Transposable elements

II MSc Botany

Dr Giby Kuriakose

Transposable elements

- General features of transposable elements
- Prokaryotic transposable elements
- Eukaryotic transposable elements

Transposable element: mobile genetic elements of a chromosome that have the capacity to move from one location to another in the genome.

- **Normal and ubiquitous** components of prokaryote and eukaryote genomes.

  Prokaryotes-transpose to/from cell's chromosome, plasmid, or a phage chromosome.

  Eukaryotes-transpose to/from same or a different chromosome.

- **Nonhomologous recombination**: transposable elements insert into DNA that has no sequence homology with the transposon.

- Transposable elements **cause genetics changes** and make important contributions to the **evolution of genomes**:

  - Insert into genes.

  - Insert into regulatory sequences; **modify gene expression**.

  - Produce chromosomal mutations.

Transposable Elements (Transposons)

- DNA elements capable of moving ("transposing") about the genome
- Discovered by Barbara McClintock (1940), largely from cytogenetic studies in maize, but since found in most organisms
- She was studying "variegation" or sectoring in leaves and seeds
- She liked to call them "controlling elements“ because they affected gene expression in myriad ways

1. Nobelprize.org

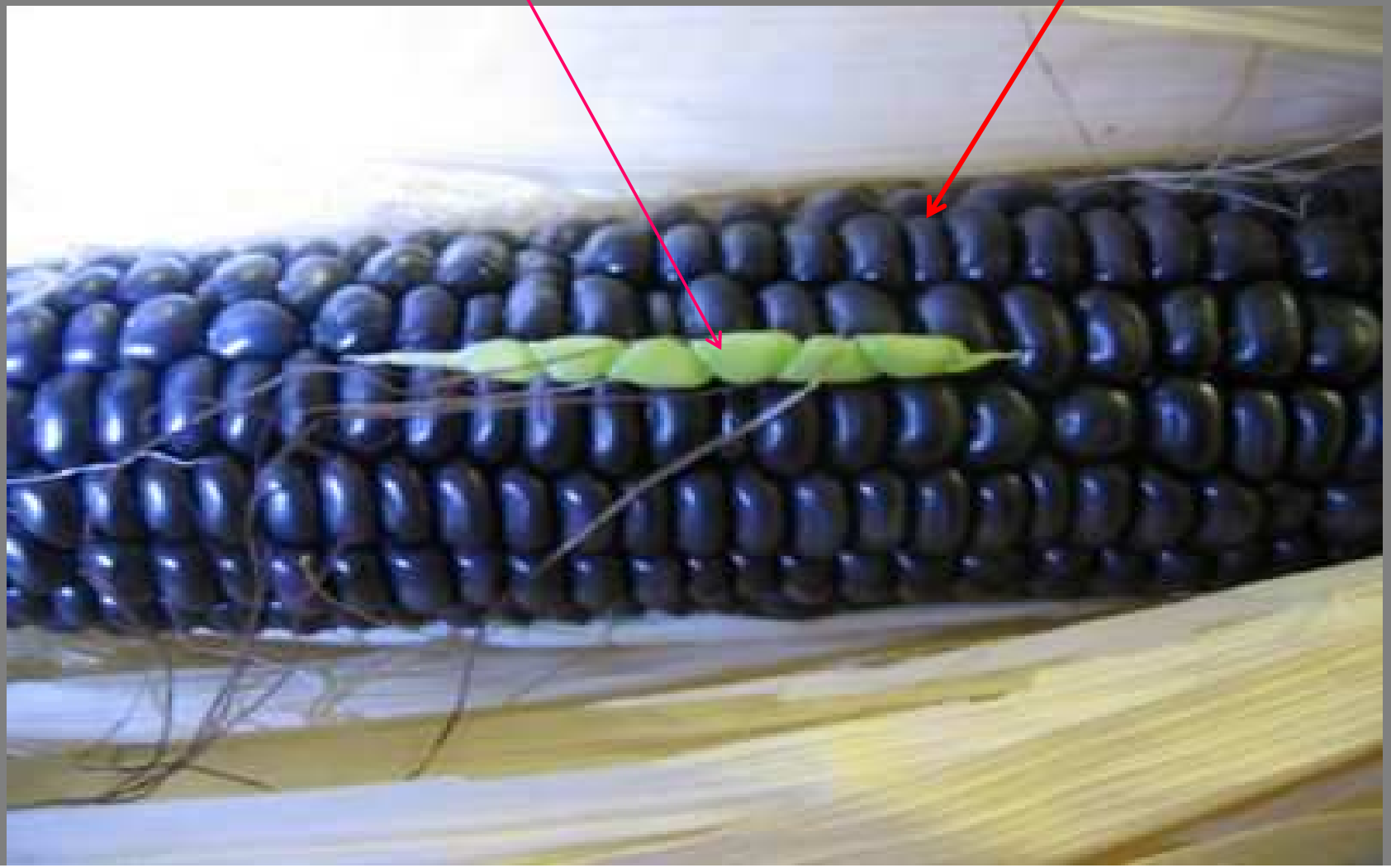
(1983 Nobel Prize in Physiology and Medicine)

2. profiles.nlm.nih.gov/LL/

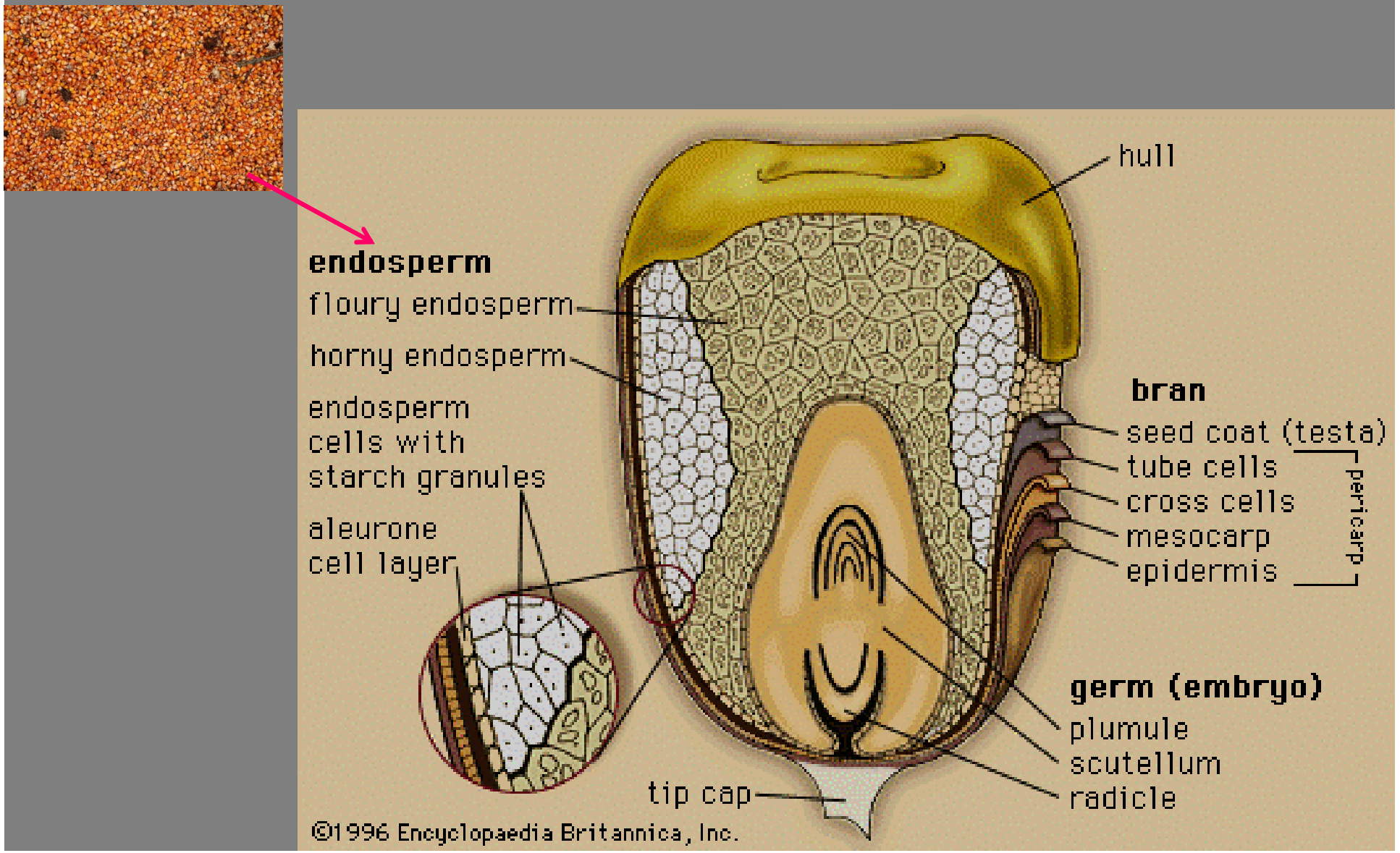
Corn (maize) varieties

Barbara McClintock 1902-1992

Corn evolution in 7000 yrs of domestication



Maize (domesticated corn) kernel structure



Other Characteristics of McClintock's Elements

- **Unstable mutations that revert frequently but often partially, giving new phenotypes.**
- **Some elements (e.g., Ds) correlated with chromosome breaks.**
- **Elements often move during meiosis and mitosis.**
- **Element movement accelerated by genome damage.**

Transposable elements:

Two classes of transposable elements/mechanisms of movement:

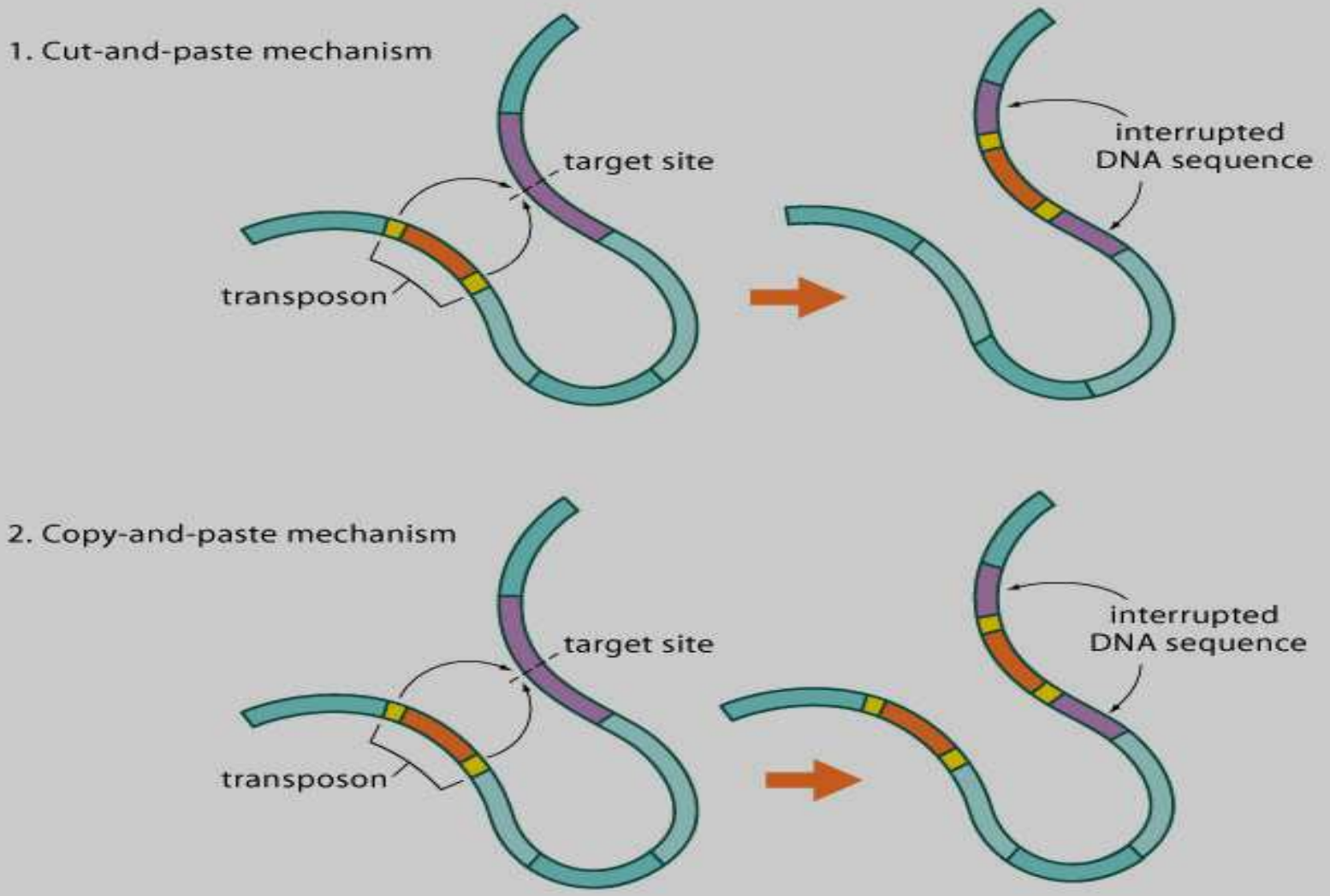
1. **Encode proteins that (1) move DNA directly to a new position or (2) replicate DNA and integrate replicated DNA elsewhere in the genome (prokaryotes and eukaryotes).**

    a) **Cut-and-paste mechanism**

    b) **Copy-and-paste mechanism**

2. **Retro-transposons encode *reverse transcriptase* and make DNA copies of RNA transcripts; new DNA copies integrate at different sites (in eukaryotes only).**

