

LECTURE PRESENTATIONS

For CAMPBELL BIOLOGY, NINTH EDITION

Jane B. Reece, Lisa A. Urry, Michael L. Cain, Steven A. Wasserman, Peter V. Minorsky, Robert B. Jackson

Chapter 21

Genomes and Their Evolution



Lectures by
Erin Barley
Kathleen Fitzpatrick

Overview: Reading the Leaves from the Tree of Life

- Complete genome sequences exist for a human, chimpanzee, *E. coli*, brewer's yeast, corn, fruit fly, house mouse, rhesus macaque, and other organisms
- Comparisons of genomes among organisms provide information about the evolutionary history of genes and taxonomic groups

- **Genomics** is the study of whole sets of genes and their interactions
- **Bioinformatics** is the application of computational methods to the storage and analysis of biological data

Figure 21.1



Concept 21.1: New approaches have accelerated the pace of genome sequencing

- The most ambitious mapping project to date has been the **sequencing of the human genome**
- Officially begun as the **Human Genome Project** in 1990, the sequencing was largely completed by 2003

Whole-Genome Shotgun Approach to Genome Sequencing

- The **whole-genome shotgun approach** was developed by **J. Craig Venter in 1992**
- This approach **skips genetic and physical mapping and sequences random DNA fragments directly**
- Powerful computer programs **are used to order fragments into a continuous sequence**

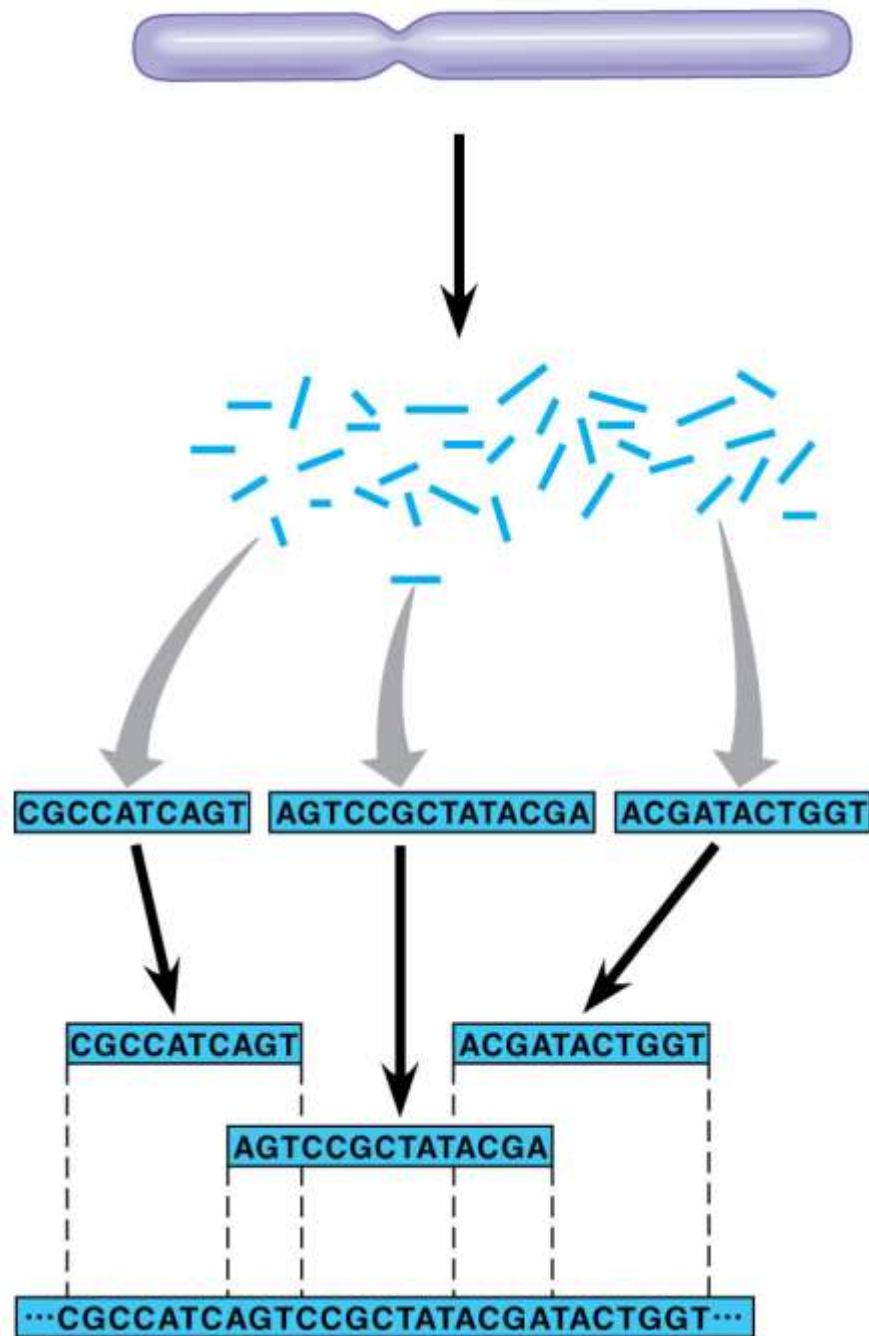
Figure 21.3-3

1 Cut the DNA into overlapping fragments short enough for sequencing.

2 Clone the fragments in plasmid or phage vectors.

3 Sequence each fragment.

4 Order the sequences into one overall sequence with computer software.



- Both the **three-stage process** and the **whole-genome shotgun approach** were used for the Human Genome Project and for genome sequencing of other organisms
- At first many scientists were skeptical about the whole-genome shotgun approach, but **it is now widely used as the sequencing method of choice**
- The development of newer sequencing techniques has **resulted in massive increases in speed and decreases in cost**

- Technological advances have also facilitated **metagenomics**, in which DNA from a group of species (**a metagenome**) is collected from an environmental sample and sequenced
- This technique has been used **on microbial communities**, allowing **the sequencing of DNA of mixed populations, and eliminating the need to culture species in the lab**

Concept 21.2 Scientists use **bioinformatics** to analyze genomes and their functions

- The Human Genome Project established databases and refined analytical software to make **data available on the Internet**
- **This has accelerated progress in DNA sequence analysis**

Identifying Protein-Coding Genes and Understanding Their Functions

- Using available DNA sequences, geneticists can study genes directly in an approach called **reverse genetics**
- The identification of protein coding genes within DNA sequences in a database is called **gene annotation**

- Gene annotation is largely **an automated process**
- **Comparison of sequences of previously unknown genes** with those of **known genes in other species may help provide clues about their function**

Understanding Genes and Gene Expression at the Systems Level

- **Proteomics** is the systematic study of all proteins encoded by a genome
- Proteins, not genes, carry out most of the activities of the cell

How Systems Are Studied: An Example

- A **systems biology** approach can be **applied to define gene circuits and protein interaction networks**
- Researchers working on the yeast ***Saccharomyces cerevisiae*** used sophisticated techniques to disable pairs of genes one pair at a time, creating double mutants
- Computer software then mapped genes to produce a network-like “functional map” of their interactions
- The systems biology approach is possible because of advances in bioinformatics

