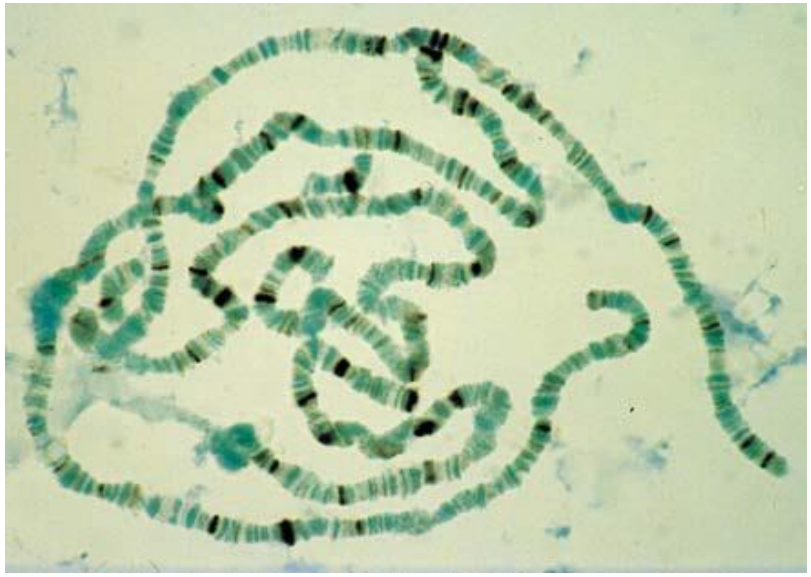

1 mm long
head
egg
tail


NBRP No. 0273
Wistar-TgN(CAG-GFP)184Ys
13 週齢

- The fruitfly *Drosophila melanogaster* has been used as a model genetic organism for longer than any other; in fact, the foundations of classical genetics were built to a large extent on studies of this insect.

- *Drosophila have giant chromosomes, with characteristic banded appearance, that are visible in some of its cells.*

- Specific changes in the hereditary information, manifest in families of mutant flies, were found to correlate exactly with the loss or alteration of specific giant-chromosome bands.



- Drosophila mutants with body parts strangely misplaced or mispatterned provided the key to the identification and characterization of the genes required to make a properly structured body, with gut, limbs, eyes, and all the other parts in their correct places.

- Once these Drosophila genes were sequenced, the genomes of vertebrates could be scanned for homologs.

- These were found, and their functions in vertebrates were then tested by analyzing mice in which the genes had been mutated.

- The results, reveal an astonishing degree of similarity.

- Drosophila requires only 9 days to progress from a fertilized egg to an adult; it is vastly easier and cheaper to breed than any vertebrate, and its genome is much smaller

- Almost every gene in the vertebrate genome has paralogs—other genes in the same genome that are unmistakably related and must have arisen by gene duplication.

- In many cases, a whole cluster of genes is closely related to similar clusters present elsewhere in the genome, suggesting that genes have been duplicated in linked groups rather than as isolated individuals.

