Repetitive DNA

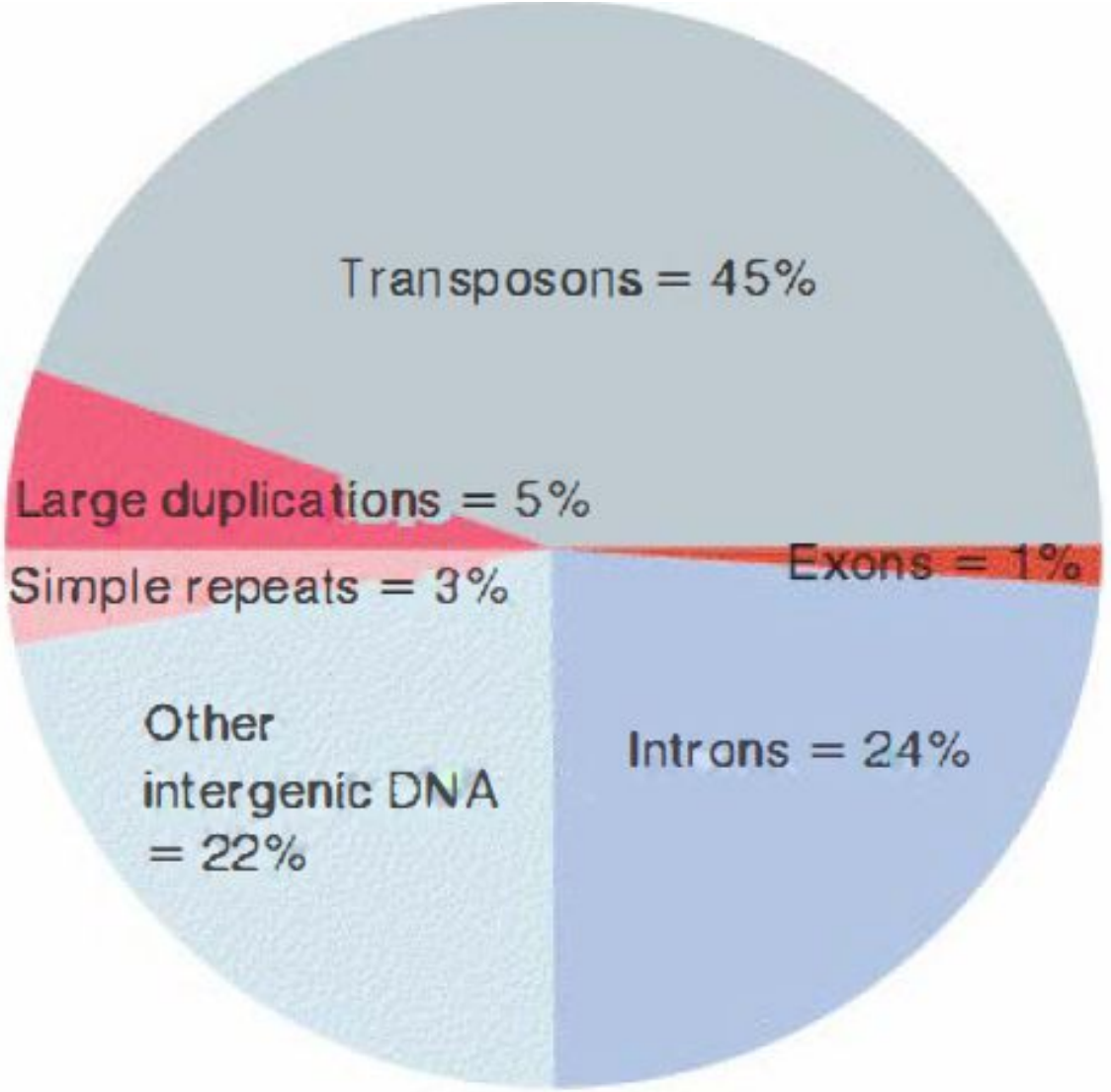
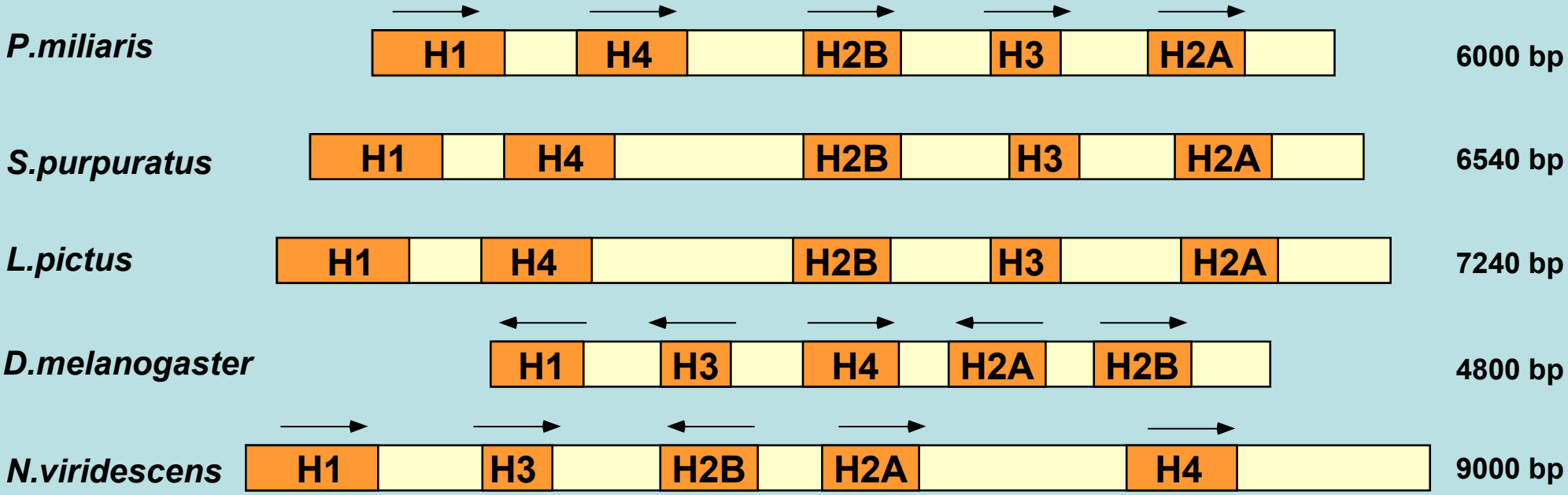


FIGURE 6.14 The largest component of the human genome consists of transposons. Other repetitive sequences include large duplications and simple repeats.

Tandemly Repeated Genes

Organization of the repeating units for histone genes in different species



Multiple copies of rRNA and tRNA genes are present in ALL genomes

Species	18S/28S rRNA genes	5S RNA genes	tRNA genes
E.coli	7	7	60
S.cerevisiae	140	140	250
D.discoideum	180	180	?
D.melanogaster		165	850
X	250		
Y	150		
X.laevis	450	24,000	1,150
H.sapiens	280	2,000	1,300

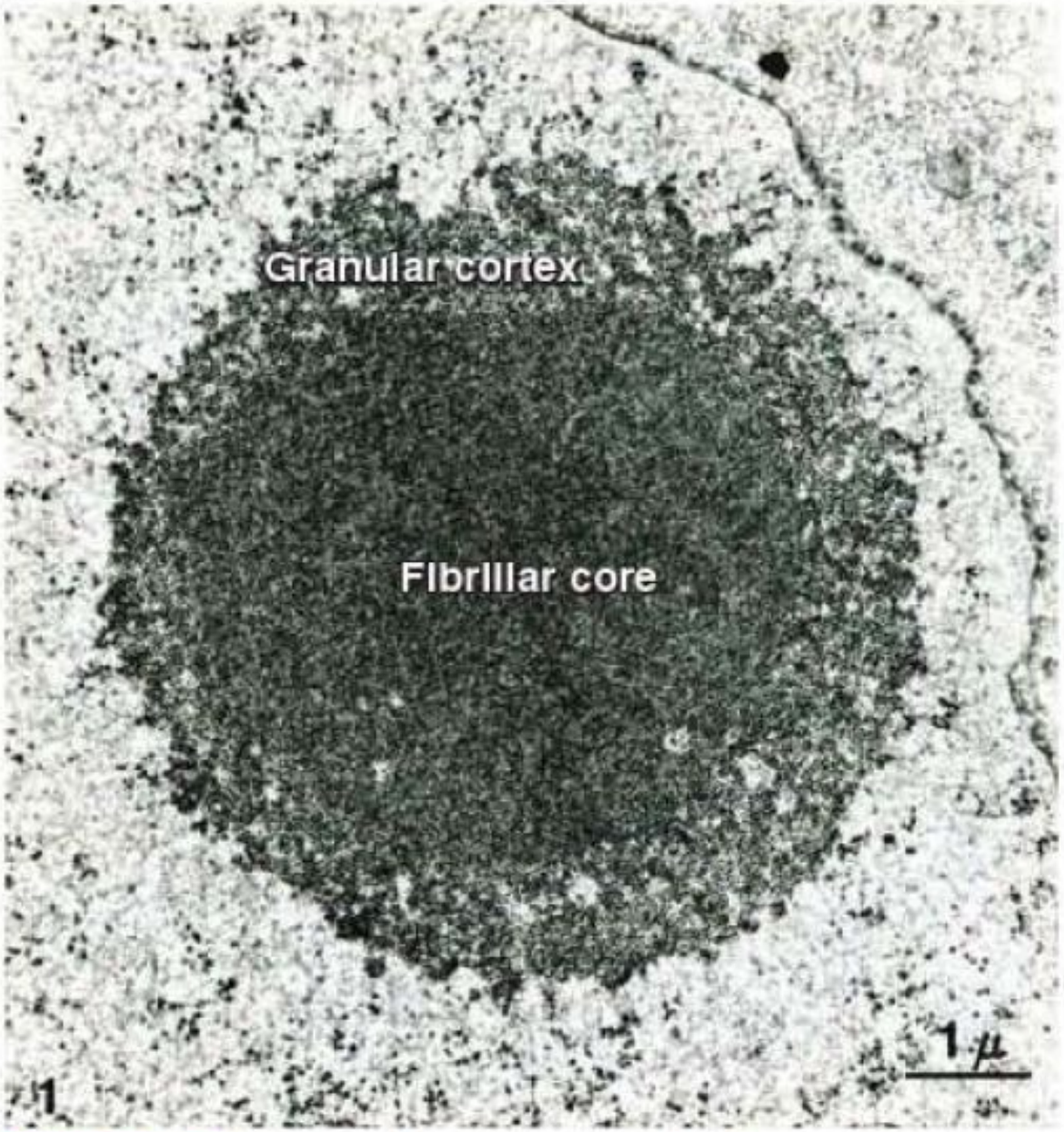


FIGURE 7.8 The nucleolar core identifies rDNA under transcription and the surrounding granular cortex consists of assembling ribosomal subunits. This thin section shows the nucleolus of the newt *Notophthalmus viridescens*. Photo courtesy of Oscar Miller.