Two methods of transposition:



Broad Institute, MIT

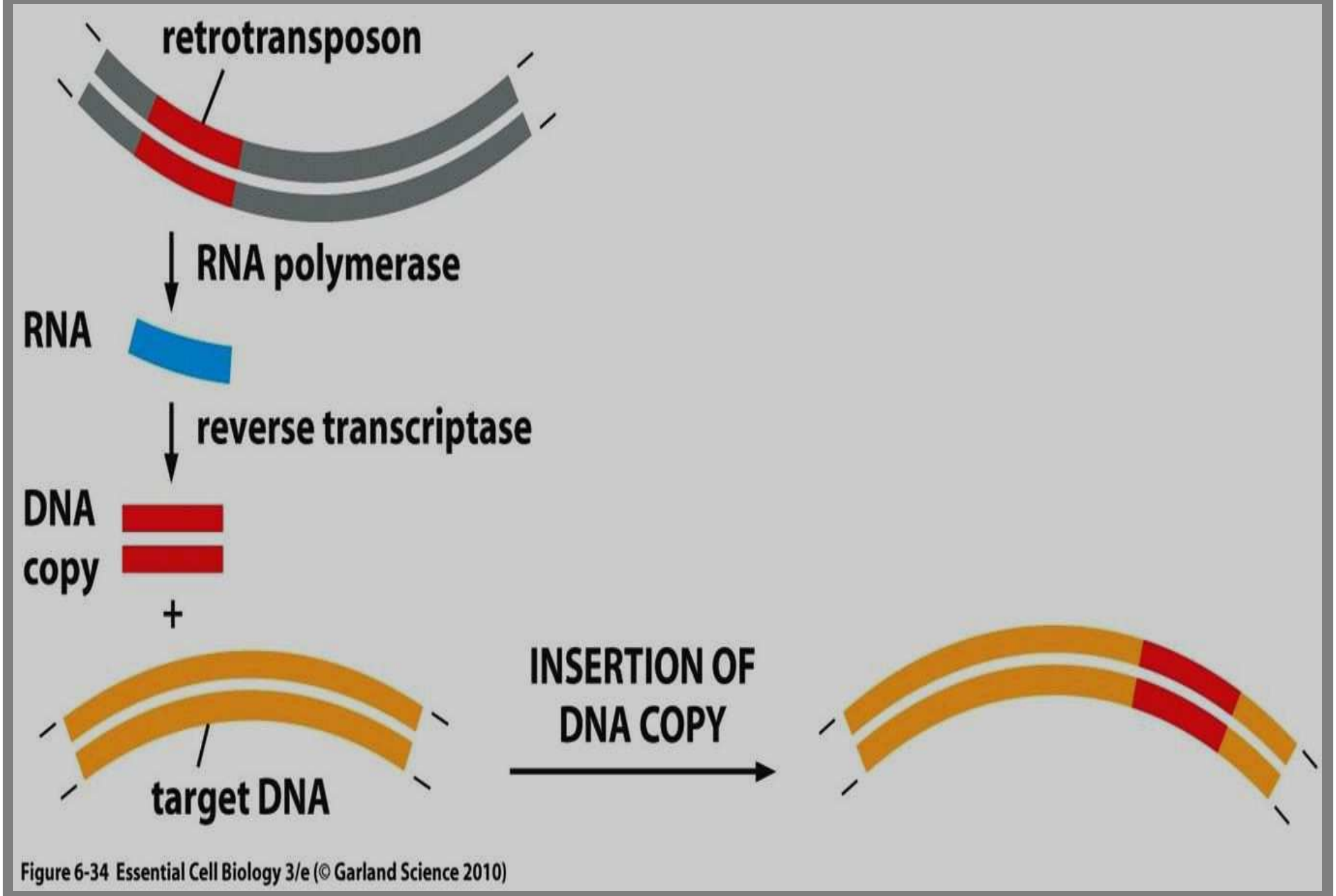


Figure 6-34 Essential Cell Biology 3/e (© Garland Science 2010)

Transposable elements in prokaryotes:

Two examples:

1. **Insertion sequence (IS) elements**

2. **Transposons (Tn)**

Insertion sequence (IS) elements:

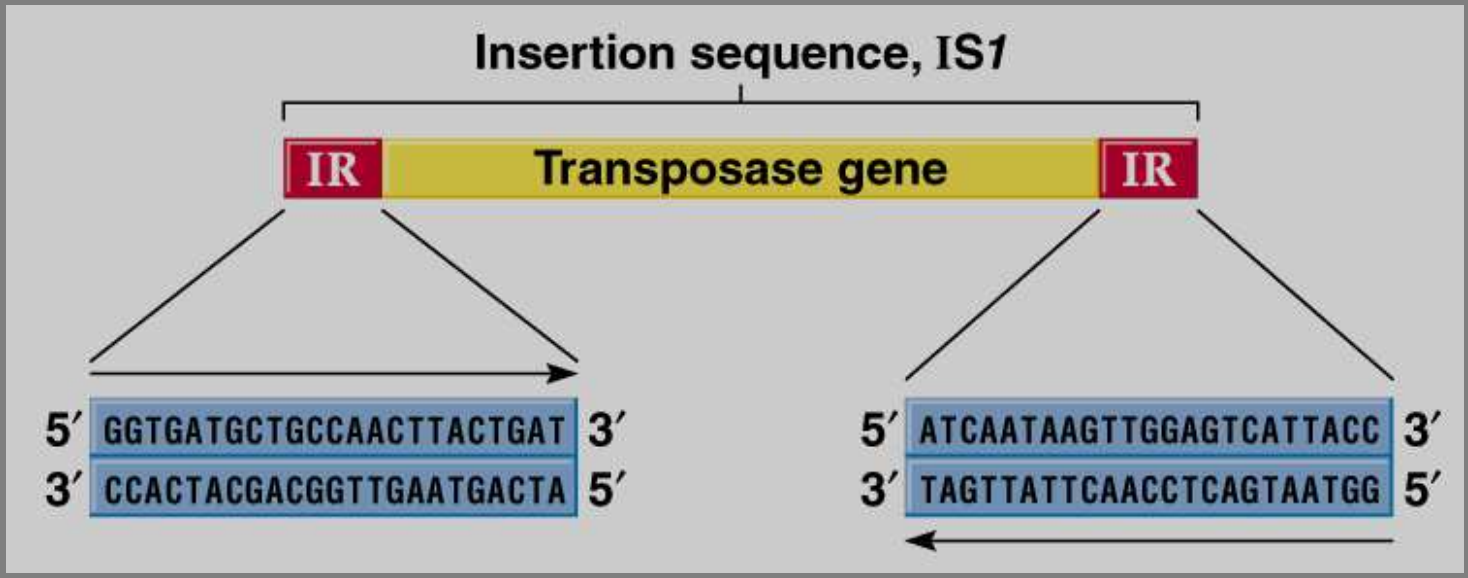
1. **Simplest type of transposable element found in bacterial chromosomes and plasmids.**

2. **Encode gene (transposase) for mobilization and insertion.**

3. **Range in size from 768 bp to 5 kb.**

4. **IS1 first identified in *E. coli*'s glactose operon is 768 bp long and is present with 4-19 copies in the *E. coli* chromosome.**

5. **Ends of all known IS elements show inverted terminal repeats (ITRs).**



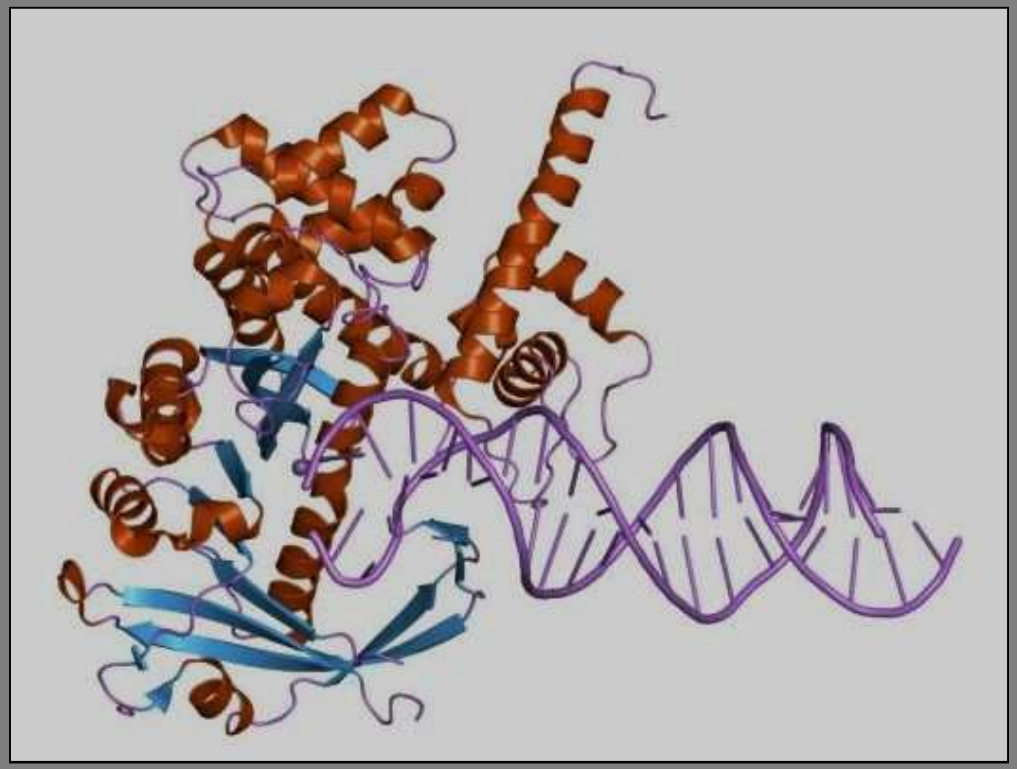
Insertion sequence (IS) elements:

Integration of an IS element may:

- **Disrupt coding sequences or regulatory regions.**
- **Alter expression of nearby genes.**
- **Cause deletions and inversions in adjacent DNA.**
- **Result in crossing-over.**

Transposition of insertion sequence (IS) elements:

1. **Original copy remains in place; new copy inserts randomly.**
2. **Transposition requires transposase, coded by the IS element.**
3. **IS element otherwise uses host enzymes for replication.**
4. **Transposition initiates when transposase recognizes ITRs.**



