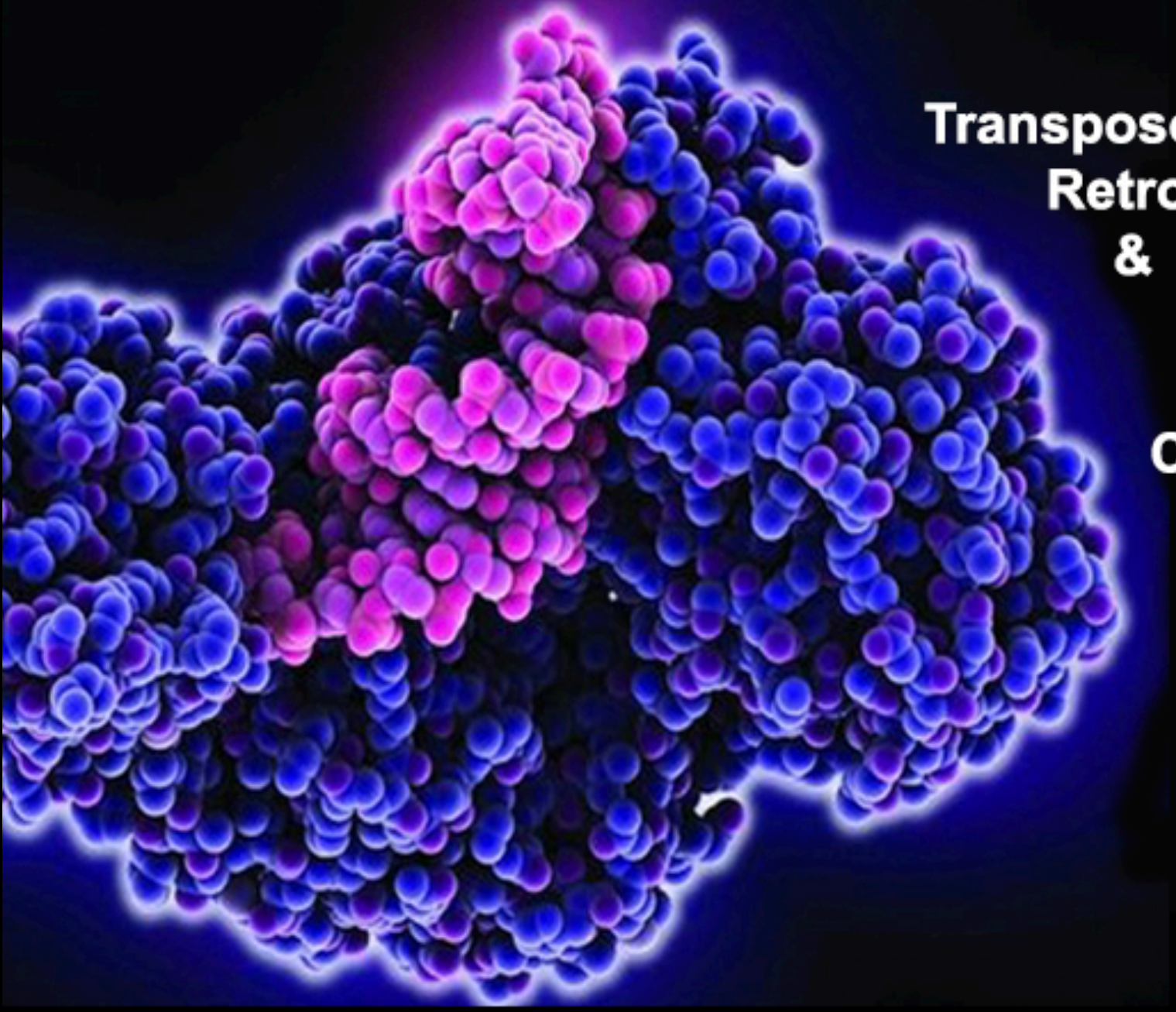
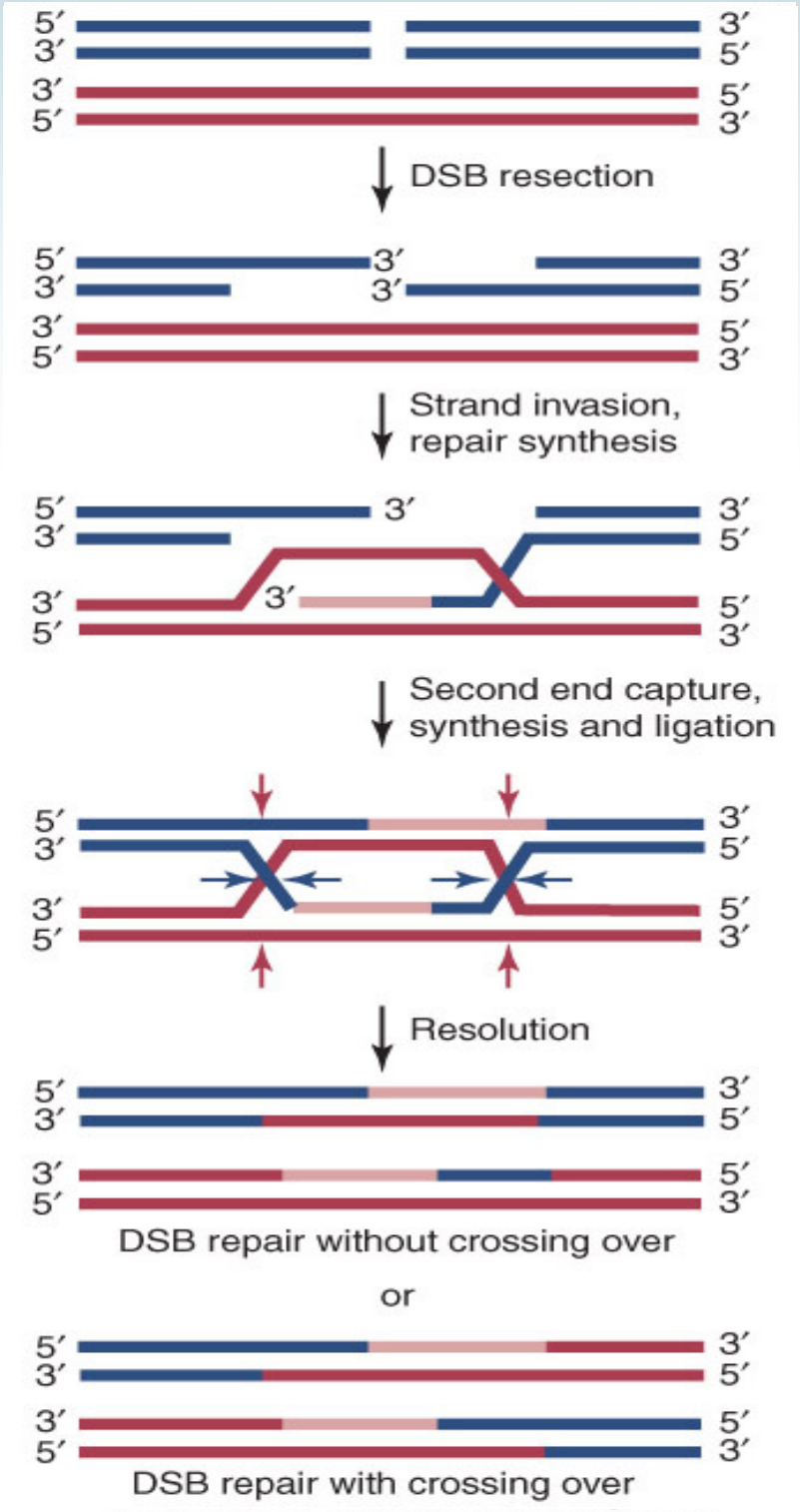
BIOL8620 Eukaryotic Genetics



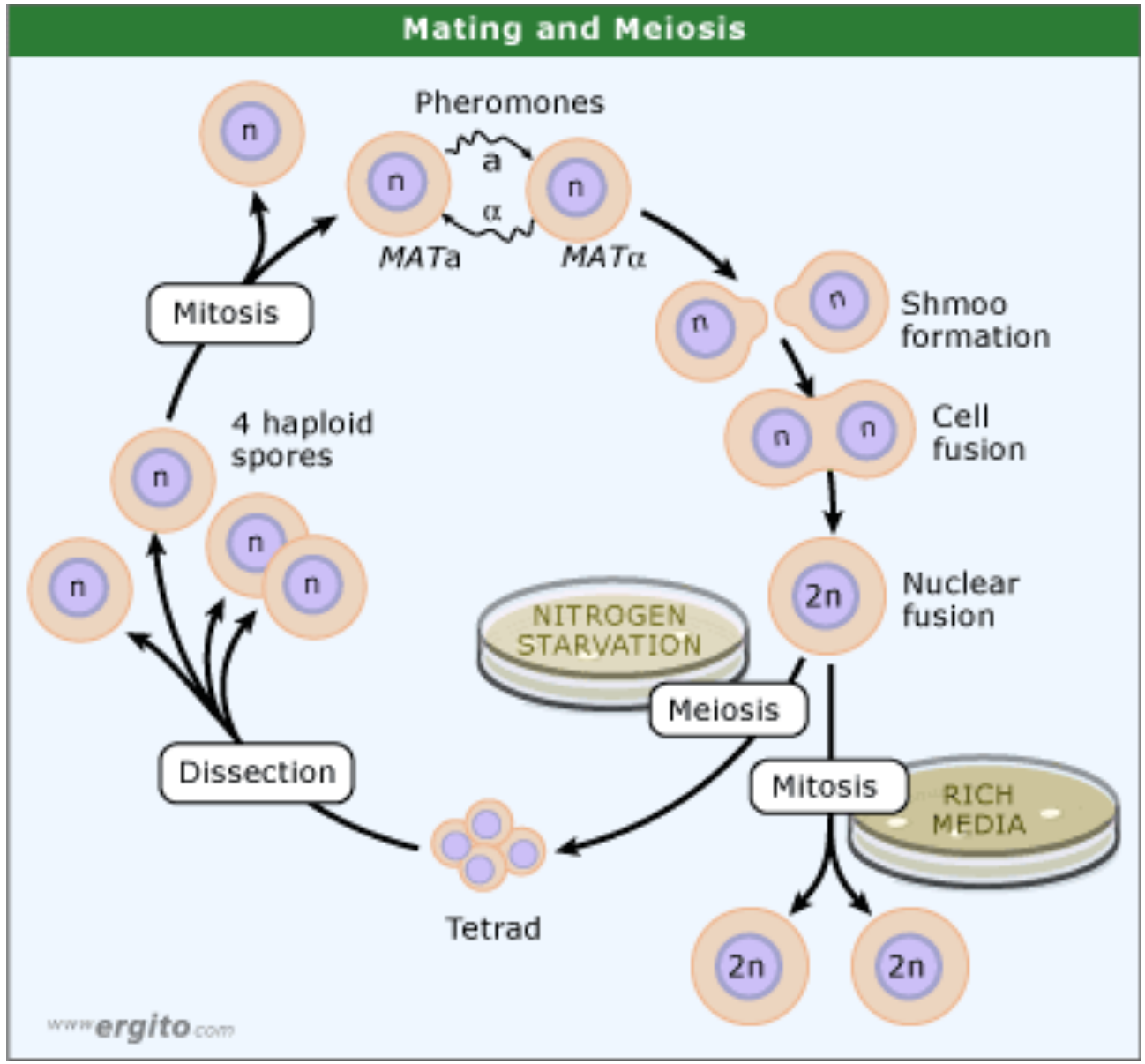
Transposons, Retrotransposons & Retroposons

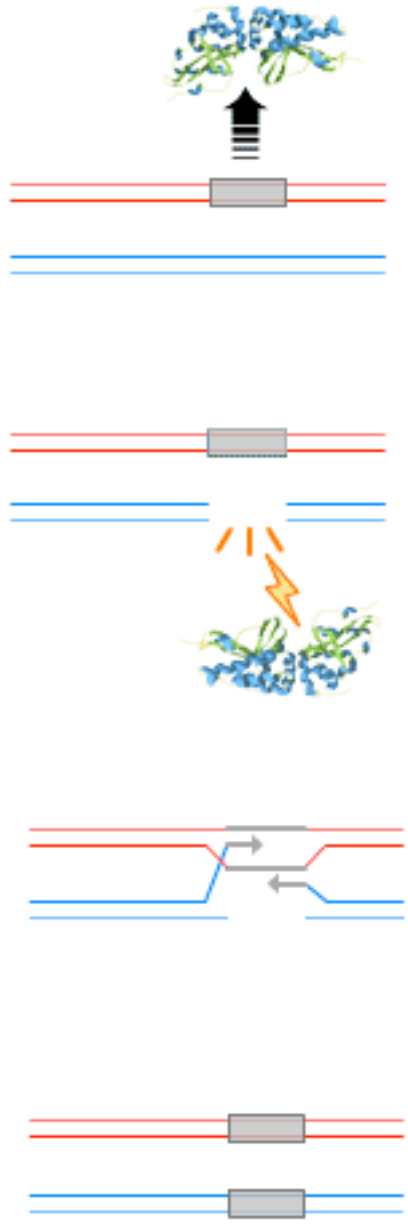
Chapter 15 &
parts of 14



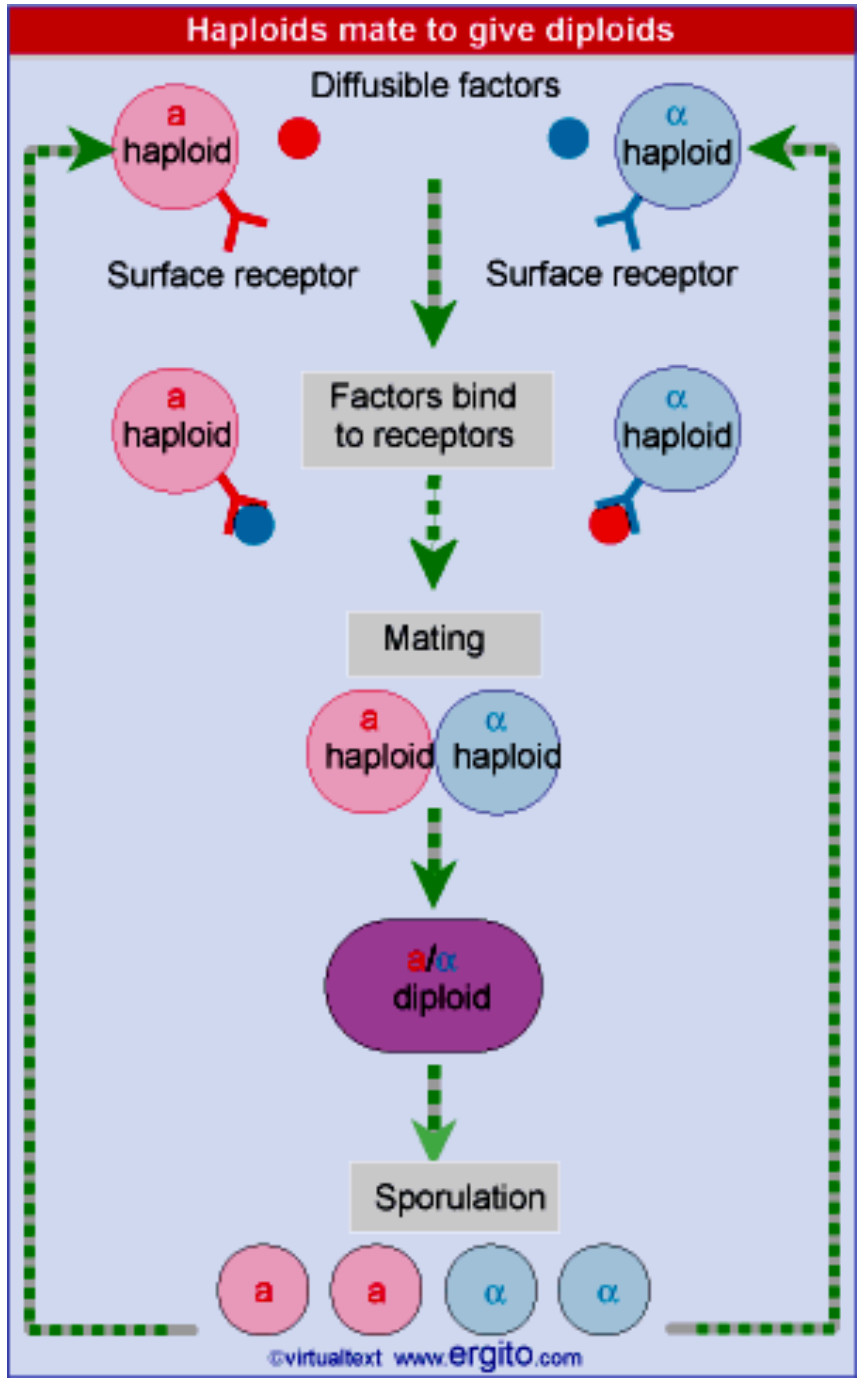
The Double-strand break-repair model and the **Synthesis-Dependent Strand-Annealing Model**

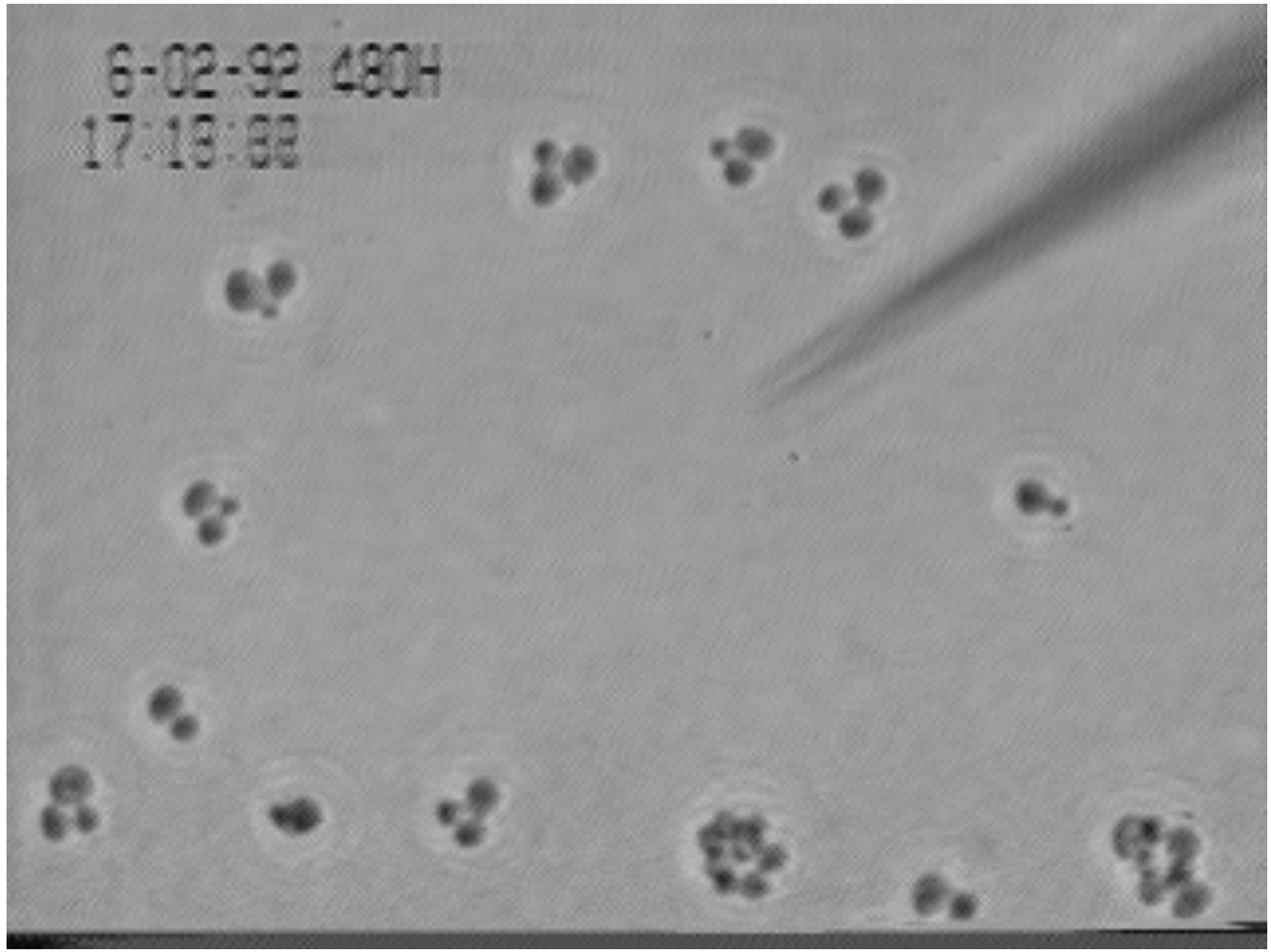
- The synthesis-dependent strand-annealing model (SDSA) is relevant for **DNA repair** and **mitotic recombination**, as can be modified produce **gene conversions** from double-strand breaks without having associated crossovers.