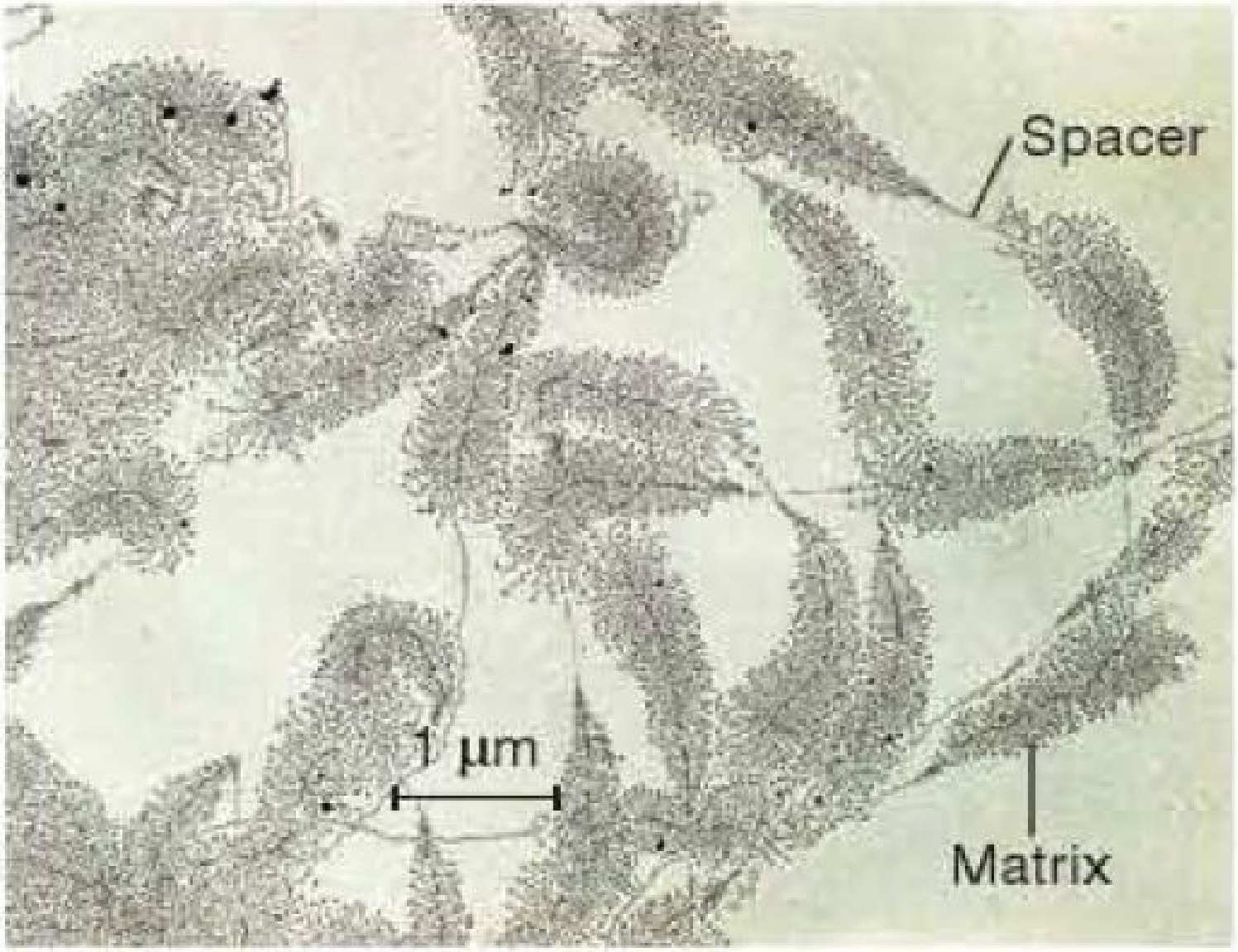


FIGURE 7.9 Transcription of rDNA clusters generates a series of matrices, each corresponding to one transcription unit and separated from the next by the nontranscribed spacer. © Don W. Fawcett/Photo Researchers, Inc.

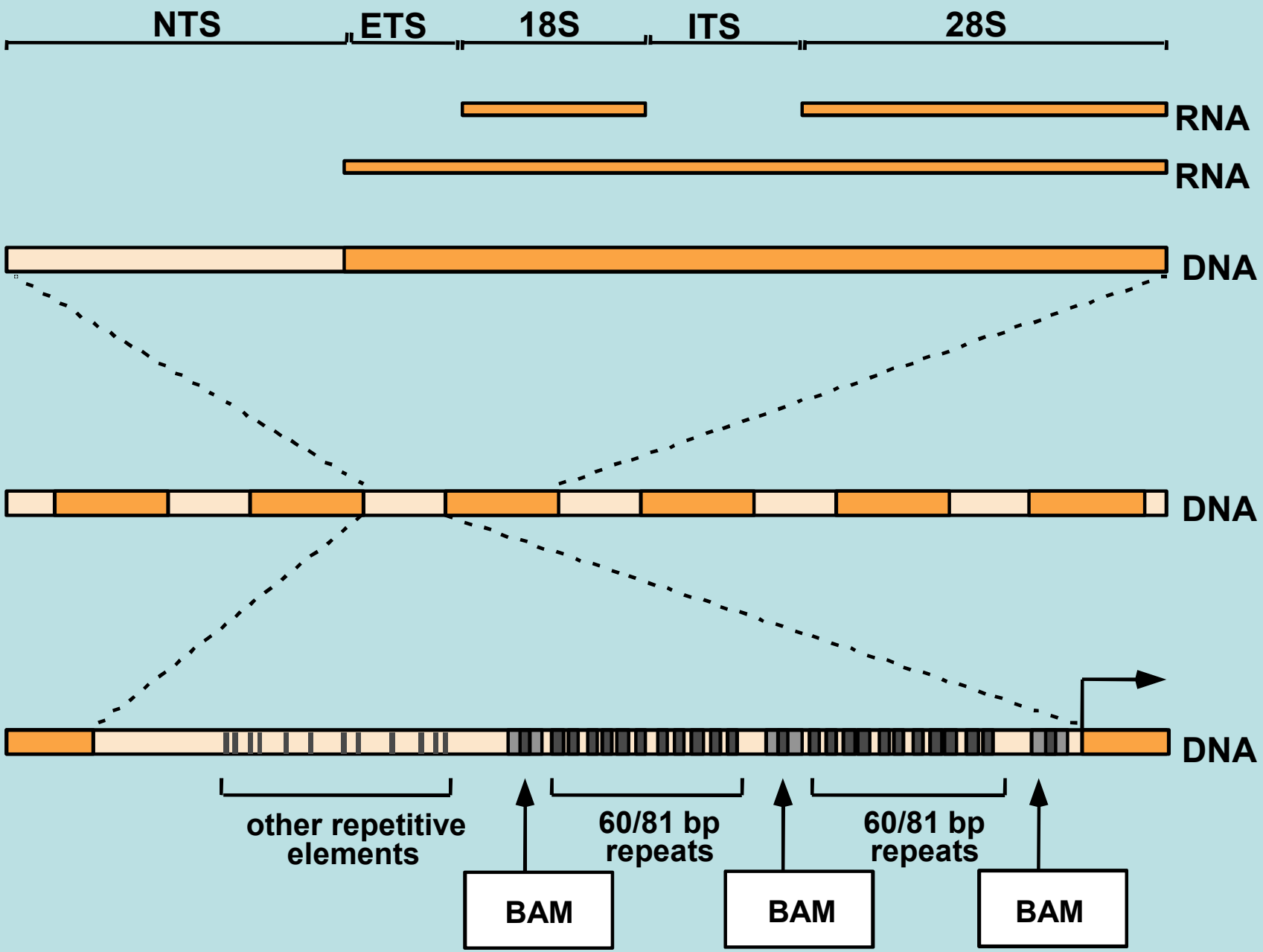
The length of the repeat units varies a great deal in rRNA genes

Species	Repeat unit length (bp)	Non-transcribed spacer length (bp)	Length of the transcript (bp)
S.cerevisiae	8,950	1,750	7,200
D.melanogaster	11,500-14,200	3,750-6,450	7,750
X.laevis	10,500-13,500	2,300-5,300	7,875
M.musculus	44,000	30,000	13,400

rRNA PRECURSORS ARE LONGER THAN THE MATURE rRNA TAKEN TOGETHER

Species	Transcript length (nt)	Length of rRNA		rRNA % of the transcript
		Long	Short	
E.coli	5,600	2,904	1,542	80 %
S.cerevisiae	7,200	3,750	2,000	80 %
D.melanogaster	7,750	4,100	2,000	78 %
X.laevis	7,875	4,475	1,925	79 %
N.tabaccum	7,900	3,700	1,900	71 %
G.domesticus	11,250	4,625	1,800	57 %
M.musculus	13,400	4,712	1,950	52 %