Figure 21.5

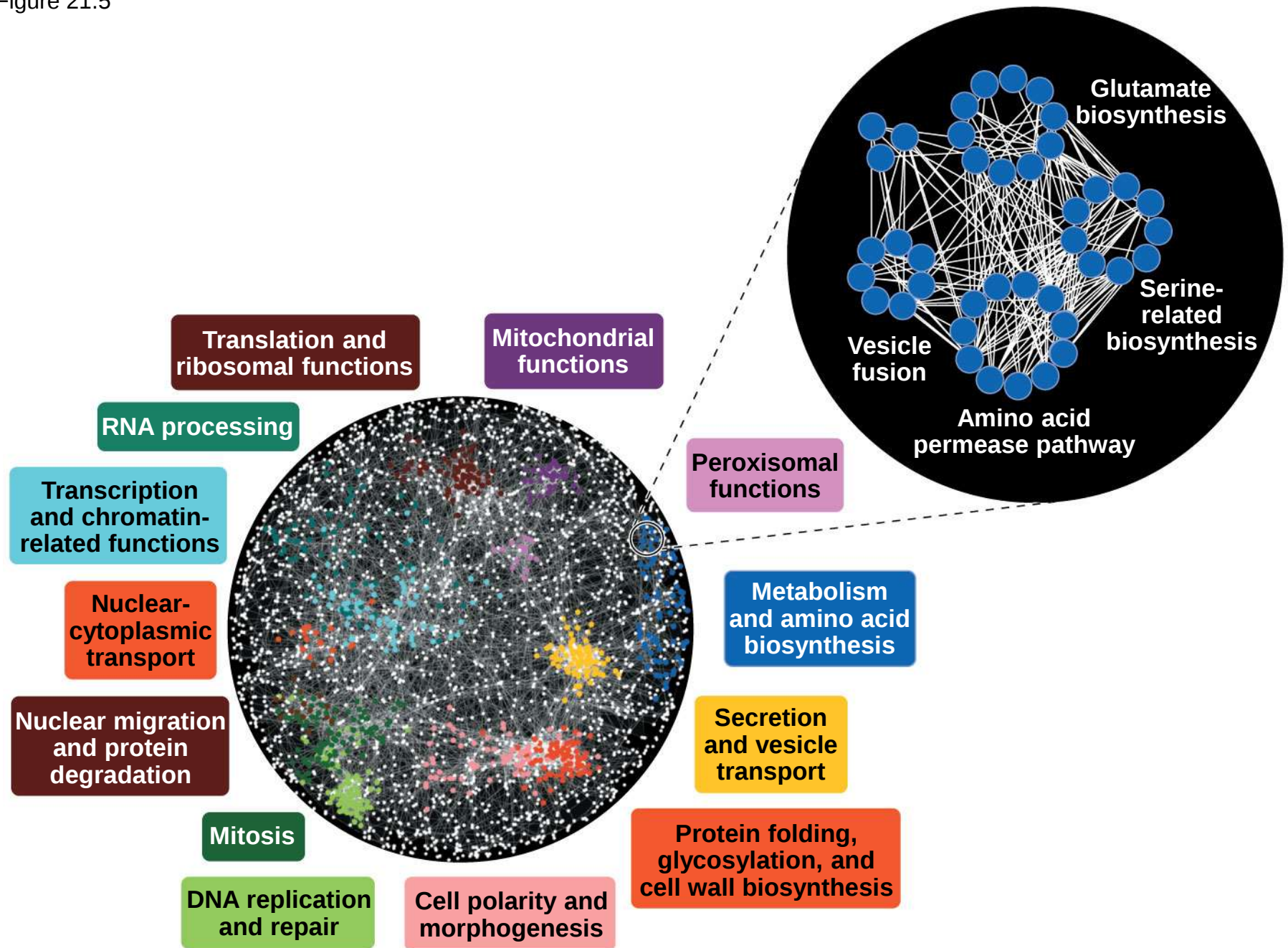


Figure 21.5a

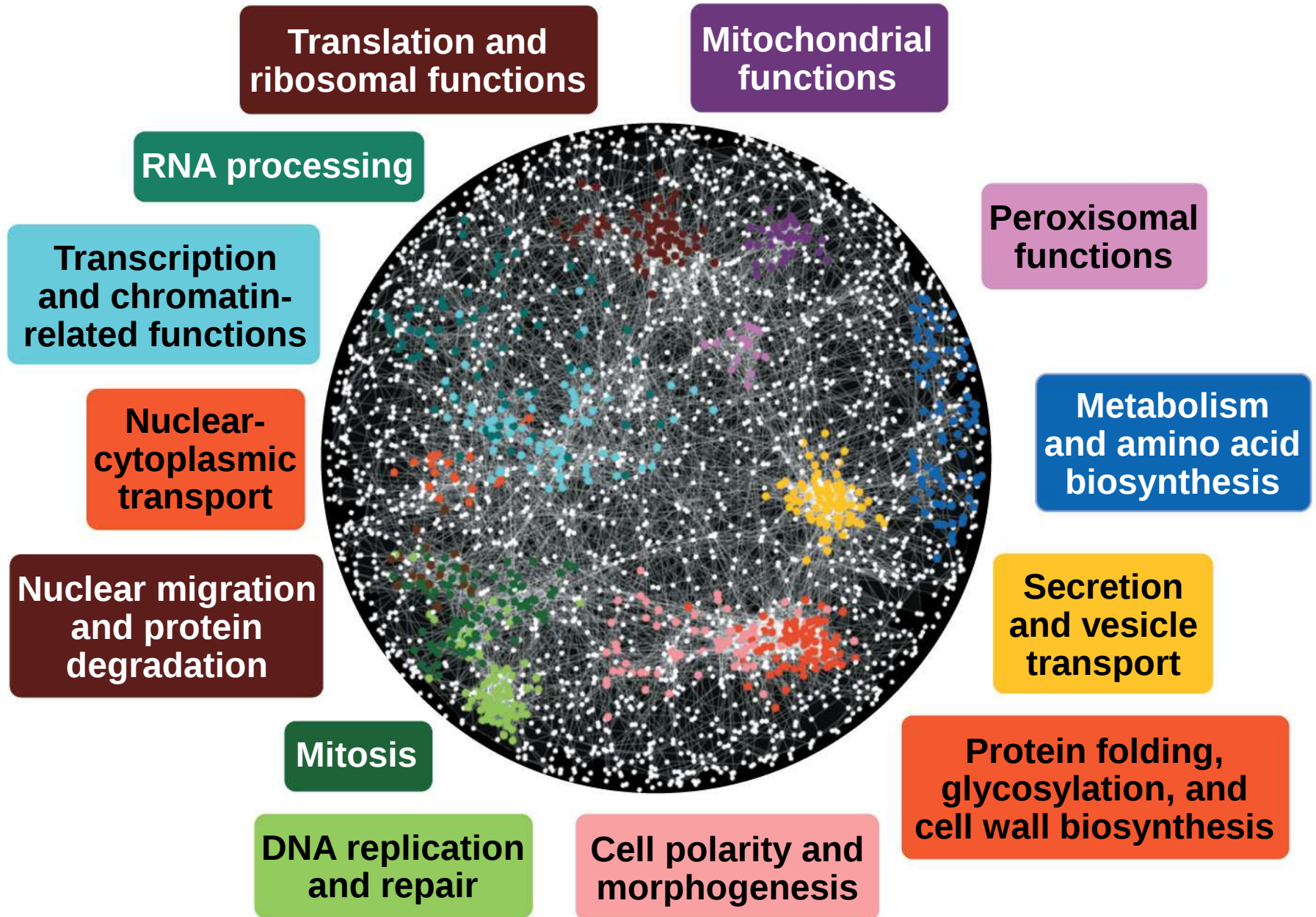
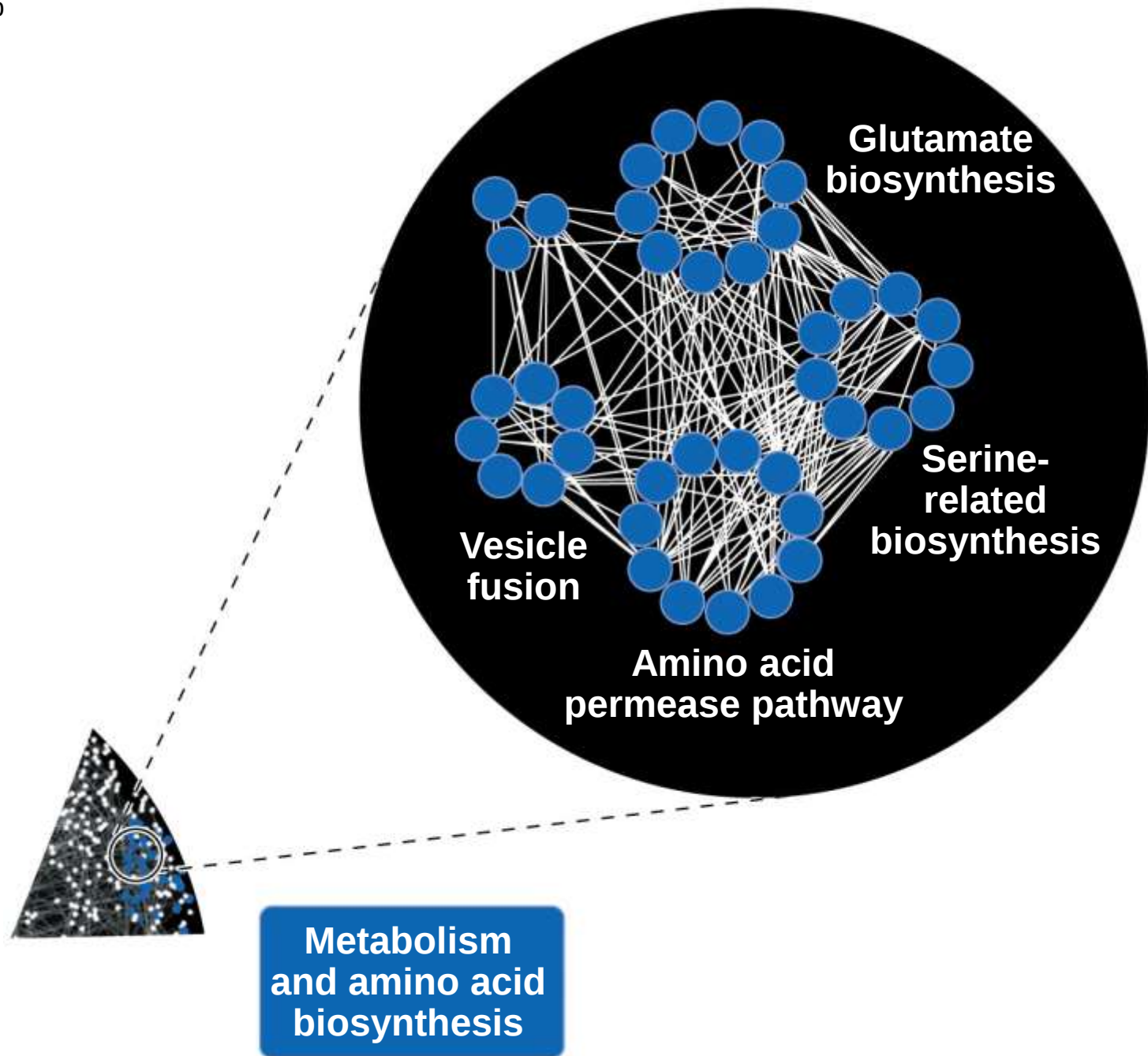


Figure 21.5b



Application of Systems Biology to Medicine

- A systems biology approach has **several medical applications**
 - The Cancer Genome Atlas project is currently **seeking all the common mutations in three types of cancer** by **comparing gene sequences and expression in cancer versus normal cells**
 - This has been so fruitful, it **will be extended to ten other common cancers**
 - Silicon and glass “chips” have been produced that hold a **microarray** of most known human genes

Figure 21.6



© 2011 Pearson Education, Inc.

Concept 21.3 Genomes vary in size, number of genes, and gene density

- By early 2010, over 1,200 genomes were completely sequenced, including 1,000 bacteria, 80 archaea, and 124 eukaryotes
- Sequencing of over 5,500 genomes and over 200 metagenomes is currently in progress
- بحلول أوائل عام 2010 ، تم تسلسل أكثر من 1200 جينوم بالكامل ، بما في ذلك 1000 بكتيريا و 80 عتائق و 124 حقيقيات النوى يجري حاليًا تنفيذ تسلسل أكثر من 5500 جينوم وأكثر من 200 ميتاجينوم

Genome Size

- Genomes of most bacteria and archaea range **from 1 to 6 million base pairs (Mb)**; genomes of eukaryotes are usually larger
- **Most plants and animals have genomes greater than 100 Mb**; **humans have 3,000 Mb**
- Within each domain there is no systematic relationship between genome size and phenotype
- تتراوح جينومات معظم البكتيريا والعتائق من 1 إلى 6 ملايين زوج قاعدي (ميغابايت) ؛ عادة ما تكون جينومات حقيقيات النوى أكبر تحتوي معظم النباتات والحيوانات على جينومات أكبر من 100 ميغا بايت ؛ البشر لديهم 3000 ميغا بايت داخل كل مجال لا توجد علاقة منهجية بين حجم الجينوم والنمط الظاهري

Table 21.1

Table 21.1 Genome Sizes and Estimated Numbers of Genes*			
Organism	Haploid Genome Size (Mb)	Number of Genes	Genes per Mb
Bacteria			
<i>Haemophilus influenzae</i>	1.8	1,700	940
<i>Escherichia coli</i>	4.6	4,400	950
Archaea			
<i>Archaeoglobus fulgidus</i>	2.2	2,500	1,130
<i>Methanosarcina barkeri</i>	4.8	3,600	750
Eukaryotes			
<i>Saccharomyces cerevisiae</i> (yeast, a fungus)	12	6,300	525
<i>Caenorhabditis elegans</i> (nematode)	100	20,100	200
<i>Arabidopsis thaliana</i> (mustard family plant)	120	27,000	225
<i>Drosophila melanogaster</i> (fruit fly)	165	13,700	83
<i>Oryza sativa</i> (rice)	430	42,000	98
<i>Zea mays</i> (corn)	2,300	32,000	14
<i>Mus musculus</i> (house mouse)	2,600	22,000	11
<i>Ailuropoda melanoleuca</i> (giant panda)	2,400	21,000	9
<i>Homo sapiens</i> (human)	3,000	<21,000	7
<i>Fritillaria assyriaca</i> (lily family plant)	124,000	ND	ND
*Some values given here are likely to be revised as genome analysis continues. Mb = million base pairs. ND = not determined.			