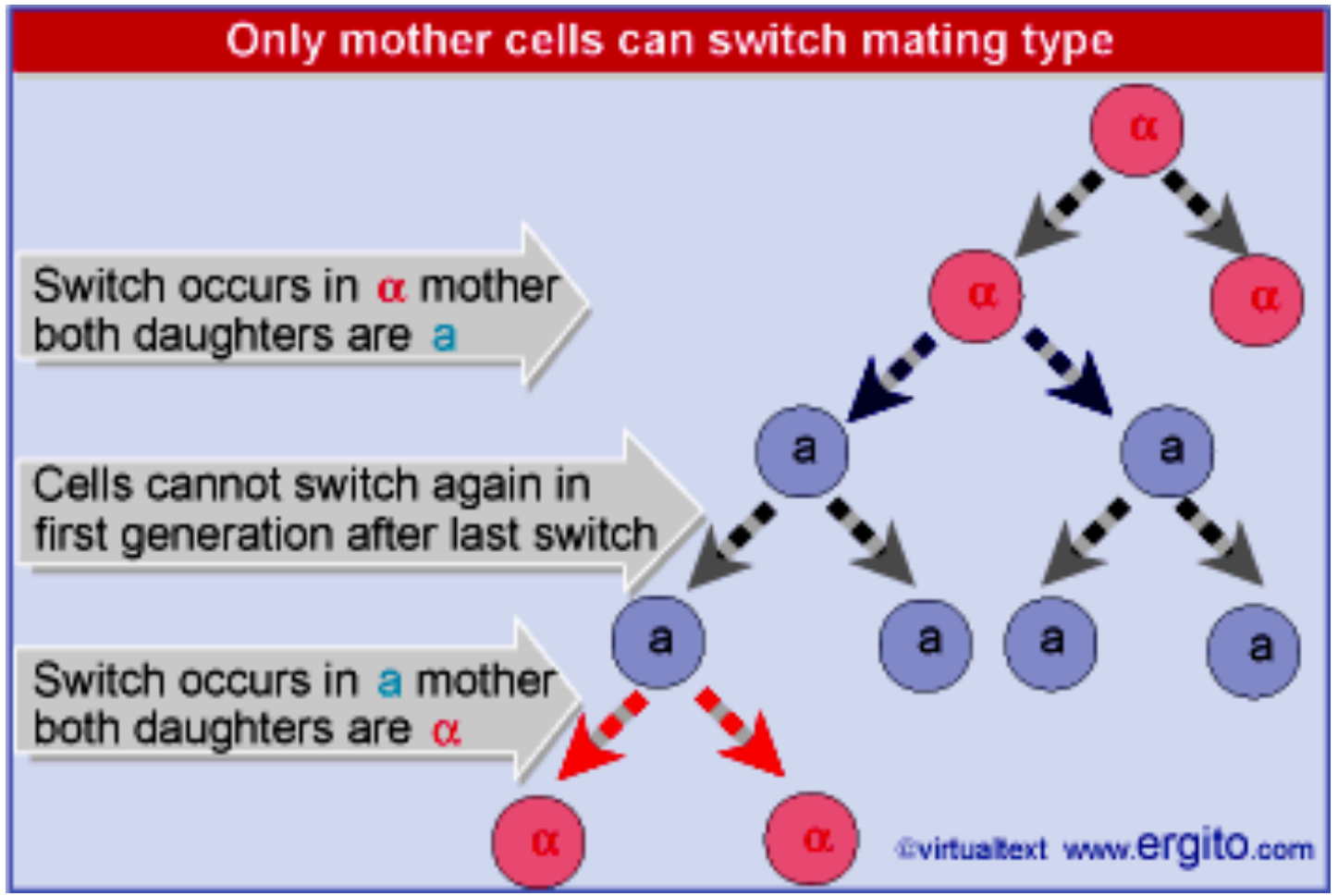

Mating and Meiosis
Pheromones
a
α
n
MATa
MATα
Shmoo formation
Cell fusion
2n
Nuclear fusion
NITROGEN STARVATION
Meiosis
Mitosis
RICH MEDIA
Tetrad
Dissection
4 haploid spores
Mitosis
www.ergito.com



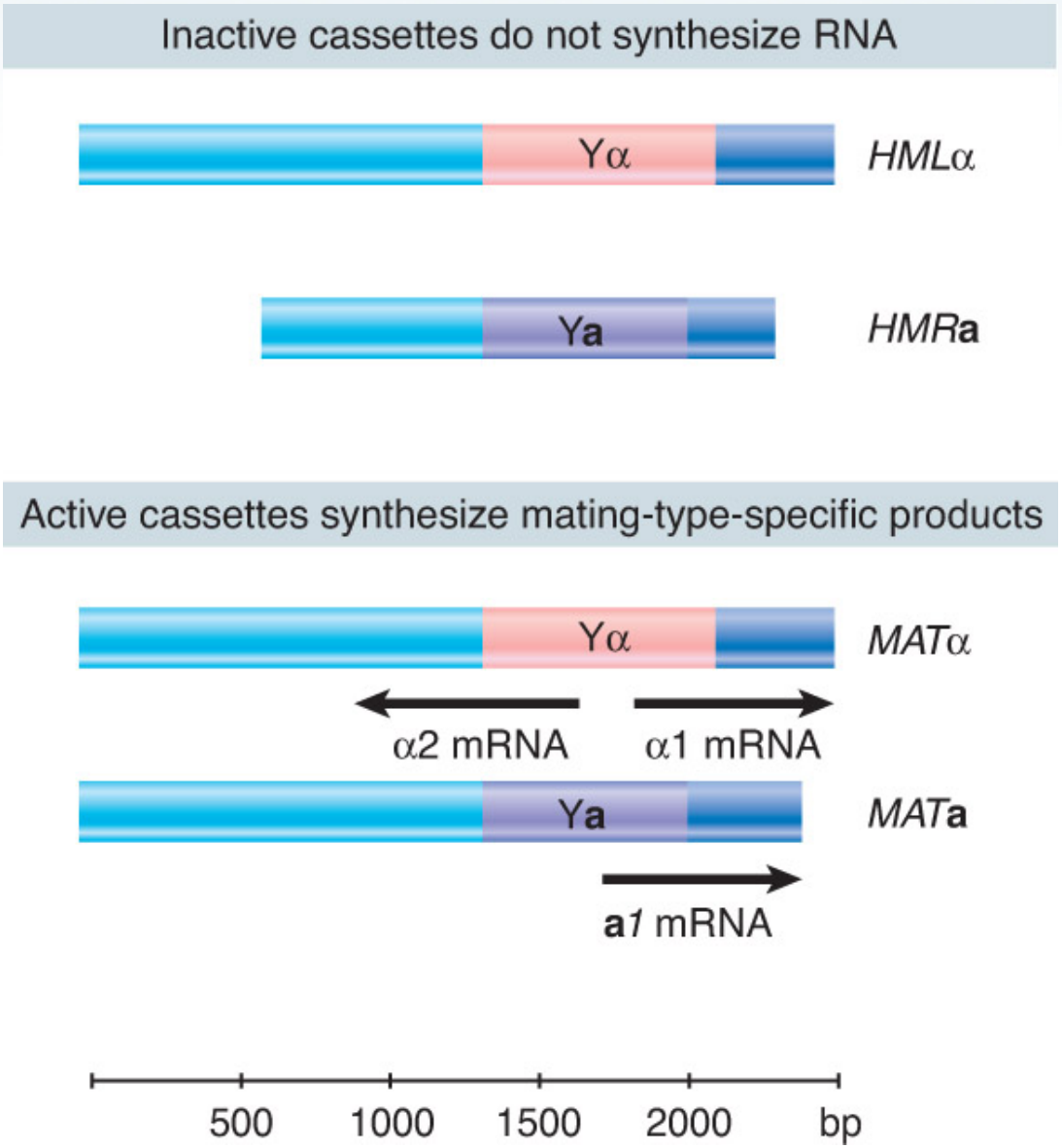
a unidirectional genetic exchange..... involves an HO (homing) endonuclease.


Haploids mate to give diploids
Diffusible factors
a haploid
α haploid
Surface receptor
Surface receptor
a haploid
Factors bind to receptors
α haploid
Mating
a haploid
α haploid
a/α diploid
Sporulation
a
a
α
α
©virtualtext www.ergito.com


6-02-92 48OH
17:13:08

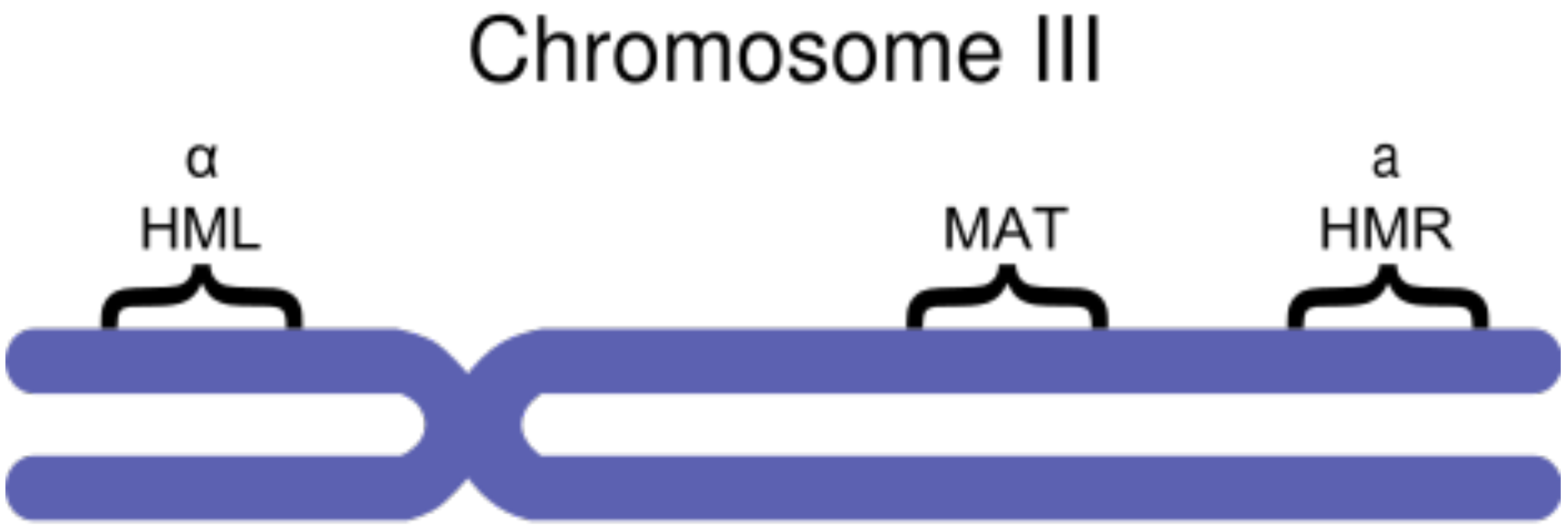

Only mother cells can switch mating type
Switch occurs in α mother both daughters are a
Cells cannot switch again in first generation after last switch
Switch occurs in a mother both daughters are α
α
α
α
a
a
a
a
a
α
α
©virtualtext www.ergito.com

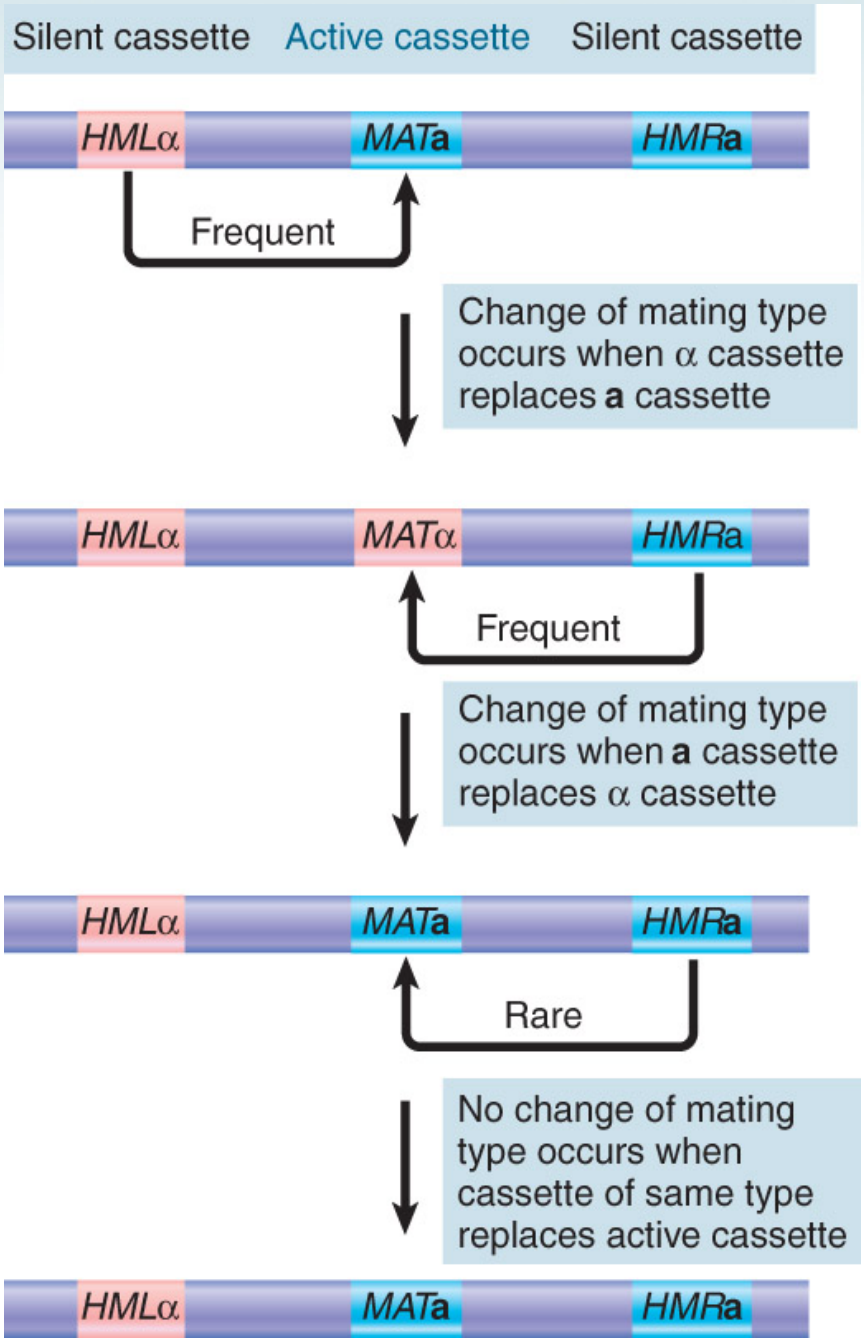
Yeast Can Switch Silent and Active Loci for Mating Type



- Switching occurs if *MAT***a** is replaced by *HMR*α or *MAT*α is replaced by *HMR***a**.

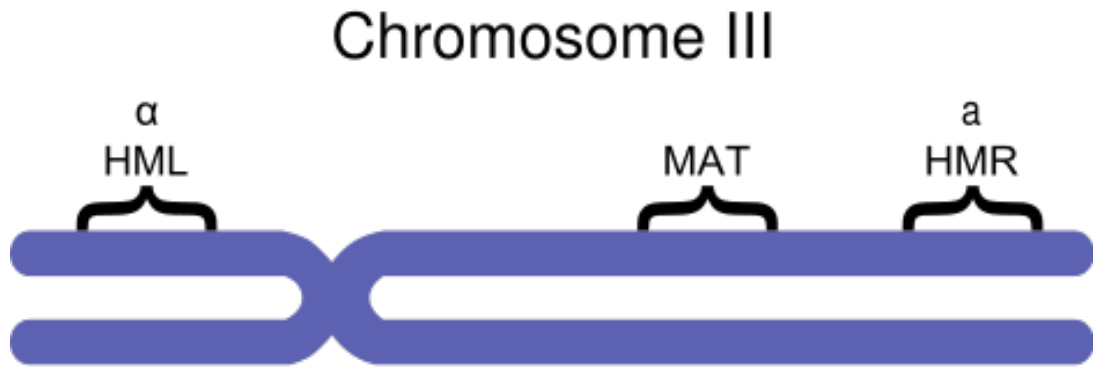
Mating type loci organization


Chromosome III
α
HML
MAT
a
HMR



- The yeast mating type locus *MAT*, a **mating type cassette**, has either the *MAT***a** or *MAT*α genotype.
- Yeast with the dominant allele *HO* switch their mating type at a frequency ~10^{-6}.
- The allele at *MAT* is called the active cassette.
- There are also two silent cassettes, *HML*α and *HMR***a**.

Cassette model for mating type



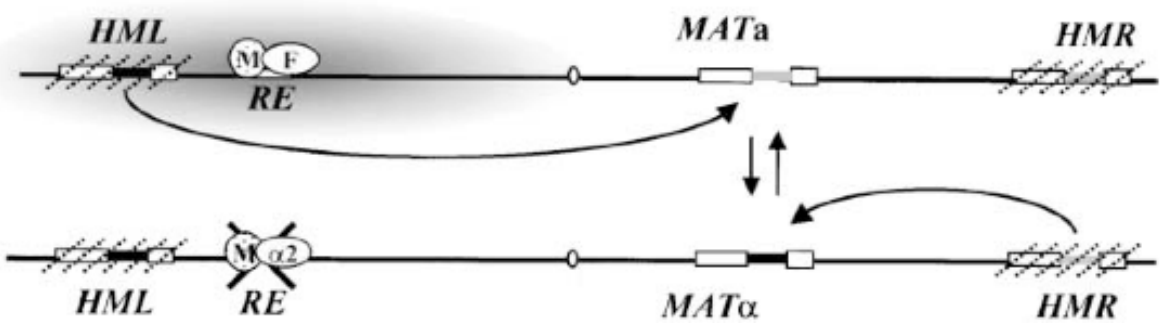
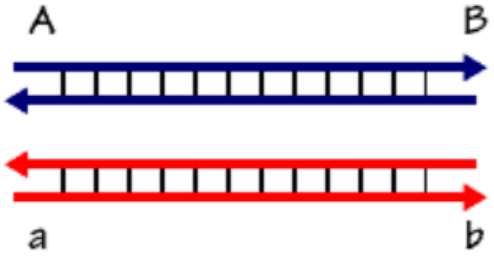
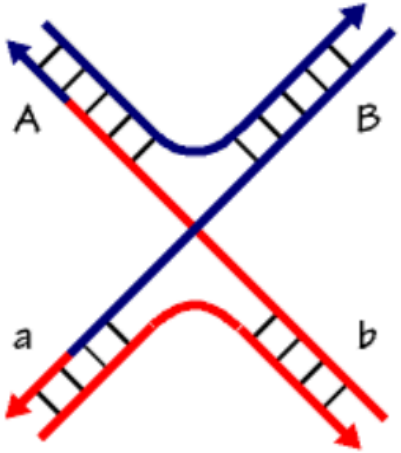
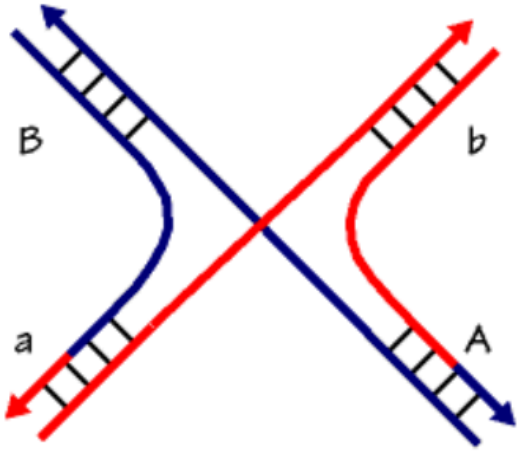


FIGURE 1.—Mating-type switching exhibits cell type-dependent donor preference. A diagram of chromosome III indicating the relative positions of the donor loci, *HML* and *HMR*, the *MAT* locus, and the recombination enhancer, *RE*, is shown. *HML* and *HMR* are transcriptionally silent, indicated by the hatched lines, while *MAT* is transcriptionally active, giving rise to the mating type of the cell. Open rectangles at the three mating loci indicate blocks of homology while the thick line indicates the allele-specific region (α allele is black, **a** allele is gray). In **a** cells (top), Mcm1 (M) and Fkh1 (F), a forkhead transcription factor, occupy *RE* and promote enhanced recombination potential (shaded area) extending over *HML*, rendering it the preferred donor during mating-type switching and resulting in conversion from *MAT***a** to *MAT*α. In α cells (bottom), α2 binds to *RE*, precluding occupation by Fkh1, to suppress enhanced recombination potential, rendering *HMR* the preferred donor through *RE*-independent mechanisms and resulting in conversion from *MAT*α to *MAT***a**.