rRNA genes are tandemly repeated in eukaryotic genomes



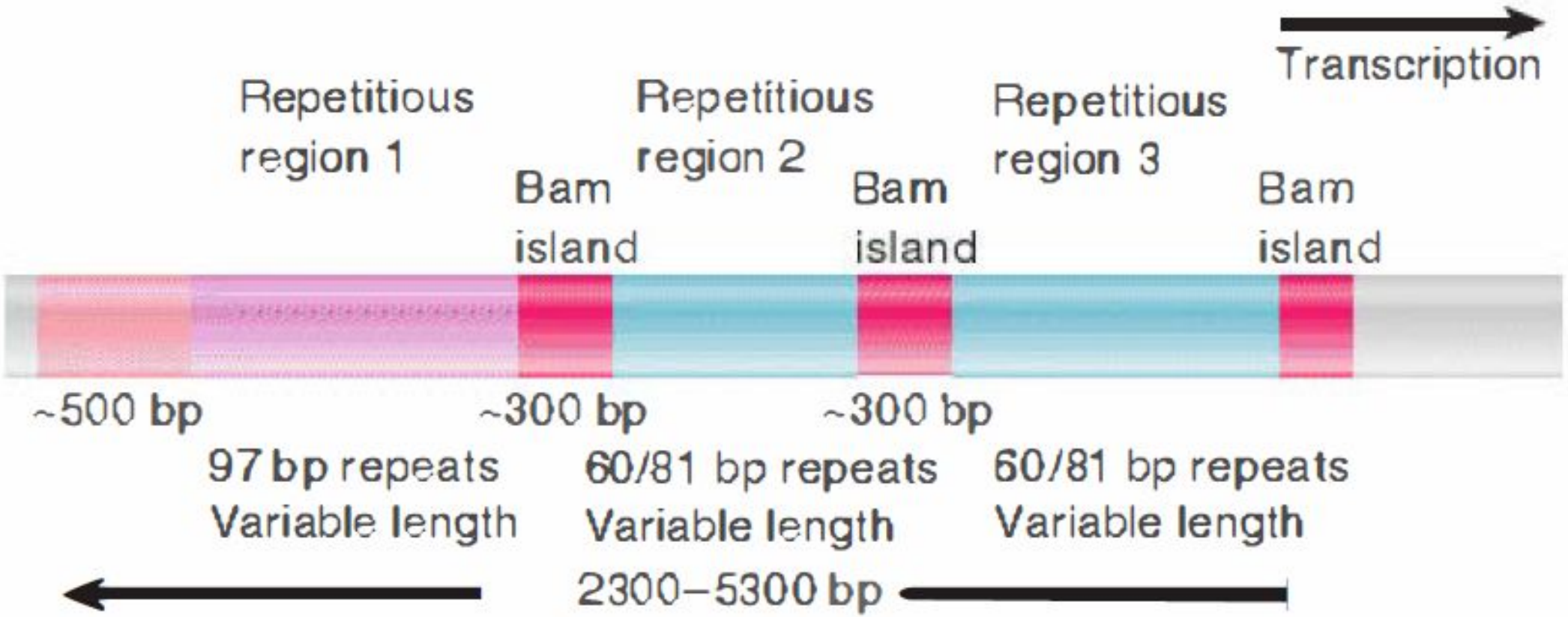


FIGURE 7.10 The nontranscribed spacer of *X. laevis* rDNA has an internally repetitious structure that is responsible for its variation in length. The Bam islands are short, constant sequences that separate the repetitious regions.

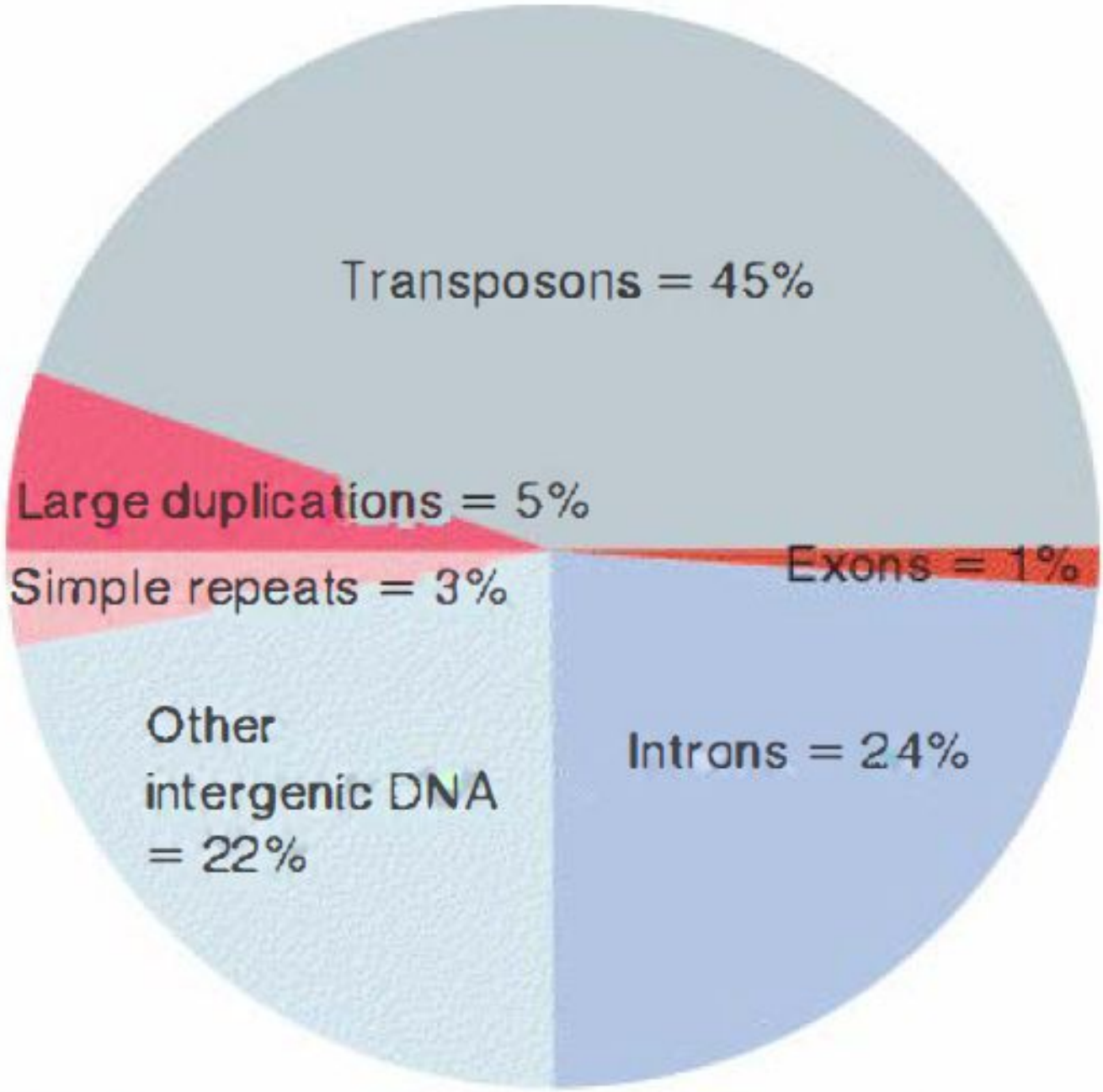
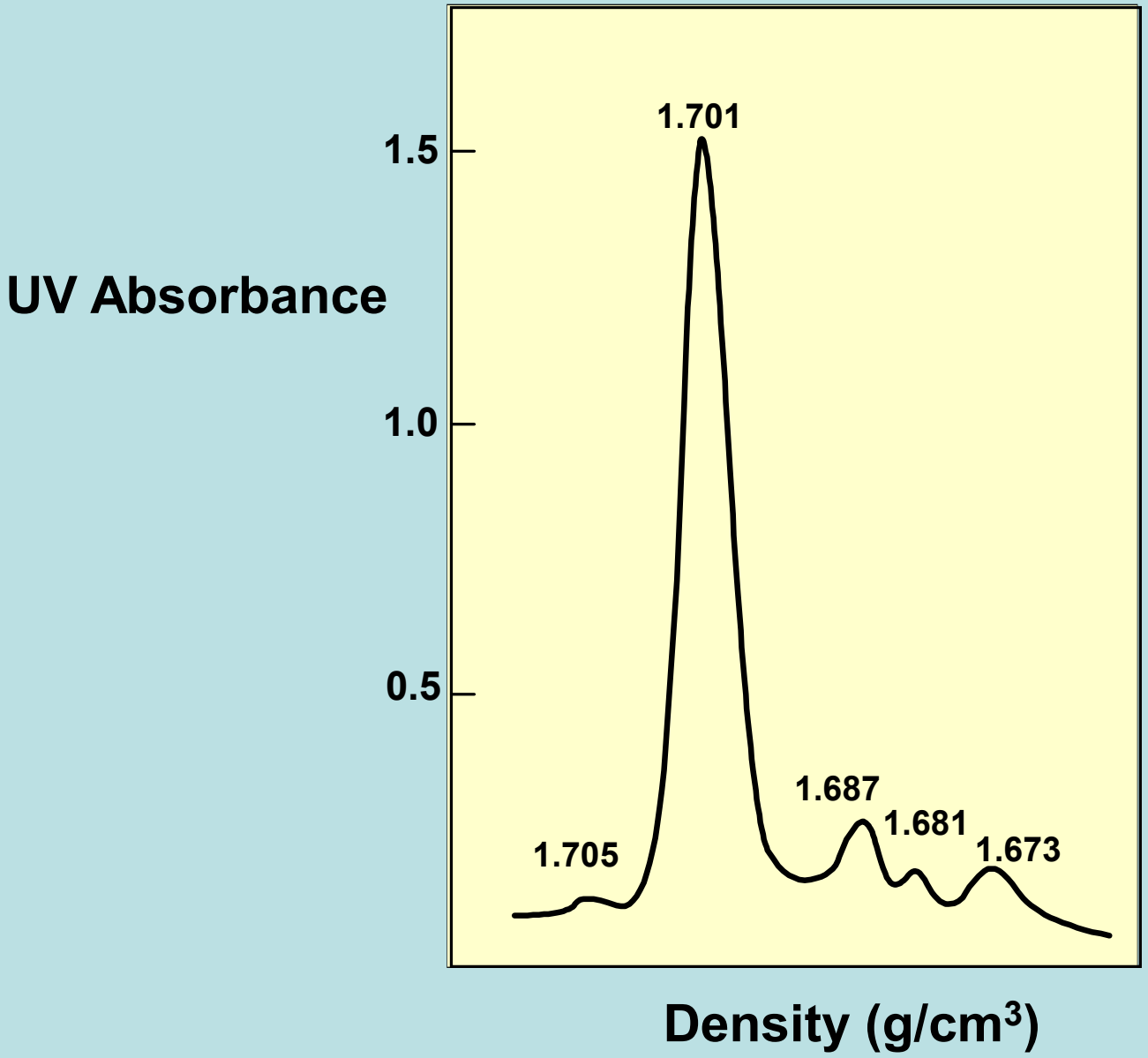


FIGURE 6.14 The largest component of the human genome consists of transposons. Other repetitive sequences include large duplications and simple repeats.