Number of Genes

- Free-living bacteria and archaea have 1,500 to 7,500 genes
- Unicellular fungi have from about 5,000 genes and multicellular eukaryotes up to at least 40,000 genes

- Number of genes is not correlated to genome size
- For example, it is estimated that the nematode *C. elegans* has 100 Mb and 20,000 genes, while *Drosophila* has 165 Mb and 13,700 genes
- Vertebrate genomes can produce more than one polypeptide per gene because of alternative splicing of RNA transcripts
- لا يرتبط عدد الجينات بحجم الجينوم على سبيل المثال ، تشير التقديرات بها 100 ميغا بايت و 20 ألف *C. elegans* إلى أن الديدان النيماتودا جين ، بينما تحتوي ذبابة الفاكهة على 165 ميغا بايت و 13700 جين يمكن أن تنتج جينومات الفقاريات أكثر من عديد بيتيد واحد لكل جين بسبب التضفير البديل لنصوص الحمض النووي الريبي

Gene Density and Noncoding DNA

- **Humans and other mammals** have the lowest gene density, or number of genes, in a given length of DNA
- Multicellular eukaryotes have many **introns within genes and noncoding DNA between genes**

Concept 21.4: Multicellular eukaryotes have much noncoding DNA and many multigene families

- The **bulk of most eukaryotic genomes neither encodes proteins nor functional RNAs**
- Much evidence indicates that **noncoding DNA (previously called “junk DNA”) plays important roles in the cell**
- For example, **genomes of humans, rats, and mice show high sequence conservation for about 500 noncoding regions**

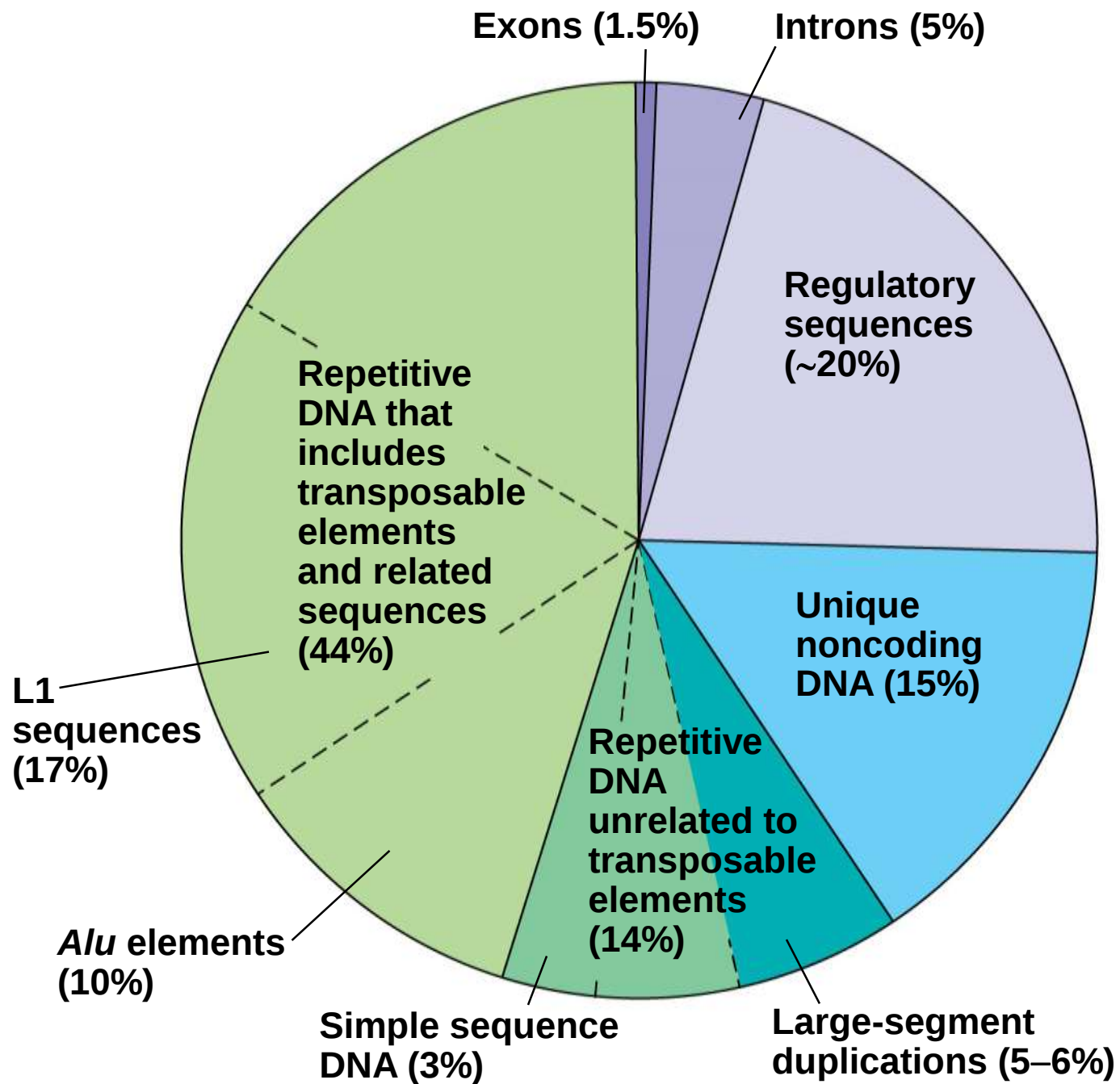
- لجزء الأكبر من الجينومات حقيقية النواة لا يشفر البروتينات ولا الحمض النووي الريبي الوظيفي
- تشير الكثير من الأدلة إلى أن الحمض النووي غير غير المرغوب DNA المشفر (المعروف سابقًا باسم "فيه") يلعب أدوارًا مهمة في الخلية
- على سبيل المثال ، تُظهر جينومات البشر والجرذان والفئران حفظًا عالي التسلسل لحوالي 500 منطقة غير مشفرة

- Sequencing of the human genome reveals that **98.5% does not code for proteins, rRNAs, or tRNAs**
- About a **quarter of the human genome codes for introns and gene-related regulatory sequences**
- حوالي ربع أكواد الجينوم البشري للإنترونات والتسلسلات التنظيمية المتعلقة بالجينات

- **Intergenic DNA** is noncoding DNA found between genes
 - **Pseudogenes** are former genes that have accumulated mutations and are nonfunctional
 - **Repetitive DNA** is present in multiple copies in the genome
- About **three-fourths of repetitive DNA** is made up of **transposable elements** and sequences related to them

- غير مشفر DNA الحمض النووي بين الجينات هو موجود بين الجينات
- الجينات الكاذبة هي جينات سابقة تراكتت عليها الطفرات وهي غير وظيفية
- يوجد الحمض النووي المتكرر في نسخ متعددة في الجينوم
- يتكون حوالي ثلاثة أرباع الحمض النووي المتكرر من عناصر قابلة للنقل ومتواليات مرتبطة بها

Figure 21.7



Transposable Elements and Related Sequences

- The first evidence for mobile DNA segments came from geneticist **Barbara McClintock's breeding experiments with Indian corn**
- McClintock identified **changes in the color of corn kernels** that made sense only by postulating that some genetic elements move from other genome locations into the genes for kernel color
- These **transposable elements** move from one site to another in a cell's DNA; they are present in both prokaryotes and eukaryotes

- جاء الدليل الأول على مقاطع الحمض النووي المتنقلة من تجارب التكاثر التي أجرتها عالمة الوراثة باربرا مكلينتوك مع الذرة الهندية
- حددت مكلينتوك التغيرات في لون حبات الذرة التي لا معنى لها إلا من خلال افتراض أن بعض العناصر الجينية تنتقل من مواقع الجينوم الأخرى إلى جينات لون النواة.
- تنتقل هذه العناصر القابلة للنقل من موقع إلى آخر في الحمض النووي للخلية ؛ هم موجودون في كل من بدائيات النوى وحقيقيات النوى

Figure 21.8



© 2011 Pearson Education, Inc.



Figure 21.8a



© 2011 Pearson Education, Inc.