Matthew P Scott, Paul Matsudaira, Harvey Lodish, James Darnell, Lawrence Zipursky, Chris A Kaiser, Arnold Berk, Monty Krieger (2004). *Molecular Cell Biology, Fifth Edition*. WH Freeman and Col, NY


A
B
a
b


A
B
a
b


B
b
a
A

Nature. 2008 Nov 20;456(7220):357-61. doi: 10.1038/nature07470.

Identification of Holliday junction resolvases from humans and yeast.

Ip SC[1], Rass U, Blanco MG, Flynn HR, Skehel JM, West SC.

Author information

Abstract

Four-way DNA intermediates, also known as Holliday junctions, are formed during homologous recombination and DNA repair, and their resolution is necessary for proper chromosome segregation. Here we identify nucleases from *Saccharomyces cerevisiae* and human cells that promote Holliday junction resolution, in a manner analogous to that shown by the *Escherichia coli* Holliday junction resolvase RuvC. The human Holliday junction resolvase, **GEN1**, and its yeast orthologue, **Yen1**, were independently identified using two distinct experimental approaches: GEN1 was identified by mass spectrometry following extensive fractionation of HeLa cell-free extracts, whereas Yen1 was detected by screening a yeast gene fusion library for nucleases capable of Holliday junction resolution. The eukaryotic Holliday junction resolvases represent a new subclass of the Rad2/XPG family of nucleases. Recombinant GEN1 and Yen1 resolve Holliday junctions by the introduction of symmetrically related cuts across the junction point, to produce nicked duplex products in which the nicks can be readily ligated.

PMID: 19020614 DOI: 10.1038/nature07470

[Indexed for MEDLINE]

Nucleic Acids Res. 2012 Jul; 40(13): 5795–5818. PMCID: PMC3401455

Published online 2012 Mar 30. doi: 10.1093/nar/gks270

Homologous recombination and its regulation

Lumir Krejci,[1,2,3,*] Veronika Altmannova,[1,2,3] Mario Spirek,[1] and Xiaolan Zhao[4]

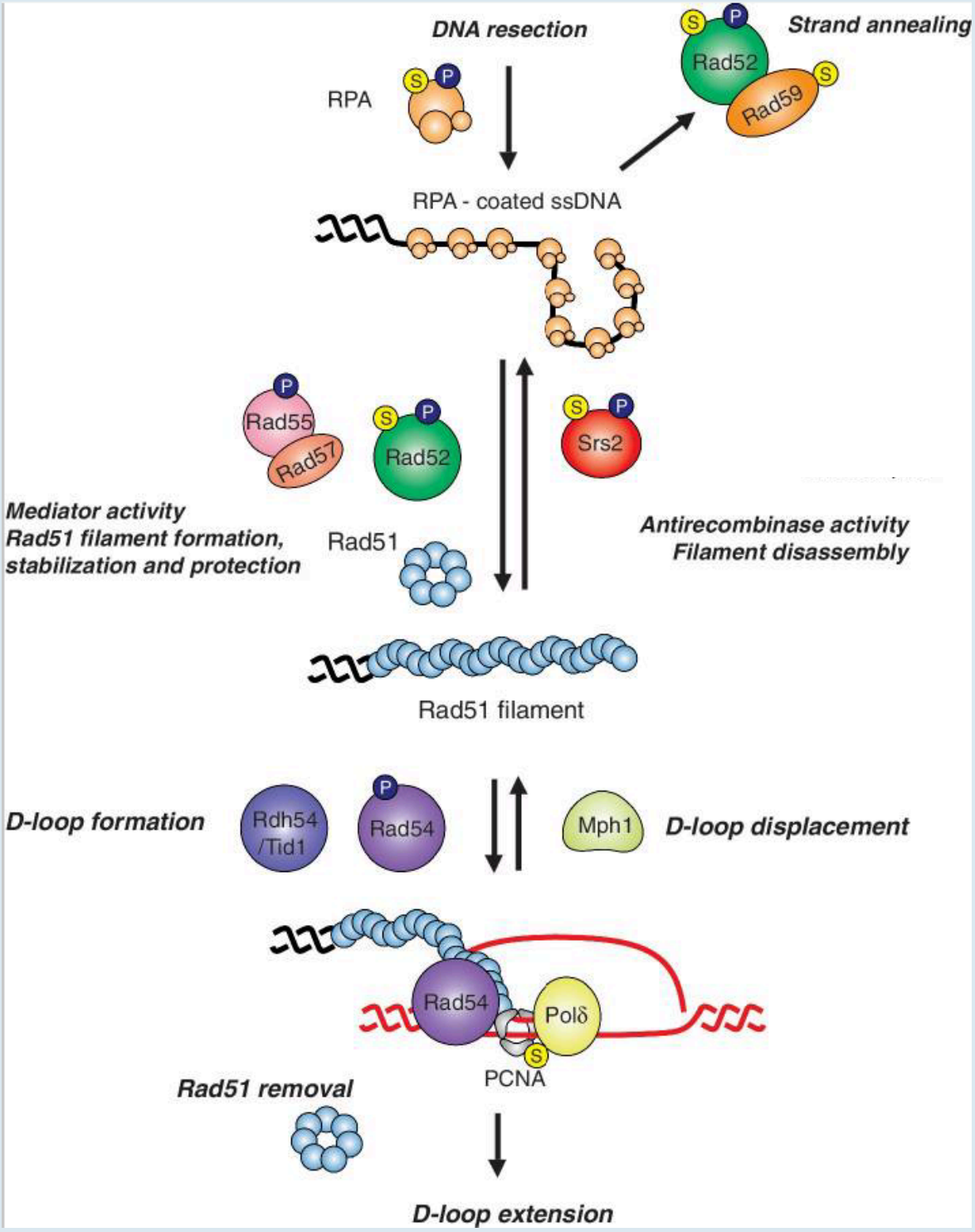
Author information ▶ Article notes ▶ Copyright and License information ▶

ABSTRACT

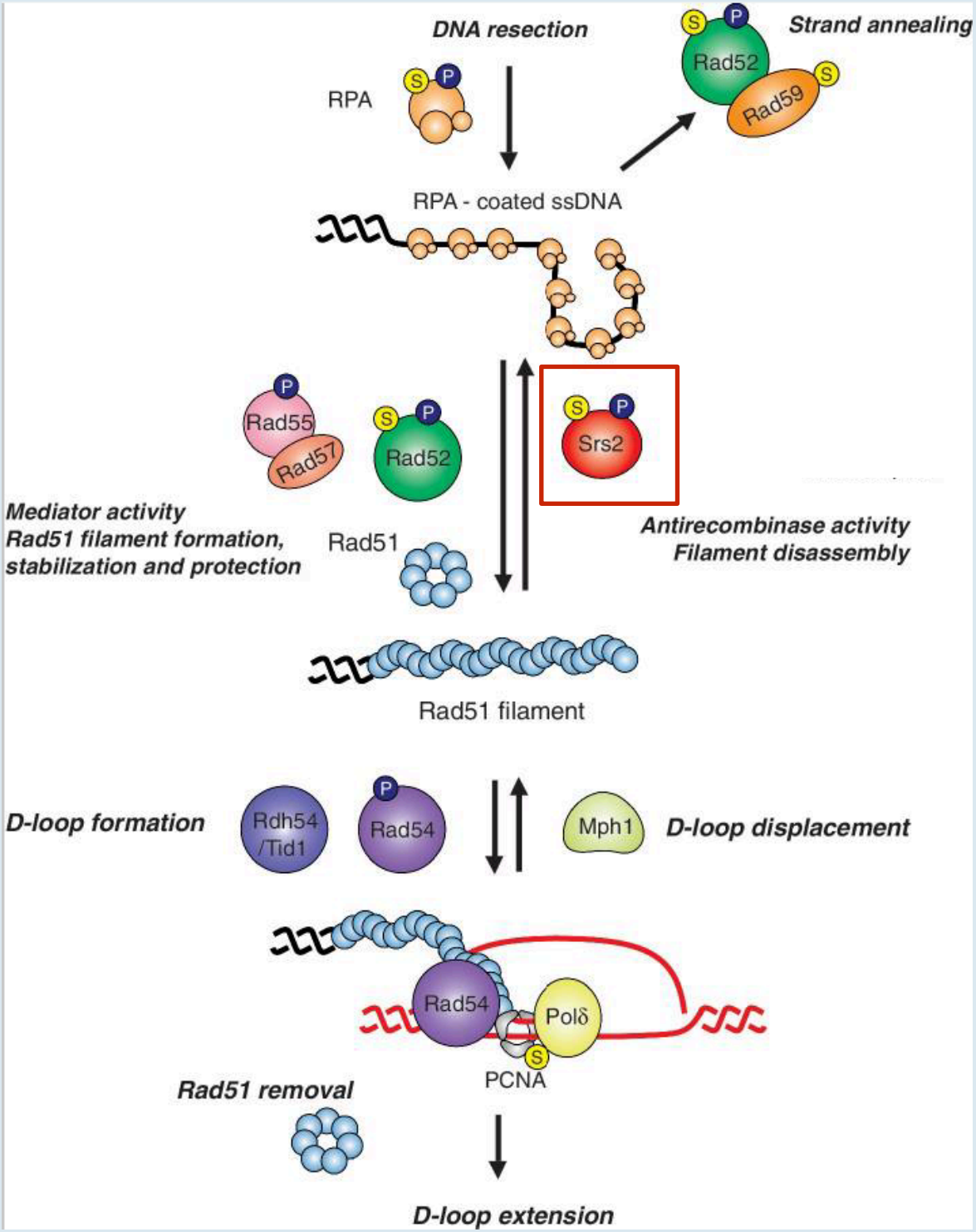
Recommended Reading

Go to:

Homologous recombination (HR) is critical both for repairing DNA lesions in mitosis and for chromosomal pairing and exchange during meiosis. However, some forms of HR can also lead to undesirable DNA rearrangements. Multiple regulatory mechanisms have evolved to ensure that HR takes place at the right time, place and manner. Several of these impinge on the control of Rad51 nucleofilaments that play a central role in HR. Some factors promote the formation of these structures while others lead to their disassembly or the use of alternative repair pathways. In this article, we review these mechanisms in both mitotic and meiotic environments and in different eukaryotic taxa, with an emphasis on yeast and mammal systems. Since mutations in several proteins that regulate Rad51 nucleofilaments are associated with cancer and cancer-prone syndromes, we discuss how understanding their functions can lead to the development of better tools for cancer diagnosis and therapy.



Rad51 filament formation and regulation. RPA can be replaced by Rad51 from ssDNA with the help of recombination mediators, including **Rad52** and **Rad55/57**.
These proteins can promote both the formation and stabilization of Rad51 presynaptic filaments



Rad51 filament formation and regulation. RPA can be replaced by Rad51 from ssDNA with the help of recombination mediators, including **Rad52** and **Rad55/57**.
These proteins can promote both the formation and stabilization of Rad51 presynaptic filaments

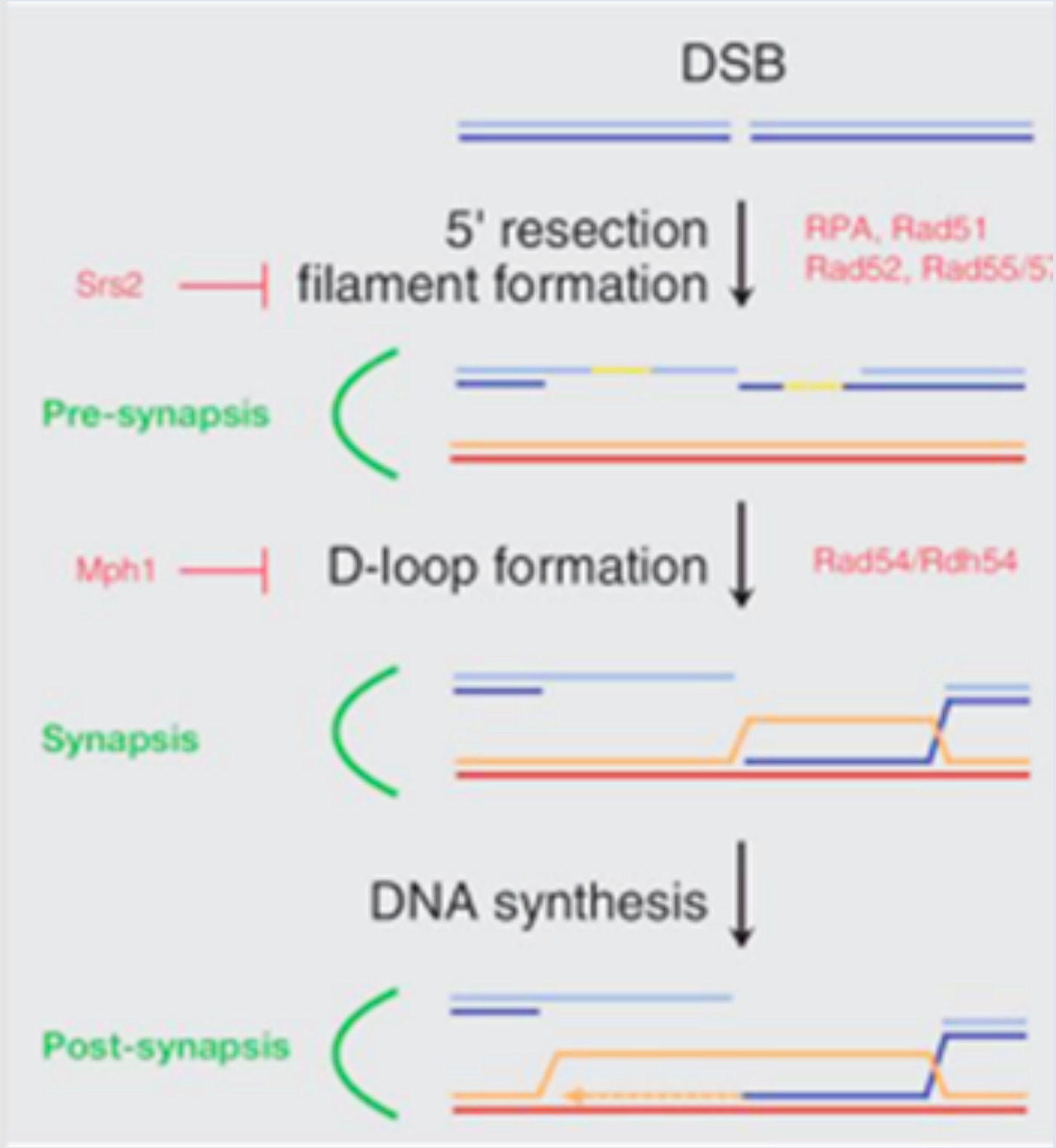


Figure 1. Models for the repair of DNA double-strand breaks. DNA DSBs are resected to generate 3'-protruding ends followed by formation of Rad51 filaments that invade into homologous template to form D-loop structures.

(A) After priming DNA synthesis, three pathways can be invoked. In the DSBR pathway, the second end is captured and a dHJ intermediate is formed.

(B) Resolution of dHJs can occur in either plane to generate crossover or non-crossover products. Alternatively, dHJs can be dissolved by the action of Sgs1–Top1–Rmi1 complex to generate only non-crossovers. In the SDSA pathway

(C) the extended nascent strand is displaced, followed by pairing with the other 3'-single-stranded tail, and DNA synthesis completes repair. Nucleolytic trimming might be also required.

(D) In the third pathway of BIR which can act when the second end is absent, the D-loop intermediate turns into a replication fork capable of both lagging and leading strand synthesis. Two other Rad51-independent recombinational repair pathways are also depicted.

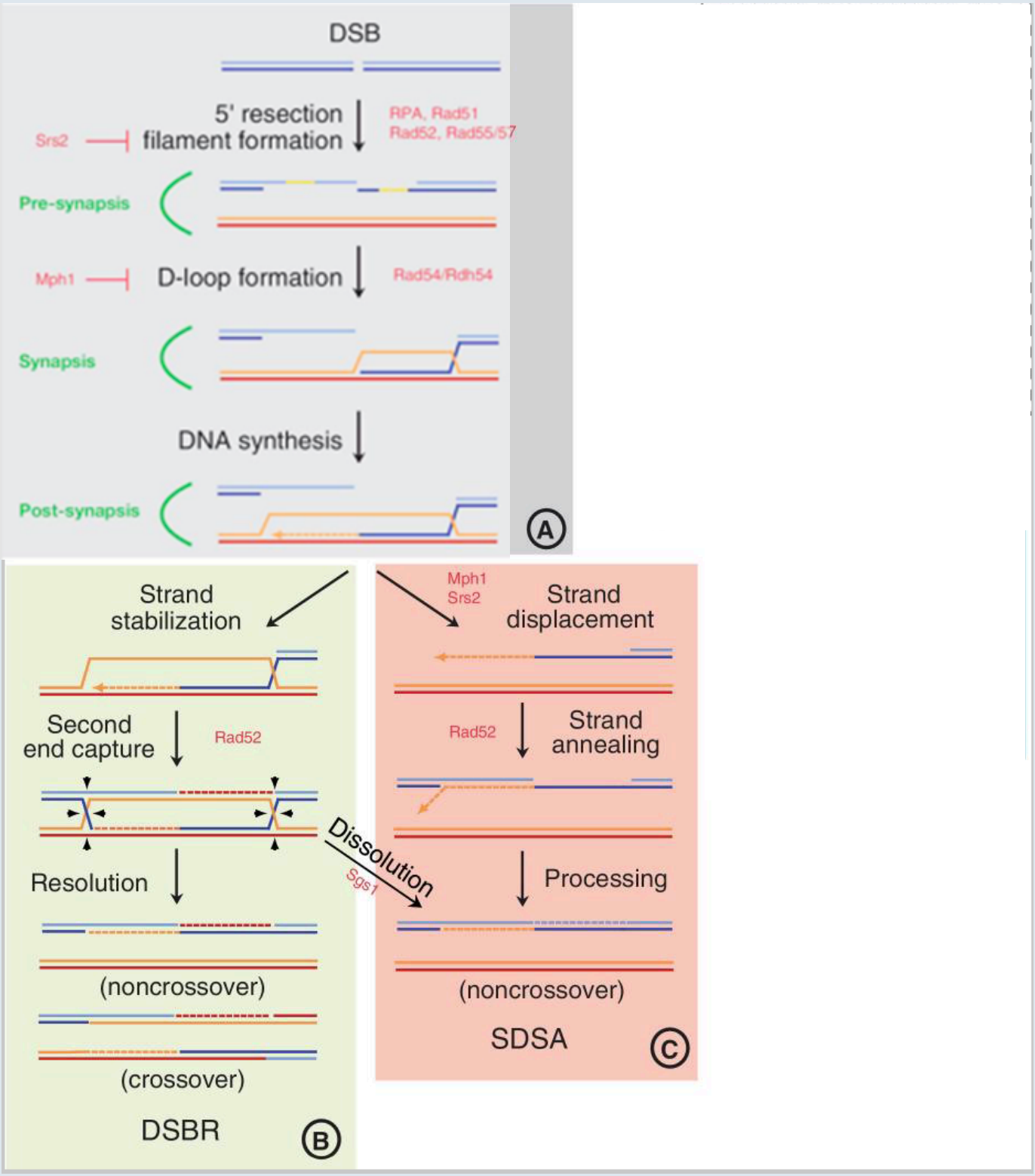


Figure 1. Models for the repair of DNA double-strand breaks. DNA DSBs are resected to generate 3'-protruding ends followed by formation of Rad51 filaments that invade into homologous template to form D-loop structures.

(A) After priming DNA synthesis, three pathways can be invoked. In the **DSBR** pathway, the second end is captured and a dHJ intermediate is formed.