https://en.wikipedia.org/wiki/Transposase

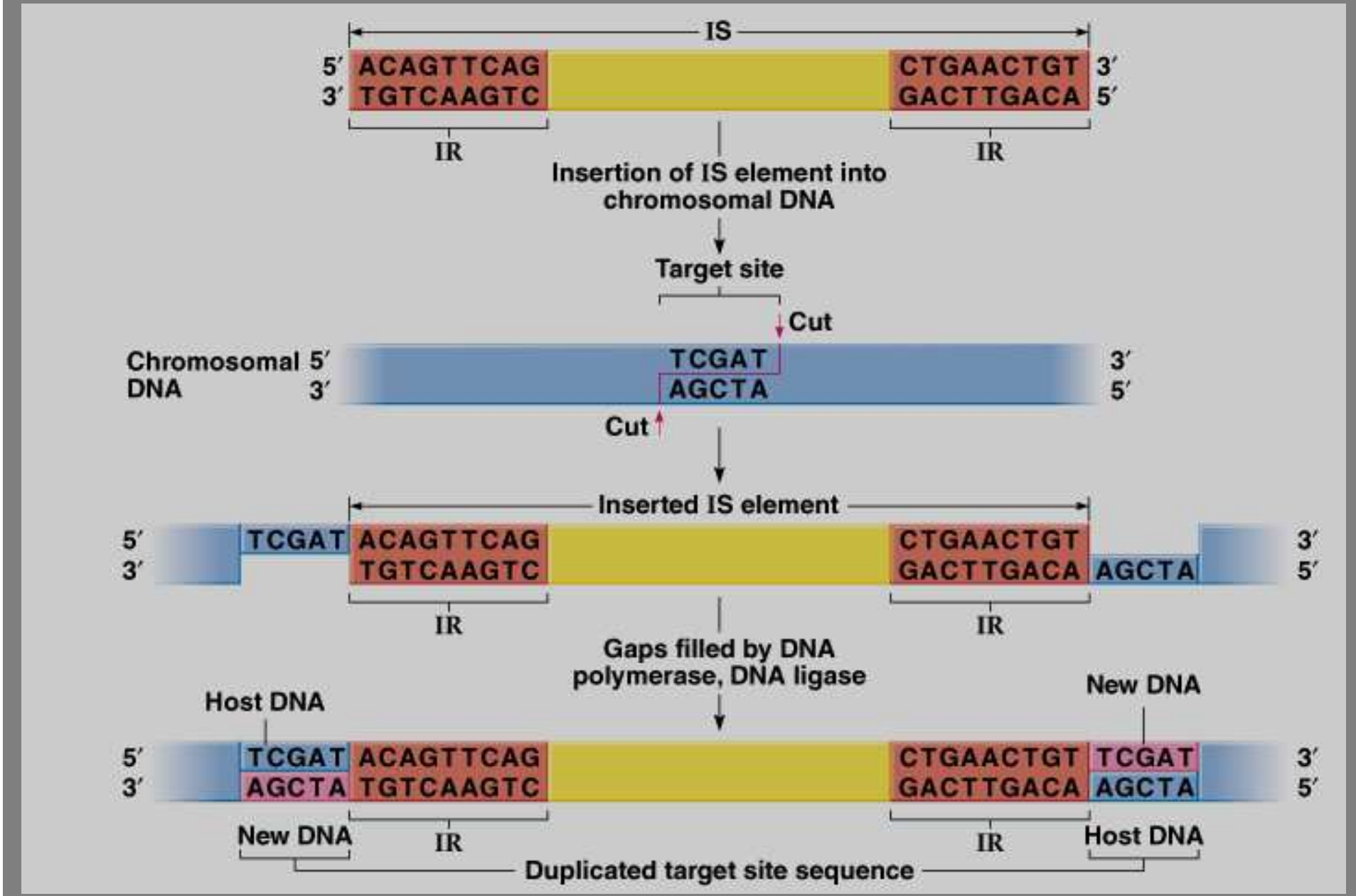
Transposition of insertion sequence (IS) elements:

1. **Site of integration = target site.**

2. **Staggered cuts are made in DNA at target site by transposase, IS element inserts, DNA polymerase and ligase fill the gaps (note--- transposase behaves like a restriction enzyme).**

3. **Small direct repeats (~5 bp) flanking the target site are created.**

Fig. 7.20, Integration of IS element in chromosomal DNA.

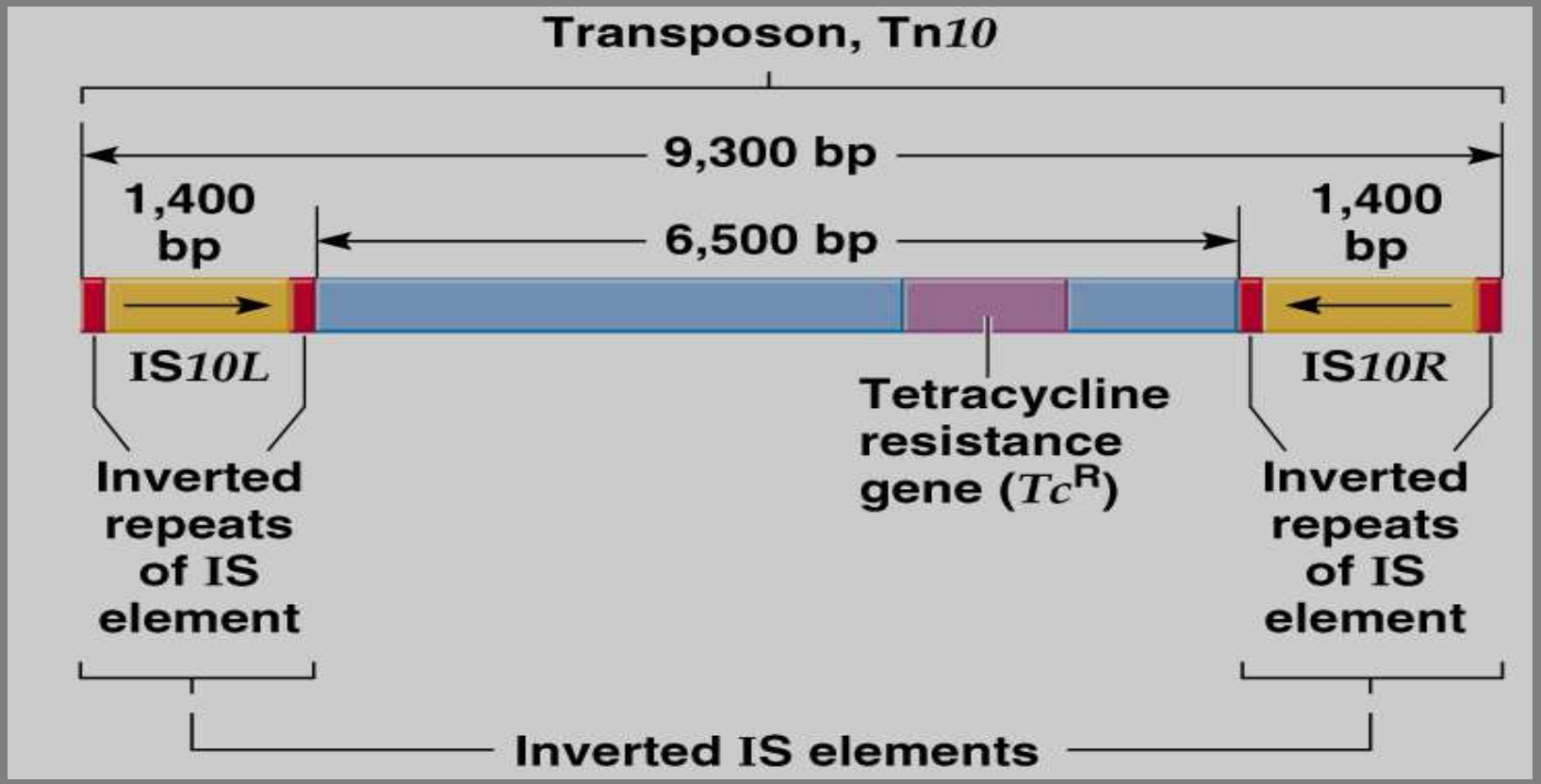


Transposons (Tn):

- **Similar to IS elements but are more complex structurally and carry additional genes**
- **2 types of transposons:**
  1. **Composite transposons**
  2. **Non-composite transposons**

Composite transposons (Tn):

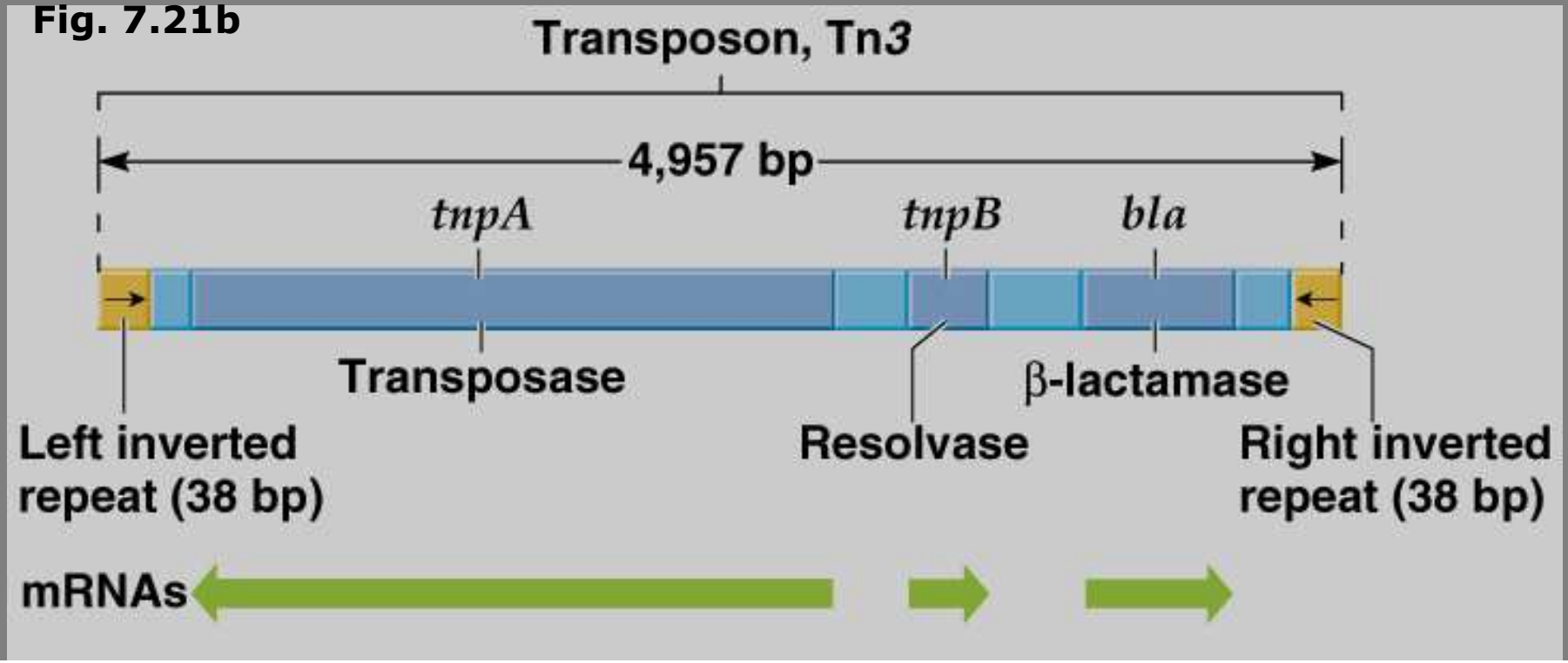
- **Carry genes (example might be a gene for antibiotic resistance) flanked on both sides by IS elements.**
- **Tn10 is 9.3 kb and includes 6.5 kb of central DNA (includes a gene for tetracycline resistance) and 1.4 kb inverted IS elements.**
- **IS elements supply transposase and ITR recognition signals.**



Noncomposite transposons (Tn):

- **Carry genes (example might be a gene for antibiotic resistance) but do not terminate with IS elements.**
- **Ends are non-IS element repeated sequences.**
- **Tn3 is 5 kb with 38-bp ITRs and includes 3 genes; *bla* (β-lactamase), *tnpA* (transposase), and *tnpB* (resolvase, which functions in recombination).**

Fig. 7.21b



IS elements

- **IS (Insertion sequence) elements: The simplest transposons, are autonomous units, each of which codes only for the proteins needed to sponsor its own transposition.**
- **Inverted repeats + transposase genes**

Keywords

4) **Transposase:** an enzyme that binds to ends of transposon and catalyses the movement of the transposon to another part of the genome by a cut and paste mechanism or a replicative transposition mechanism.

5) **Resolvase or Recombinase (nuclease enzyme):** According to the binding residue, the recombinases are grouped to Tyr- and Ser-recombinase. The enzyme activity involved in site-specific recombination between two transposons present as direct repeats in a co-integrate structure.

co-integrate structure

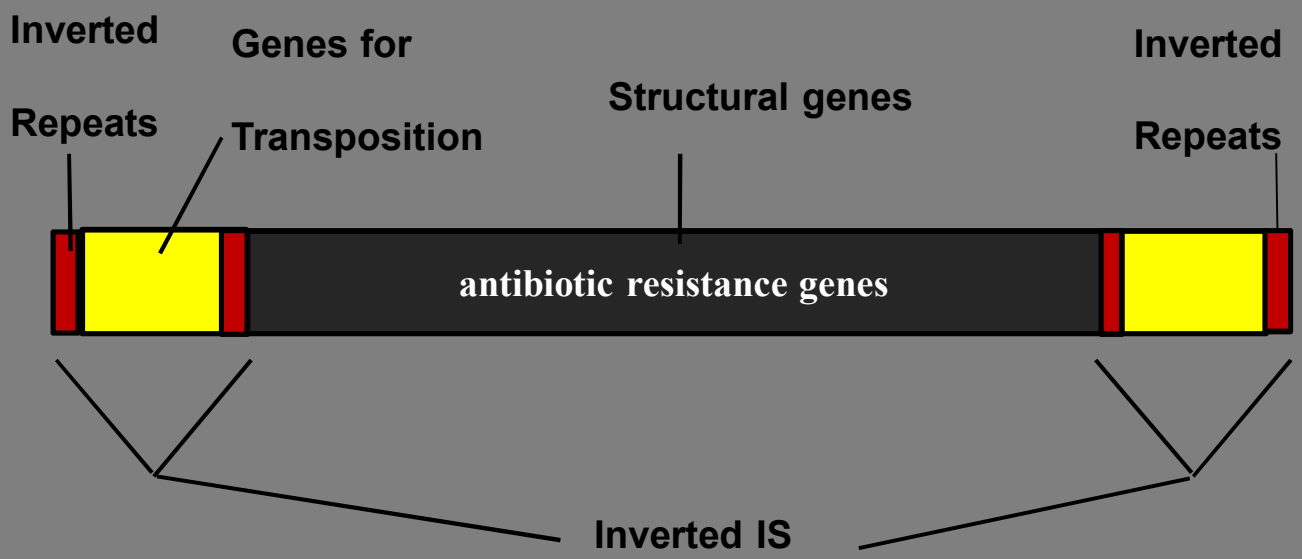
A number of bacterial transposons encode two distinct recombinases that participate in their transposition to other DNA molecules. In the initial step mediated by the element's transposase, a cointegrate is formed between the transposon-containing donor DNA and the target molecule. In this transpositional intermediate, the donor and target DNAs are joined together by copies of the duplicated transposon, one copy occurring at each donor–target junction. The second step in the pathway, cointegrate resolution, is a site-specific recombination performed by the transposon's resolvase protein, acting at a site, called *res*, located within the transposon.