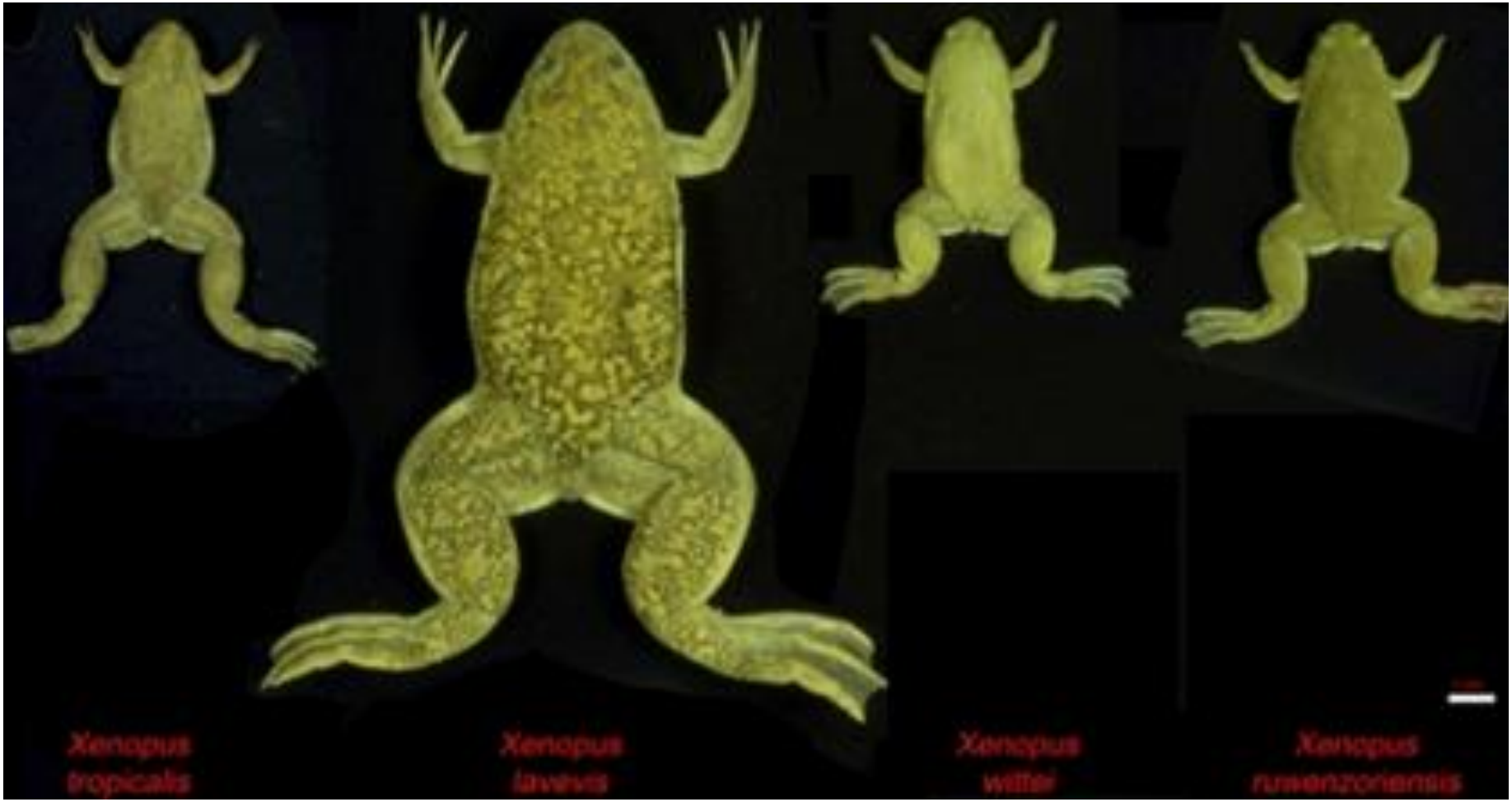
- According to one hypothesis, at an early stage in the evolution of the vertebrates, the entire genome underwent duplication twice in succession, giving rise to four copies of every gene.

- There is, however, no doubt that such whole-genome duplications do occur from time to time in evolution, for we can see recent instances in which duplicated chromosome sets are still clearly identifiable as such.

- *Xenopus*, for example, comprises a set of closely similar species related to one another by repeated duplications or triplications of the whole genome.

- Among these frogs are *X. Tropicalis* (above), with an ordinary diploid genome; the common laboratory species *X. Laevis* (below), with a duplicated genome and twice as much DNA per cell; and *X. ruwenzoriensis*, with a sixfold reduplication of the original genome and six times as much DNA per cell.

- These species are estimated to have diverged from one another within the past 120 million years


Xenopus tropicalis
Xenopus laevis
Xenopus wittei
Xenopus ruwenzoriensis

- Whatever the details of the evolutionary history, it is clear that most genes in the vertebrate genome exist in several versions that were once identical.