Figure 21.8b

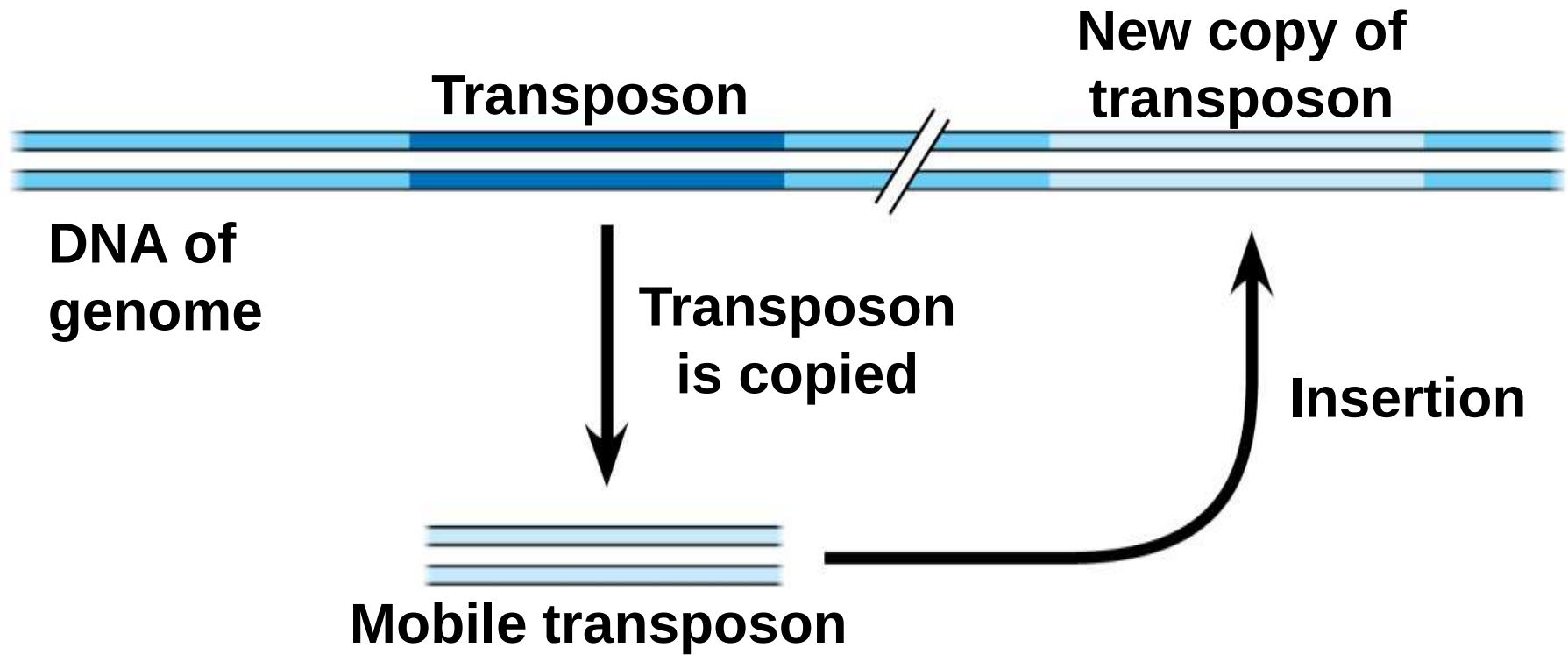


© 2011 Pearson Education, Inc.

Movement of Transposons and Retrotransposons

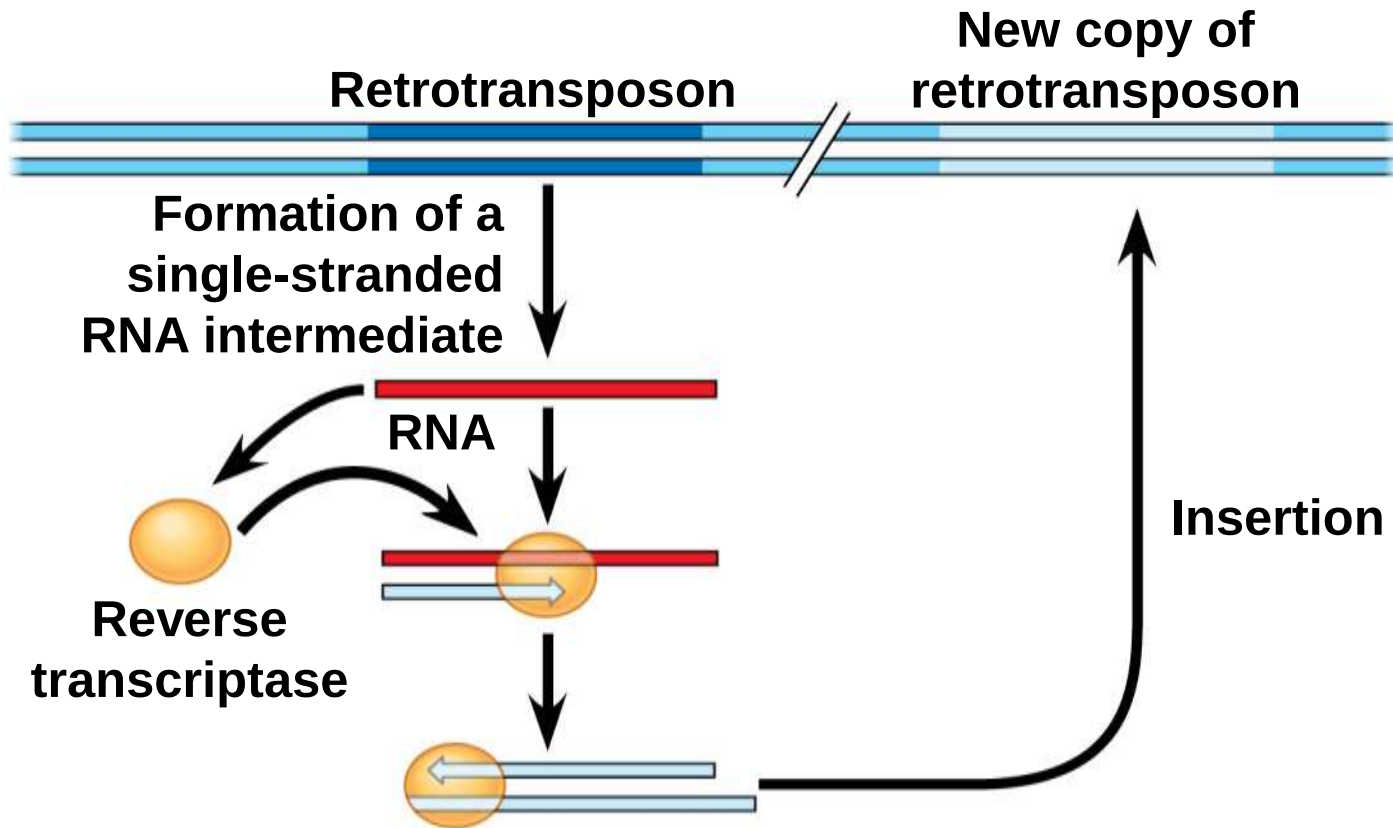
- Eukaryotic transposable elements are of two types
 - **Transposons**, which move by means of a DNA intermediate
 - **Retrotransposons**, which move by means of an RNA intermediate
 - لعناصر القابلة للنقل حقيقية النواة من نوعين
DNA الينقولات ، التي تتحرك عن طريق وسيط
RNA الينقولات العكسية ، التي تتحرك عن طريق وسيط

Figure 21.9



© 2011 Pearson Education, Inc.

Figure 21.10

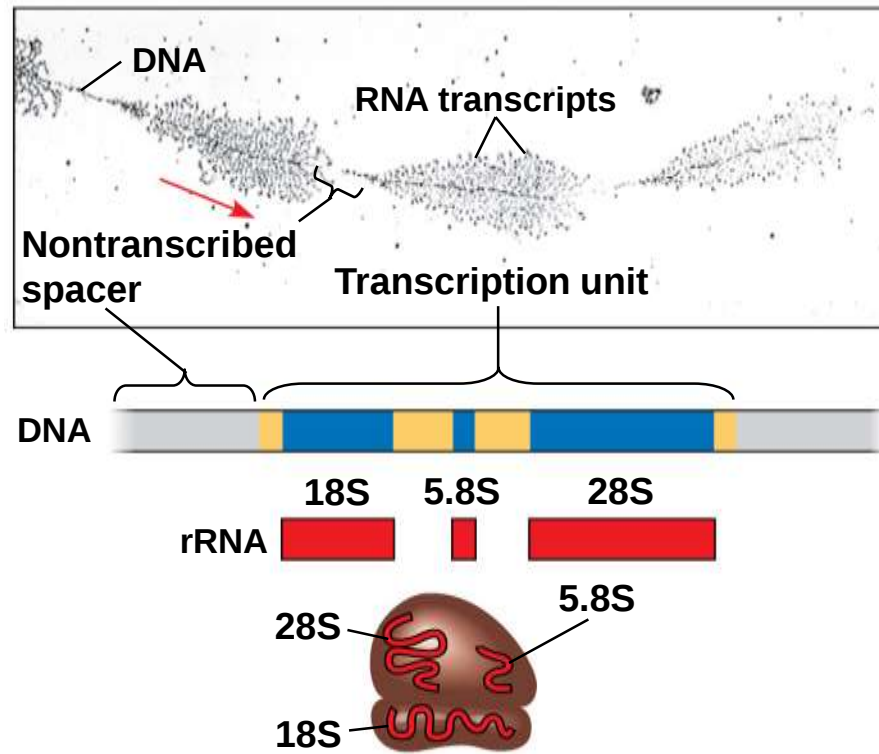


© 2011 Pearson Education, Inc.

Genes and Multigene Families

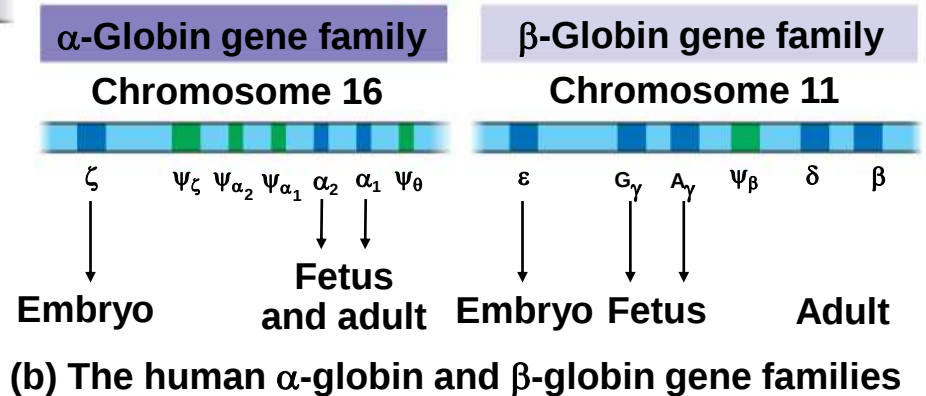
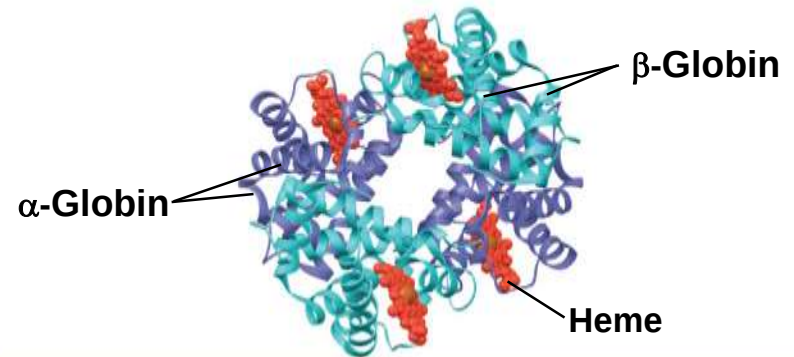
- Many eukaryotic genes are present in one copy per haploid set of chromosomes
- The rest of the genes occur in **multigene families, collections of identical or very similar genes**
- Some multigene families consist **of identical DNA sequences, usually clustered tandemly, such as those that code for rRNA products**