Composite transposons

- Composite transposons: Composite genetic elements are larger than IS elements and contain one or more protein-coding genes in addition to those required for transposition e.g. Tn5, Tn9, Tn10.

- Two IS elements + antibiotic resistance gene(s).

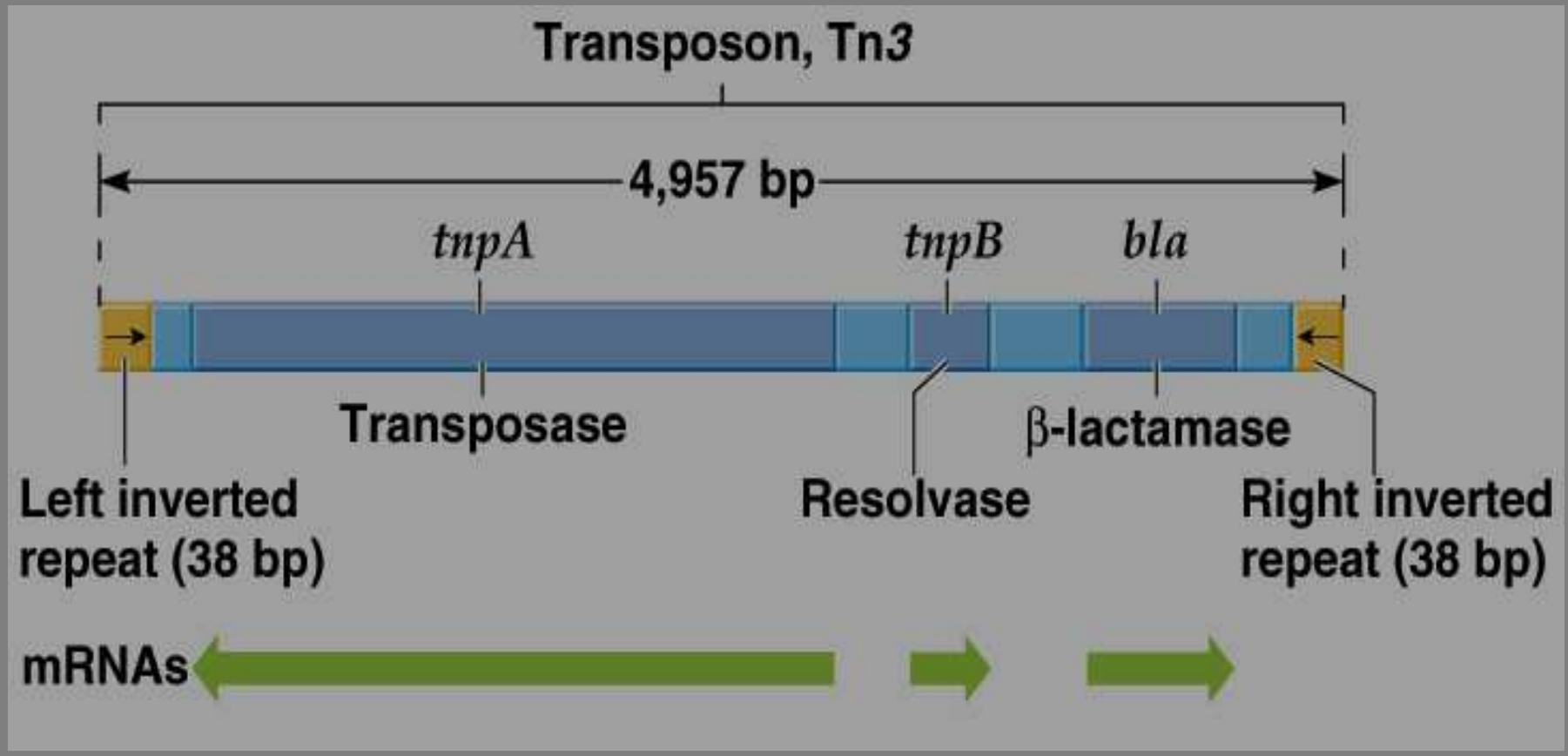
Structure of Composite transposons



Non composite transposons

- Non-composite mobile genetic elements are those which lack IS elements on its ends e.g. Tn3 and Tn7.
- Inverted repeats + transposase gene + antibiotic resistance gene (s).

Structure of Non-composite transposons



Mechanism

- All transposons use a common mechanism in which staggered nicks are made in target DNA, the transposon is joined to the protruding ends, and the gaps are filled.

- The order of events and exact nature of the connections between transposon and target DNA determine whether transposition is replicative or nonreplicative.

Replicative Transposon

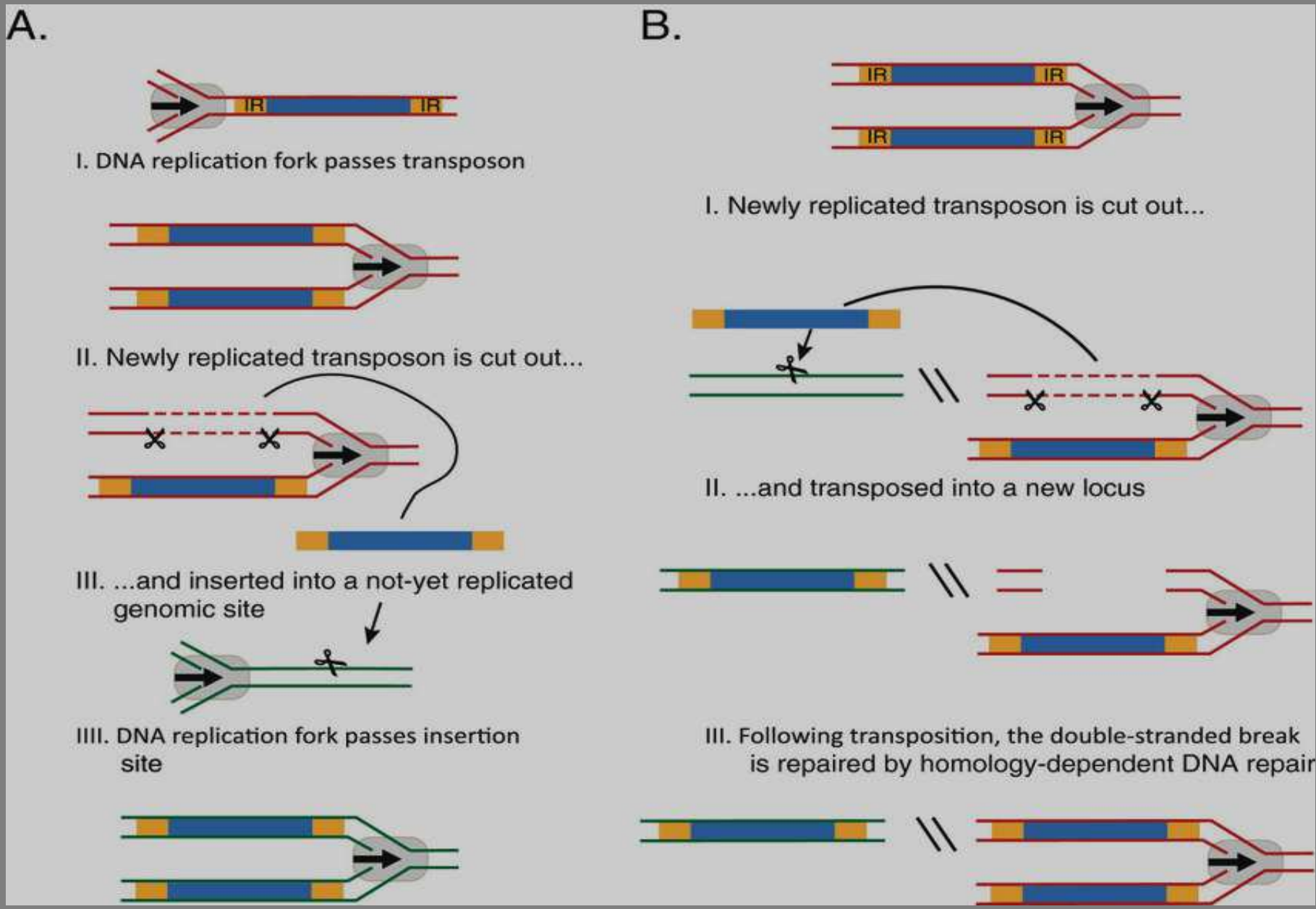
- Replicative transposon is first replicated and then one of the copy will move to the another location in the genome. Thus, the transposon will remain on its original position. "**Copy and Paste**"

- Replicative transposition involves two types of enzymatic activity:

  a) Transposase that acts on the ends of the original transposon and b) Resolvase that acts on the duplicated copies.

- A group of transposons related to TnA move only by replicative transposition.

Replicative Transposon



Replicative Transposon

- Replicative transposition occurs through a cointegrate formation.
- A co-integrate structure is produced by fusion of two replicons, one originally possessing a transposon, the other lacking it; the co-integrate has copies of the transposon present at both junctions of the replicons, oriented as direct repeats.
- Resolution occurs by a homologous recombination reaction between the two copies of the transposon in a co-integrate.
- The reaction generates the donor and target replicons, each with a copy of the transposon.
- Resolvase is the enzyme activity involved in site-specific recombination between two transposons present as direct repeats in a co-integrate structure.

Replicative Transposon

- The reactions involved in generating a cointegrate have been defined in detail for phage Mu.

- The process starts with the formation of the strand transfer complex (sometimes also called a crossover complex).

- The donor and target strands are ligated so that each end of the transposon sequence is joined to one of the protruding single strands generated at the target site.

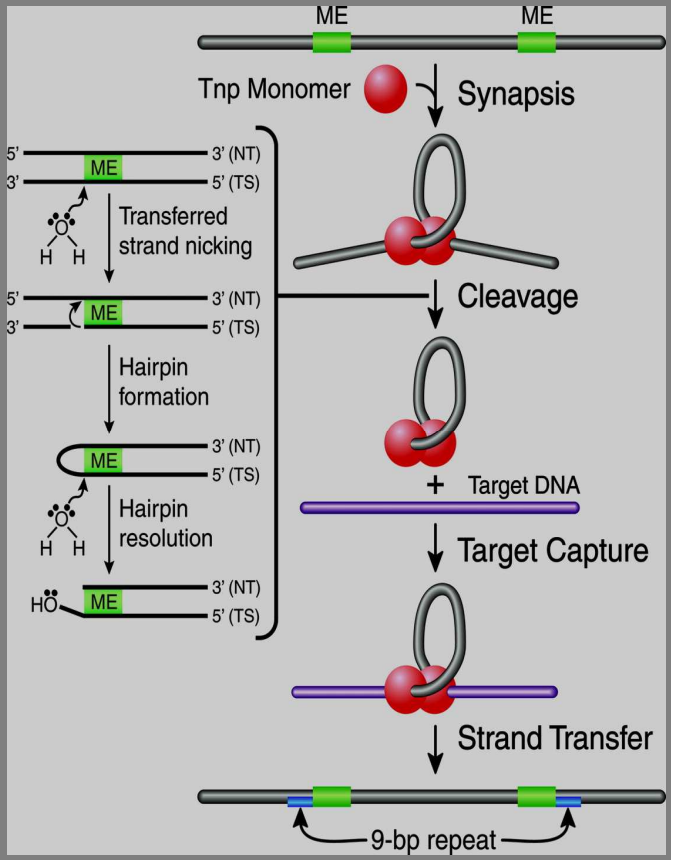
- The crossover structure contains a single-stranded region at each of the staggered ends. These regions are pseudoreplication forks that provide a template for DNA synthesis. (Use of the ends as primers for replication implies that the strand breakage must occur with a polarity that generates a 3′–OH terminus at this point.)