(B) Resolution of dHJs can occur in either plane to generate crossover or non-crossover products. Alternatively, dHJs can be dissolved by the action of **Sgs1–Top1–Rmi1** complex to generate primarily **non-crossovers.** In the **SDSA** pathway

(C) the extended nascent strand is displaced, followed by pairing with the other 3'-single-stranded tail, and DNA synthesis completes repair. Nucleolytic trimming might be also required.

(D) In the third pathway of **BIR** which can act when the second end is absent, the D-loop intermediate turns into a replication fork capable of both lagging and leading strand synthesis.

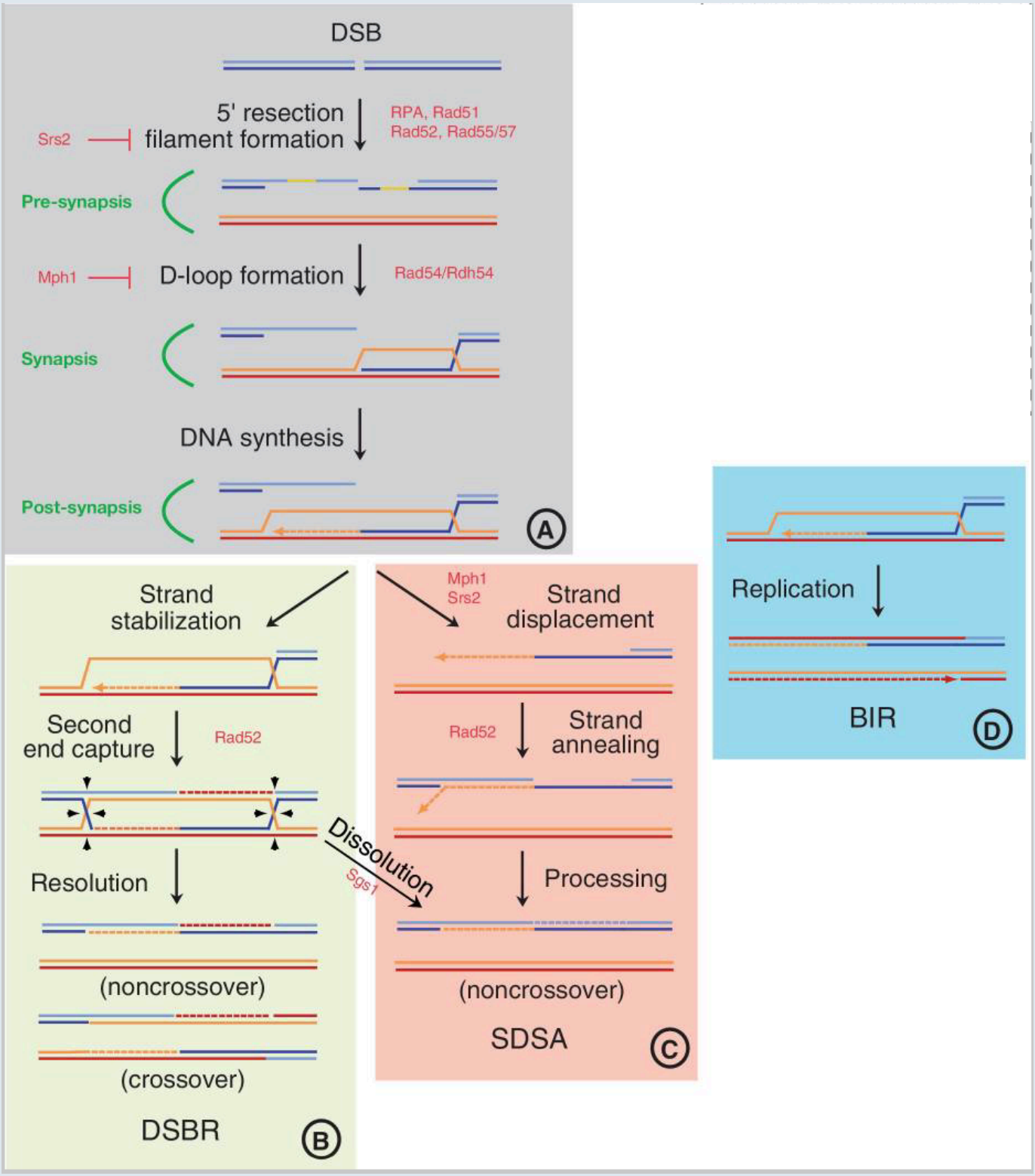


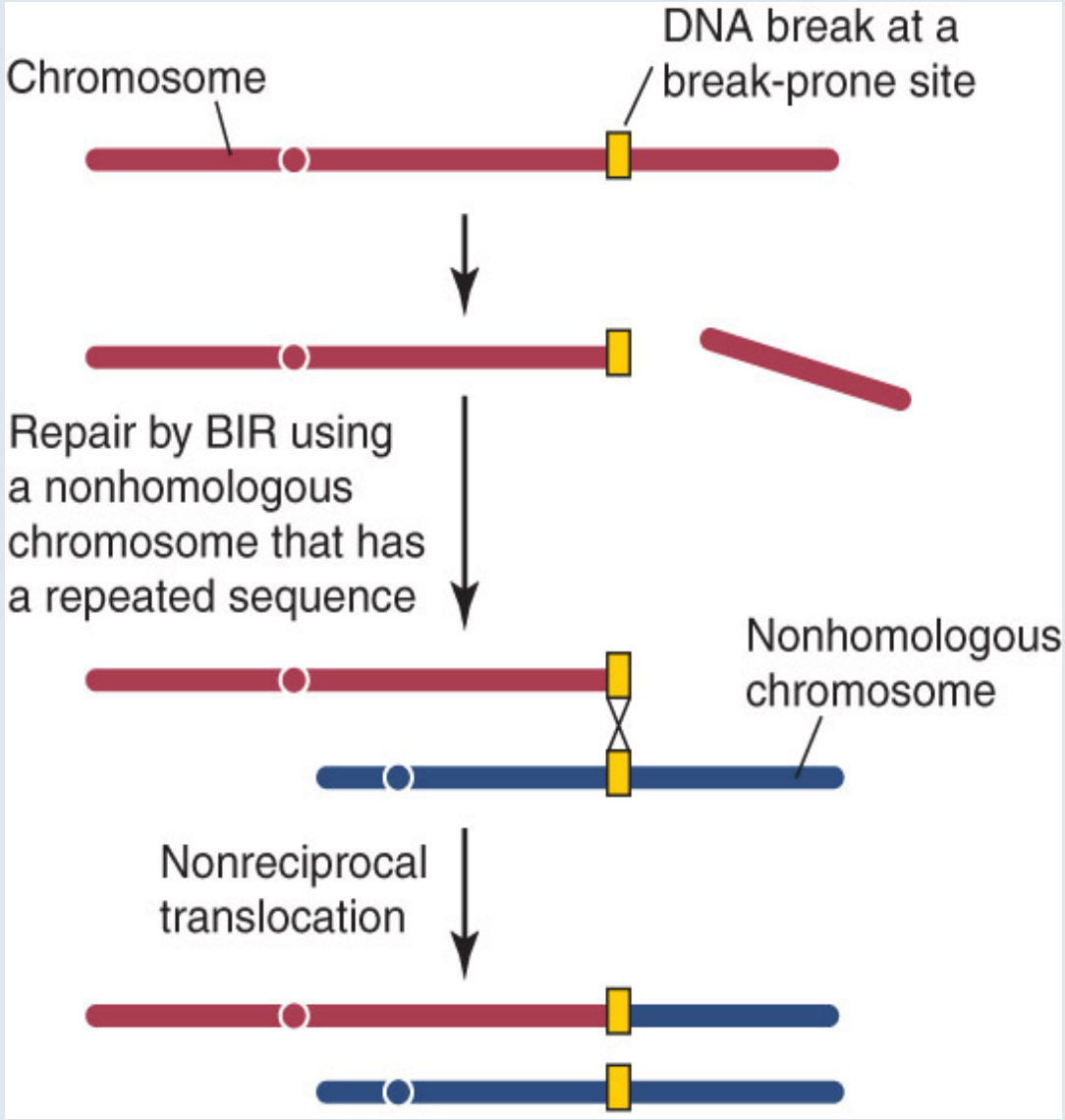
Figure 1. Models for the repair of DNA double-strand breaks. DNA DSBs are resected to generate 3ʾ-protruding ends followed by formation of Rad51 filaments that invade into homologous template to form D-loop structures.

(A) After priming DNA synthesis, three pathways can be invoked. In the **DSBR** pathway, the second end is captured and a dHJ intermediate is formed.

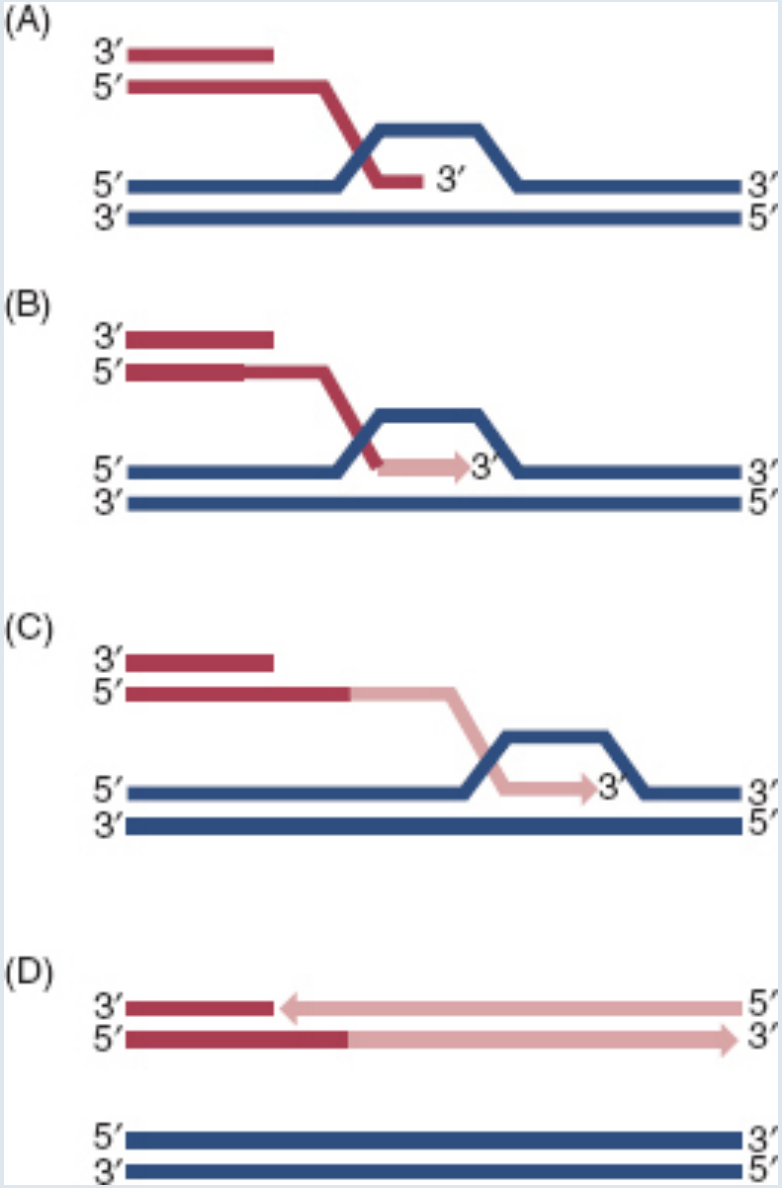
(B) Resolution of dHJs can occur in either plane to generate crossover or non-crossover products. Alternatively, dHJs can be dissolved by the action of **Sgs1–Top1–Rmi1** complex to generate primarily **non-crossovers.** In the **SDSA** pathway

(C) the extended nascent strand is displaced, followed by pairing with the other 3ʾ-single-stranded tail, and DNA synthesis completes repair. Nucleolytic trimming might be also required.

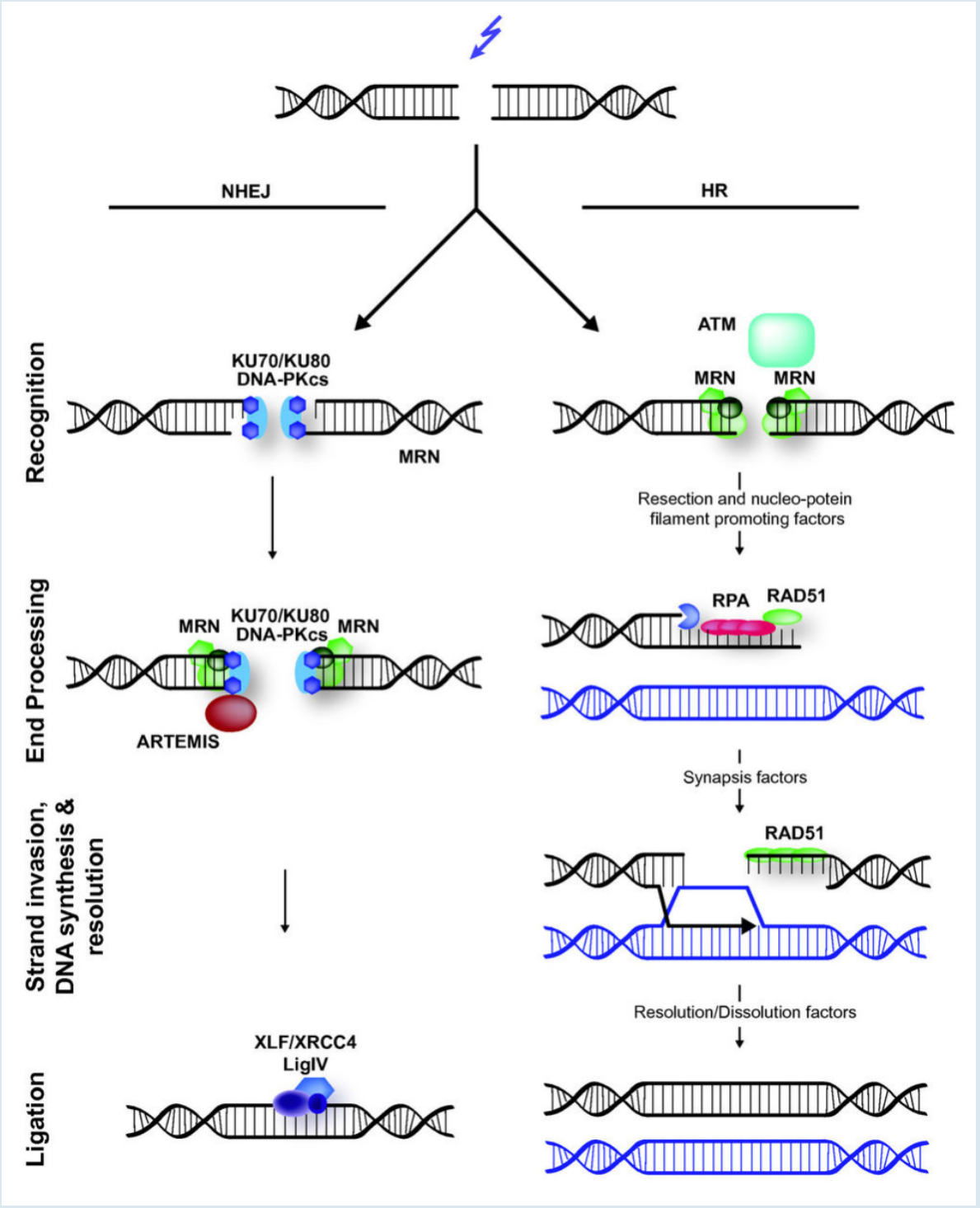
(D) In the third pathway of **BIR** which can act when the second end is absent, the D-loop intermediate turns into a replication fork capable of both lagging and leading strand synthesis.

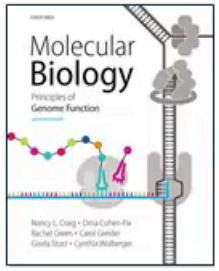


Break **I**nduced **R**ecombination (**BIR**) initiating translocations



Break **I**nduced **R**ecombination (BIR) initiating translocations, sometimes through regional homology with DNA potentially from other chromosomes


NHEJ
HR
Recognition
KU70/KU80
DNA-PKcs
MRN
ATM
MRN
MRN
Resection and nucleo-potein
filament promoting factors
End Processing
MRN
KU70/KU80
DNA-PKcs
MRN
ARTEMIS
RPA
RAD51
Synapsis factors
Strand invasion,
DNA synthesis &
resolution
RAD51
Resolution/Dissolution factors
Ligation
XLF/XRCC4
LigIV



Molecular Biology: Principles of Genome Function

Second Edition

OXFORD UNIVERSITY PRESS

Animation 12: **Non-homologous end joining**

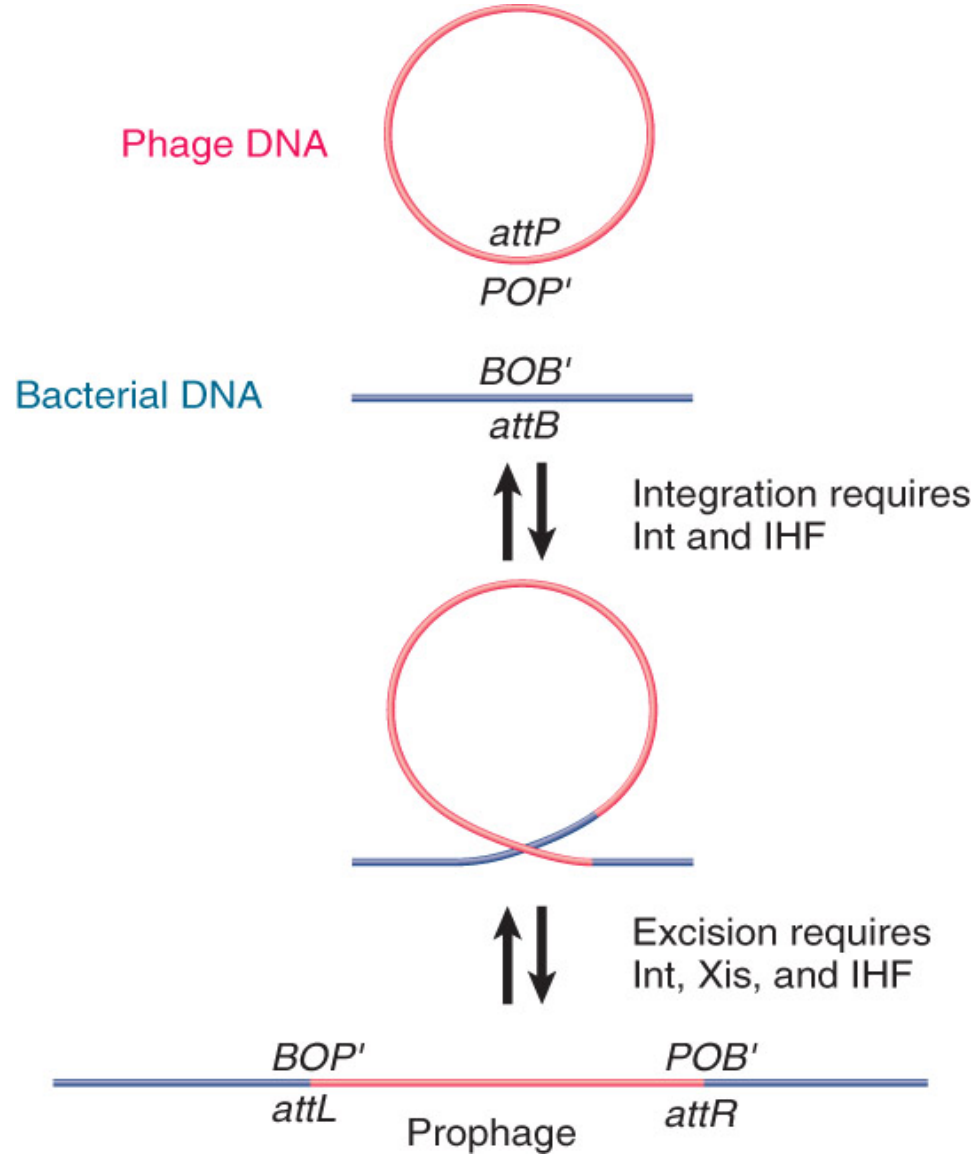
https://www.youtube.com/watch?v=31stiofJjYw

- Recombination occurs in a manner which prevents the loss or insertion of DNA bases. Three types of recombination:
  - **Homologous recombination** – also known as generalized; occurs in DNA repair and at meiosis.
  - **Non homologous-** Error prone repair
  - **Site specific homologous**(specialized) recombination – typically in bacteria and viruses; enzymes involved act only on a particular pair of target sequences in an intermolecular reaction
  - **Transposition (including gene rearrangements) – where DNA sequences can be inserted into another sequence without relying on sequence homology between the two.**

- Recombination occurs in a manner which prevents the loss or insertion of DNA bases. Three types of recombination:
  - **Homologous recombination** – also known as generalized; occurs in DNA repair and at meiosis.
  - **Non homologous-** Error prone repair.
  - **Site specific homologous**(specialized) recombination – typically in bacteria and viruses; enzymes involved act only on a particular pair of target sequences in an intermolecular reaction
  - **Transposition (including gene rearrangements) – where DNA sequences can be inserted into another sequence without relying on sequence homology between the two.**

- Recombination occurs in a manner which prevents the loss or insertion of DNA bases. Three types of recombination:
  - **Homologous recombination** – also known as generalized; occurs in DNA repair and at meiosis.
  - **Non homologous-** Error prone repair.
  - **Site specific homologous**(specialized) recombination – typically in bacteria and viruses; enzymes involved act only on a particular pair of target sequences in an intermolecular reaction
  - **Transposition (including gene rearrangements) – where DNA sequences can be inserted into another sequence without relying on sequence homology between the two.**

- Specialized recombination involves reaction between specific sites that are not necessarily homologous.
- **recombinase (integrase)–** which is an enzyme that catalyzes site-specific recombination.
- Phage lambda **integrates** into the bacterial chromosome by recombination between a site on the phage and the ***att*** **site** on the *E. coli* chromosome.