- توجد العديد من جينات حقيقيات النوى في نسخة واحدة لكل مجموعة أحادية الصبغيات من الكروموسومات
- تحدث بقية الجينات في عائلات متعددة الجينات ، وهي مجموعات من جينات متطابقة أو متشابهة جدًا
- تتكون بعض العائلات متعددة الجينات من تسلسلات دنا متطابقة ، وعادة ما تكون متجمعة بالترادف ، مثل تلك التي ترمز لمنتجات الرنا الريباسي



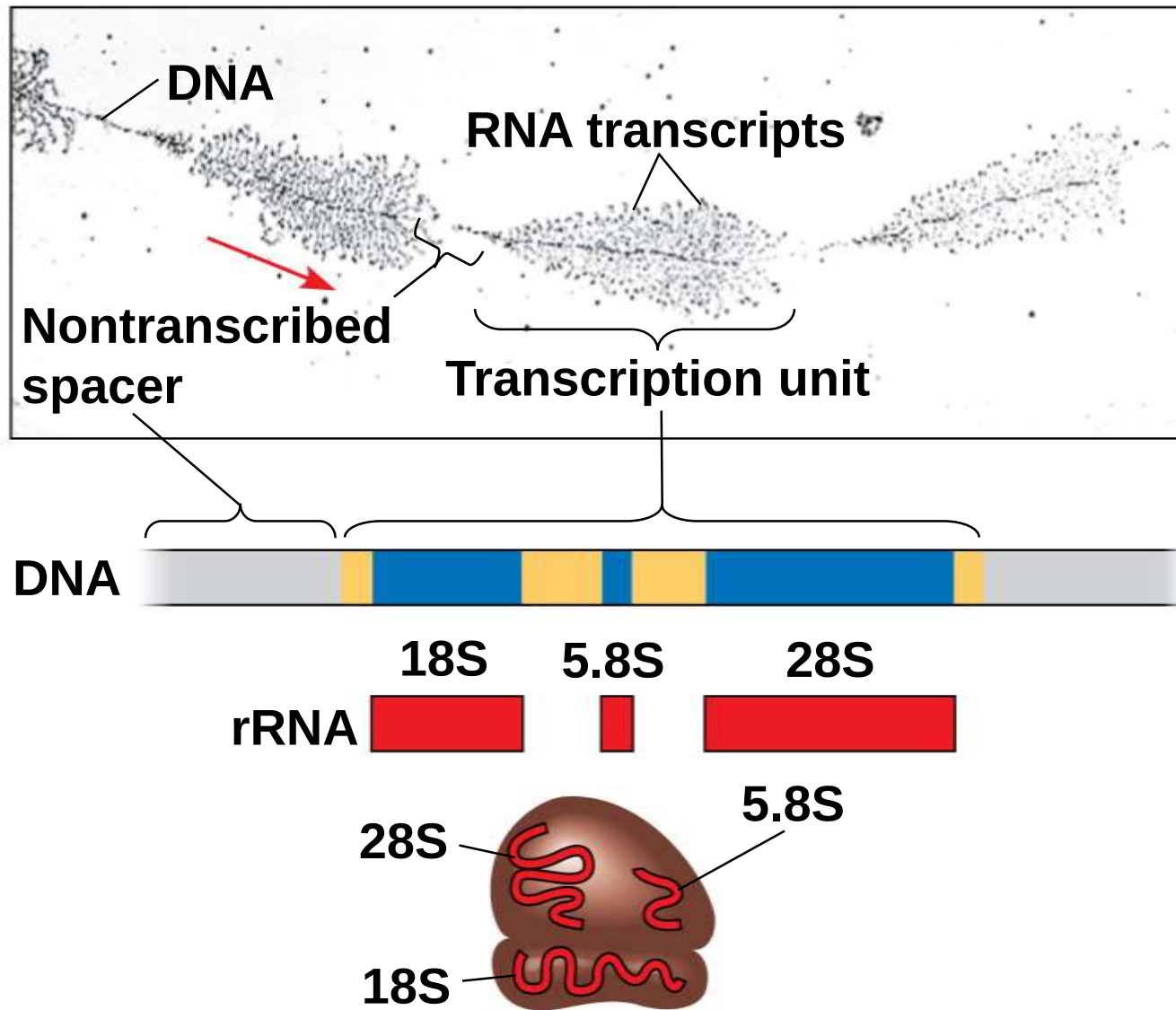
(a) Part of the ribosomal RNA gene family

© 2011 Pearson Education, Inc.



(b) The human α -globin and β -globin gene families

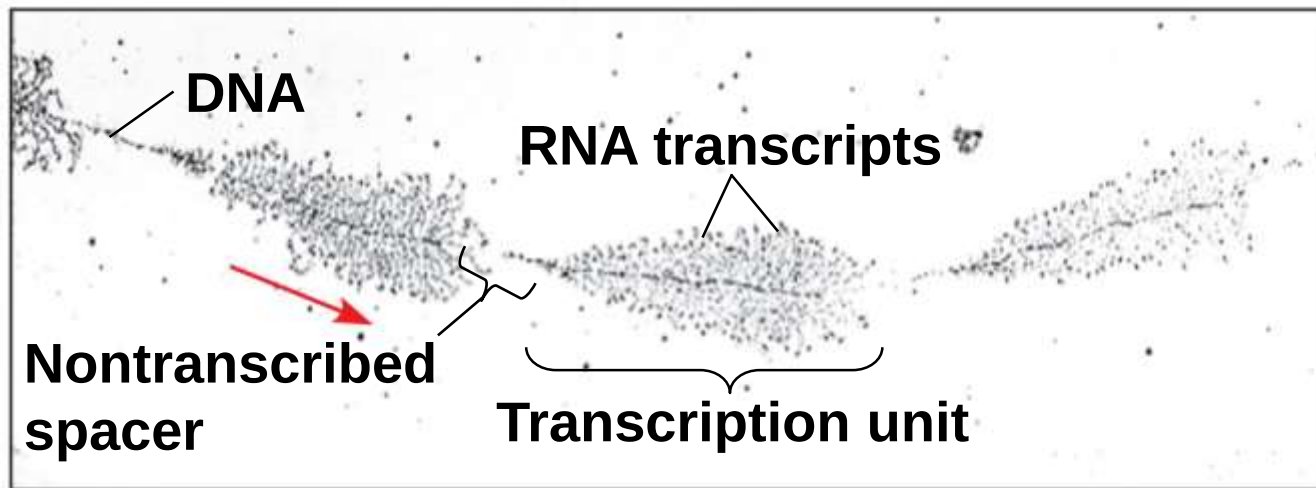
Figure 21.11a



(a) Part of the ribosomal RNA gene family

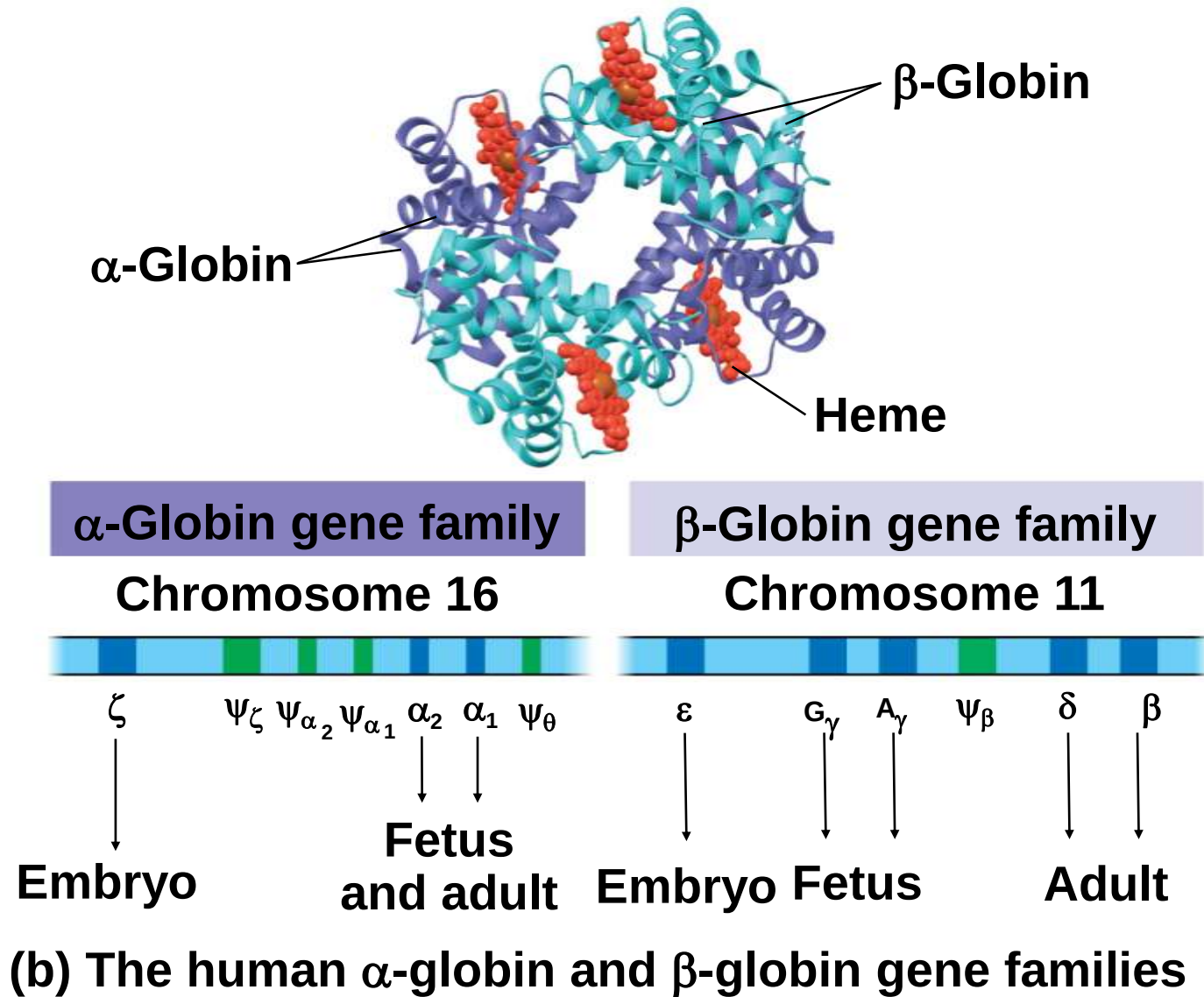
© 2011 Pearson Education, Inc.