REPETITIVE DNA IN GENOME

- Multiple tandem copies: satellite DNA
- Dispersed: mobile elements (retrovirus, retrotransposons and transposons)

Analysis of the *Drosophila melanogaster* DNA using a CsCl density gradient ultracentrifugation



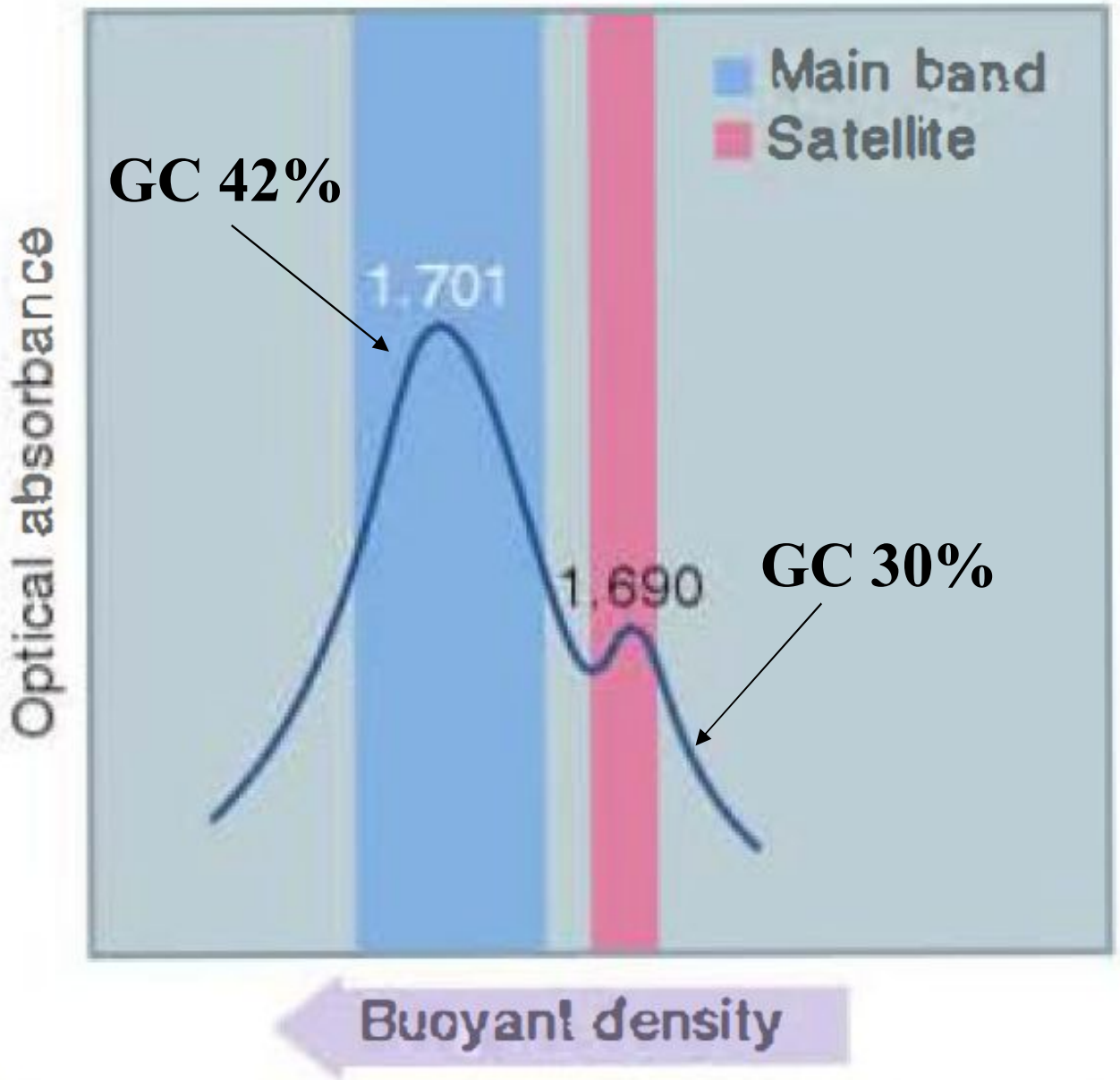


FIGURE 7.12 Mouse DNA is separated into a main band and a satellite band by centrifugation through a density gradient of CsCl.

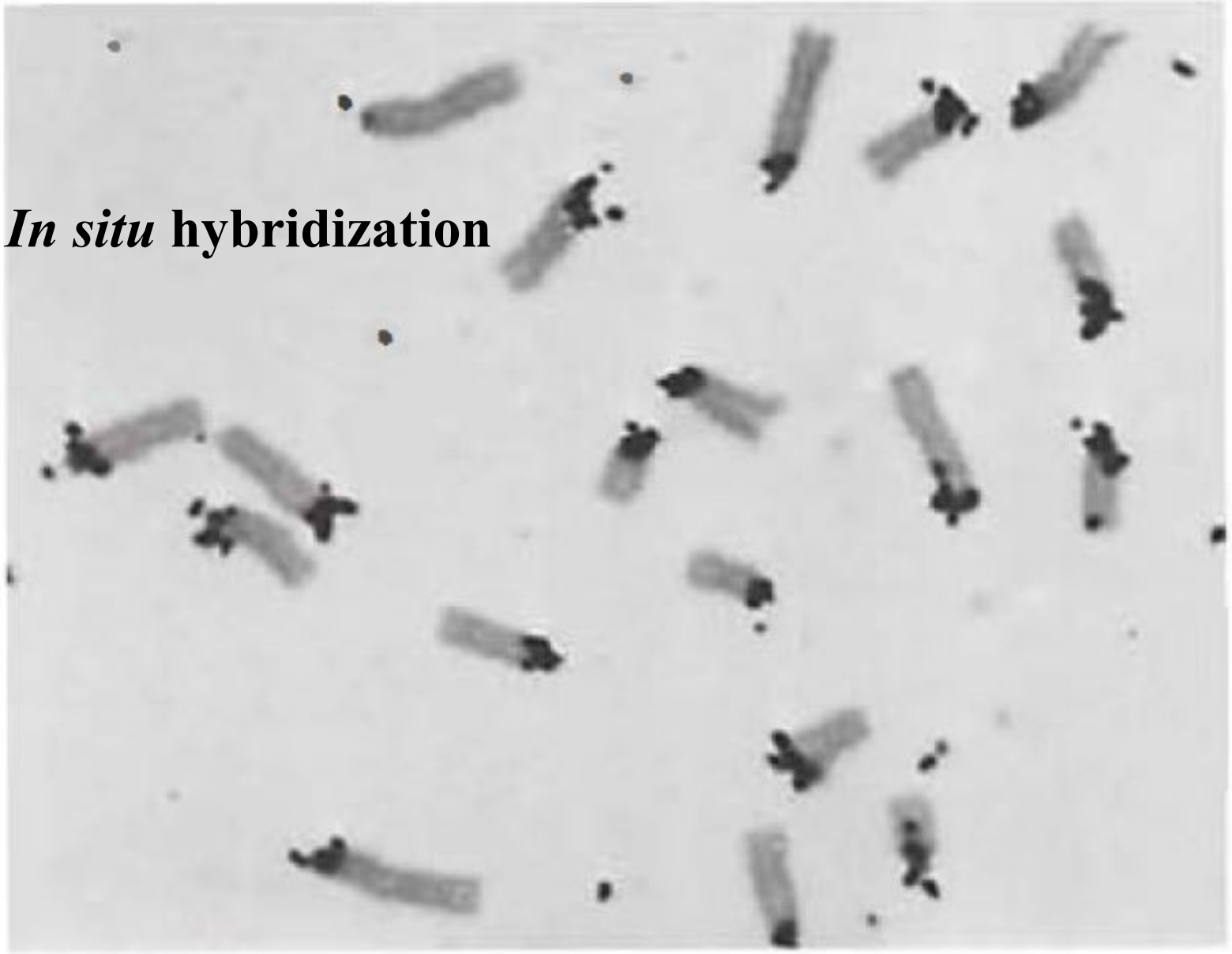


FIGURE 7.13 Cytological hybridization shows that mouse satellite DNA is located at the centromeres. Photo courtesy of Mary Lou Pardue and Joseph G. Gall, Carnegie Institution.

The repeating units of arthropod satellite DNAs are only a few nucleotides long. Most of the copies of the sequence are identical.

Satellite	Predominant Sequence	Total Length	Genome Proportion
I	A C A A A C T T G T T T G A	1.1×10^7	25%
II	A T A A A C T T A T T T G A	3.6×10^6	8%
III	A C A A A T T T G T T T A A	3.6×10^6	8%
Cryptic	A A T A T A G T T A T A T C		

The three major satellites have closely related sequences.

FIGURE 7.14 Satellite DNAs of *D. virilis* are related. More than 95% of each satellite consists of a tandem repetition of the predominant sequence.

In mammals the satellite DNA units are longer (200-300 bp) and have evolved by duplication/mutation starting from a short repetitive unit, forming a longer repeating unit that is itself repeated in tandem with some variation.

FIGURE 7.15 The repeating unit of mouse satellite DNA contains two half-repeats, which are aligned to show the identities (in blue).

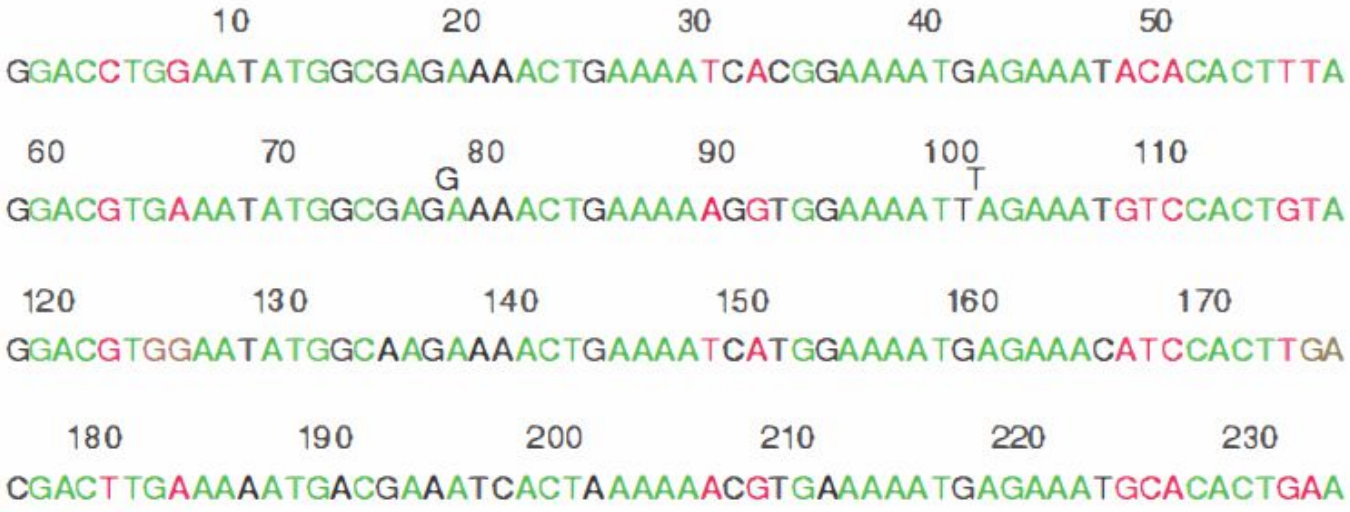
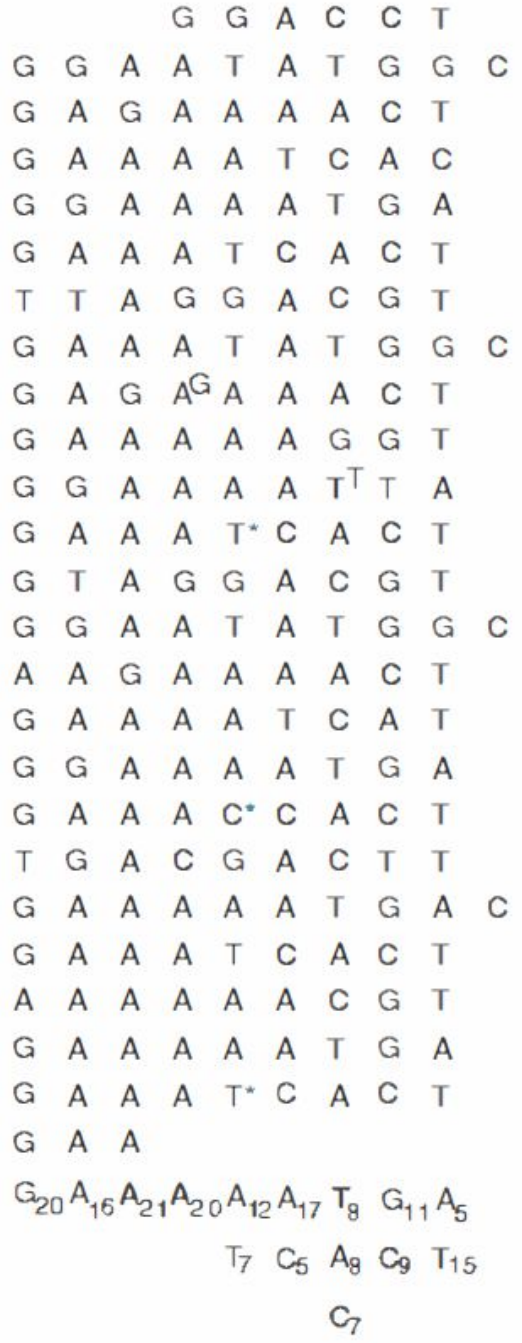