Phage integration at att sites

Site-specific recombination of yeast 2-micron DNA in vitro

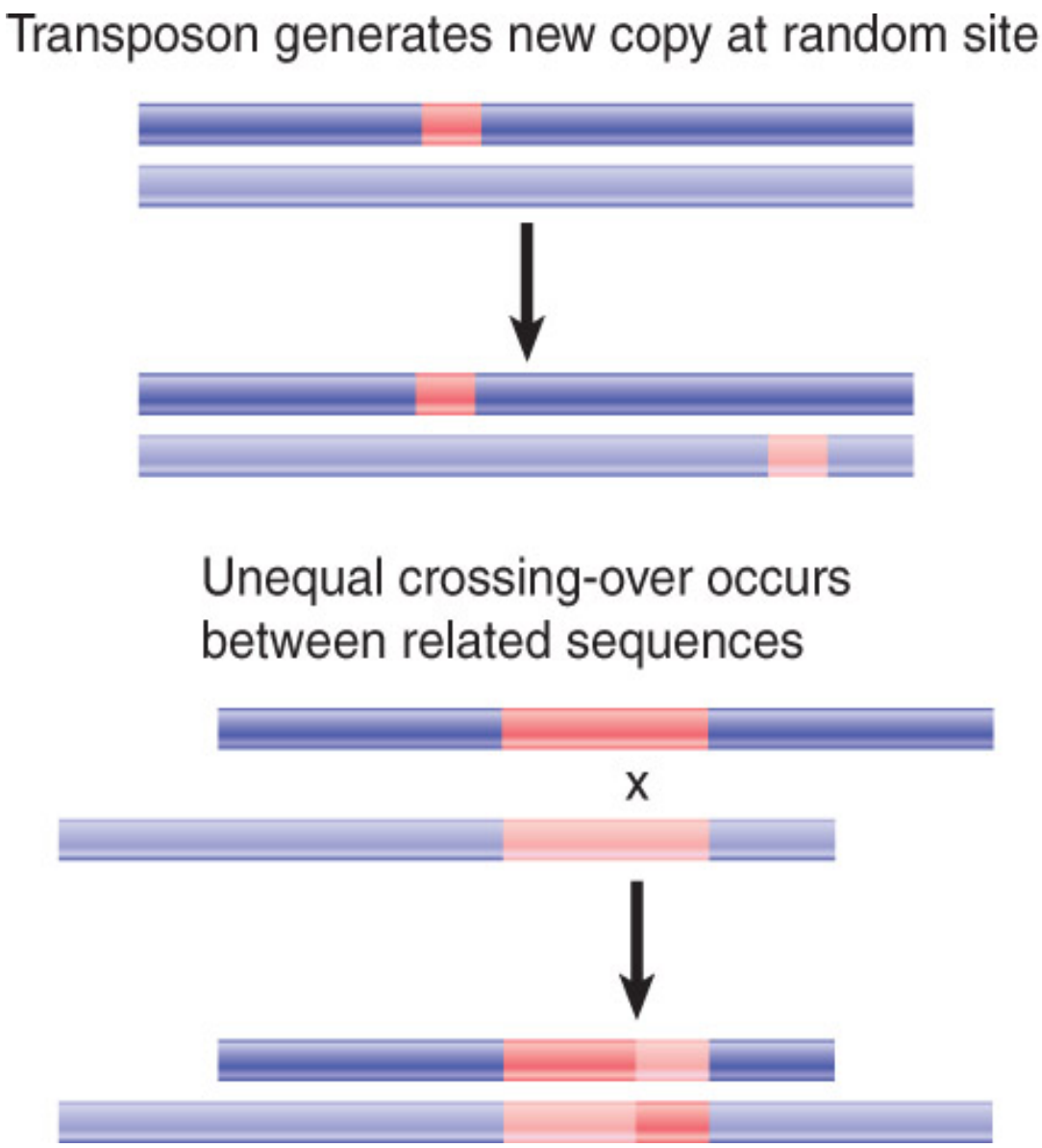
D Vetter, B J Andrews, L Roberts-Beatty, and P D Sadowski

PNAS December 1, 1983 80 (23) 7284-7288; https://doi.org/10.1073/pnas.80.23.7284

Abstract

Most strains of the yeast Saccharomyces cerevisiae harbor several copies of a 2-micron plasmid circle DNA termed "2 micron." This circular plasmid contains two 599-base-pair precise inverted repeats across which a site-specific inversion event occurs in vivo. This inversion is promoted by a plasmid-encoded function called "FLP." We have cloned the FLP gene of 2-micron DNA under control of a strong yeast promoter and transformed yeast cells with a plasmid containing the cloned FLP gene. Cell-free extracts from such a transformant promote highly efficient inversion of 2-micron DNA in vitro. The reaction requires a cation and works efficiently on supercoiled, relaxed circular, or linear DNA. The FLP activity bears certain similarities to the cre protein, a site-specific recombinase encoded by bacteriophage P1.

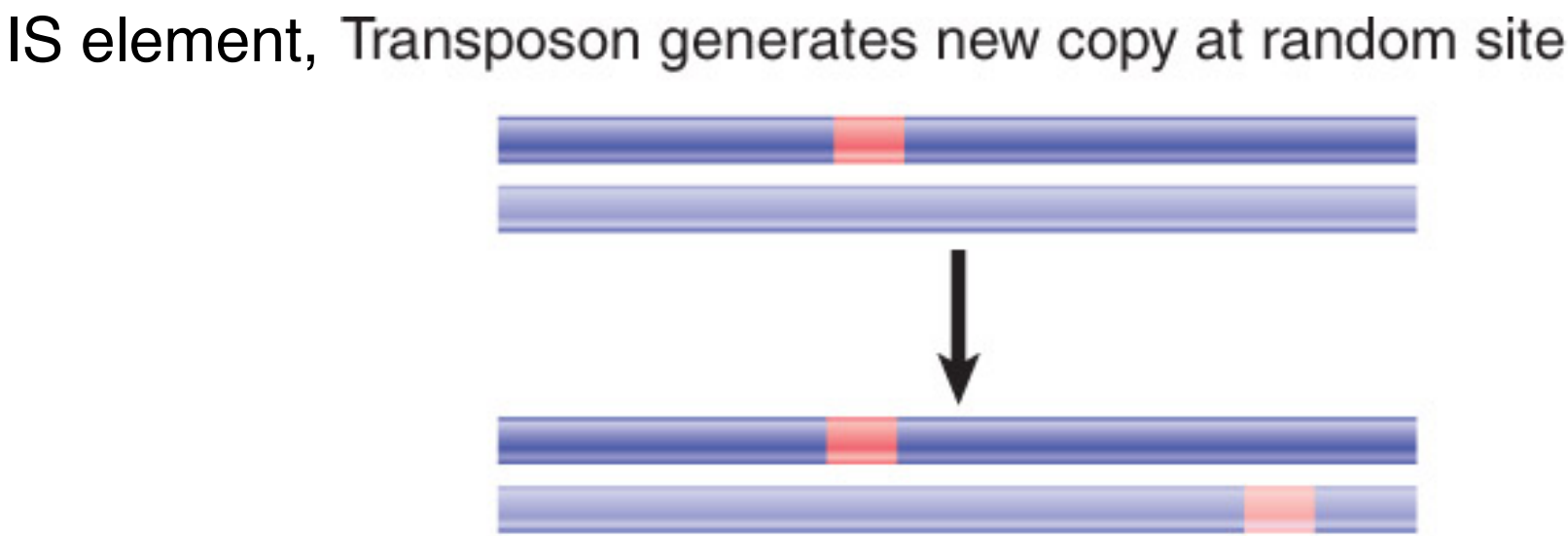
- Recombination occurs in a manner which prevents the loss or insertion of DNA bases. Three types of recombination:
  - **Homologous recombination** – also known as generalized; occurs in DNA repair and at meiosis.
  - **Non homologous-** Error prone repair.
  - **Site specific homologous**(specialized) recombination – typically in bacteria and viruses; enzymes involved act only on a particular pair of target sequences in an intermolecular reaction
  - **Transposition (including gene rearrangements) – where DNA sequences can be inserted into another sequence without relying on sequence homology between the two.**

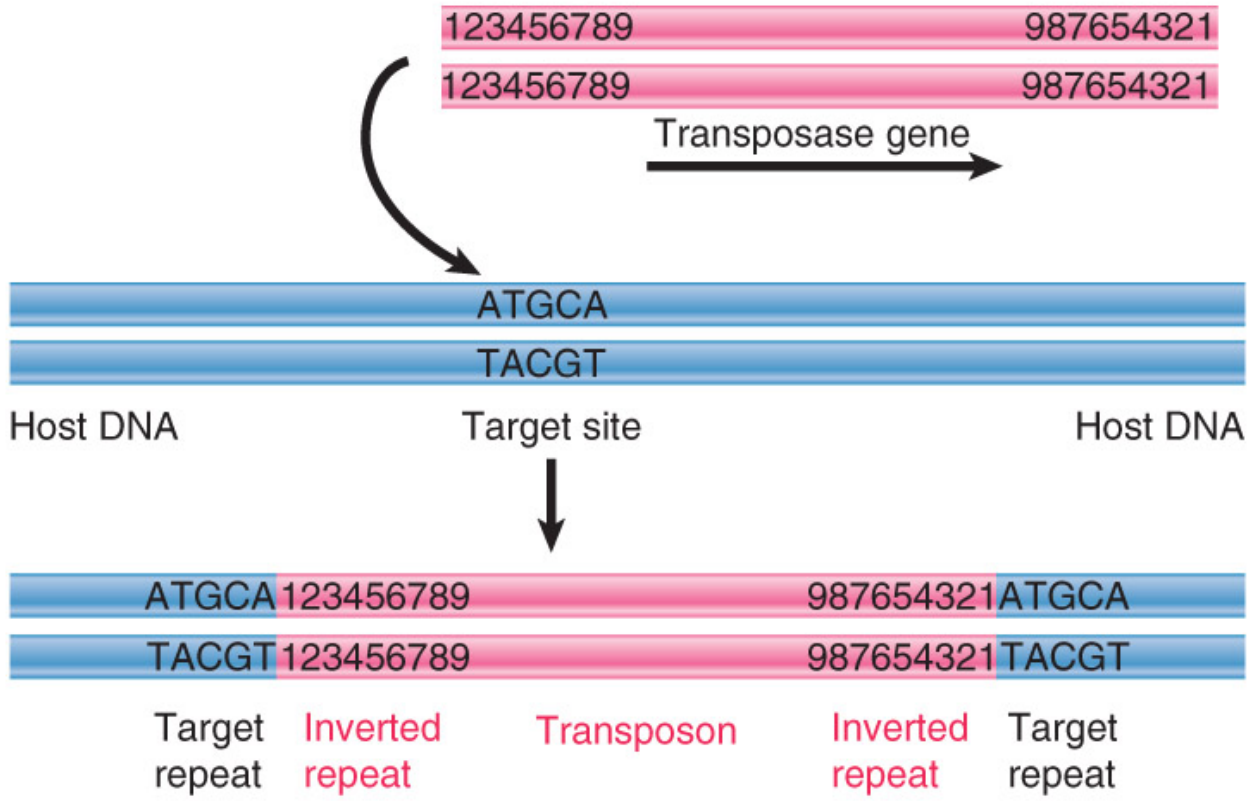


- Transposons move directly from one site in the genome to another.
- Transposons fall into two general classes: those that are mobilized directly through DNA and those that are mobilized in an RNA dependent manner.
- Transposable elements can promote rearrangements of the genome directly or indirectly
  - By causing deletions or inversions or leading to the movement of host sequences to a new location
  - By serving as substrates for cellular recombination systems

- An **insertion sequence (IS)** is a transposon that codes for the enzyme(s) needed for transposition flanked by short **inverted terminal repeats**.

- The target site at which an insertion sequence is inserted is duplicated during the insertion process to form two repeats in direct orientation at the ends of the transposon (**direct repeats**).





Transposon	Target repeat (bp)	Inverted repeat (bp)	Overall length (bp)	Target selection
IS1	9	23	768	random
IS2	5	41	1327	hotspots
IS4	11–13	18	1428	$AAAN_{20}TTT$
IS5	4	16	1195	hotspots
IS10R	9	22	1329	NGCTNAGCN
IS50R	9	9	1531	hotspots
IS903	9	18	1057	random

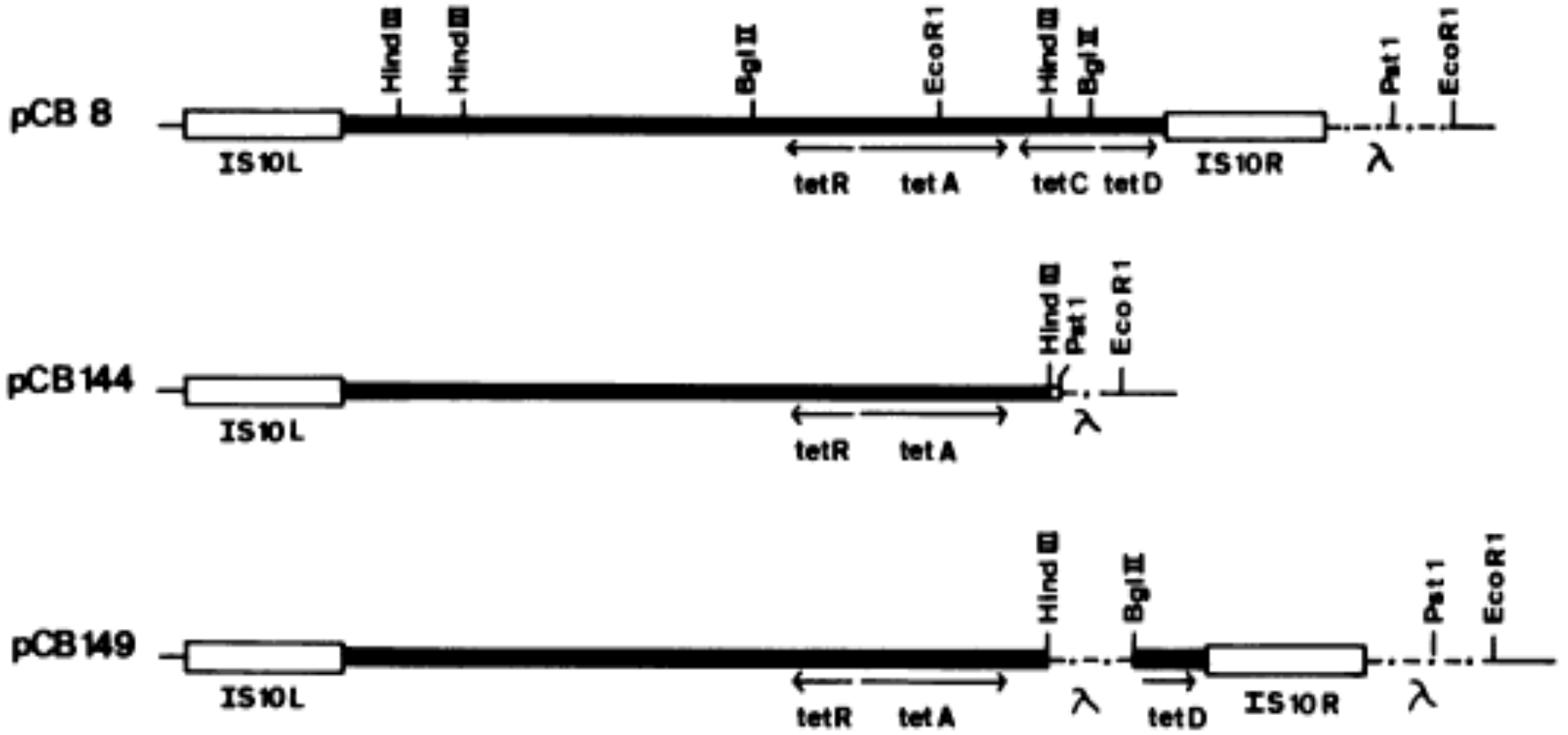
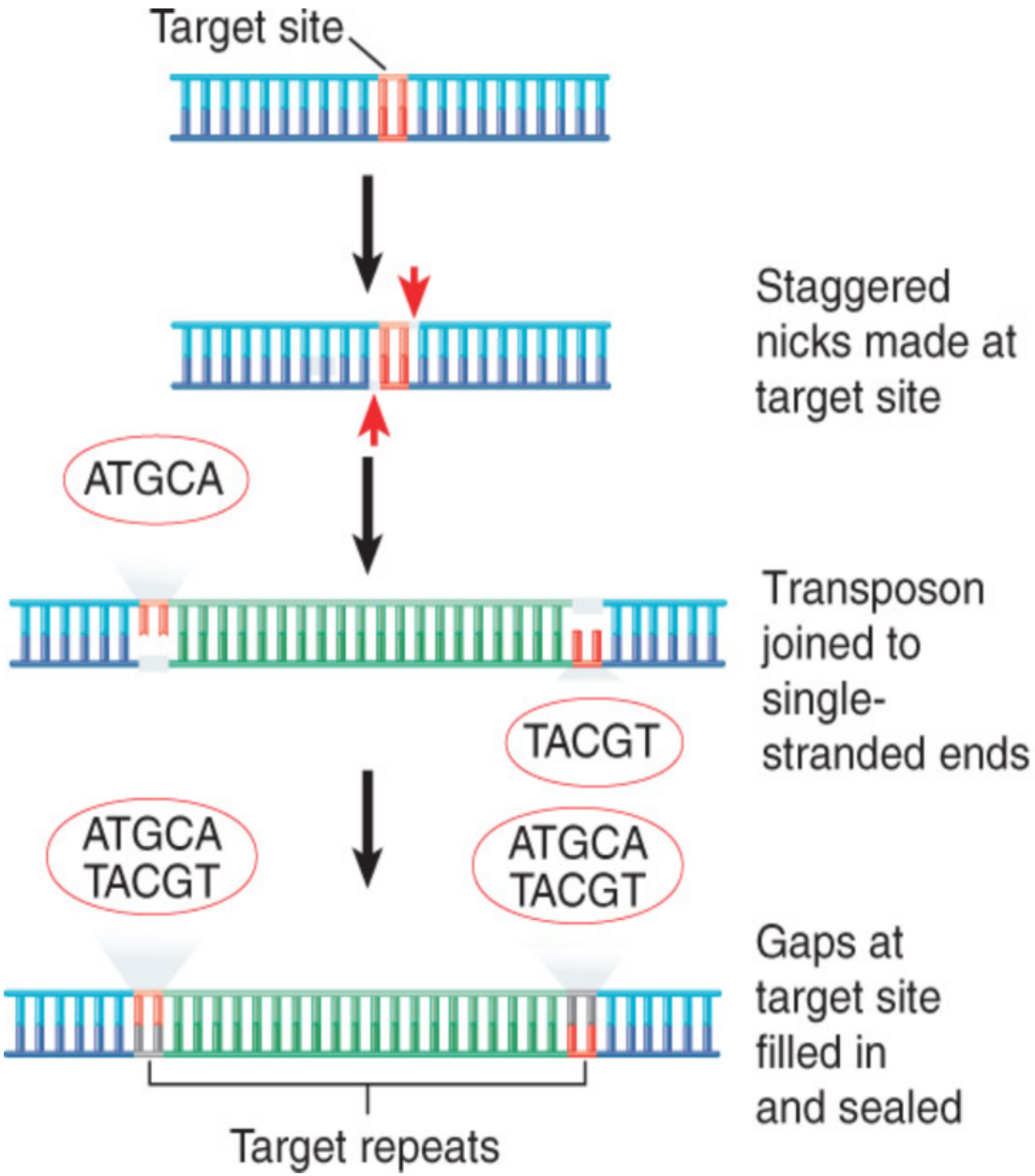