Non-replicative Transposon

- **Non-Replicative transposon leaves its original place and move to the another location in the genome. "Cut and Paste"**
- **This type of mechanism requires only a transposase.**
- **The insertion elements and composite transposons like Tn5 and Tn10 use this mechanism.**
- **Non-replicative transposons leave a break in the donor molecule which is lethal to the cell unless it is repaired.**

Non-replicative Transposon

Tn*5*. Transposition initiated by Tnp binding to the transposon-specific ESs and the formation of a highly ordered nucleoprotein complex (synaptic complex, SC) through a process called synapsis. The SC contains two protomers of Tnp, which exist as a dimer, and two ESs. Catalytic cleavage occurs when an activated H_2O coordinated by Mg^{2+} nicks the transferred DNA strand (TS) on both sides of the transposon, through a nucleophilic attack, forming a 3′-hydroxyl group. The free 3′-hydroyxl group acts as a nucleophile and cleaves the non-transferred DNA strand (NT), forming a hairpin. A second activated water molecule resolves the hairpin, resulting in a double-stranded DNA cleavage product. The post-cleavage synaptic complex is now free to bind to target DNA through target capture. The 3′-hydroxyl group of the transposon end attacks the phosphodiester backbone of target DNA during strand transfer. A 9-bp duplication in the target results, due to the staggered strand transfer reactions followed by DNA repair by host enzymes.

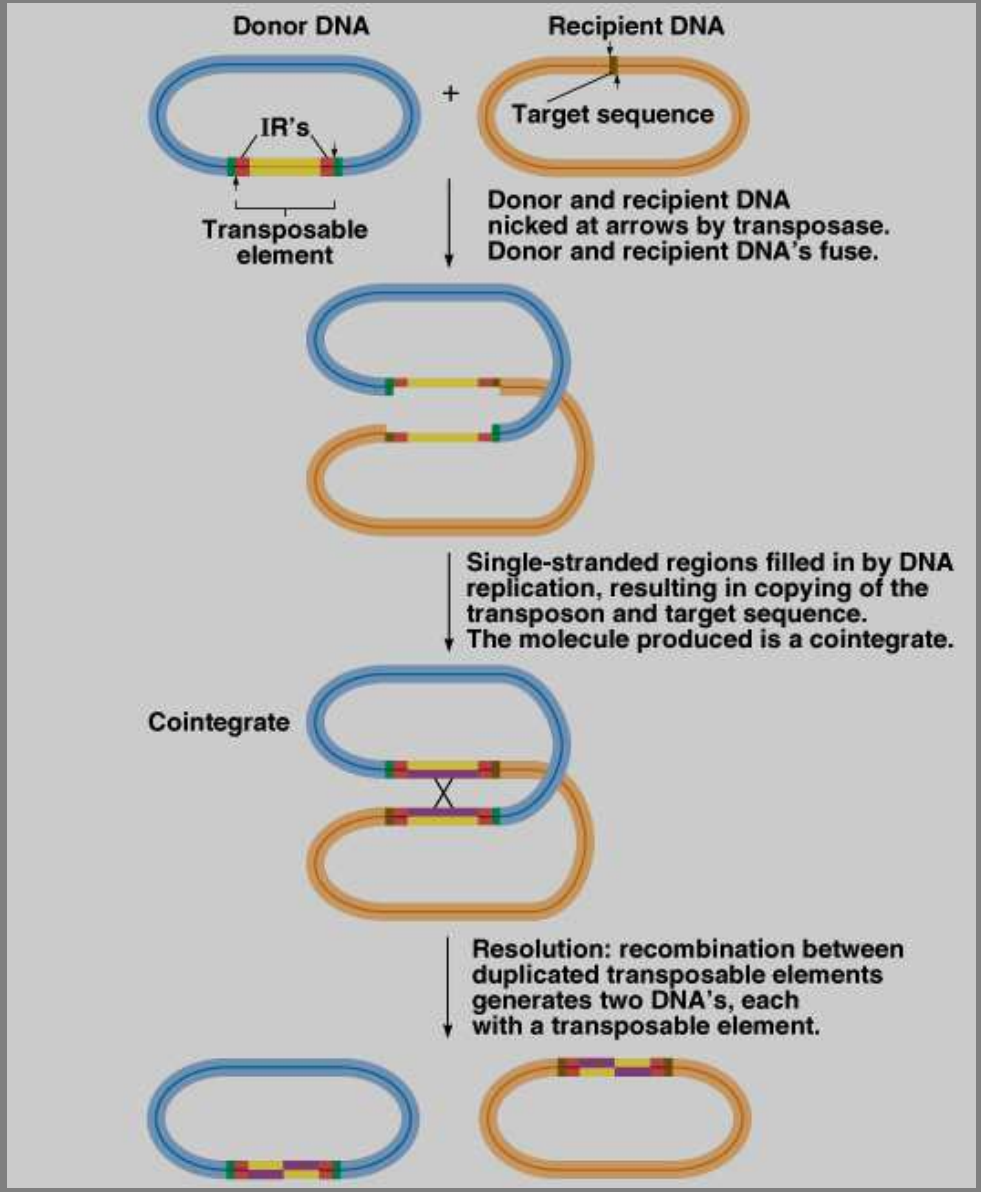


ME- Mosaic Ends

Models of transposition:

- Similar to that of IS elements; duplication at target sites occurs.
- Transposition may be replicative (duplication = copy and paste), but it can also be non-replicative (transposon lost from original site = cut and paste).
- Movement of a transposon from one genome (e.g., plasmid) to another (e.g., chromosome)can cointegrate transposon to both genomes (duplication) in the case of replicative.
- Result in same types of mutations as IS elements: insertions, deletions, changes in gene expression, or duplication.

Fig. 7.22, Recombination, crossing-over, and duplication of a transposable element.



Transposable elements in eukaryotes:

Barbara McClintock (1902-1992)
Cold Spring Harbor Laboratory, NY

Nobel Prize in Physiology and Medicine 1983

"for her discovery of mobile genetic elements"

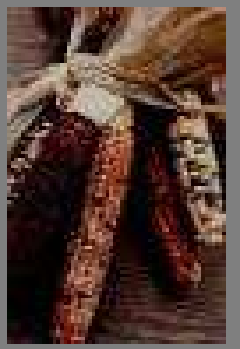
- **Studied transposable elements in corn (*Zea mays*) 1940s-1950s**
  (formerly identified as mutator genes by Marcus Rhoades 1930s)

- **Also known for work demonstrating crossing over as part of the chromosomal basis of inheritance.**

- **Biographical sketch, pp. 155-156**

General properties of plant transposons:

- **Possess ITR sequences and generate short repeats at target sites.**
- **May activate or repress target genes, cause chromosome mutations, and disrupt genes.**
- **Two types:**
  - **Autonomous elements transpose themselves; possess transposition gene.**
  - **Nonautonomous elements do not transpose themselves; lack transposition gene and rely on presence of another Tn**
- **McClintock demonstrated purple spots in otherwise white corn (*Zea mays*) kernels are results of both these types of transposable elements.**

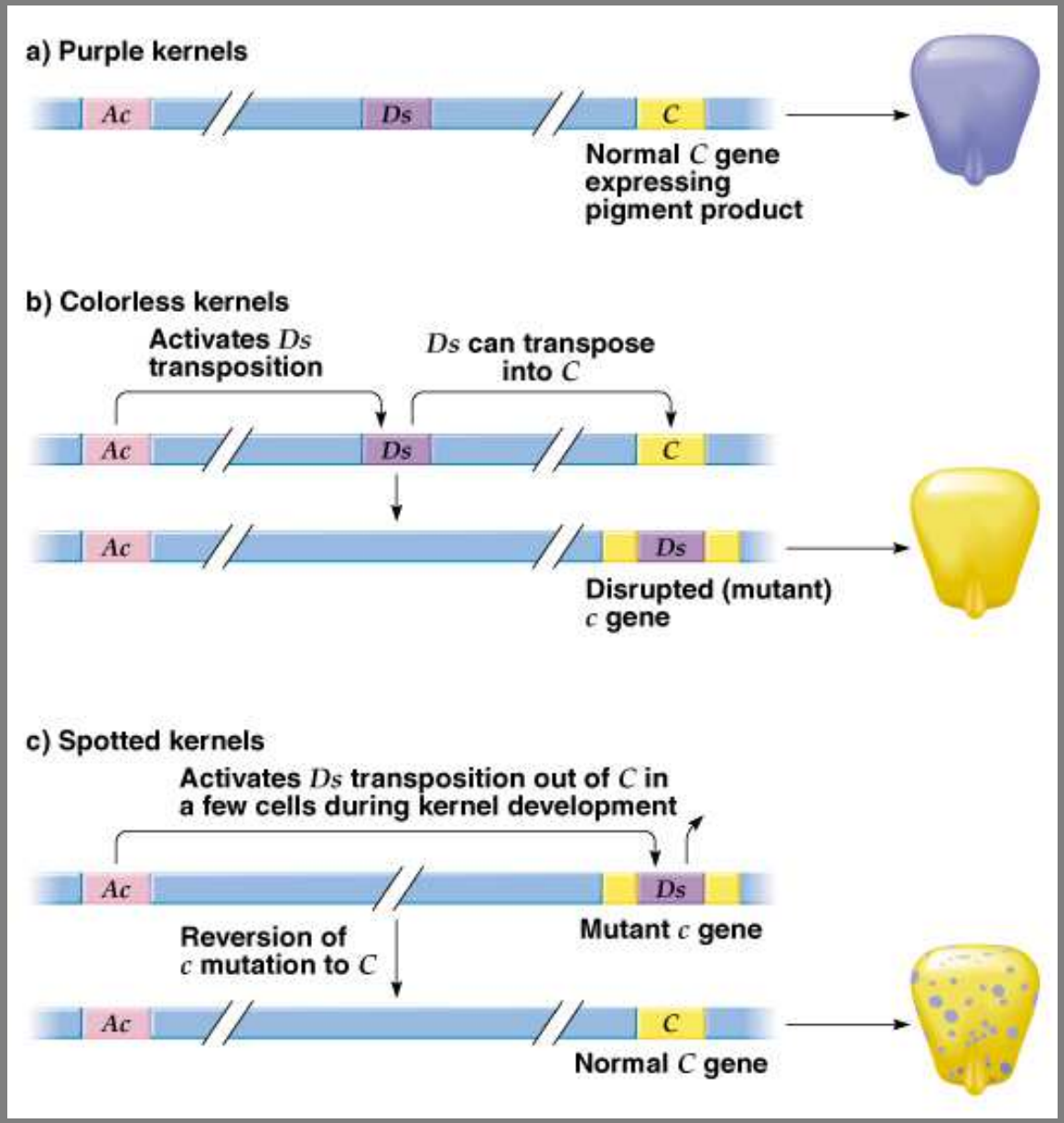


McClintock's discovery of transposons in corn:

- **c/c = white kernels and C/- = purple kernels**
- **Kernal color alleles/traits are "unstable".**
- **If reversion of c to C occurs in a cell, cell will produce purple pigment and a spot.**
- **Earlier in development reversion occurs, the larger the spot.**
- **McClintock concluded "c" allele results from a non-autonomous transposon called "Ds" inserted into the "C" gene (Ds = dissassociation).**
- **Autonomous transposon "Ac" controls "Ds" transposon (Ac = activator).**

Barbara McClintock's Transposable Elements In Corn

Fig. 7.24, Transposon effects on corn kernel color.



McClintock's discovery of transposons in corn (cont.):

- ***Ac* element is autonomous/*Ds* element is nonautonomous.**
- ***Ac* is 4,563 bp with 11 bp ITRs and 1 transcription unit encoding an 807 amino acid transposase.**
- ***Ac* activates *Ds*; *Ds* varies in length and sequence, but possesses same ITRs as *Ac*.**
- **Many *Ds* elements are deleted or rearranged version of *Ac*; *Ds* element derived from *Ac*.**
- ***Ac/Ds* are developmentally regulated; *Ac/Ds* transpose only during chromosome replication and do not leave copies behind.**

Fig. 20.12 2nd edition, Structure of *Ac* autonomous and *Ds* non-autonomous transposable elements in corn.