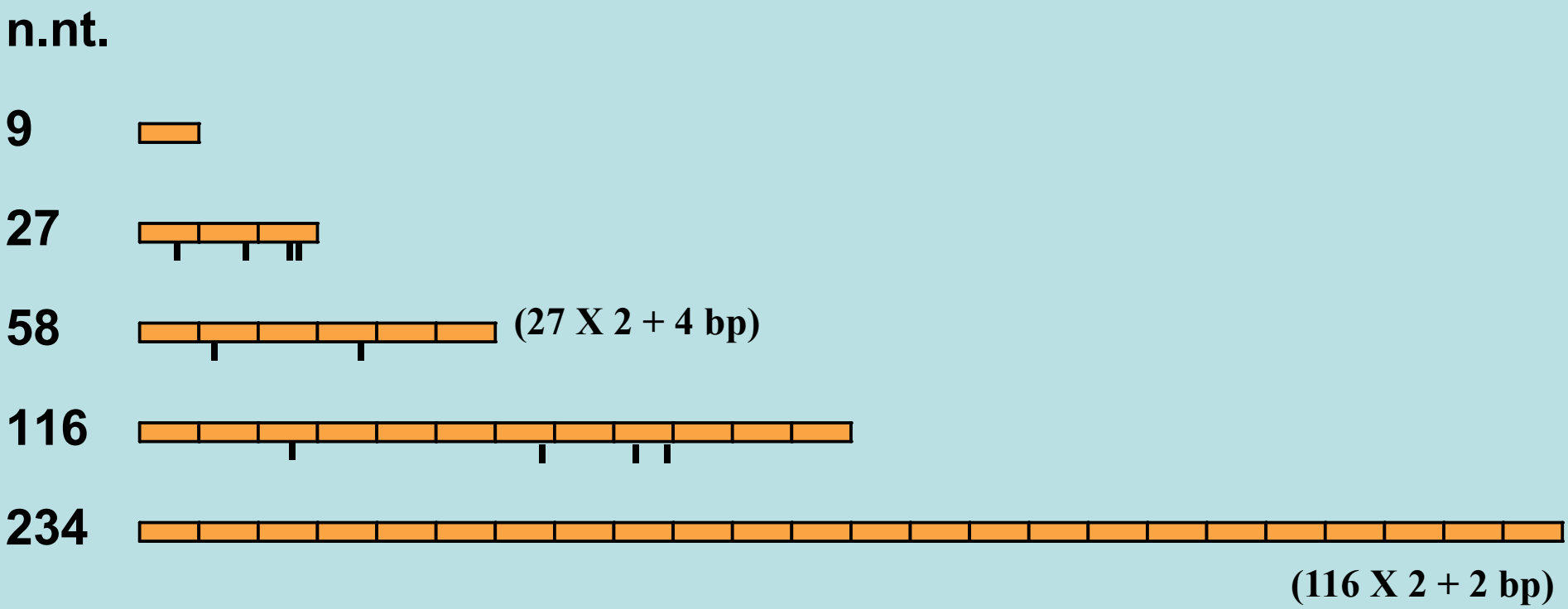
FIGURE 7.16 The alignment of quarter-repeats identifies homologies between the first and second half of each half-repeat. Positions that are the same in all four quarter-repeats are shown in green. Identities that extend only through three-quarters of the quarter repeats are in black, with the divergent sequences in red.



* indicates inserted triplet in β sequence
C in position 10 is extra base in α sequence

FIGURE 7.18 The existence of an overall consensus sequence is shown by writing the satellite sequence as a 9-bp repeat.

SATELLITE DNA EVOLUTION THROUGH DUPLICATION AND DIVERGENCY OF A SIMPLE SEQUENCE



Tandemly repeated DNA

Satellites → repeat unit with variable length, with a high number of repeats.

Minisatellites and microsatellites → short repeated units (<10 bp (micro), 10-100 (mini), number of repeats between 5 –50.

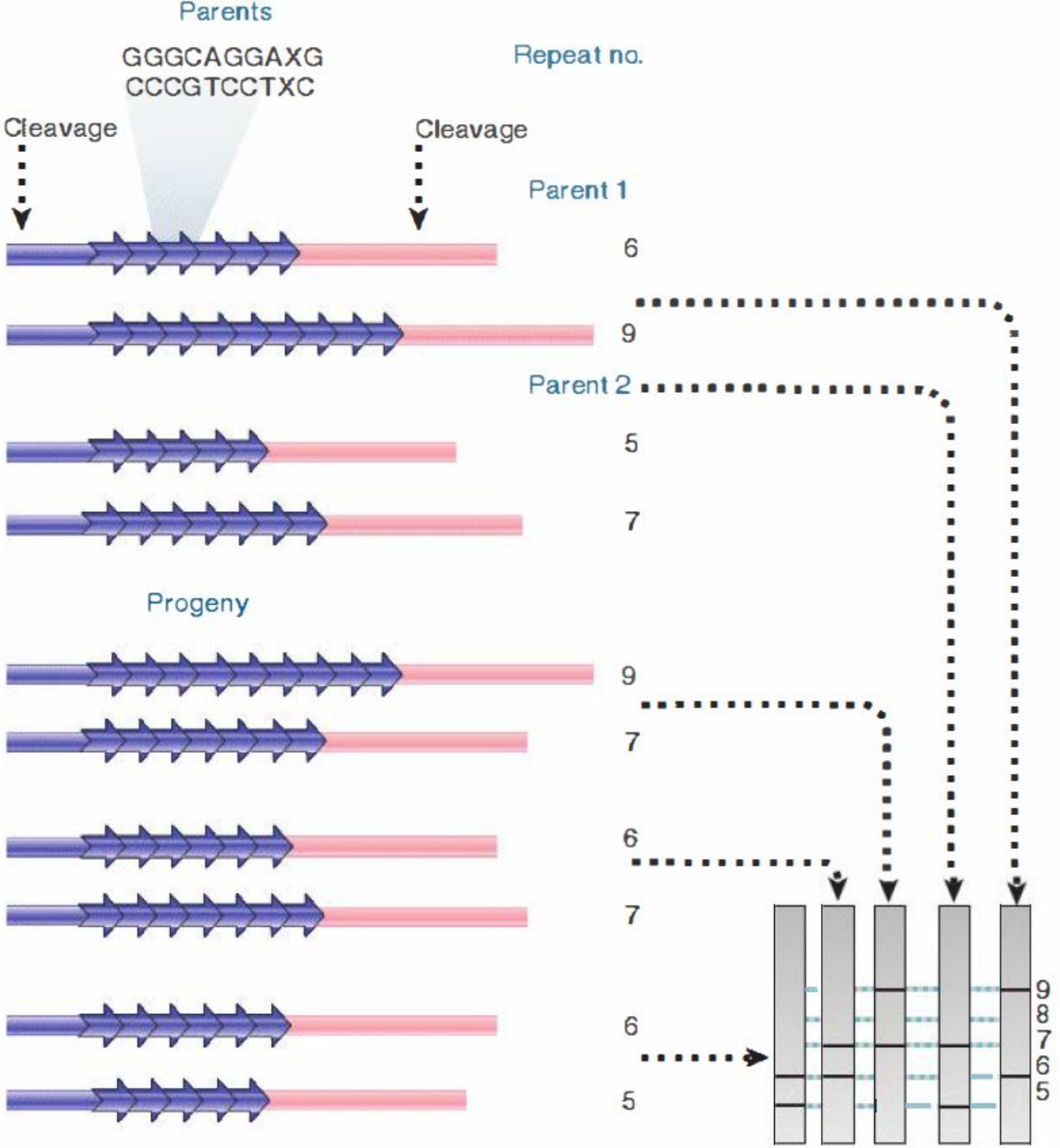


FIGURE 7.20 Alleles may differ in the number of repeats at a minisatellite locus, so that digestion on either side generates restriction fragments that differ in length. By using a minisatellite with alleles that differ between parents, the pattern of inheritance can be followed.