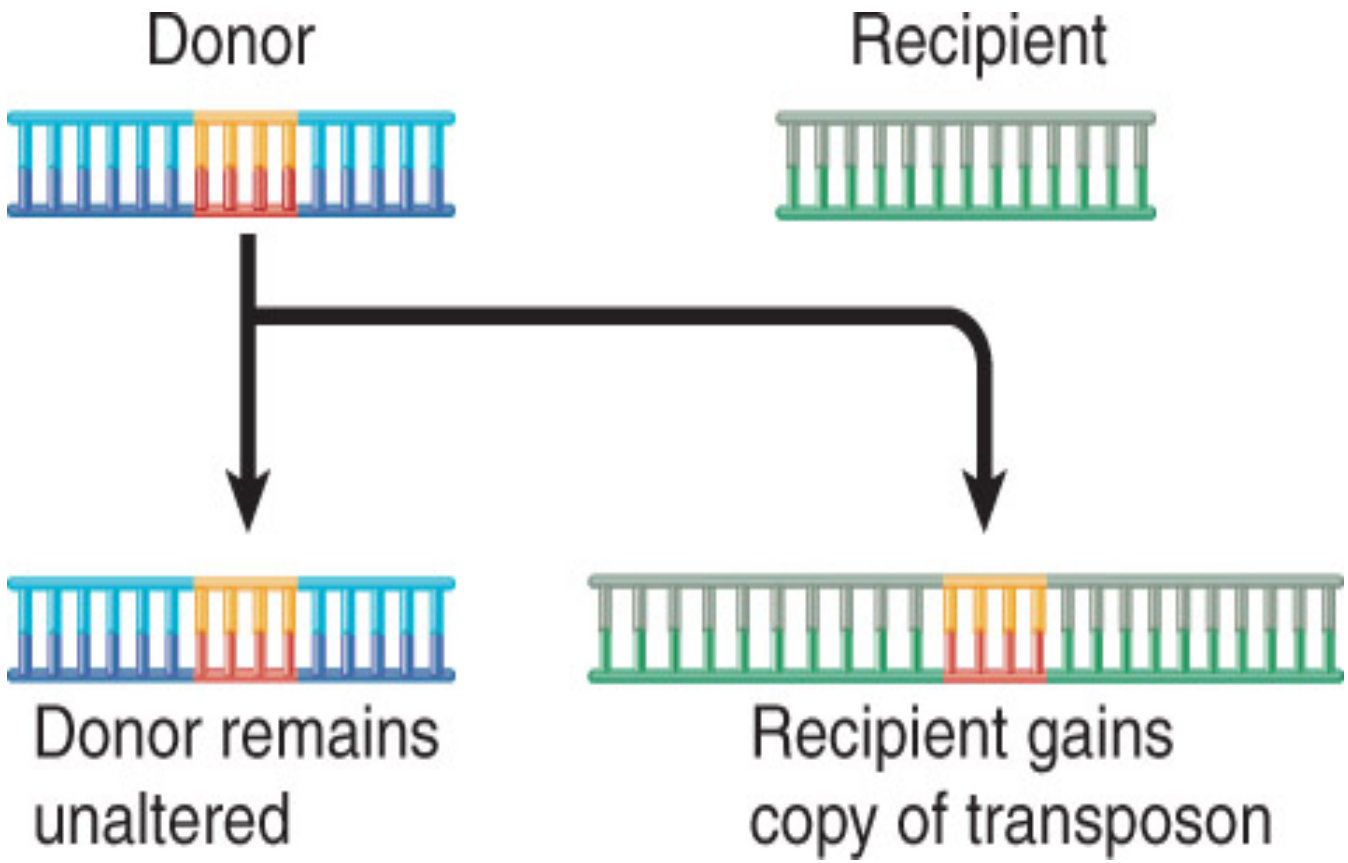
FIG. 1. Map of transposon Tn*10* on plasmid pCB8 and of mutant derivatives. The thick solid line represents the Tn*10* core region. The genes identified and the relevant restriction sites are shown. The construction of plasmids pCB144 and pCB149 is described in the text.

Transposition Occurs by Both **Replicative** and **Non replicative** Mechanisms



- All transposons use a common mechanism in which:
  - staggered nicks are made in target DNA
  - the transposon is joined to the protruding ends
  - the gaps are filled
- The generation and filling of the staggered ends explain the occurrence of direct repeats of target DNA at the site of insertion.
- Transposons move by three different mechanisms, which are also differentiated by the order of events and the nature of the connections between the target and donor.

- In **replicative transposition**, a copy of the original element is moved into the recipient. The original donor molecule is unchanged.
- This type of transposition involves a transposase activity and a resolvase activity.



- In **nonreplicative transposition**, the transposing element moves as a physical entity directly from one site to another and is conserved.
- For some elements, this type of transposition requires only a transposase. For other elements it requires a connection between the donor and target sequences and shares some steps with replicative transposition.

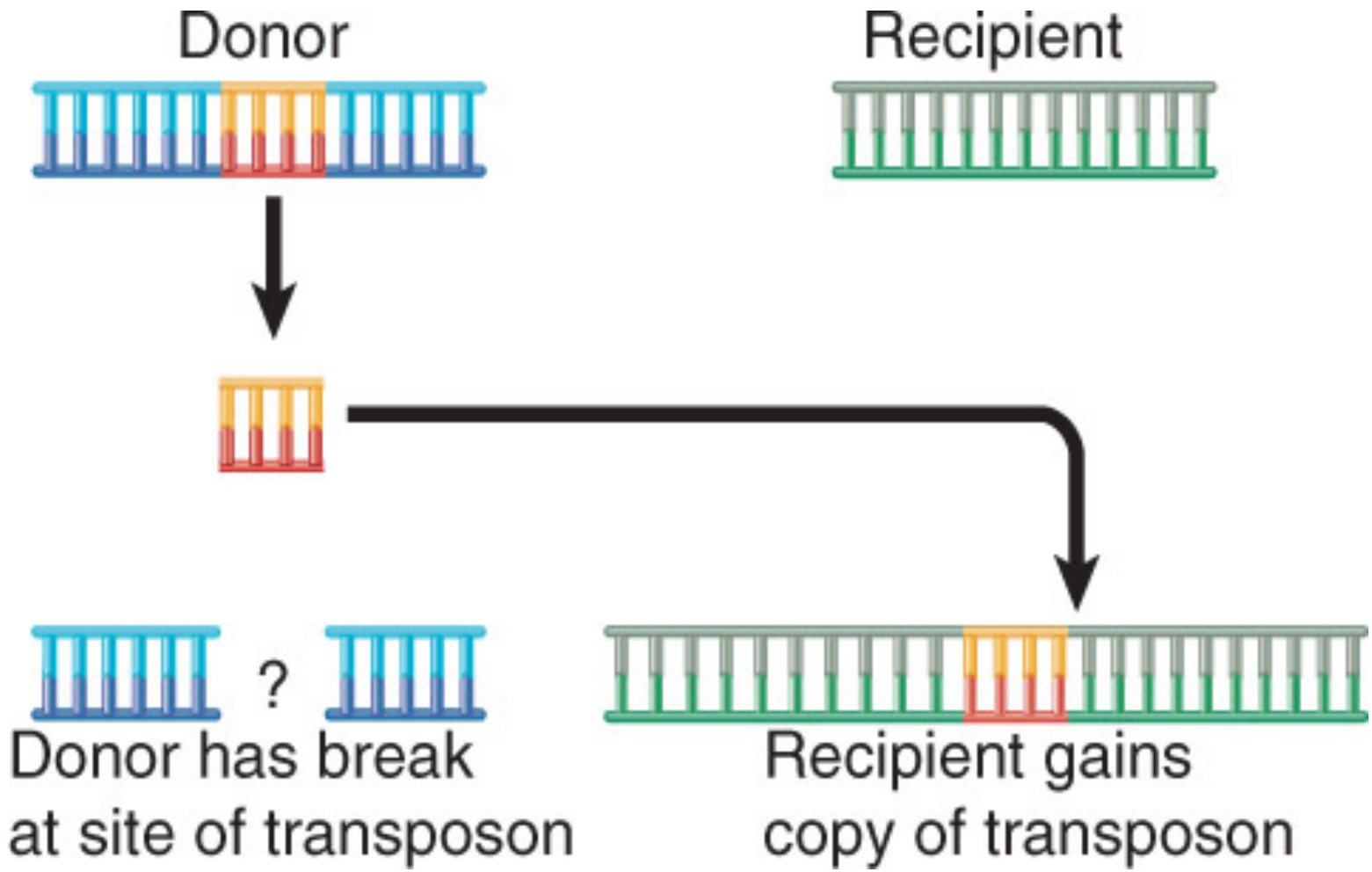


Table 1

Classification and characteristics of eukaryotic DNA transposons

Superfamily	Related IS	TSD	Length[1] (kb)	TIRs[1] (bp)	Terminal motif (5′-3′)	TPase[1] (aa)	Catalytic motif	DNA-binding motif	Additionnal proteins
Tc1/mariner	**IS630**	TA	1.2–5.0	17–1100	variable	300–550	DD(30–41)D/E	HTH (cro/paired)	
hAT	**nd**	8 bp	2.5–5	10–25	YARNG	600–850	D(68)D(324)E[2]	ZnF (BED)	
P element	**nd**	7/8 bp	3–11	13–150	CANRG	800–900	D(83)D(2)E(13)D[3]	ZnF (THAP)	
MuDR/Foldback	**IS256**	7–10 bp	1.3–7.4	0-sev. Kb	variable	450–850	DD(~110)E	ZnF (WRKY/GCM1)	
CACTA	**nd**	2/3 bp	4.5–15	10–54	CMCWR	500–1,200	Nd	nd	TNPA (DNA-binding protein)
PiggyBac	**IS1380**	TTAA	2.3–6.3	12–19	CCYT	550–700	DDE?	nd	
PIF/Harbinger	**IS5**	TWA	2.3–5.5	15–270	GC-rich	350–550	DD(35–37/47–48)E	HTH	PIF2p (Myb/SANT domain)
Merlin	**IS1016**	8/9 bp	1.4–3.5	21–462	GGNRM	270–330	DD(36–38)E	nd	
Transib	**nd**	5-bp	3–4	9–60	CACWATG	650–700	DD(206–214)E	nd	
Banshee	**IS481**	4/15 bp	3.5	41–950	TGT	300–400[4]	DD(34)E	HTH	
Helitron	**IS91**	none	5.5–17	none	5′-TC…CTAR-3′	1,400–3,000[5]	HHYY (‘REP motif’)	ZnF-like	RPA (in Plants)
Maverick	**none**	5/6 bp	15–25	150–700	simple repeat	350–450[4]	DD(33–35)E	ZnF (HHCC)	4–10 DNA virus-like proteins

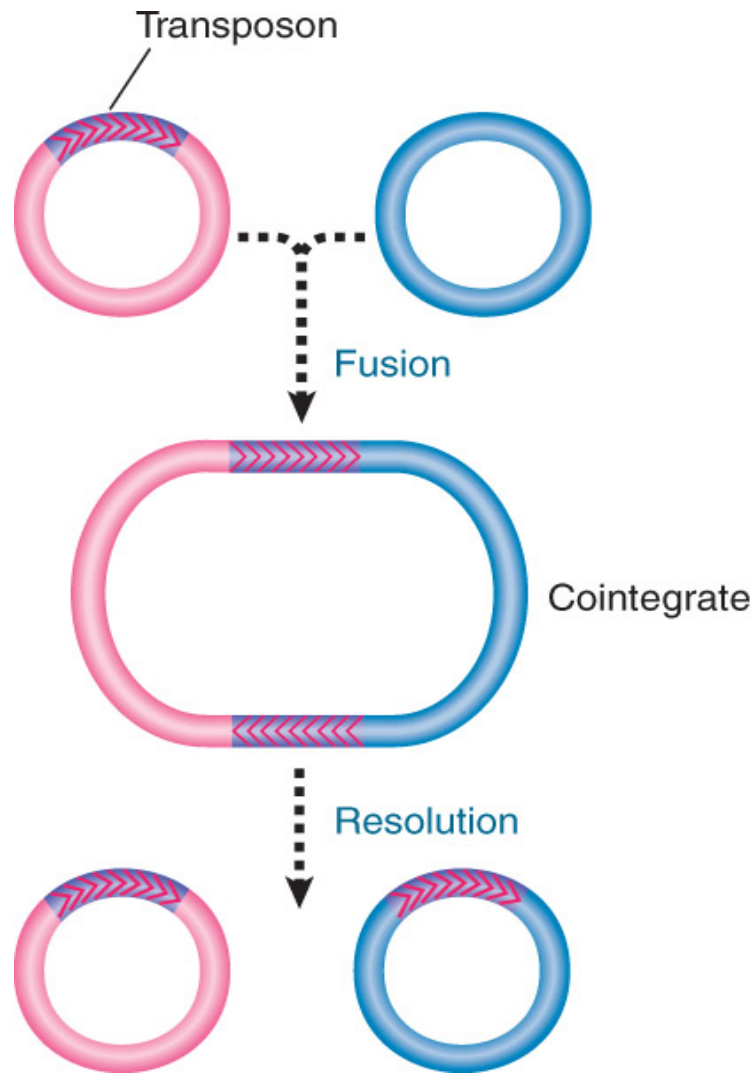
[1]refers to potentially complete, autonomous element
[2]motif in *Hermes* Tpase
[3]motif in Drosophila P element Tpase
[4]RVE integrase-like
[5]REP-Helicase

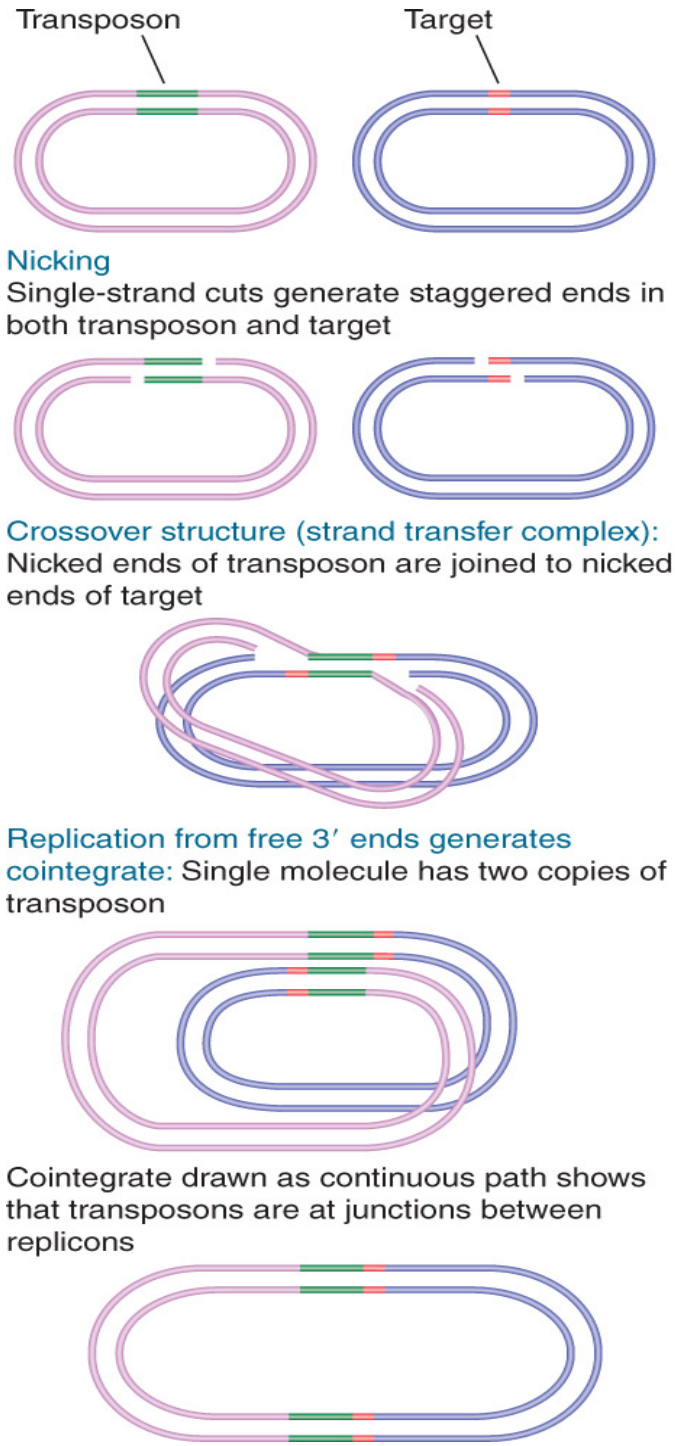
nd= not determined

Replicative transposition proceeds through a **cointegrate**

- Replication of a strand transfer complex generates a **cointegrate**, which is a fusion of the donor and target replicons.

- The **cointegrate** has two copies of the transposon, which lie between the original replicons.
- Recombination between the transposon copies regenerates the original replicons, but the recipient has gained a copy of the transposon.
- The recombination reaction is catalyzed by a **resolvase** encoded by the transposon.





- In **Replicative transposition**, Recombination between the transposon copies regenerates the original replicons, but the recipient has gained a copy of the transposon.

- The recombination reaction (**resolution**) is catalyzed by a resolvase, almost always encoded by the transposon.