- The related genes often remain functionally interchangeable for many purposes.

- This phenomenon is called genetic redundancy.

- Genome duplication has clearly allowed the development of more complex life forms; it provides an organism with a cornucopia of spare gene copies, which are free to mutate to serve divergent purposes.

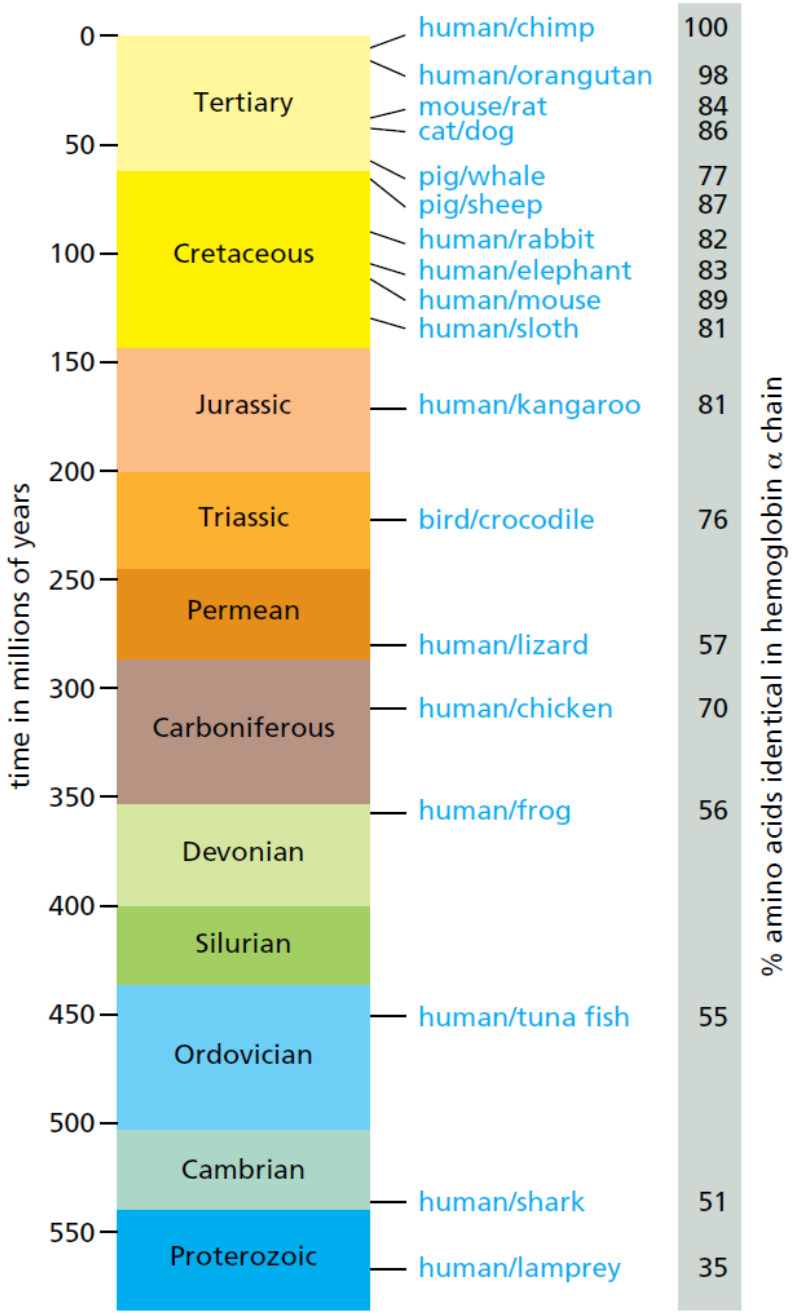
- While one copy becomes optimized for use in the liver, say, another can become optimized for use in the brain or adapted for a novel purpose.

- Mammals have typically three or four times as many genes as Drosophila, a genome that is 20 times larger, and millions or billions of times as many cells in their adult bodies.