Figure 21.11c



© 2011 Pearson Education, Inc.

- The classic examples of multigene families of **nonidentical genes** are **two related families of genes that encode globins**
- α -globins and β -globins are polypeptides of hemoglobin and are coded by genes on different human chromosomes and are expressed at different times in development



© 2011 Pearson Education, Inc.

Concept 21.5: Duplication, rearrangement, and mutation of DNA contribute to genome evolution

- The **basis of change at the genomic level** is **mutation**, which underlies much of genome evolution
- The **earliest forms of life** likely **had a minimal number of genes, including only those necessary for survival and reproduction**
- **The size of genomes has increased over evolutionary time**, with the extra genetic material **providing raw material for gene diversification**

Duplication of Entire Chromosome Sets

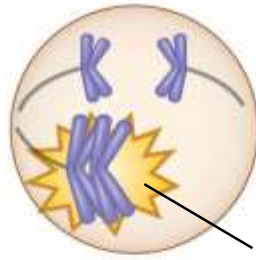
- Accidents in meiosis can lead to one or more extra sets of chromosomes, a condition known as **polyploidy**
- The **genes in one or more of the extra sets can diverge by accumulating mutations**; these **variations may persist if the organism carrying them survives and reproduces**

Abnormal Chromosome Number

- In **nondisjunction**, pairs of homologous chromosomes do not separate normally during meiosis
- As a result, one gamete receives two of the same type of chromosome, and another gamete receives no copy

Figure 15.13-1

Meiosis I



Nondisjunction

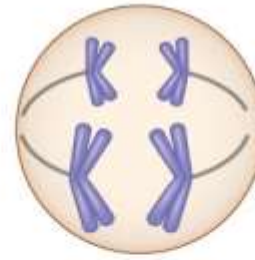
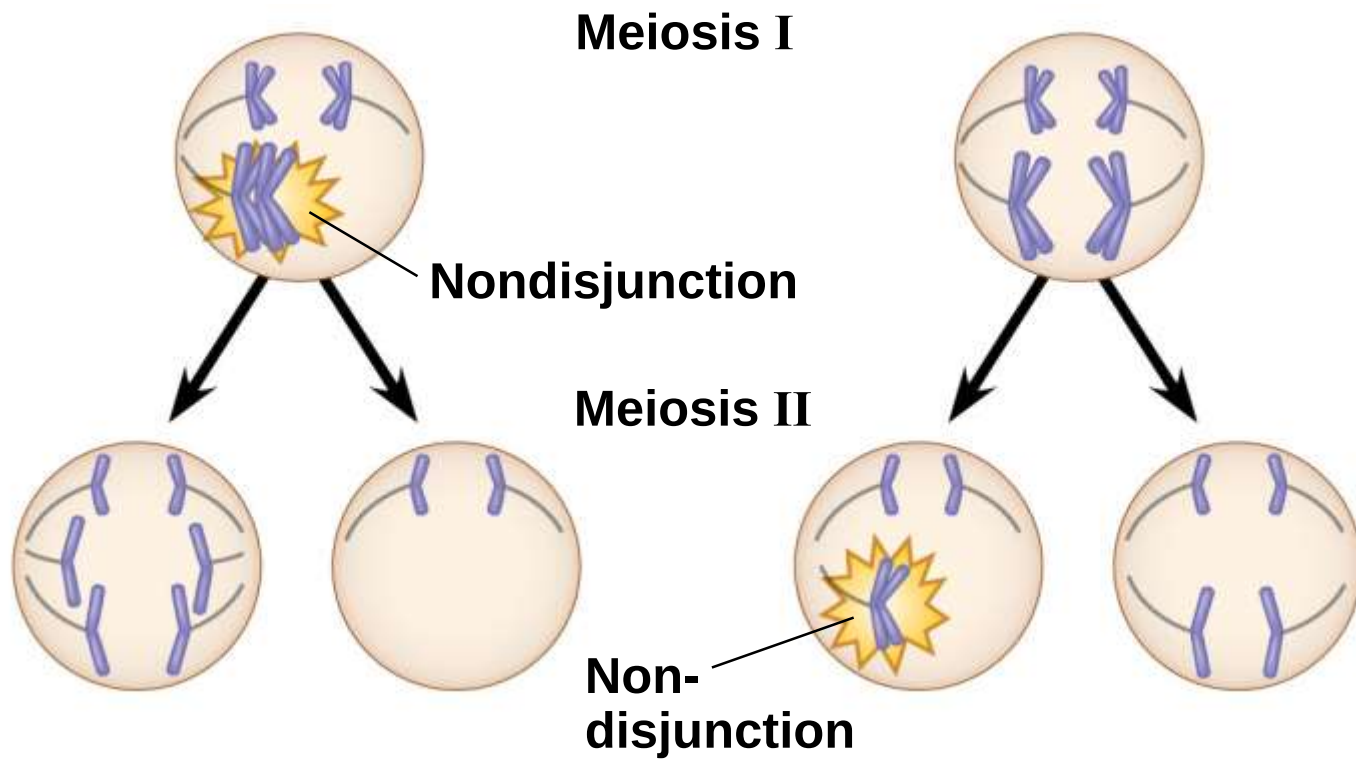
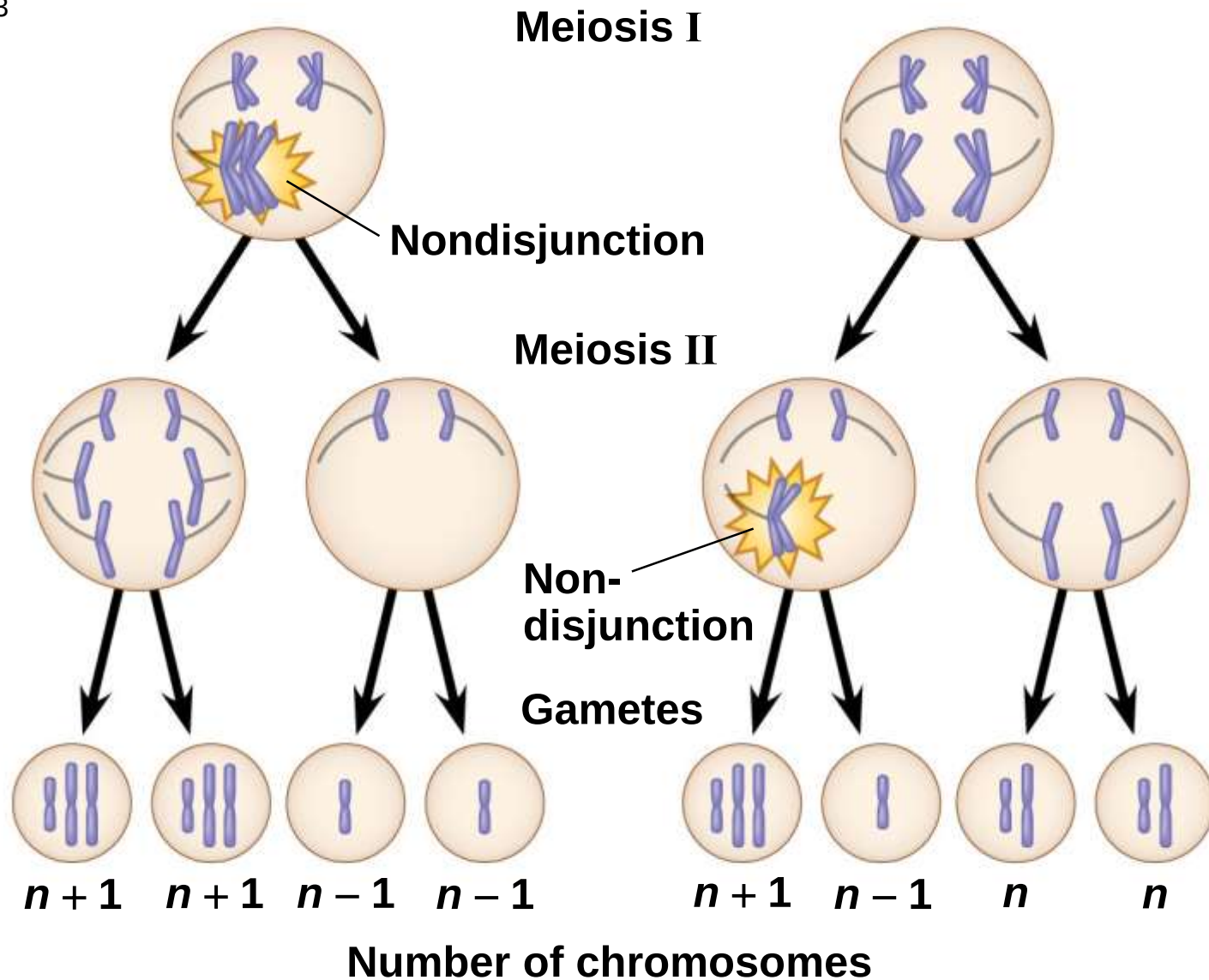