- eg. **Tn3**

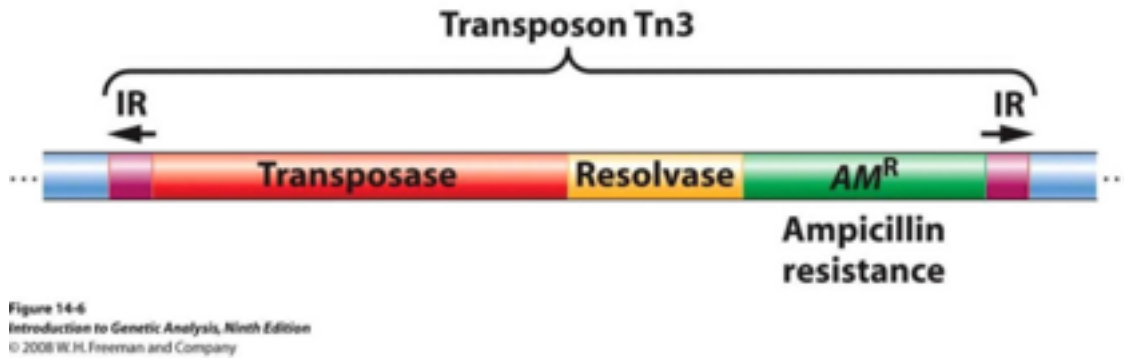
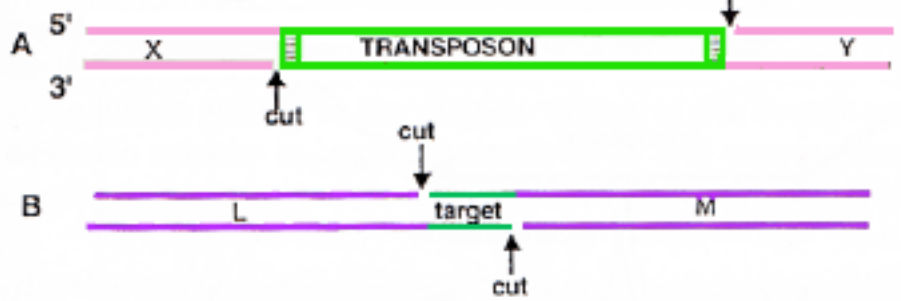
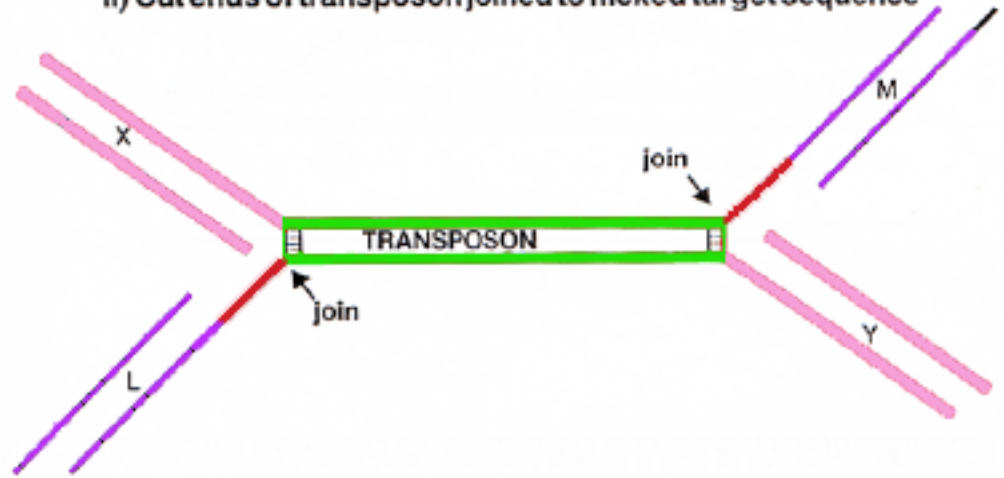


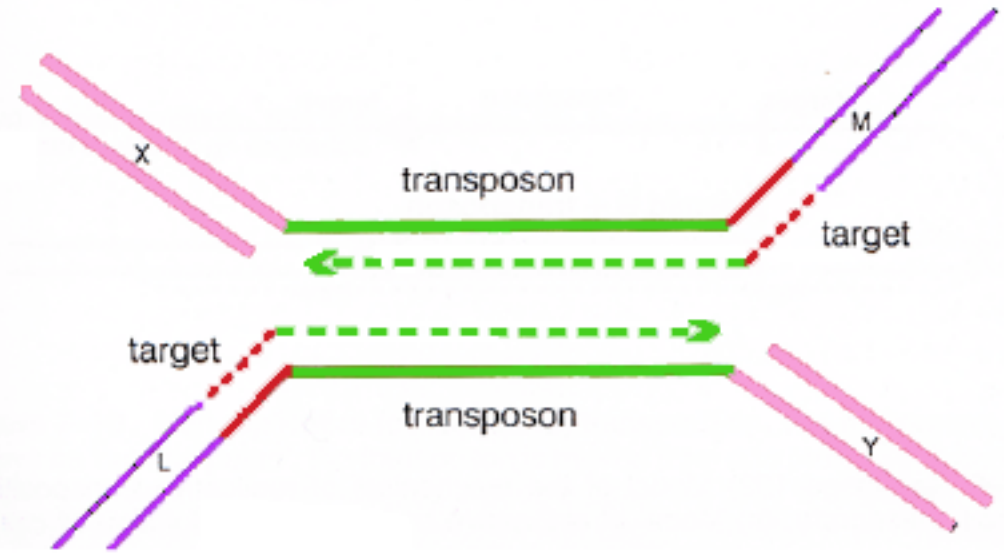
Figure 14-6
Introduction to Genetic Analysis, Ninth Edition

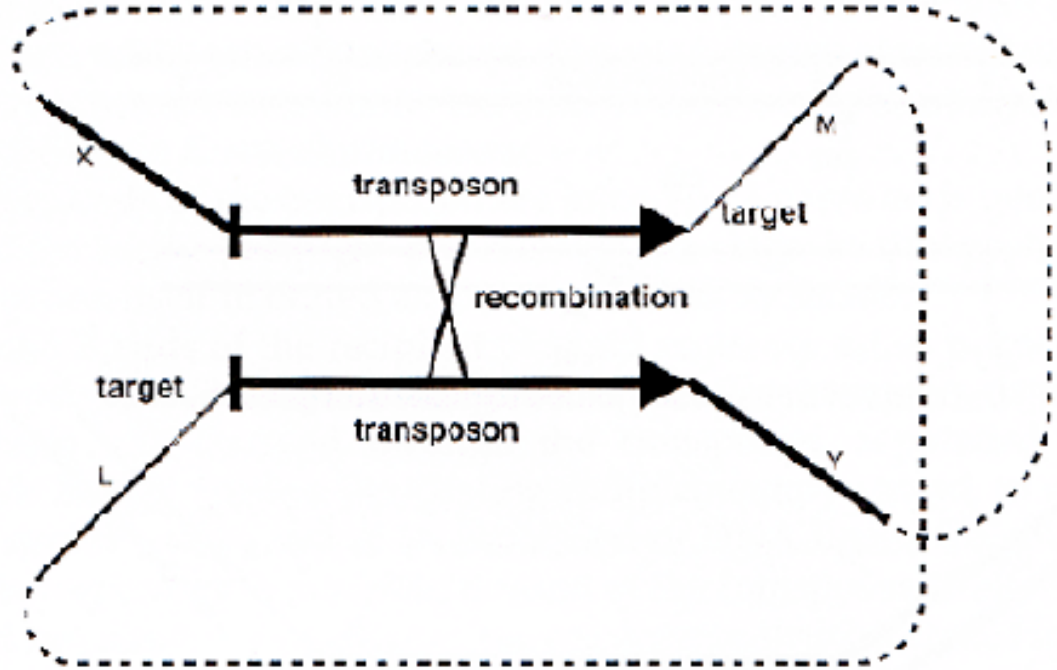


ii) Cut ends of transposon joined to nicked target sequence

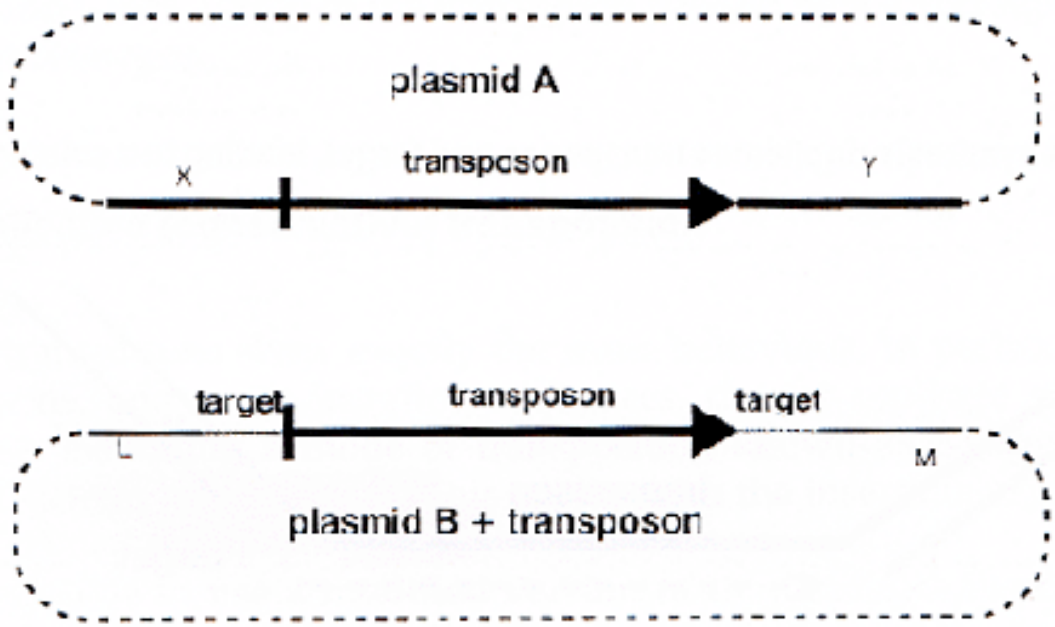


iii) DNA synthesis duplicates transposon and target, leading to a cointegrate





v) Resolution of cointegrate; both plasmids have a copy of the transposon



In **replicative transposition**, recombination between the two DNAs is catalyzed by a transposase (gene ***tnpA*** in **Tn3**) to form what's termed a **"cointegrate"**.

The term, **"illegitimate recombination"** is also often used because it requires little to no sequence similarity between the transposing element and the DNA sequence that is being modified. Resolution of the **cointegrate** occurs, when the cointegrate breaks down into **two independent** DNA sequences (each bearing one copy of the transposon). this resolution is catalyzed by a specific enzyme, a **"resolvase"** (encoded for by the *tnpR* gene in **Tn3**).

Such recombination utilizes two homologous sites on **Tn3** called "res" sites.

Curiously, **TnpR** protein is **pleiotropic** in that it also regulates its own expression, as well as that of *tnpA.* The **res sites** to which this resolvase binds lies within the promoter regions for both *tnpA* and *tnpR:* effectively killing two birds with one stone, as it were.

How might you be able to tell the difference in prokaryotes?

In **replicative transposition**, recombination between the two DNAs is catalyzed by a transposase (gene ***tnpA*** in **Tn3**) to form what's termed a **"cointegrate"**.

The term, **"illegitimate recombination"** is also often used because it requires little to no sequence similarity between the transposing element and the DNA sequence that is being modified. Resolution of the **cointegrate** occurs, when the cointegrate breaks down into **two independent** DNA sequences (each bearing one copy of the transposon). this resolution is catalyzed by a specific enzyme, a **"resolvase"** (encoded for by the *tnpR* gene in **Tn3**).

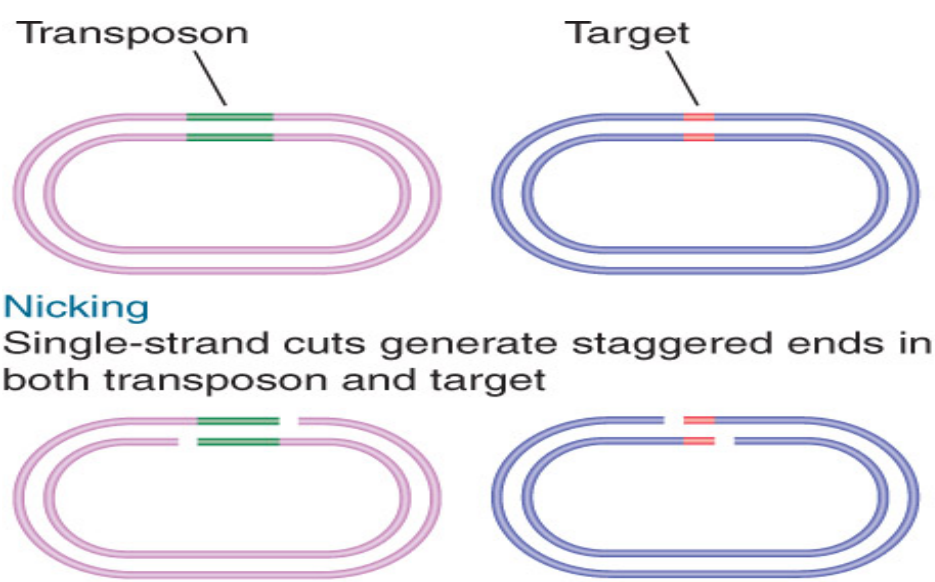
Such recombination utilizes two homologous sites on **Tn3** called "res" sites.

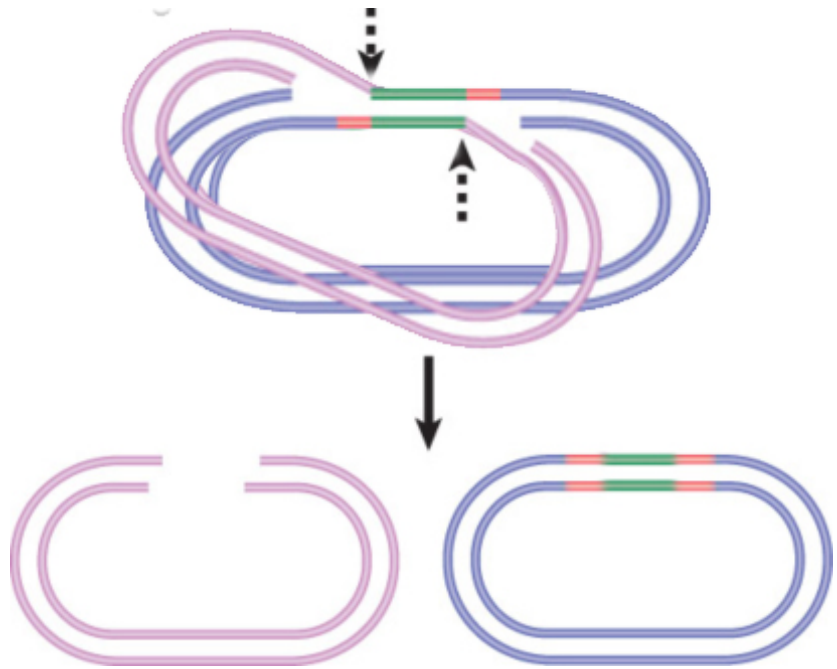
Curiously, **TnpR** protein is **pleiotropic** in that it also regulates its own expression, as well as that of *tnpA.* The **res sites** to which this resolvase binds lies within the promoter regions for both *tnpA* and *tnpR:* effectively killing two birds with one stone, as it were.

How might you be able to tell the difference in prokaryotes?

Use mutants of **tnpR**, which form the **cointegrate,** and would give the cointegrate as a final product -as it cannot be resolved - except perhaps by the generic **RuvAB and C** systems, but at a much lower frequency.

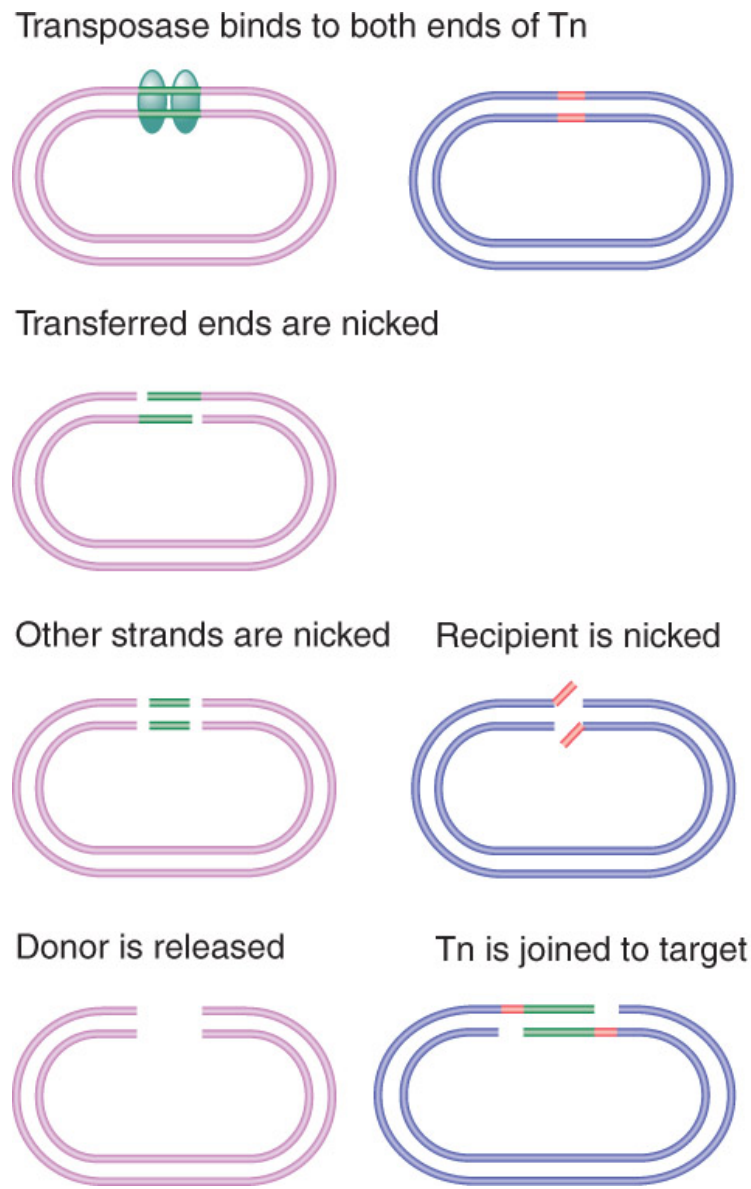
- **Nonreplicative** transposition results if a crossover structure is nicked on the unbroken pair of donor strands and the target strands on either side of the transposon are ligated.

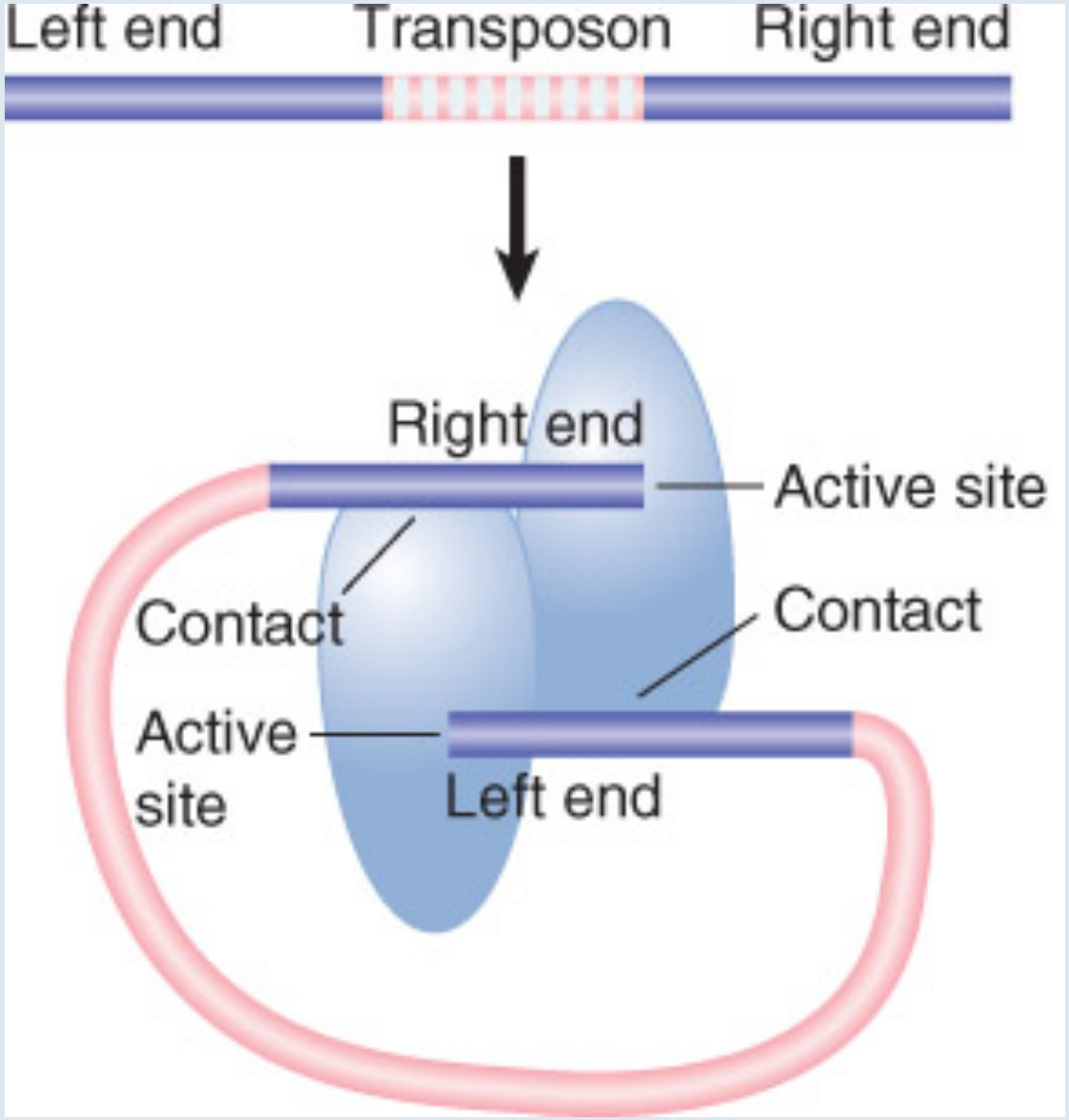




- There is an alternative pathway for nonreplicative transposition, which differs from the previous mechanism according to whether the first **pair** of transposon strands are joined to the target **before** the second pair are cut (Tn5),

- or whether all **four strands** are cut before joining to the target (Tn10).