- In terms of genome size and function, cell biology, and molecular mechanisms, mammals are nevertheless a highly uniform group of organisms.

- For a more exact measure of how closely mammalian species resemble one another genetically, we can compare the nucleotide sequences of corresponding (orthologous) genes, or the amino acid sequences of the proteins that these genes encode.

- The results for individual genes and proteins vary widely. But typically, if we line up the amino acid sequence of a human protein with that of the orthologous protein from, say, an elephant, about 85% of the amino acids are identical.


time in millions of years
0
50
100
150
200
250
300
350
400
450
500
550
Tertiary
Cretaceous
Jurassic
Triassic
Permean
Carboniferous
Devonian
Silurian
Ordovician
Cambrian
Proterozoic
human/chimp
human/orangutan
mouse/rat
cat/dog
pig/whale
pig/sheep
human/rabbit
human/elephant
human/mouse
human/sloth
human/kangaroo
bird/crocodile
human/lizard
human/chicken
human/frog
human/tuna fish
human/shark
human/lamprey
100
98
84
86
77
87
82
83
89
81
81
76
57
70
56
55
51
35
% amino acids identical in hemoglobin α chain

- The mouse, being small, hardy, and a rapid breeder, has become the foremost model organism for experimental studies of vertebrate molecular genetics.

- Many naturally occurring mutations are known, often mimicking the effects of corresponding mutations in humans.

- Methods have been developed, moreover, to test the function of any chosen mouse gene, or of any noncoding portion of the mouse genome, by artificially creating mutations in it.

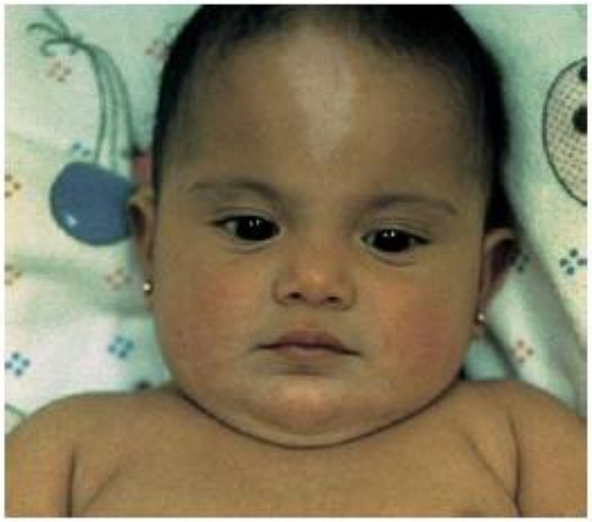


Figure 1–53 Human and mouse: similar genes and similar development. The human baby and the mouse shown here have similar white patches on their foreheads because both have mutations in the same gene (called *Kit*), required for the development and maintenance of pigment cells. (Courtesy of R.A. Fleischman.)

- What precisely do we mean when we speak of the human genome? Whose genome?

- On average, any two people taken at random differ in about one or two in every 1000 nucleotide pairs in their DNA sequence.

- The Human Genome Project has arbitrarily selected DNA from a small number of anonymous individuals for sequencing.

- The human genome—the genome of the human species—is, properly speaking, a more complex thing, embracing the entire pool of variant genes that are found in the human population and continually exchanged and reassorted in the course of sexual reproduction.

- Ultimately, we can hope to document this variation too. Knowledge of it will help us understand, for example, why some people are prone to one disease, others to another; why some respond well to a drug, others badly.

- It will also provide new clues to our history—the population movements and minglings of our ancestors, the infections they suffered, the diets they ate.

- All these things leave traces in the variant forms of genes that have survived in human communities.

- Knowledge and understanding bring the power to intervene—with humans, to avoid or prevent disease; with plants, to create better crops; with bacteria, to turn them to our own uses.

- All these biological enterprises are linked, because the genetic information of all living organisms is written in the same language.

- The new-found ability of molecular biologists to read and decipher this language has already begun to transform our relationship to the living world.