Figure 15.13-2





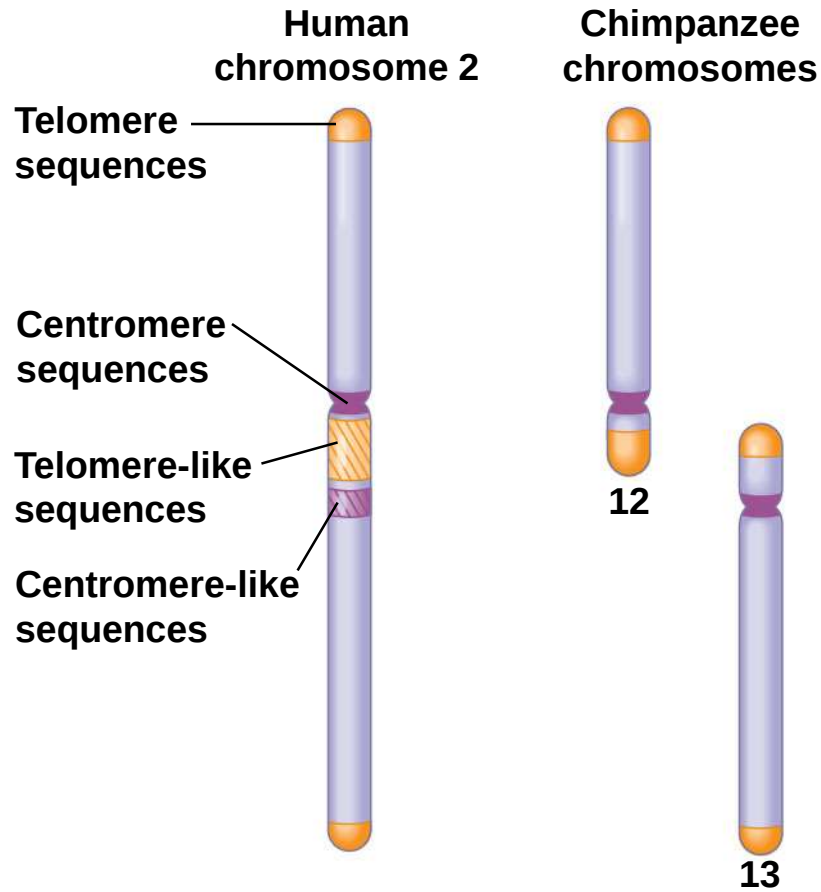
(a) Nondisjunction of homologous chromosomes in meiosis I

(b) Nondisjunction of sister chromatids in meiosis II

Alterations of Chromosome Structure

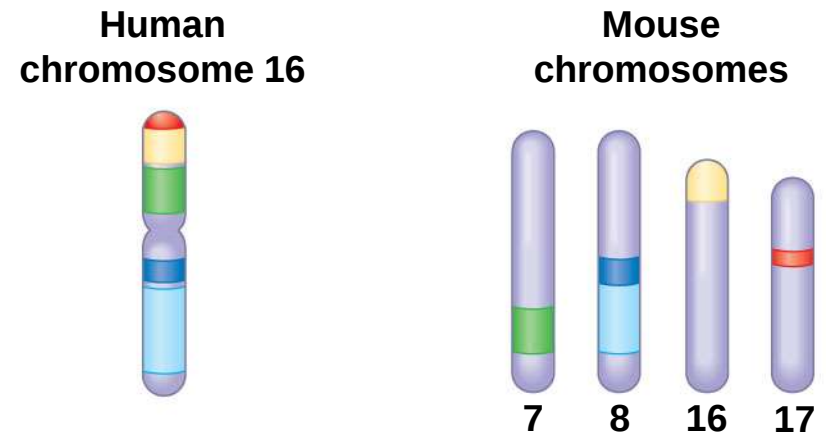
- Humans have 23 pairs of chromosomes, while chimpanzees have 24 pairs
- Following the divergence of humans and chimpanzees from a common ancestor, two ancestral chromosomes fused in the human line
- Duplications and inversions result from mistakes during meiotic recombination
- Comparative analysis between chromosomes of humans and seven mammalian species paints a hypothetical chromosomal evolutionary history

Figure 21.12

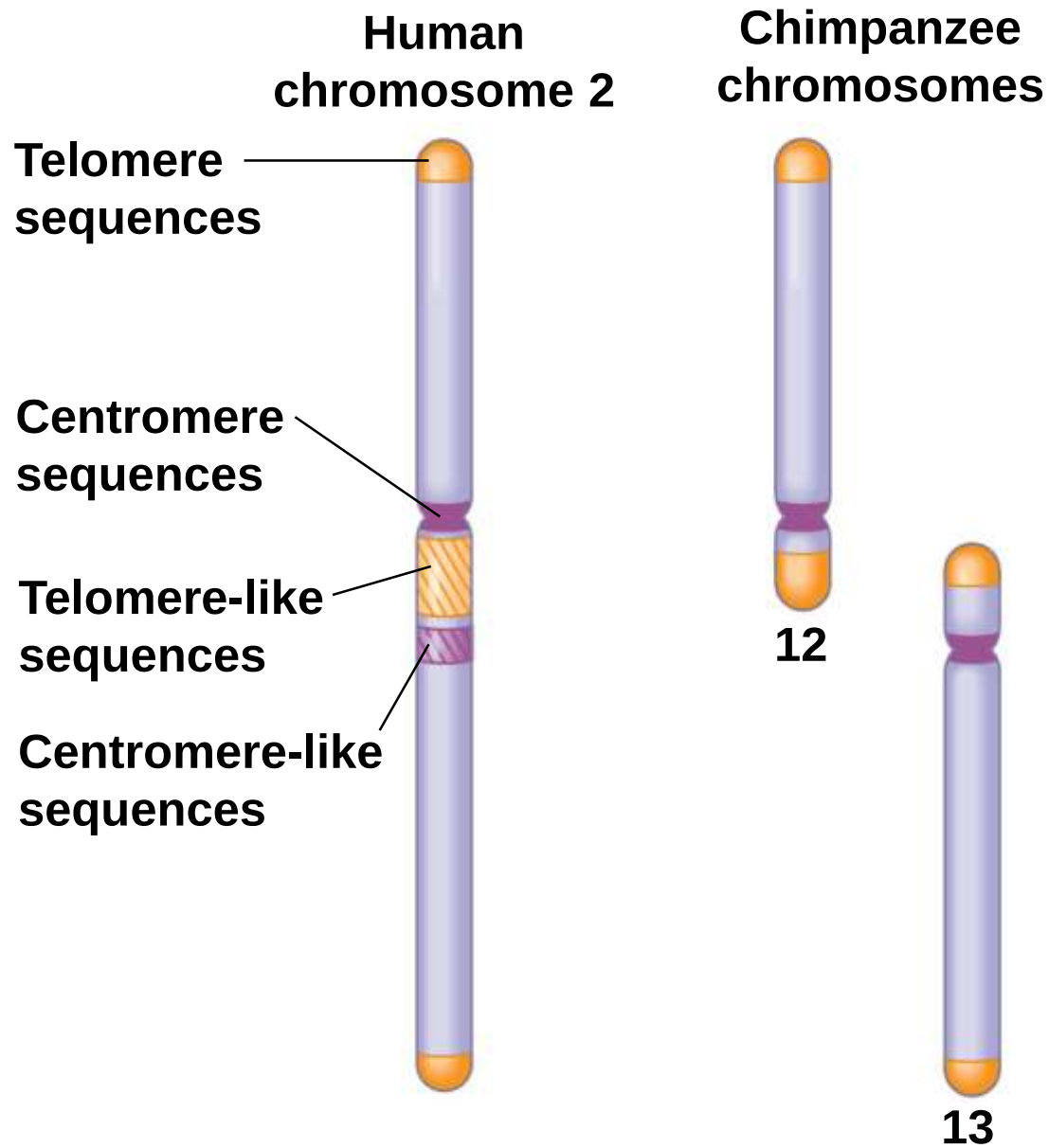


(a) Human and chimpanzee chromosomes

© 2011 Pearson Education, Inc.



(b) Human and mouse chromosomes

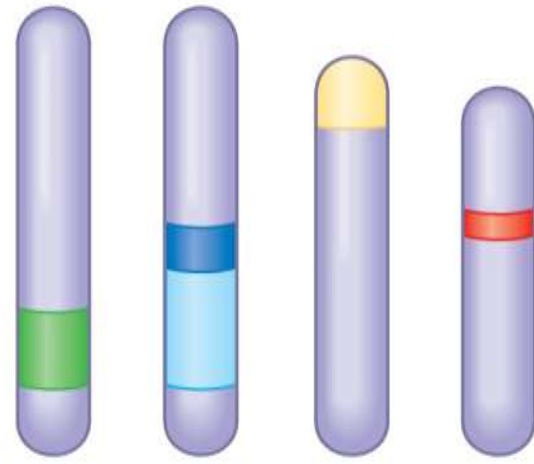


(a) Human and chimpanzee chromosomes

Human chromosome 16



Mouse chromosomes



7

8

16

17

(b) Human and mouse chromosomes

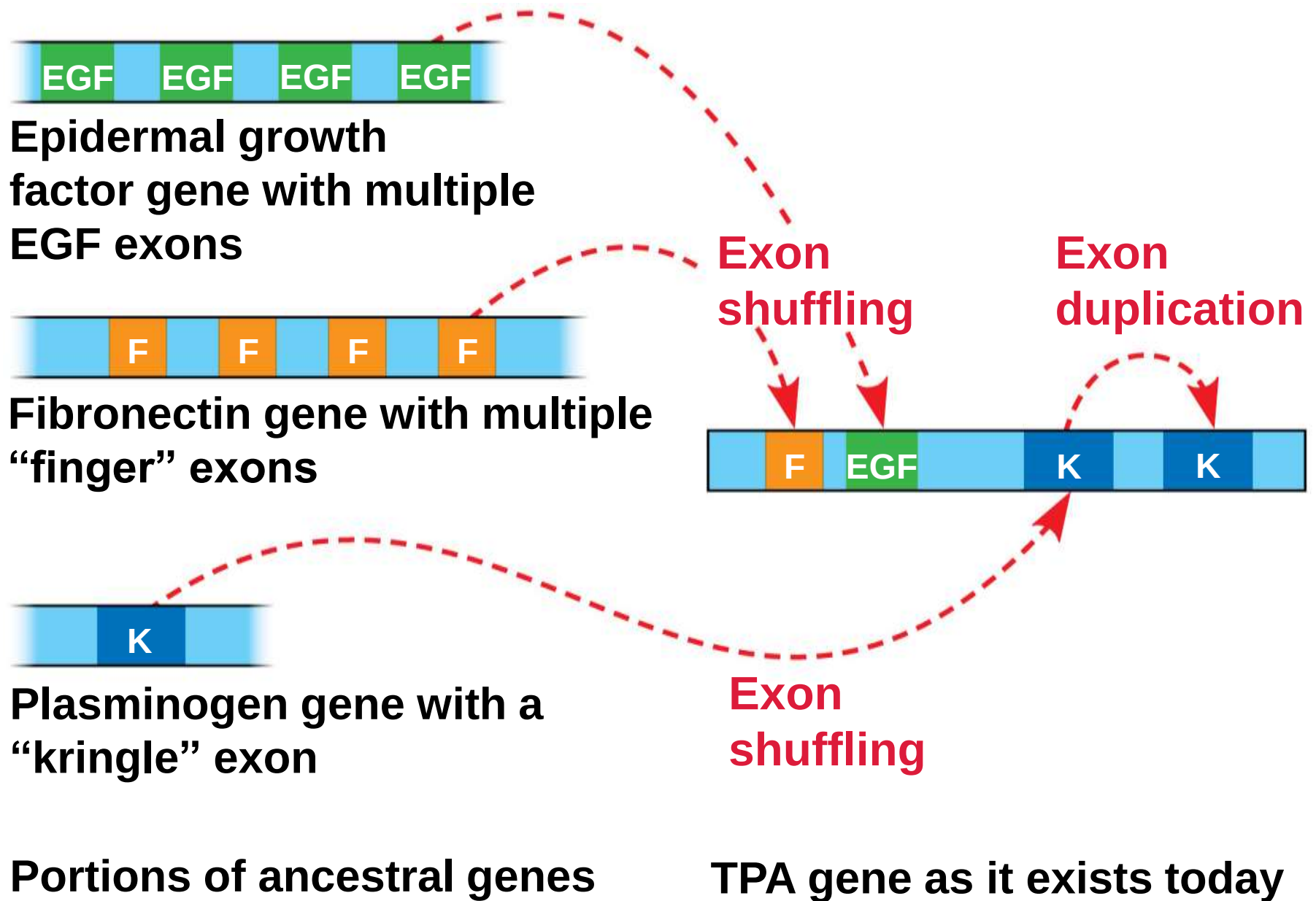
© 2011 Pearson Education, Inc.