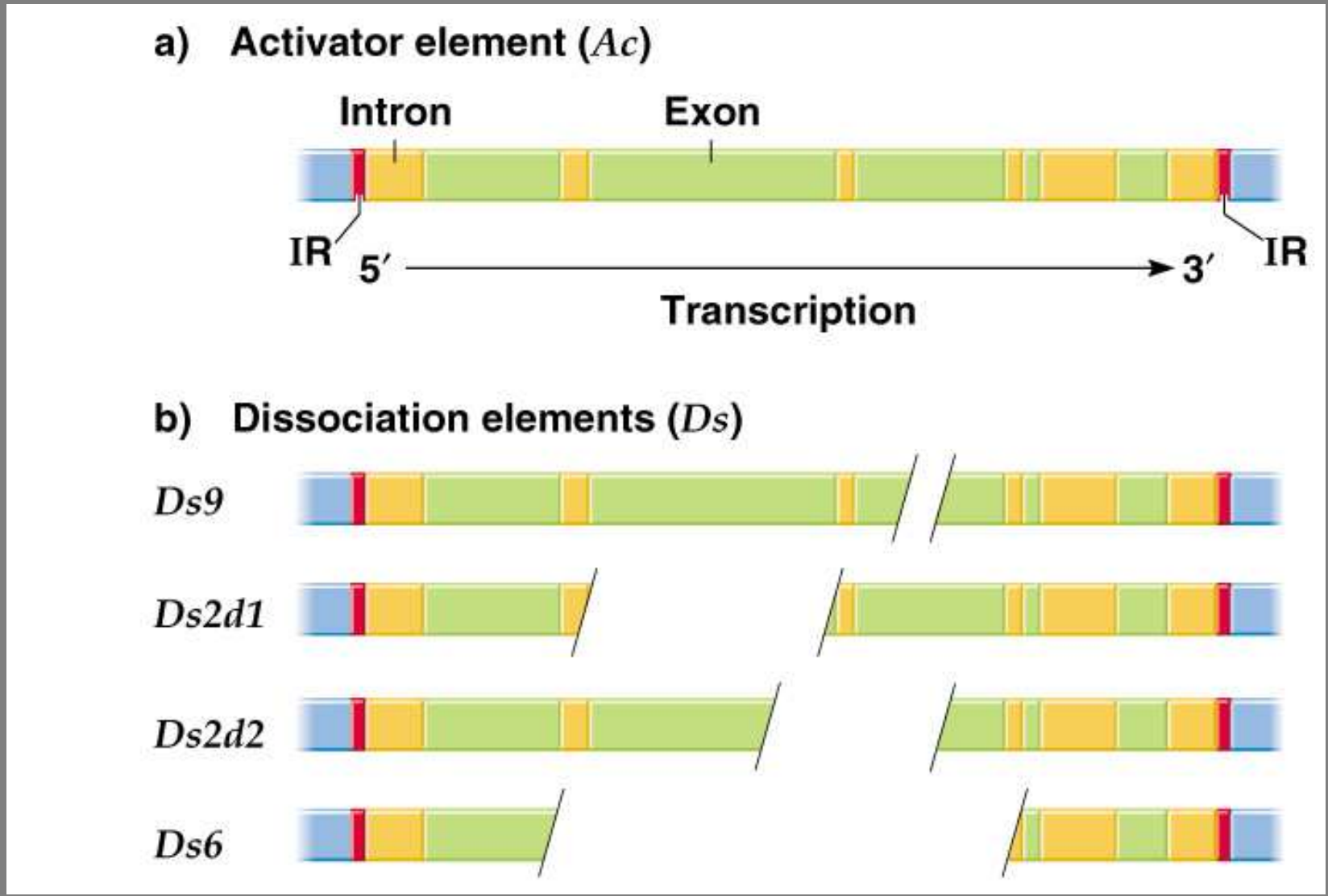
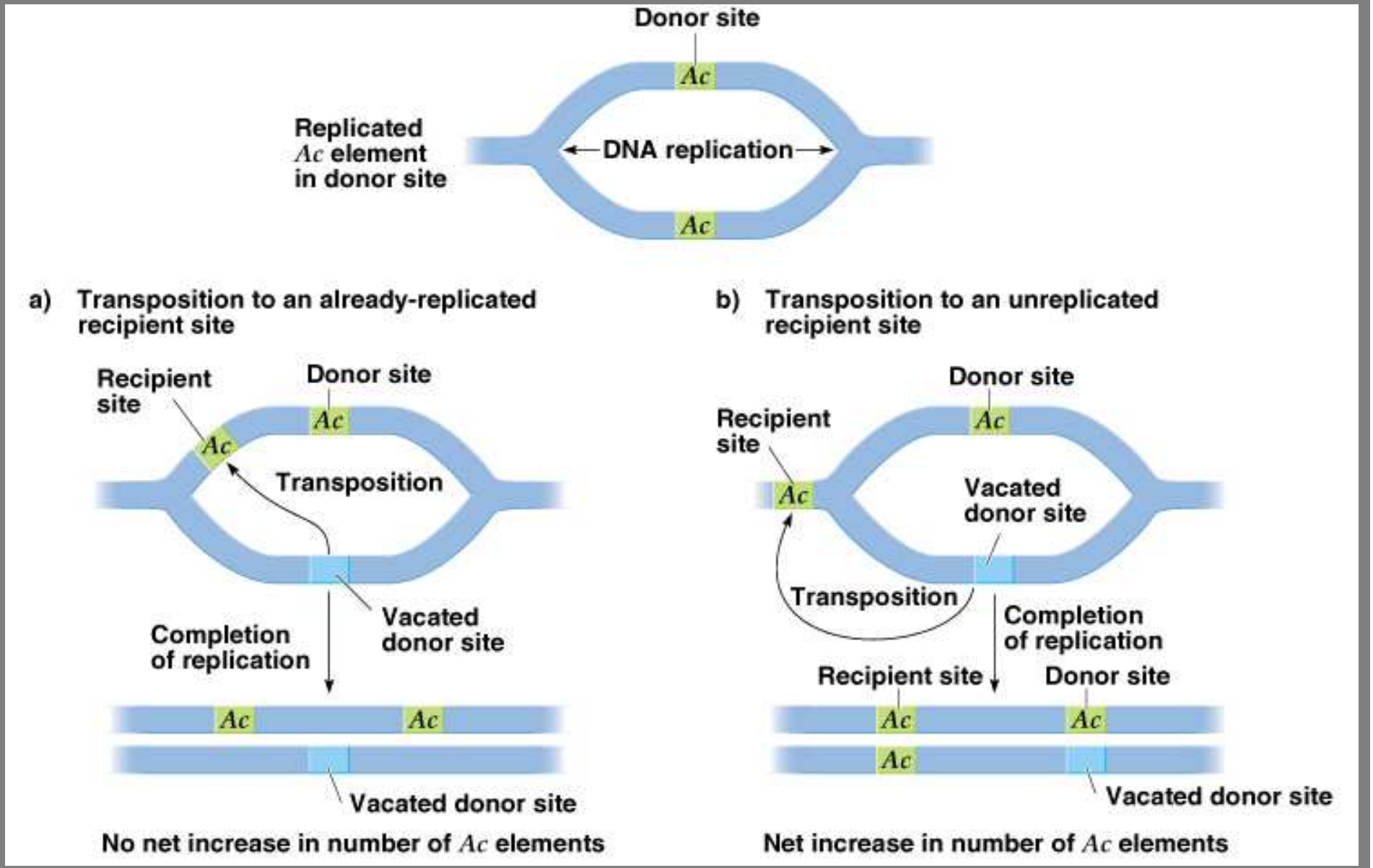


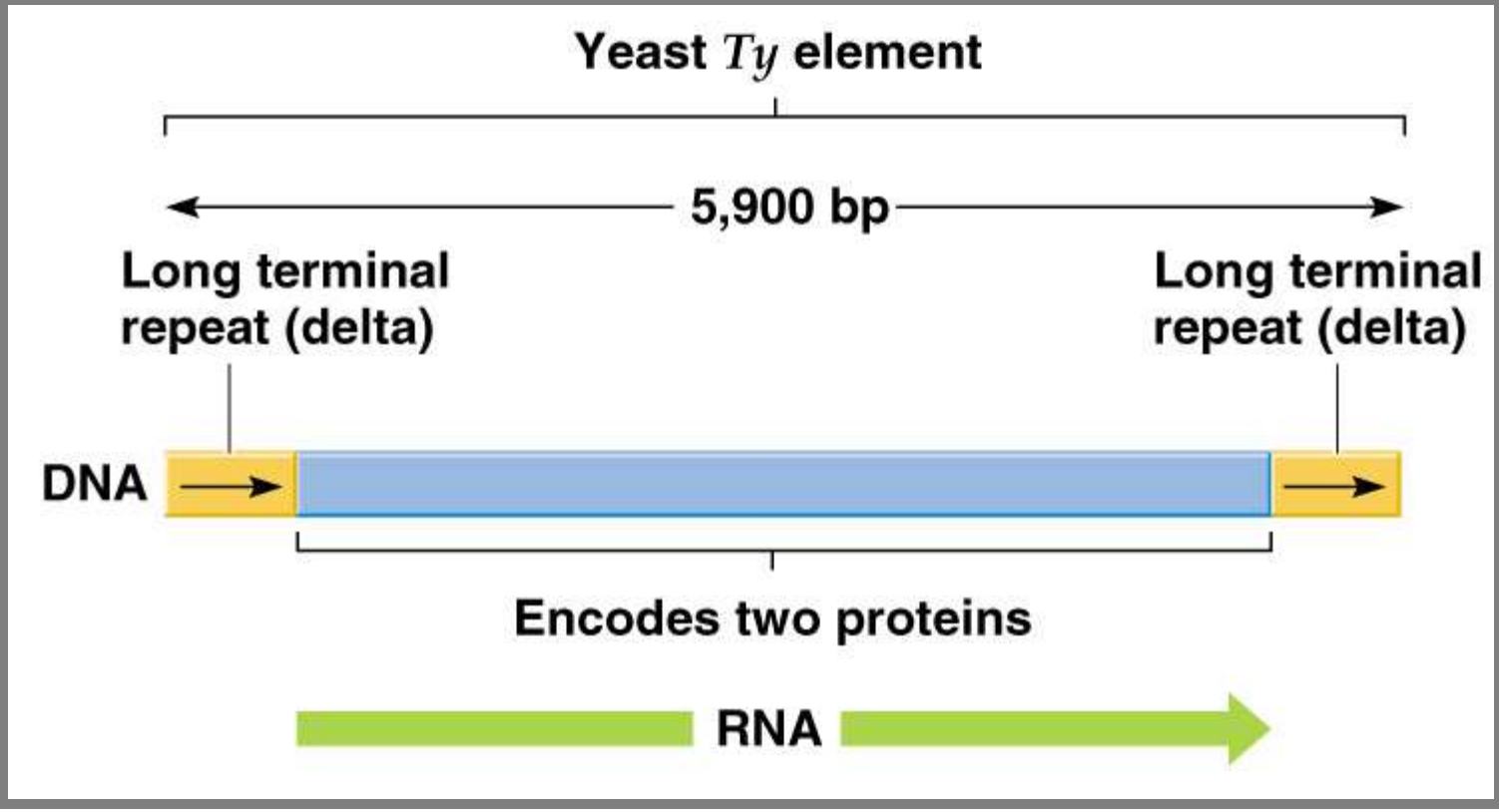
Fig. 7.25, *Ac* transposition mechanism during chromosome replication.



Ty elements in yeast:

- **Similar to bacterial transposons; terminal repeated sequences, integrate at non-homologous sites, with target site duplication.**
- **Ty elements share properties with retroviruses, retrotransposons:**
  - **Synthesize RNA copy and make DNA using reverse transcriptase.**
  - **cDNA integrates at a new chromosomal site.**