Origin of the repeated sequences

- Slippage during replication
- Unequal crossing over

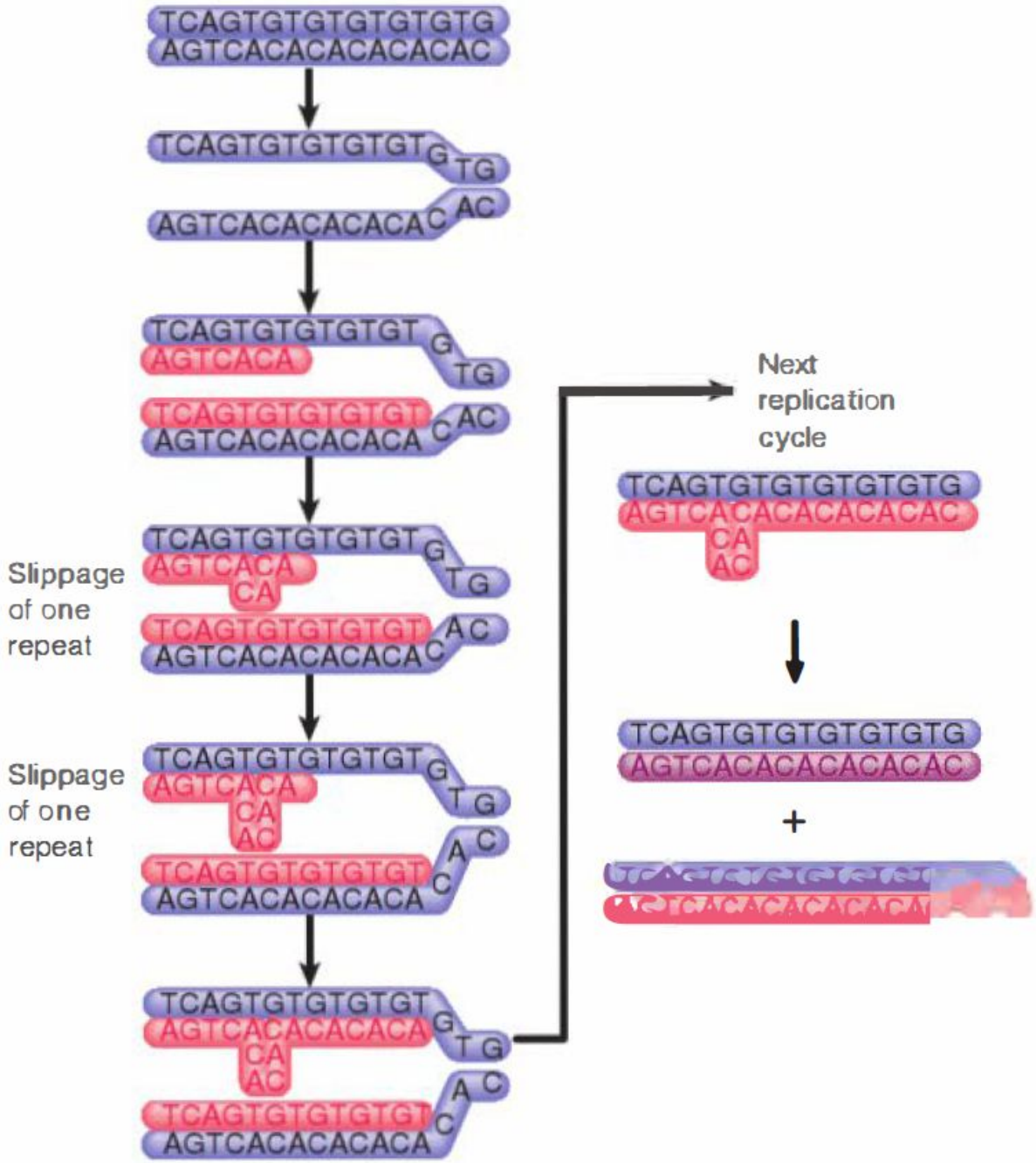
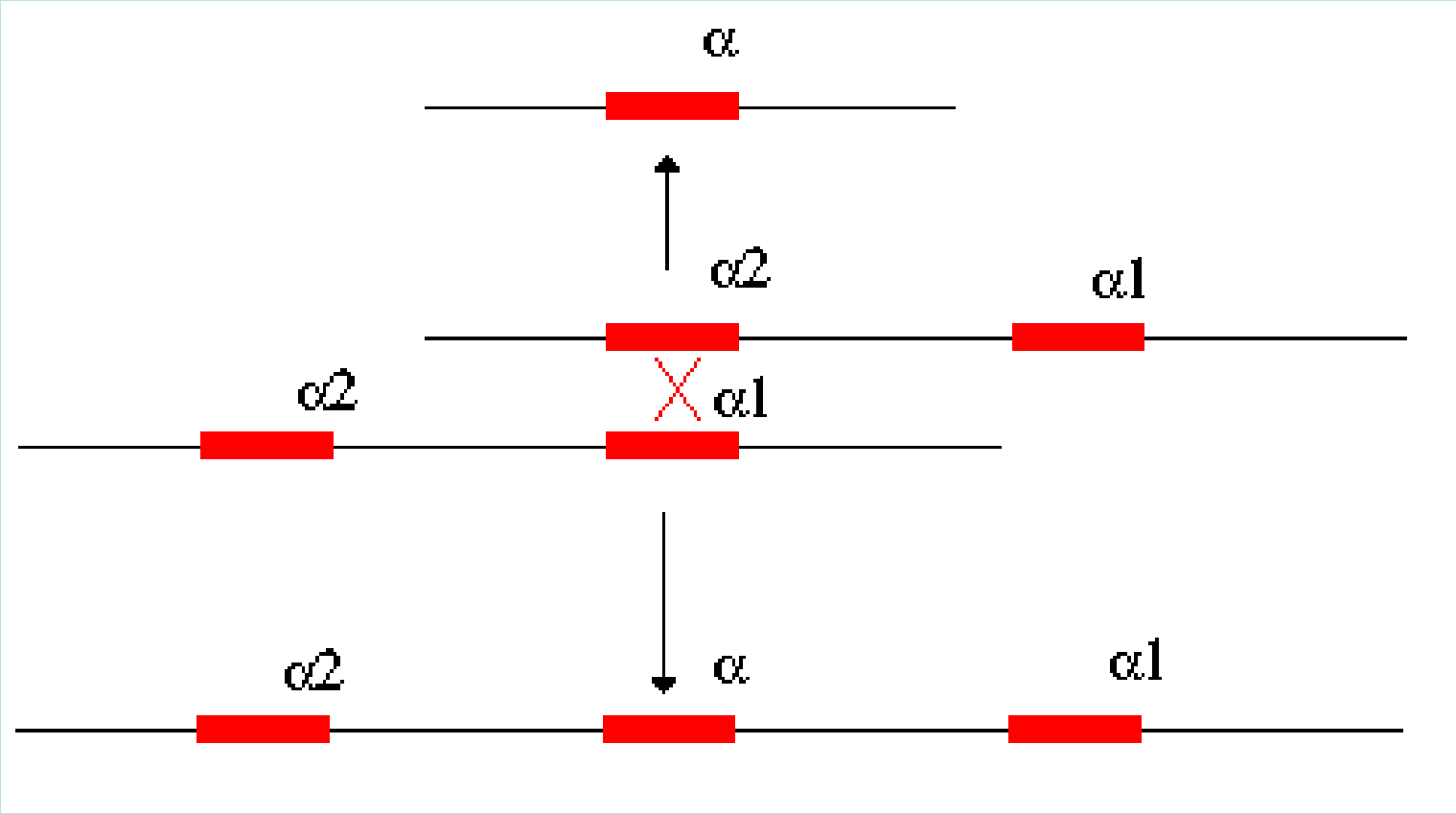


FIGURE 7.21 Replication slippage occurs when the daughter strand slips back one repeating unit in pairing with the template strand. Each slippage event adds one repeating unit to the daughter strand. The extra repeats are extruded as a single-strand loop. Replication of this daughter strand in the next cycle generates a duplex DNA with an increased number of repeats.

Equal crossing over



Transposable elements

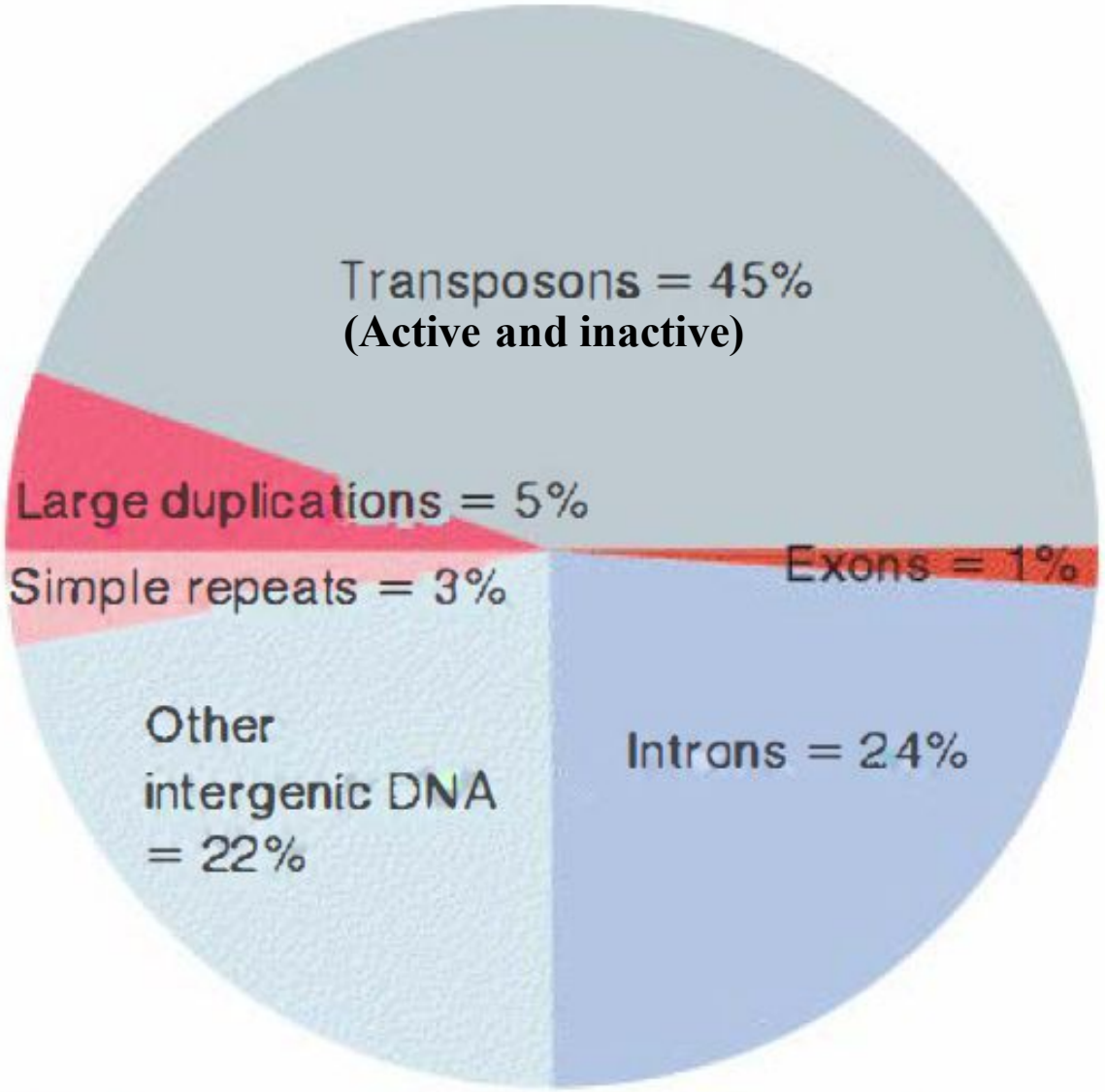


FIGURE 6.14 The largest component of the human genome consists of transposons. Other repetitive sequences include large duplications and simple repeats.