- The **rate of duplications and inversions seems to have accelerated about 100 million years ago**
- **This coincides with when large dinosaurs went extinct and mammals diversified**
- **Chromosomal rearrangements are thought to contribute to the generation of new species**
- Some of the recombination “hot spots” associated with chromosomal rearrangement are also locations that are associated with diseases

Rearrangements of Parts of Genes: Exon Duplication and Exon Shuffling

- The **duplication or repositioning of exons has contributed to genome evolution**
- Errors in meiosis can result in an **exon being duplicated on one chromosome and deleted from the homologous chromosome**
- In exon shuffling, **errors in meiotic recombination lead to some mixing and matching of exons**, either within a gene or between two nonallelic genes

Figure 21.15



© 2011 Pearson Education, Inc.

How Transposable Elements Contribute to Genome Evolution

- Multiple copies of similar transposable elements may facilitate recombination, or crossing over, between different chromosomes
- Insertion of transposable elements within a protein-coding sequence may block protein production
- Insertion of transposable elements within a regulatory sequence may increase or decrease protein production

- Transposable elements may carry a gene or groups of genes to a new position
- Transposable elements may also create new sites for alternative splicing in an RNA transcript
- In all cases, changes are usually detrimental but may on occasion prove advantageous to an organism

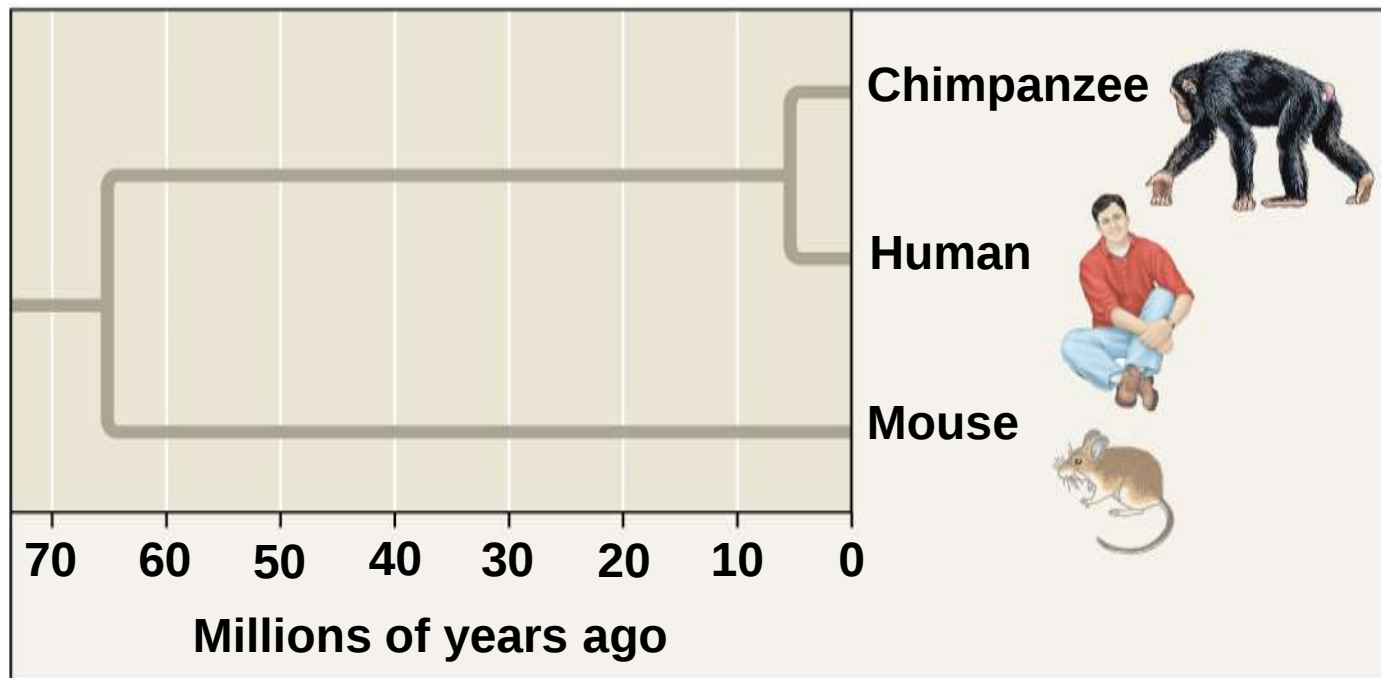
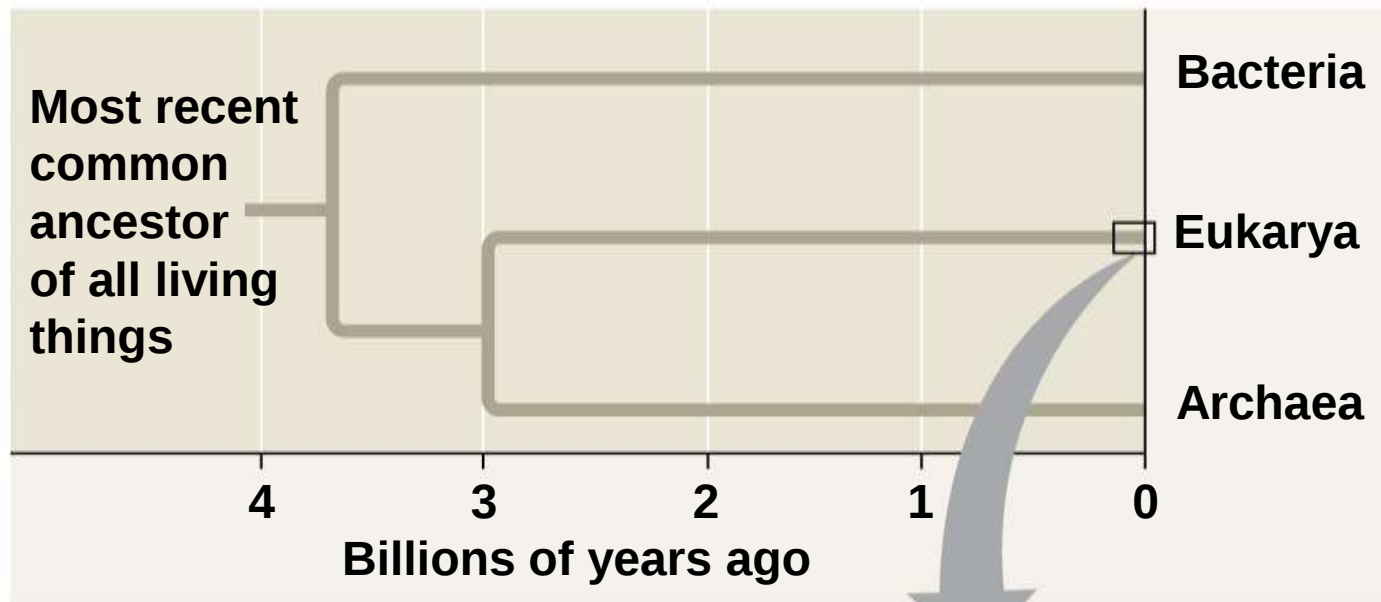
Concept 21.6: Comparing genome sequences provides clues to evolution and development

- Genome sequencing and data collection has advanced rapidly in the last 25 years
- **Comparative studies of genomes**
 - Advance our understanding of the evolutionary history of life
 - Help explain how the evolution of development leads to morphological diversity

Comparing Genomes

- Genome comparisons of **closely related species** help us understand recent evolutionary events
- Genome comparisons of **distantly related species** help us understand ancient evolutionary events
- **Relationships among species can be represented by a tree-shaped diagram**

Figure 21.16



Comparing Distantly Related Species

- **Highly conserved genes have changed very little over time**
- These help clarify relationships among species that diverged from each other long ago
- **Bacteria, archaea, and eukaryotes diverged from each other between 2 and 4 billion years ago**
- Highly conserved genes **can be studied in one model organism, and the results applied to other organisms**

Figure 21.UN01

	Bacteria	Archaea	Eukarya
Genome size	Most are 1–6 Mb		Most are 10–4,000 Mb, but a few are much larger
Number of genes	1,500–7,500		5,000–40,000
Gene density	Higher than in eukaryotes		Lower than in prokaryotes (Within eukaryotes, lower density is correlated with larger genomes.)
Introns	None in protein-coding genes	Present in some genes	Unicellular eukaryotes: present, but prevalent only in some species Multicellular eukaryotes: present in most genes
Other noncoding DNA	Very little		Can be large amounts; generally more repetitive noncoding DNA in multicellular eukaryotes

Figure 21.UN02

Human genome