Fig. 7.26

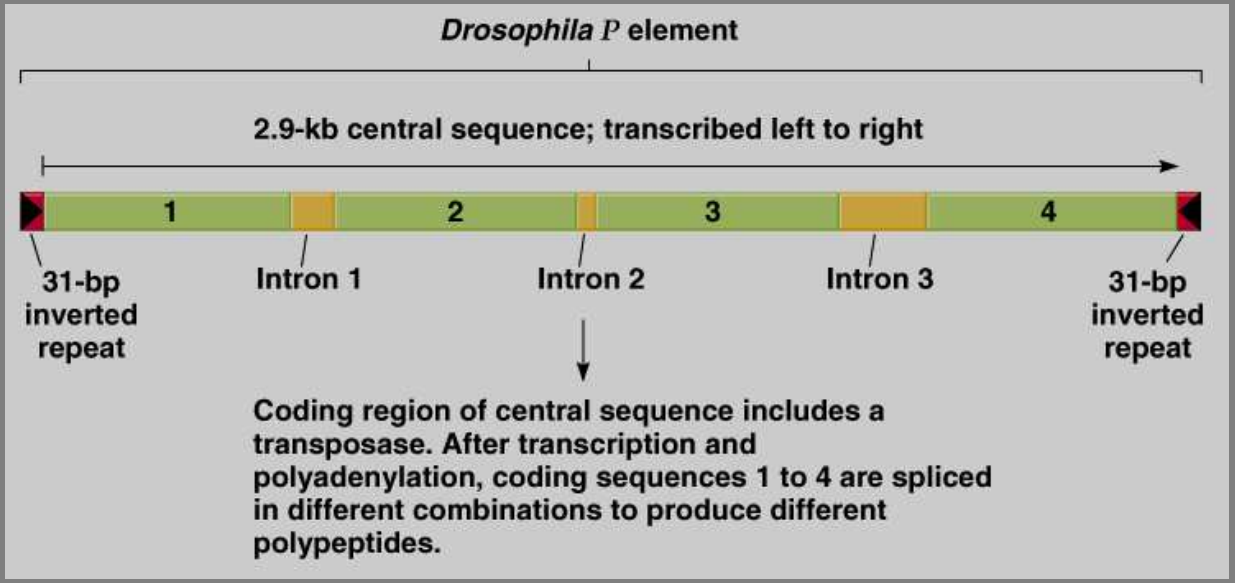


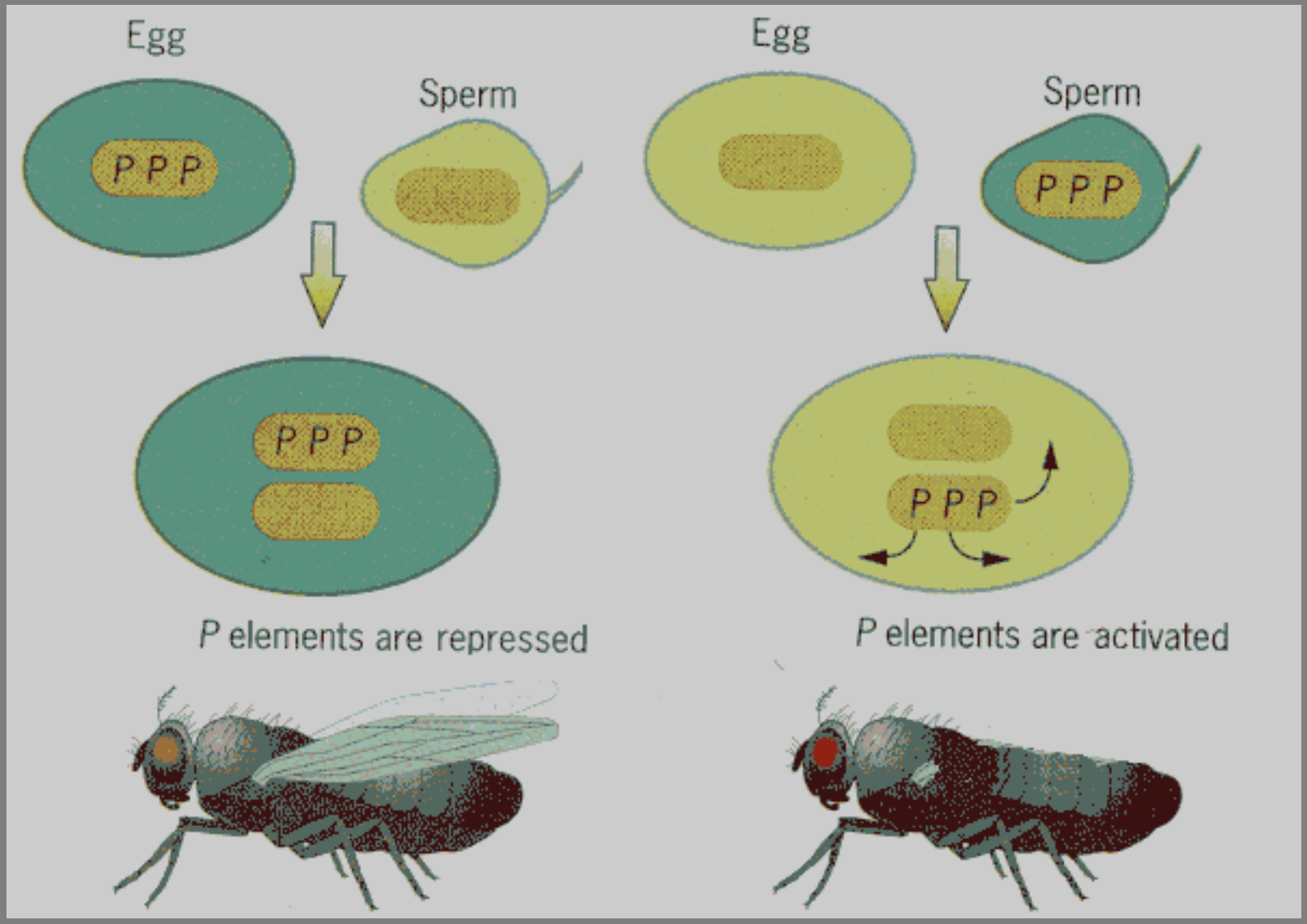
Drosophila transposons:

- **~15% of *Drosophila* genome thought to be mobile.**

P elements

- **Hybrid dysgenesis, defects arise from crossing of specific *Drosophila* strains.**
- **Occurs when haploid genome of male (P strain) possesses ~40 P elements/genome.**
- **P elements vary in length from 500-2,900 bp.**
- **P elements code a repressor present in the cytoplasm, which makes them stable in the P strain (but unstable when crossed to the wild type female; female lacks repressor in cytoplasm).**
- **Used experimentally as transformation vectors.**





http://www.mun.ca/biology/scarr/P-element_hybrid_dysgenesis.htm

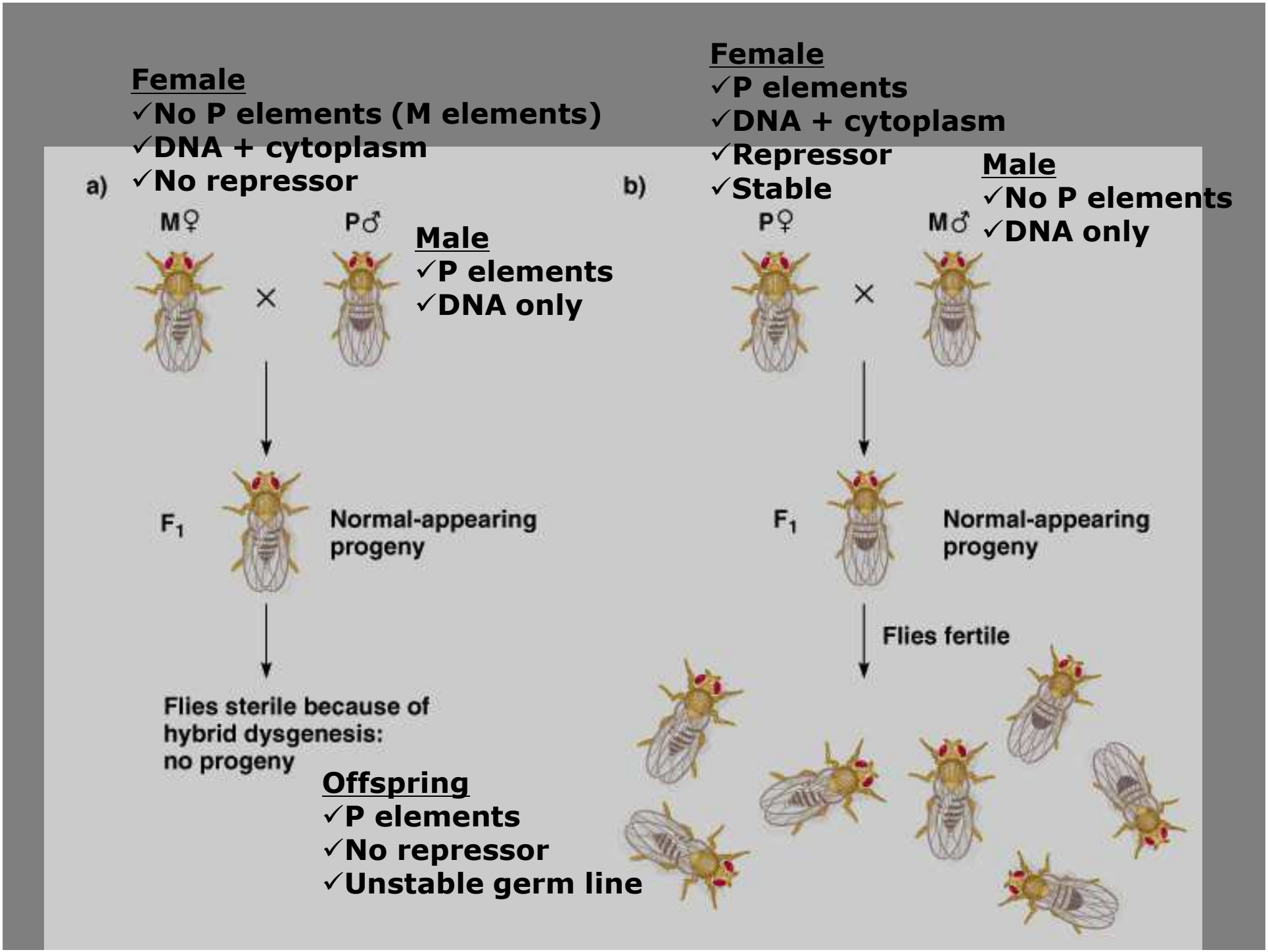
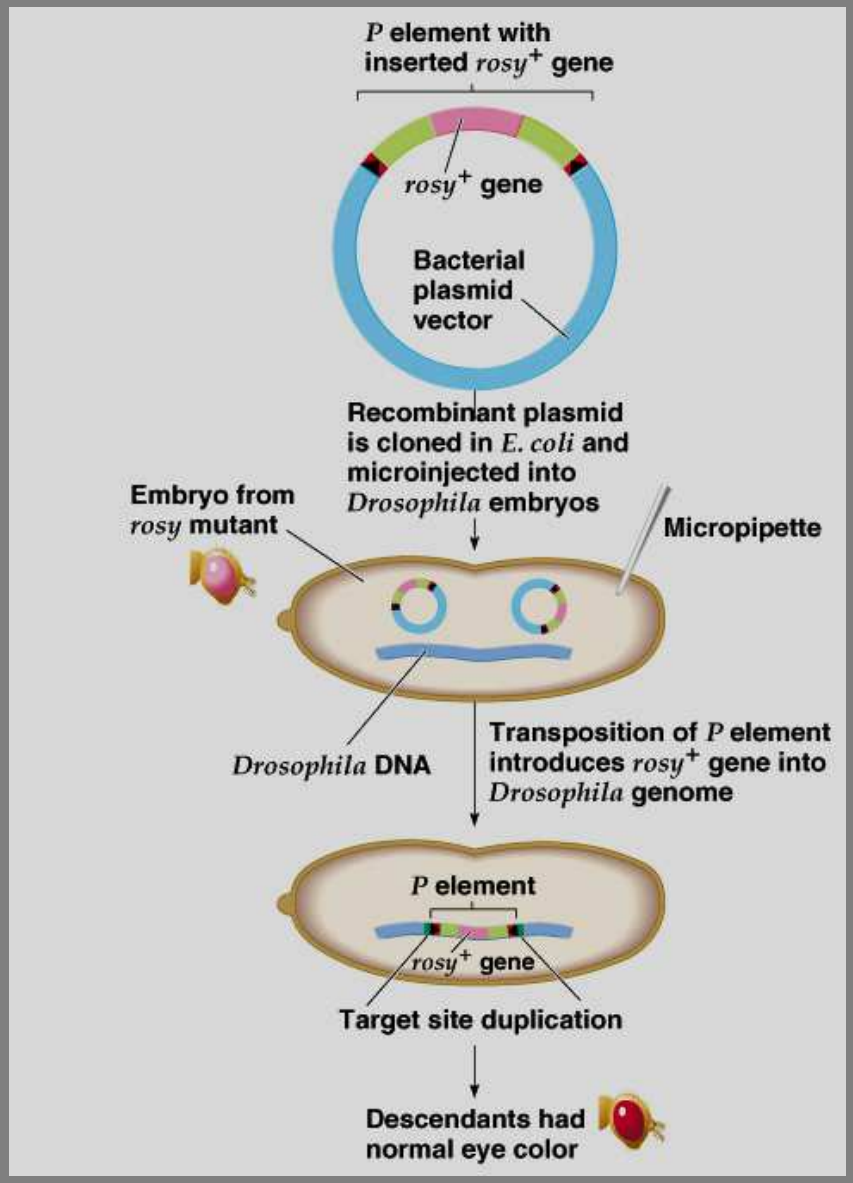
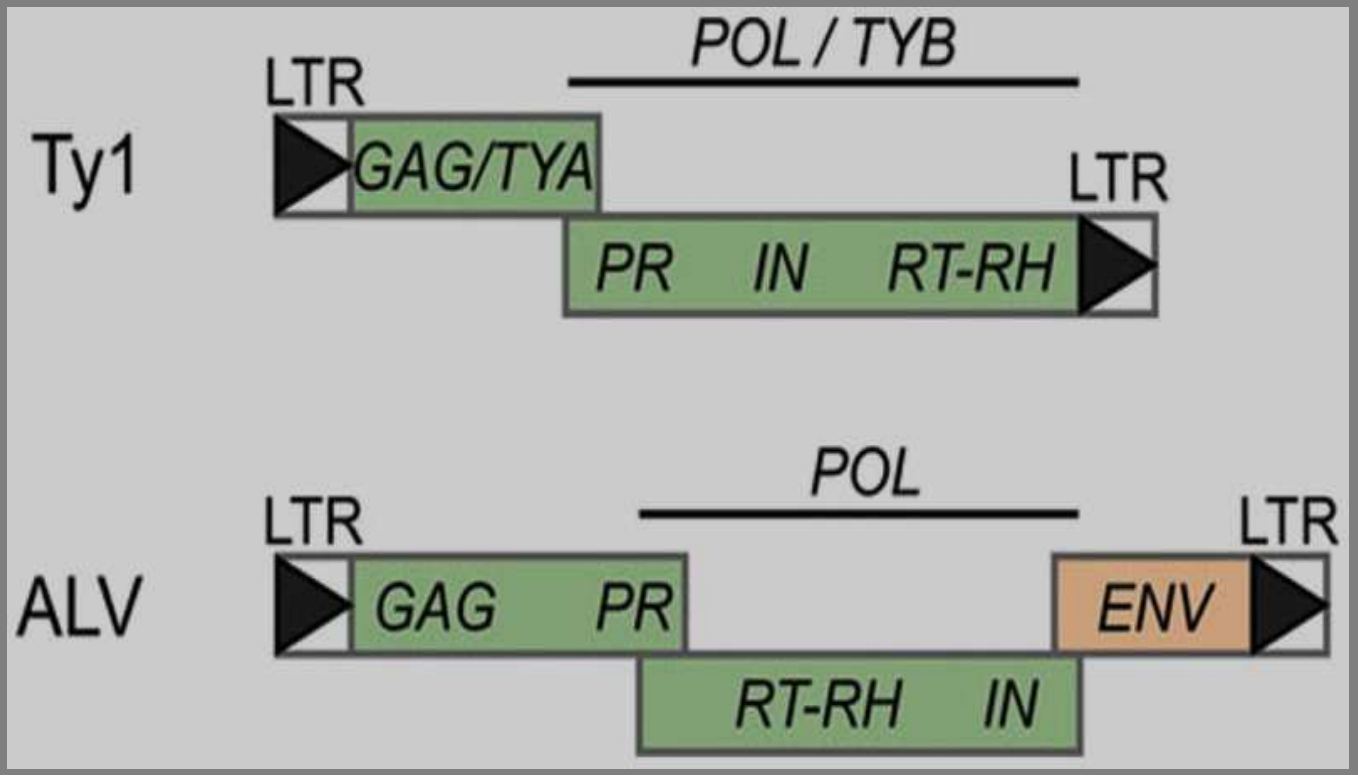