Transposable (Mobile) Elements

Short sequences of DNA (up to ~5 kb) that have the ability to move to new locations in the genome and/or to make additional copies of themselves.

Transposones → DNA intermediate

Retroelements → RNA intermediate

They were discovered by Barbara McClintock early in her career, for which she was awarded a Nobel prize in 1983.

C → colored caryopsis
c → uncolored caryopsis

Ds (dissociator) → mobile element that have the ability to convert C into c and, when removed, the C form of the allele could be reconstituted.

RETROELEMENTS

- Retroviruses
- Retrotransposons
  - Virals (Ty, copia)
  - Non virals (LINE, SINE)

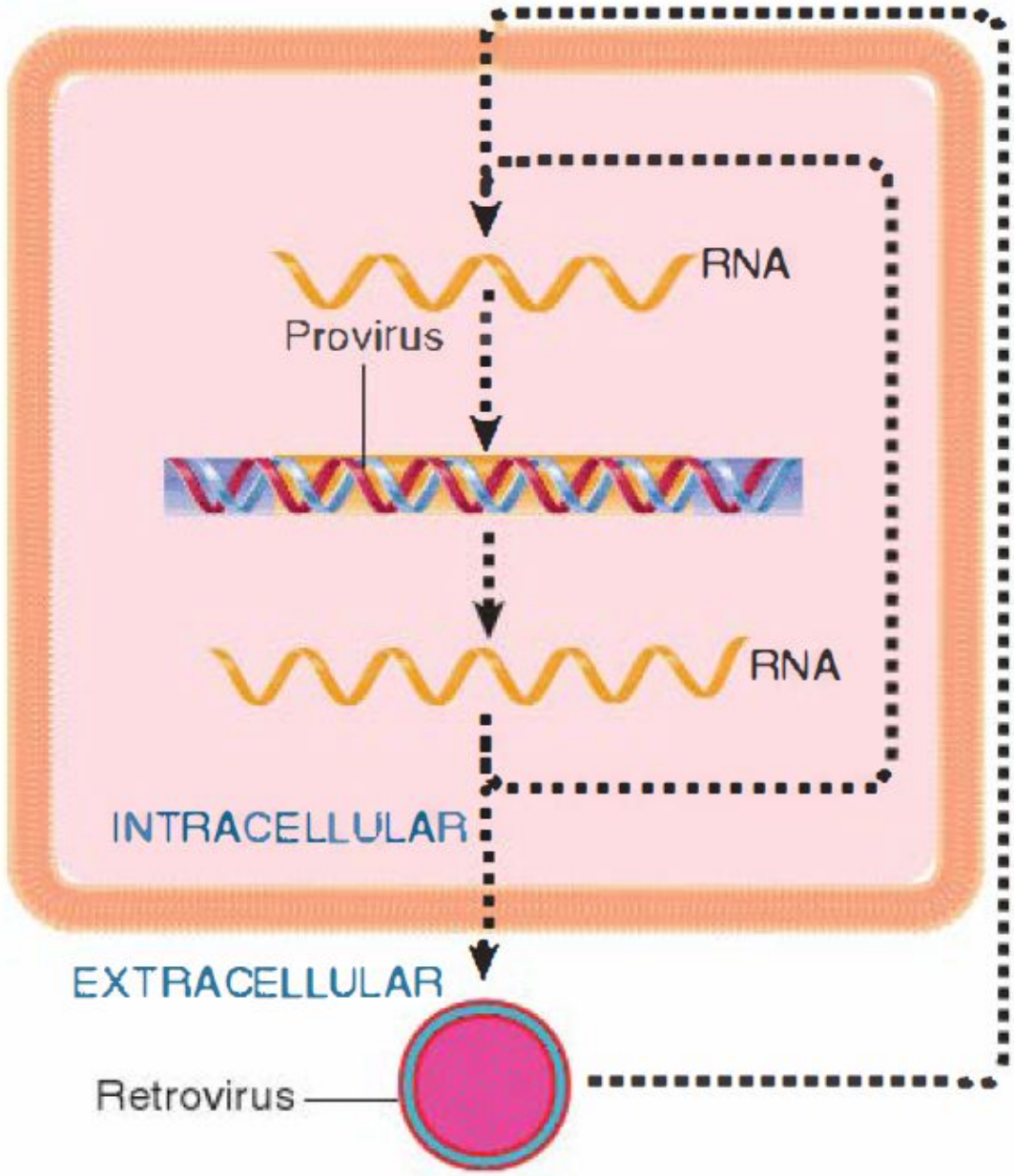


FIGURE 17.2 The reproductive cycles of retroviruses and retrotransposons alternate reverse transcription from RNA to DNA with transcription from DNA to RNA. Only retroviruses can generate infectious particles. Retrotransposons are confined to an intracellular cycle.

Long Terminal Repeats (U3-R-U5)

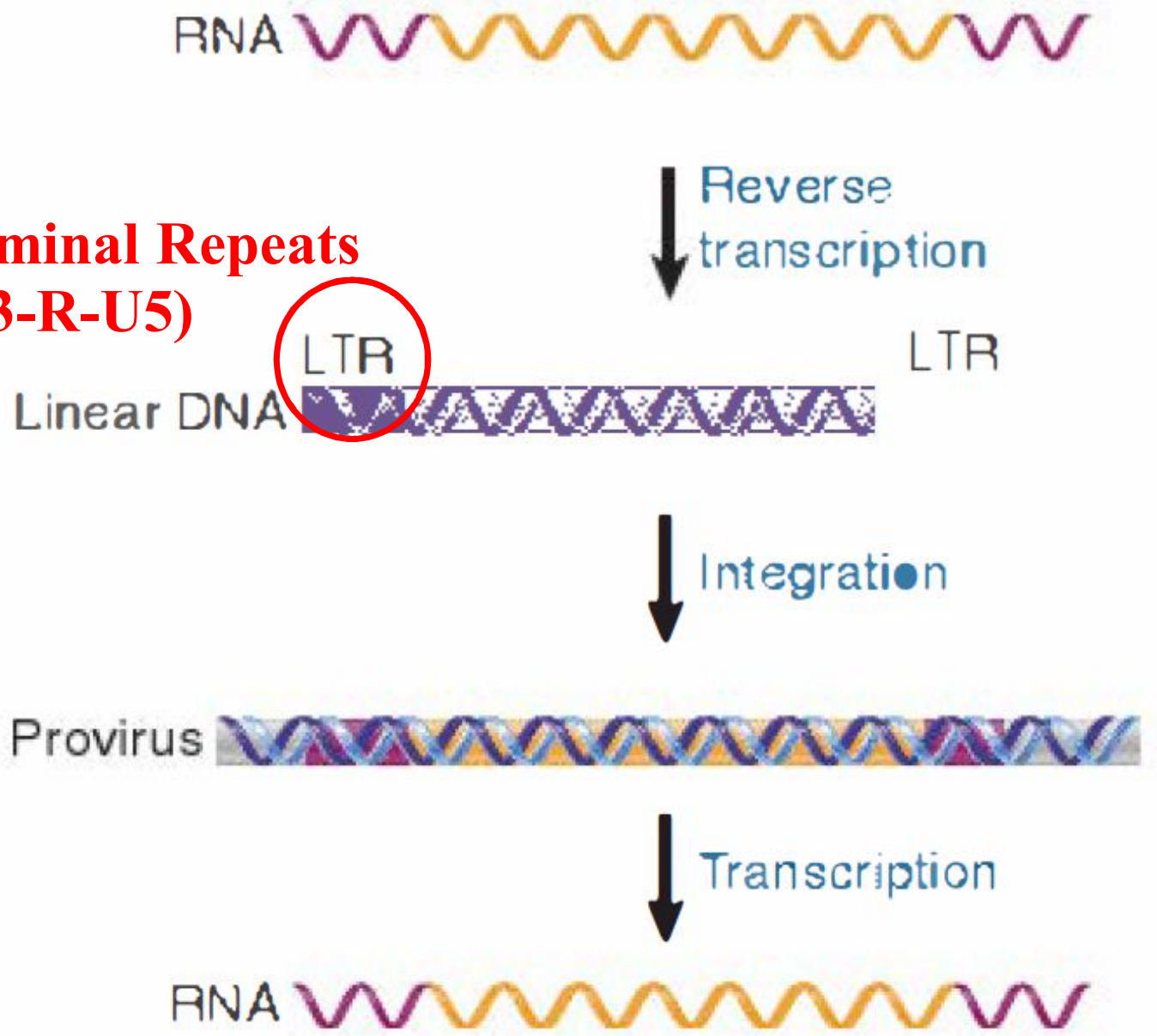


FIGURE 17.20 The retroviral lifecycle proceeds by reverse transcribing the RNA genome into duplex DNA, which is inserted into the host genome, in order to be transcribed into RNA.