a)
Female
✓No P elements (M elements)
✓DNA + cytoplasm
✓No repressor
M♀
P♂
Male
✓P elements
✓DNA only
F1
Normal-appearing progeny
Flies sterile because of hybrid dysgenesis: no progeny
Offspring
✓P elements
✓No repressor
✓Unstable germ line
b)
Female
✓P elements
✓DNA + cytoplasm
✓Repressor
✓Stable
Male
✓No P elements
✓DNA only
P♀
M♂
F1
Normal-appearing progeny
Flies fertile

Fig. 7.28 Illustration of the use of *P* elements to introduce genes into the *Drosophila* genome



- **The structure of Ty1 is analogous to that of retroviral proviruses.**
- **The most highly characterized Ty1 element is Ty1-H3, which was isolated following its retrotransposition into plasmid DNA.**
- **5918 base pairs (bp) in length with 334 bp direct repeats, or LTRs, at each end.**
- **Ty1 LTRs, have the dinucleotide inverted repeat, 5′-TG…CA-3′ at their termini, and are composed of three distinct domains-U3, R and U5.**
- **These domains are defined by their position in the major sense-strand transcript expressed from Ty1 DNA.**
- **The 38-nucleotide U5 region and 240-nucleotide U3 region are unique to the 5′ and 3′ end of the Ty1 RNA, respectively, while the R region of 56 nucleotides is repeated at both ends of the processed transcript.**
- **Functional Ty1 elements encode two partially overlapping open reading frames: *GAG* (historically known as *TYA1*) and *POL* (*TYB1*).**

- **The last 3 nucleotides of the R region of the 5′ LTR encode the first codon of *GAG*.**
- **The *GAG* ORF encodes a single functional protein with capsid and nucleic acid chaperone functions.**
- **The *POL* ORF is in the +1 frame relative to *GAG* and overlaps the last 38 base pairs of *GAG*. *POL* encodes three proteins with catalytic activity: protease (PR), integrase (IN), and reverse transcriptase/RNase H (RT/RH).**
- **Ty1 does not contain an equivalent of the retroviral *ENV* gene or any remnant of one.**



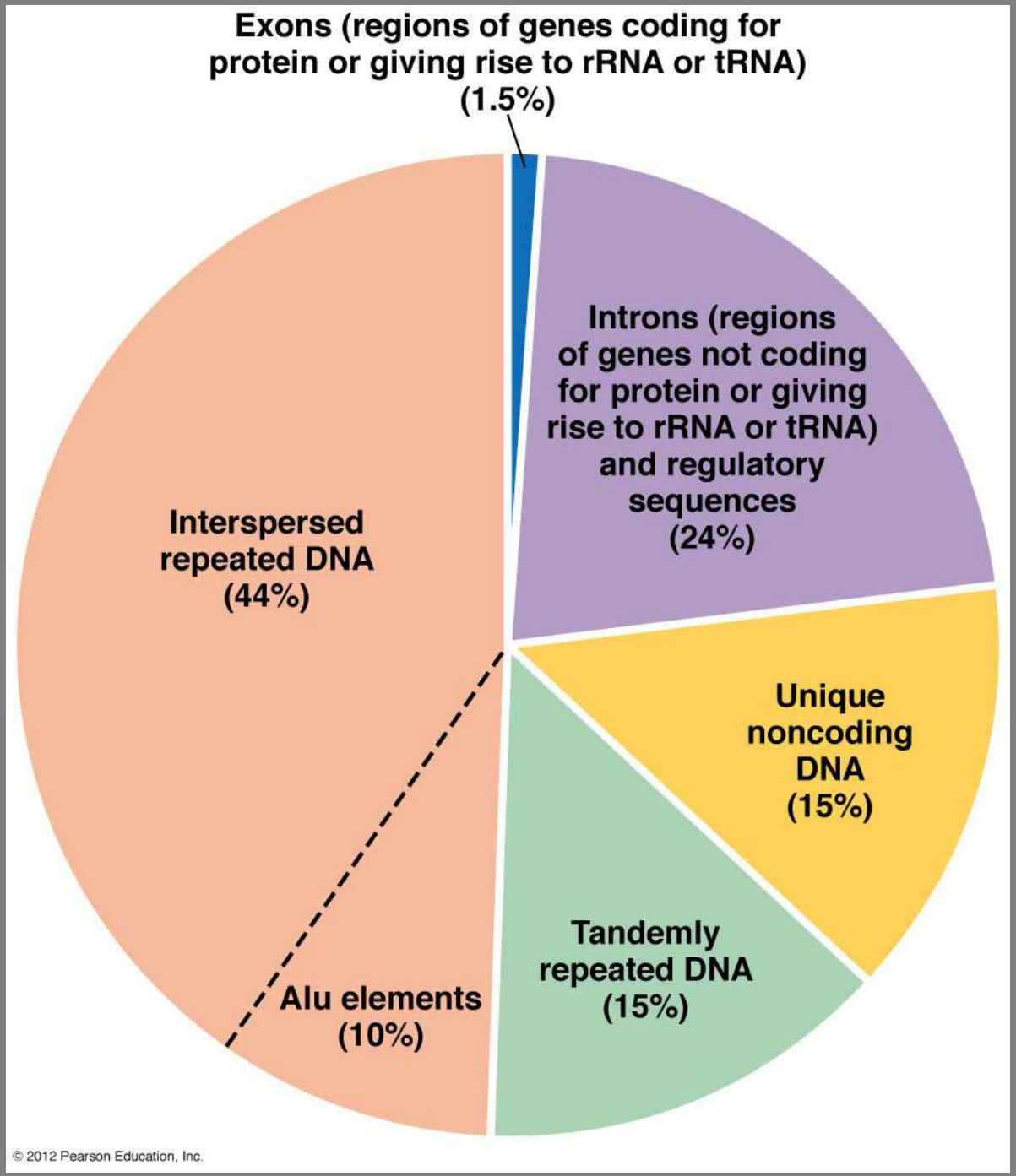
Human retrotransposons:

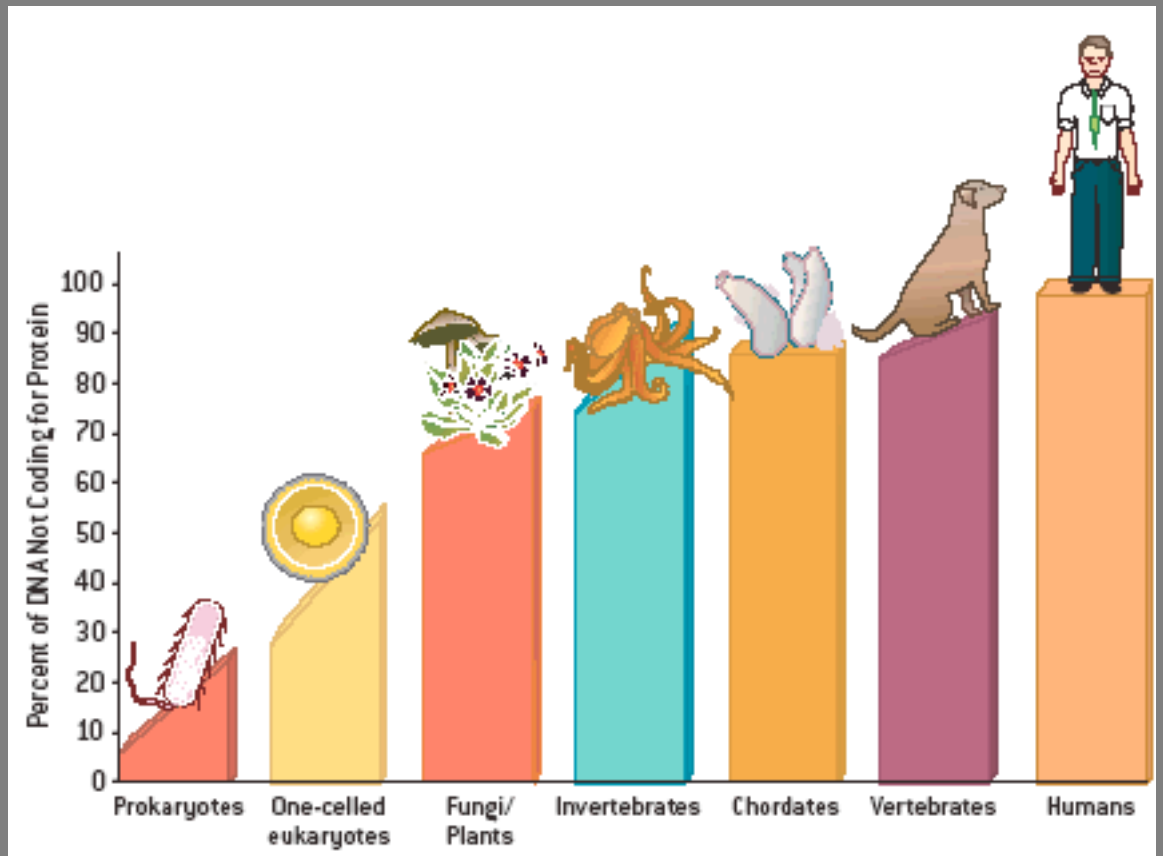
Alu1 SINEs (short-interspersed sequences)

- **~300 bp long, repeated 300,000-500,000X.**
- **Flanked by 7-20 bp direct repeats.**
- **Some are transcribed, thought to move by RNA intermediate.**
- ***AluI* SINEs detected in neurofibromatosis (OMIM1622200) intron; results in loss of an exon and non-functional protein.**

L-1 LINEs (long-interspersed sequences)

- **6.5 kb element, repeated 50,000-100,000X (~5% of genome).**
- **Contain ORFs (open reading frame) with homology to reverse transcriptases; lacks LTRs.**
- **Some cases of hemophilia (OMIM-306700) known to result from newly transposed L1 insertions.**


Exons (regions of genes coding for protein or giving rise to rRNA or tRNA) (1.5%)
Introns (regions of genes not coding for protein or giving rise to rRNA or tRNA) and regulatory sequences (24%)
Unique noncoding DNA (15%)
Tandemly repeated DNA (15%)
Alu elements (10%)
Interspersed repeated DNA (44%)
© 2012 Pearson Education, Inc.



NONPROTEIN-CODING SEQUENCES make up only a small fraction of the DNA of prokaryotes. Among eukaryotes, as their complexity increases, generally so, too, does the proportion of their DNA that does not code for protein. The noncoding sequences have been considered junk, but perhaps it actually helps to explain organisms' complexity.

http://sandwalk.blogspot.com/2008/02/theme-genomes-junk-dna.html

Biological Significance of Transposons

- **They provide a means for genomic change and variation, particularly in response to stress (McClintock's "stress" hypothesis)**

  (1983 Nobel lecture, Science 226:792)
- **or just "selfish DNA"?**
- **No known examples of an element playing a normal role in development.**