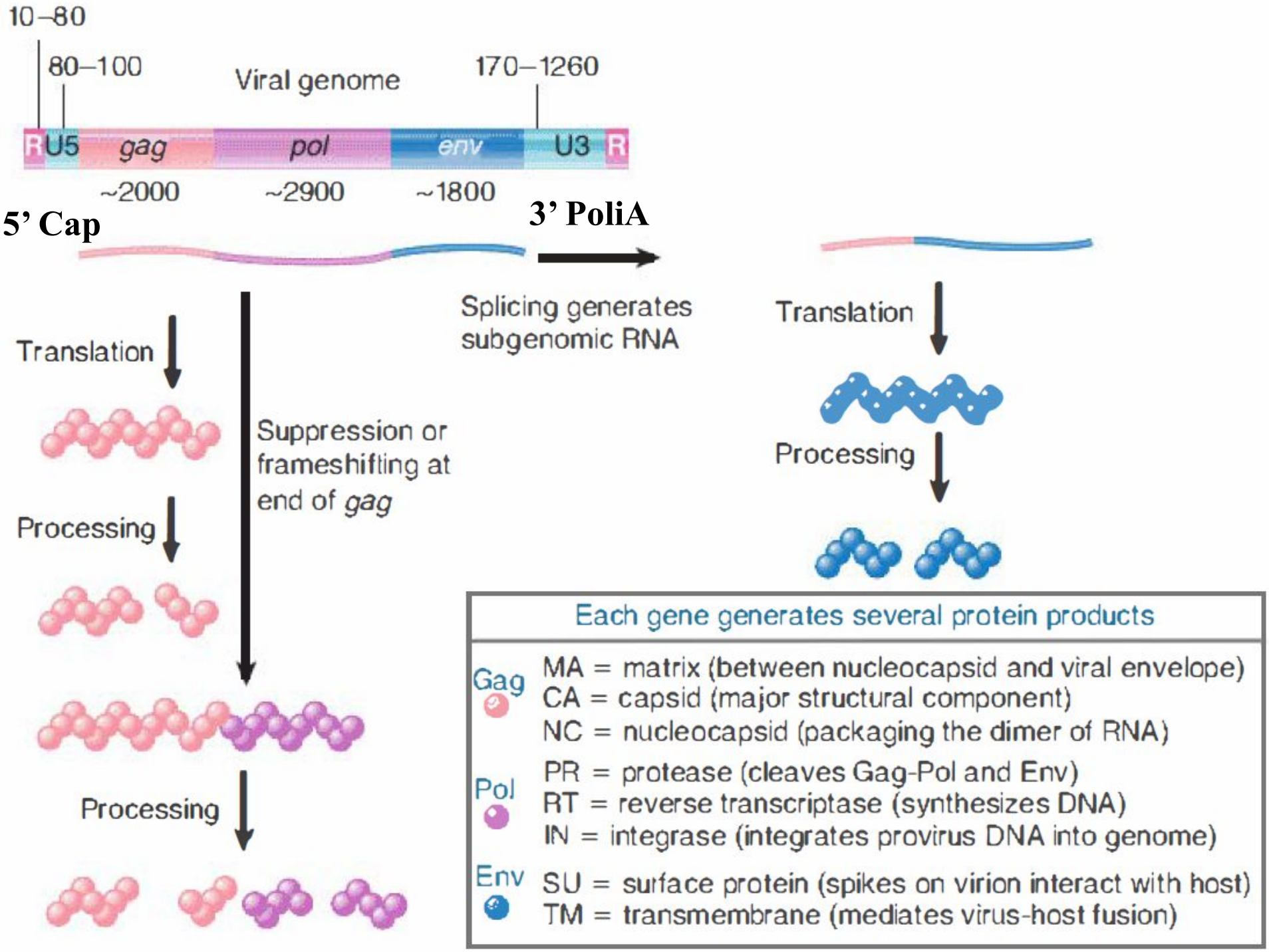


FIGURE 17.21 The genes of the retrovirus are expressed as polyproteins that are processed into individual products.


Protease Action
HHMI

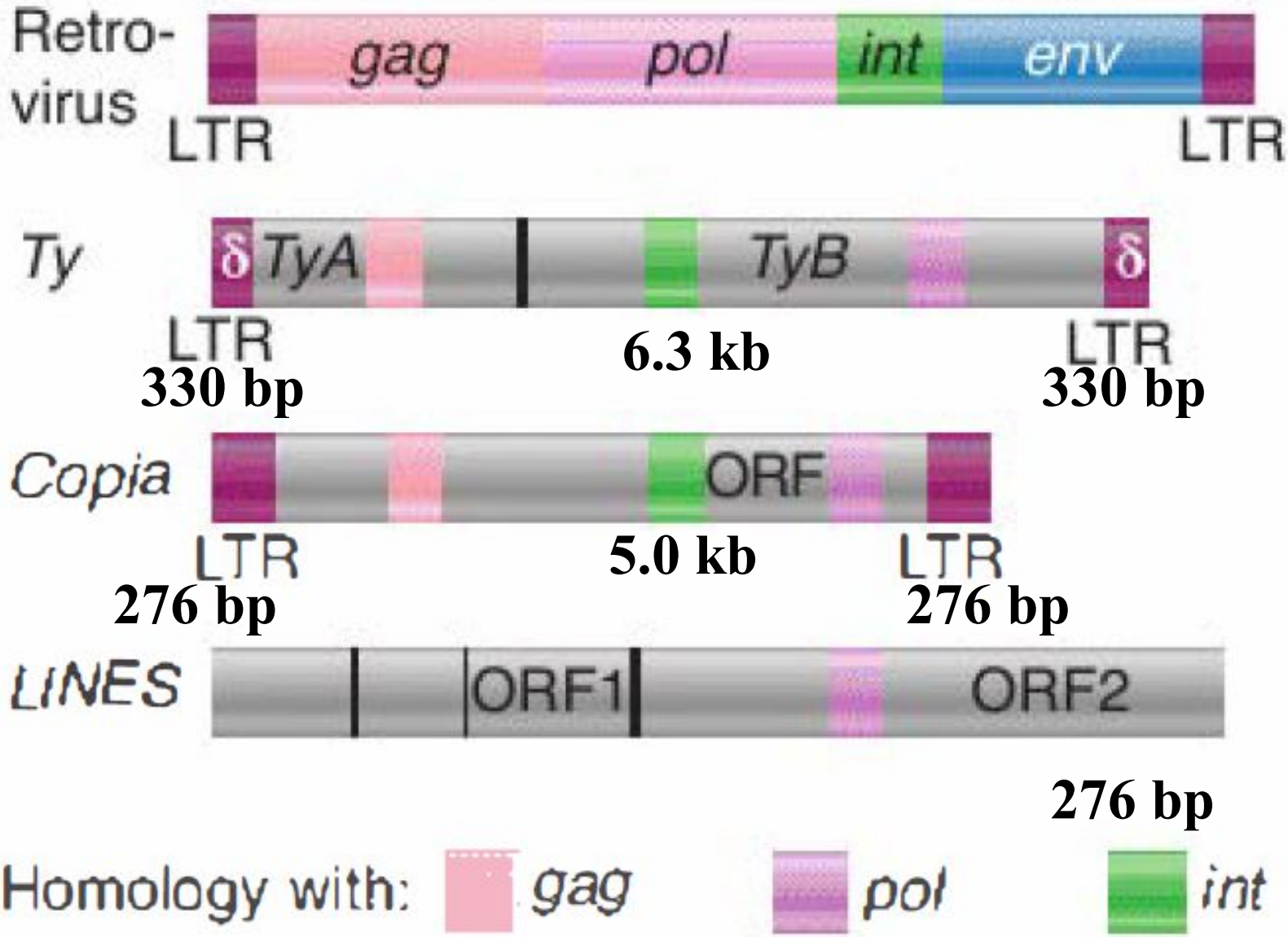
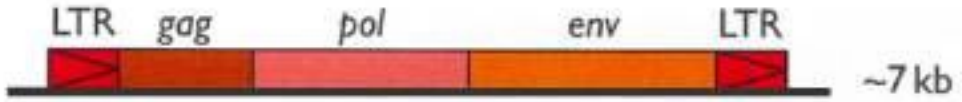
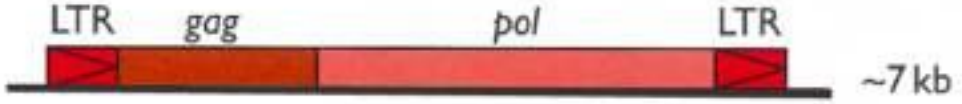


FIGURE 17.31 Retrotransposons that are closely related to retroviruses have a similar organization, but non-LTR retroposons such as LINEs share only the reverse transcriptase activity and lack LTRs.

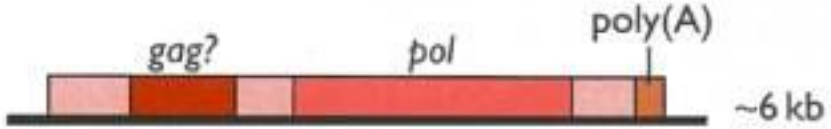
(A) Retrovirus



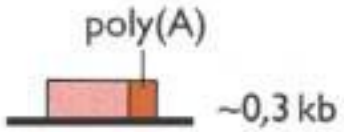
(B) Retrotrasposone *Tyl/copia*



(C) LINE



(D) SINE



LINEs (Long Interspersed Nuclear Elements)

- Have reverse transcriptase and endonuclease
- Do not have LTRs but have a Poly(A) end
- Come from transcription activity of RNA pol II
- Two Open Reading Frames (ORF1 e ORF2)
- E.g. LINE-1 (Mammals L1) 6,5 Kb

3500 whole copies and hundred thousands truncated

SINEs (Short Interspersed Nuclear Elements)

- Non coding for reverse transcriptase (nonautonomous)
- Derived from RNA pol III transcripts
- E.g. Alu sequences (ca 300 bp) about 300000 Alu sequences in the human haploid genome
- One milion copies
- Substrate for L1 active elements

	LTR retrotransposons	non-LTR retroposons	SINES
Common types	Ty (*S. cerevisiae*) copia (*D.melanogaster*) Tnt1A (*N. tabacum*)	L1 (human) B1, B2ID, B4 (mouse) Cin4 (*Z. mays*)	SINES (mammals) Pseudogenes of pol III transcripts
Termini	Long terminal repeats	No repeats	No repeats
Target repeats	4–6 bp	7–21 bp	7–21 bp
Enzyme activities	Reverse transcriptase and/or integrase	Reverse transcriptase /endonuclease	None (or none coding for transposon products)
Organization	May contain introns (removed in subgenomic mRNA)	One or two uninterrupted ORFs	No introns

FIGURE 17.30 Retroelements can be divided into LTR retrotransposons, non-LTR retroposons, and the nonautonomous SINEs.

DNA Transposons

- Few in eukaryotes (100 copies in humans)
- Ac/Ds elements in corn (Activator/Dissociator)
- Horizontal transfer of genes
- Transposition with "cut and paste" mechanism

(transposase) → non replicative mechanism

- Replicative transposition (TnA)

(transposase and resolvase)

Element	Organization	Length (kb)	Human genome	
			Number	Fraction
Retrovirus/LTR retrotransposon		1–11	450,000	8%
LINES (autonomous), e.g., L1		6–8	850,000	17%
SINES (nonautonomous), e.g., Alu		<0.3	1,500,000	15%
DNA transposon		2–3	300,000	3%

FIGURE 17.32 Four types of transposable elements constitute almost half of the human genome.

Transposable (Mobile) Elements

- Accumulates in the eukaryotic genome during evolution up to the 40% of the total content.
- They do not have apparently other function apart to hand down their own existence.
- They cause less mutations to phenotype than expected if compared to their abundance → suppression of uncontrolled transposition (E.G. cytosine metilation).
- An important role during evolution → sites for homologous recombination and have generated new genes trough new combination of exons already present (exon shuffling).