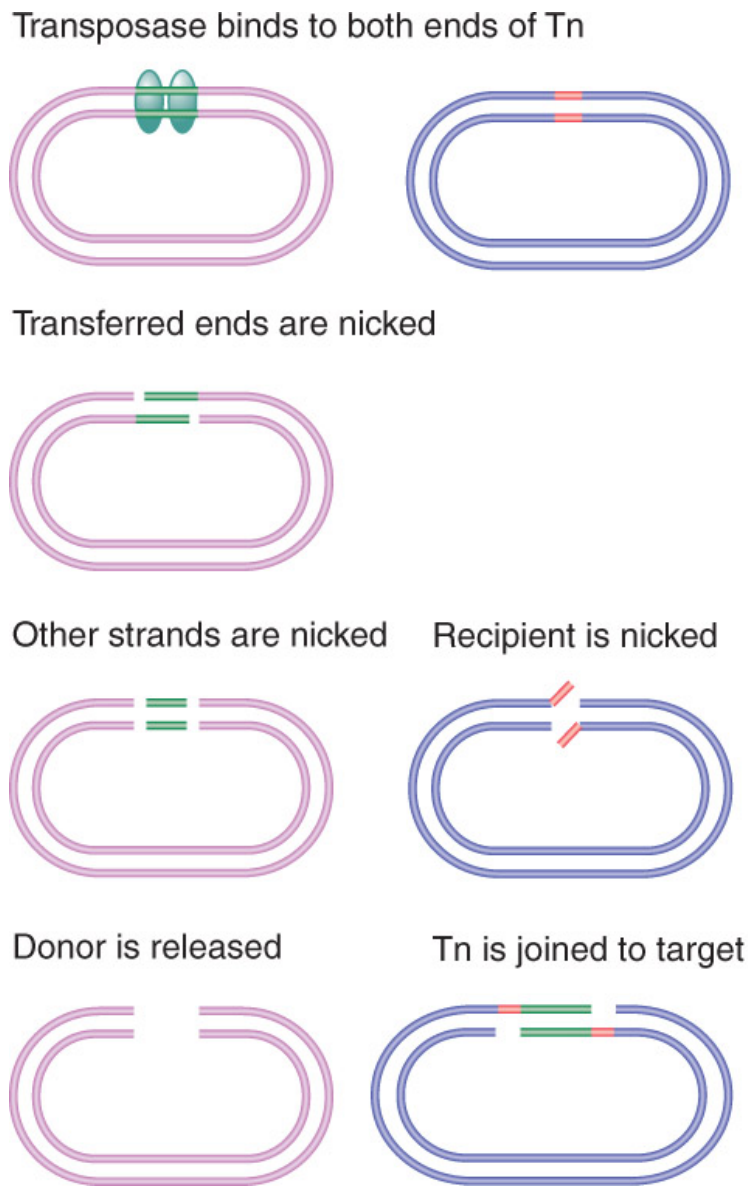


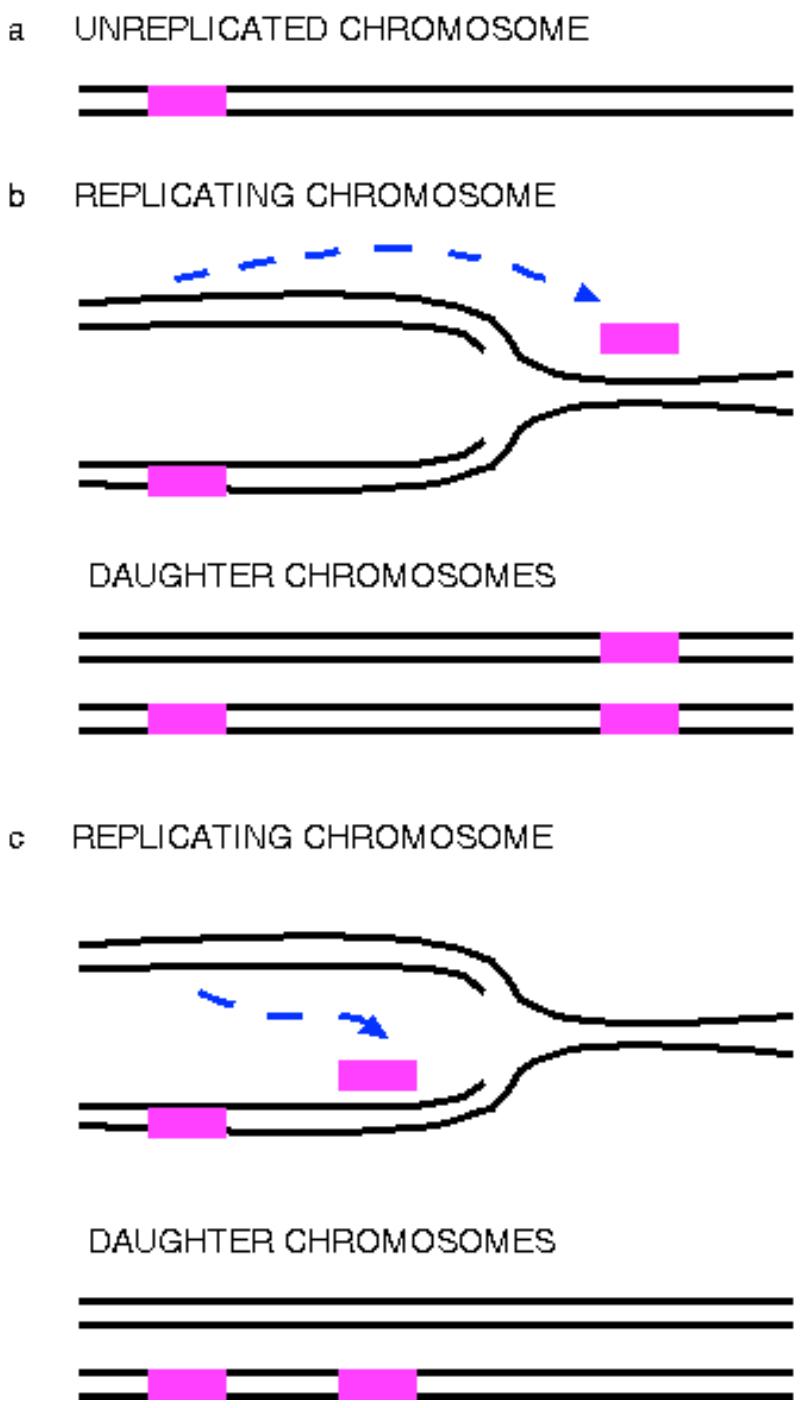

Left end
Transposon
Right end
Right end
Active site
Contact
Contact
Active
site
Left end

- There are two pathways for non-replicative transposition, which differ according to whether the first **pair** of transposon strands are joined to the target **before** the second pair are cut (Tn5),

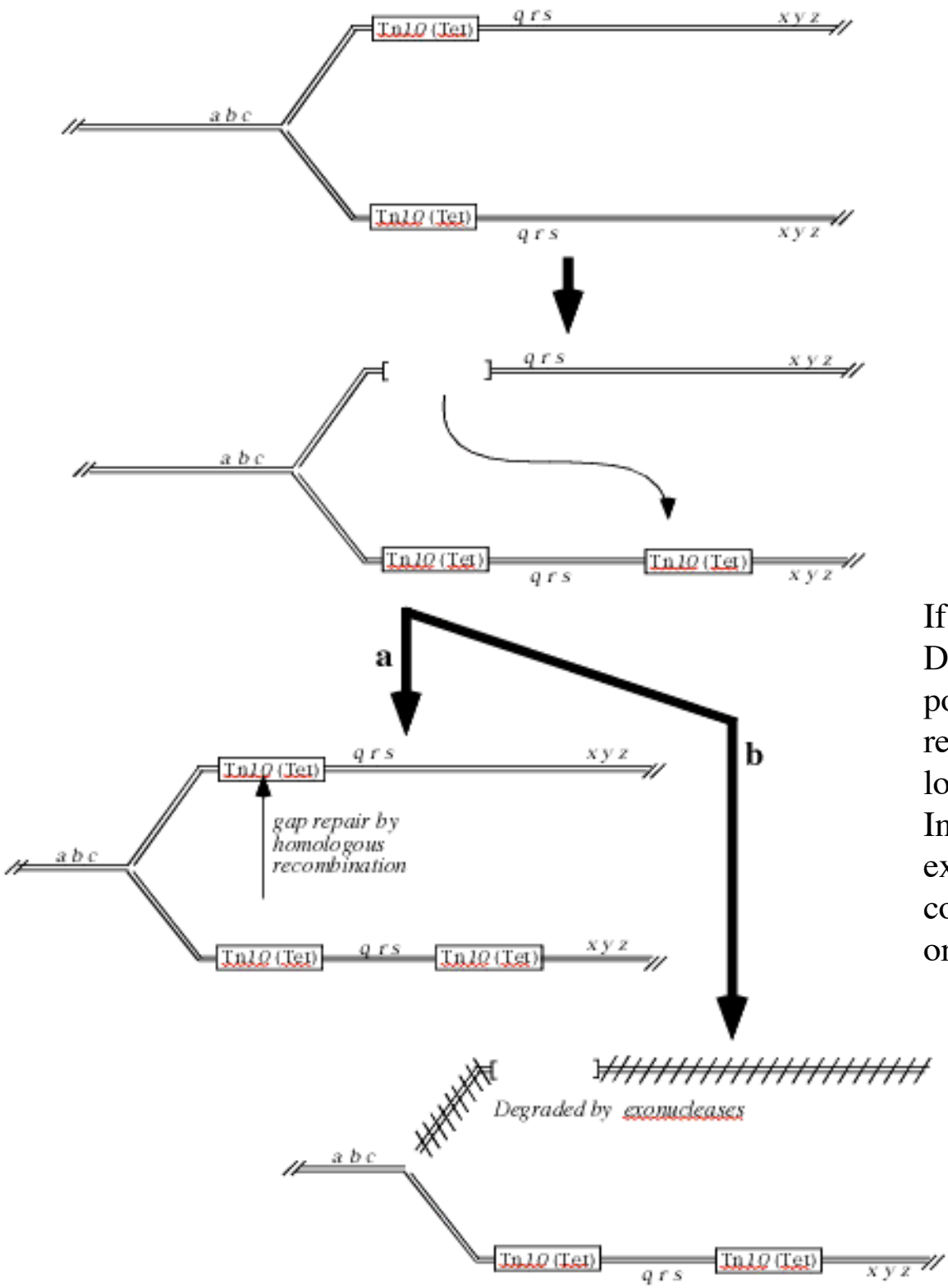
- or whether all **four strands** are cut before joining to the target (Tn10).

OK, But… How would you know?
Moreover, how might you discriminate between **replicative** and **non replicative** transposition

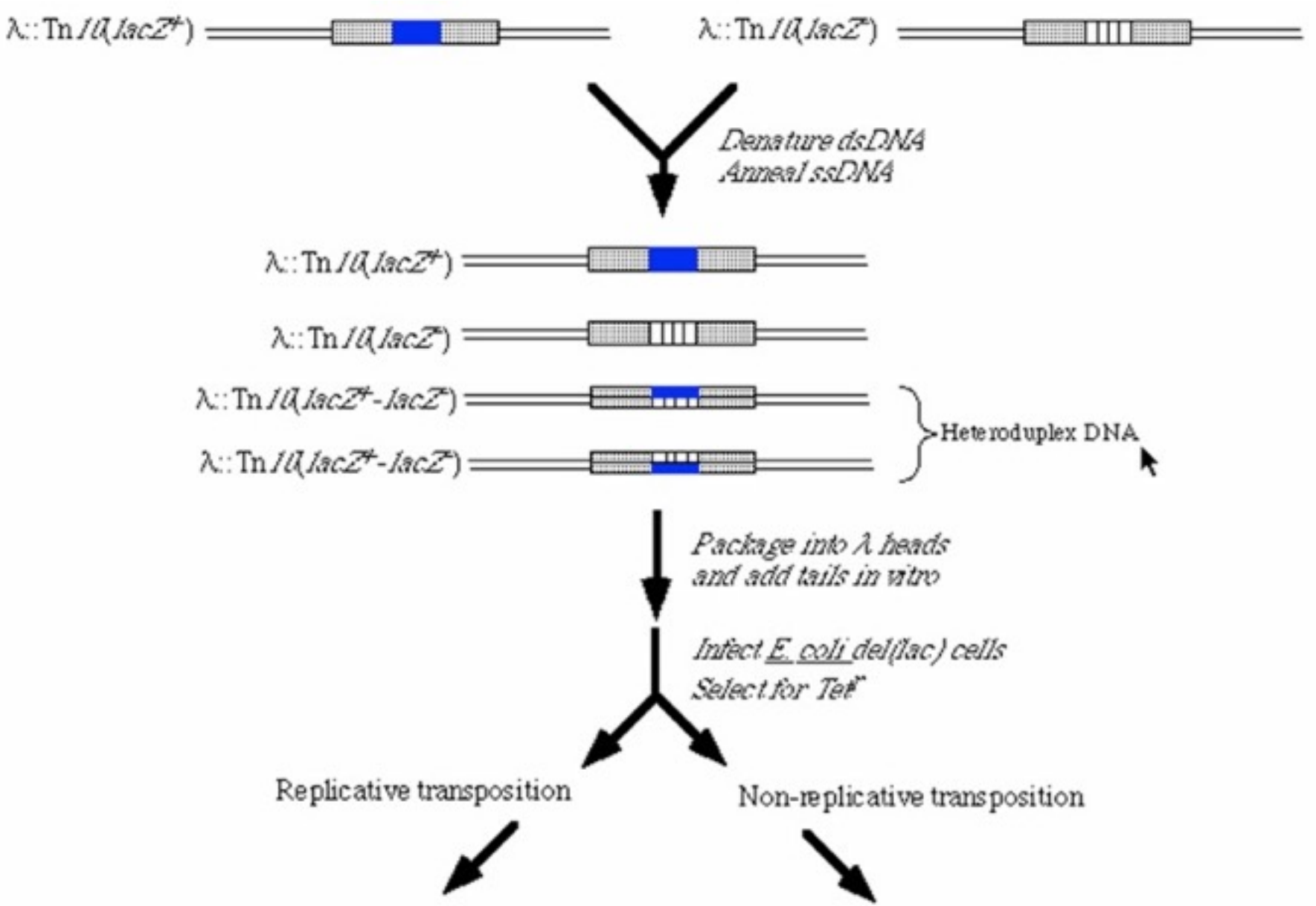




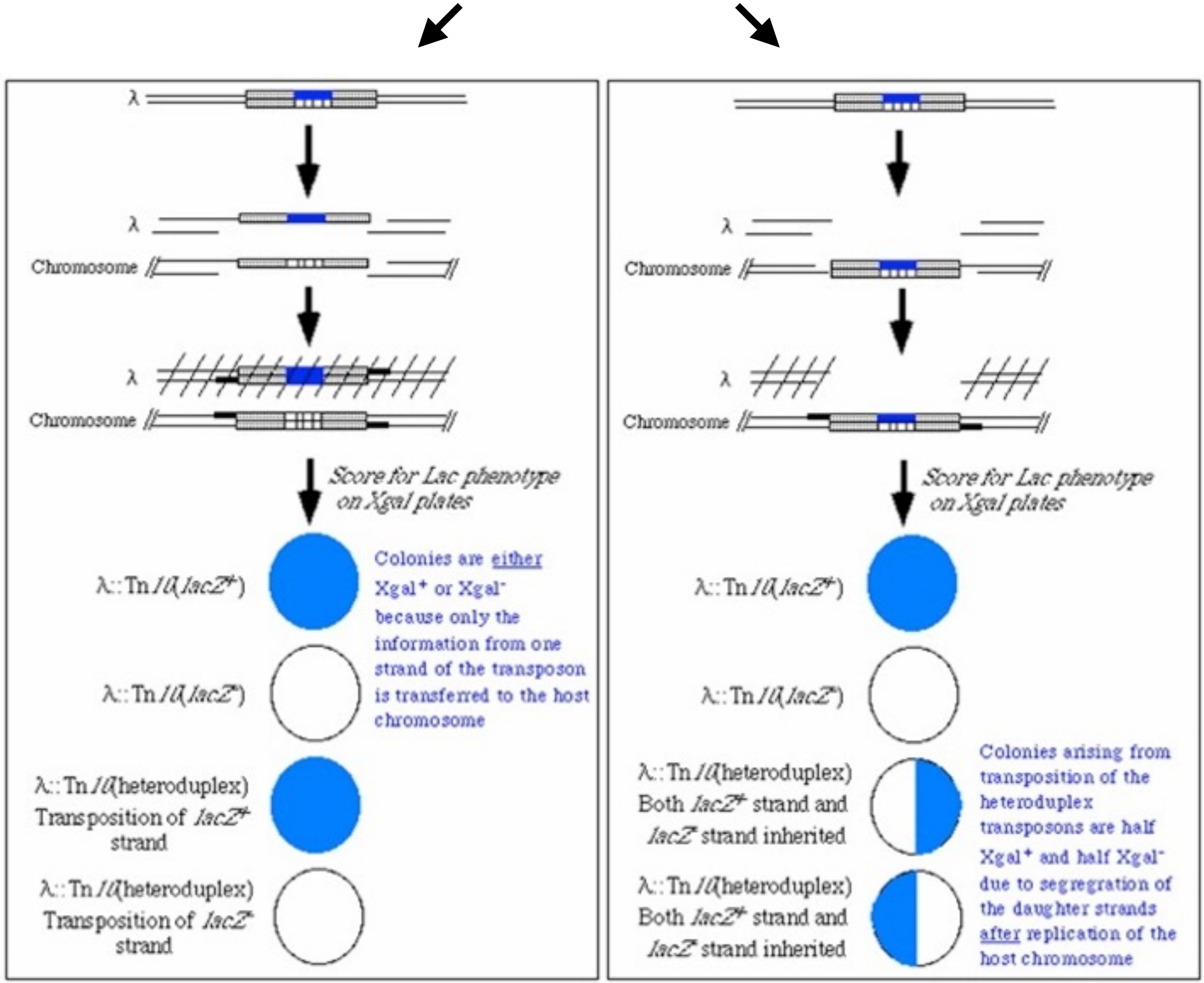
The timing of non-replicative transposition during DNA replication (for example) can affect the number and location of transposon



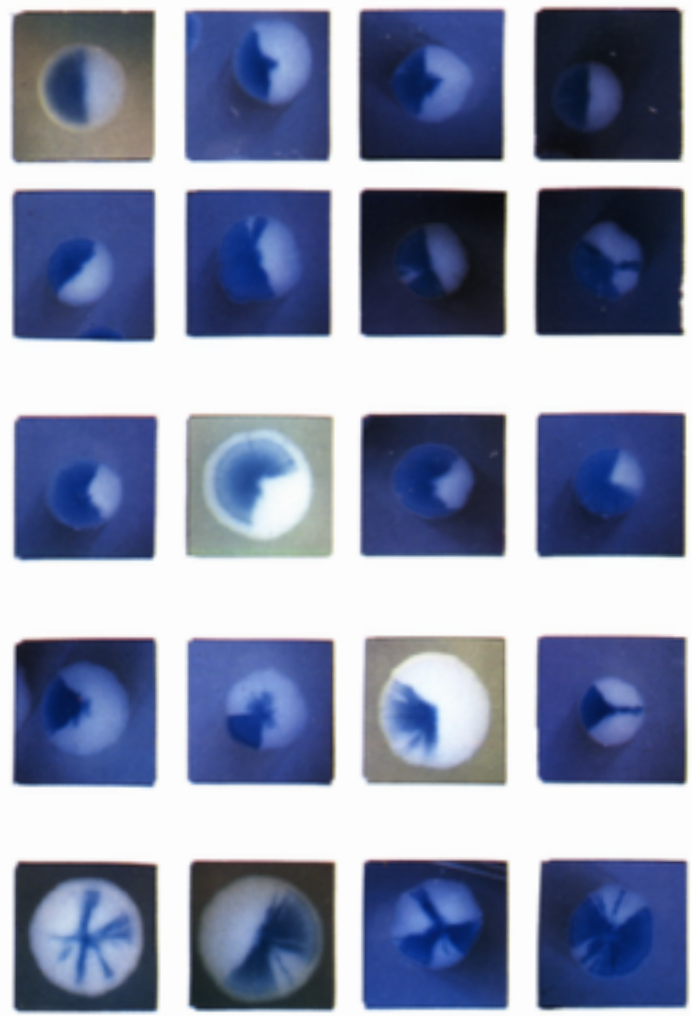
If transposition occurs after replication of the donor DNA, which is often the case, it is not always possible to determine if a transposon moves by a replicative or **non-replicative** mechanism simply by looking for loss of the transposon at the original site. In the non-replicative transposition of Tn10 (for example), because of DNA repair or recombination a copy of the transposon can still be observed in the original site.

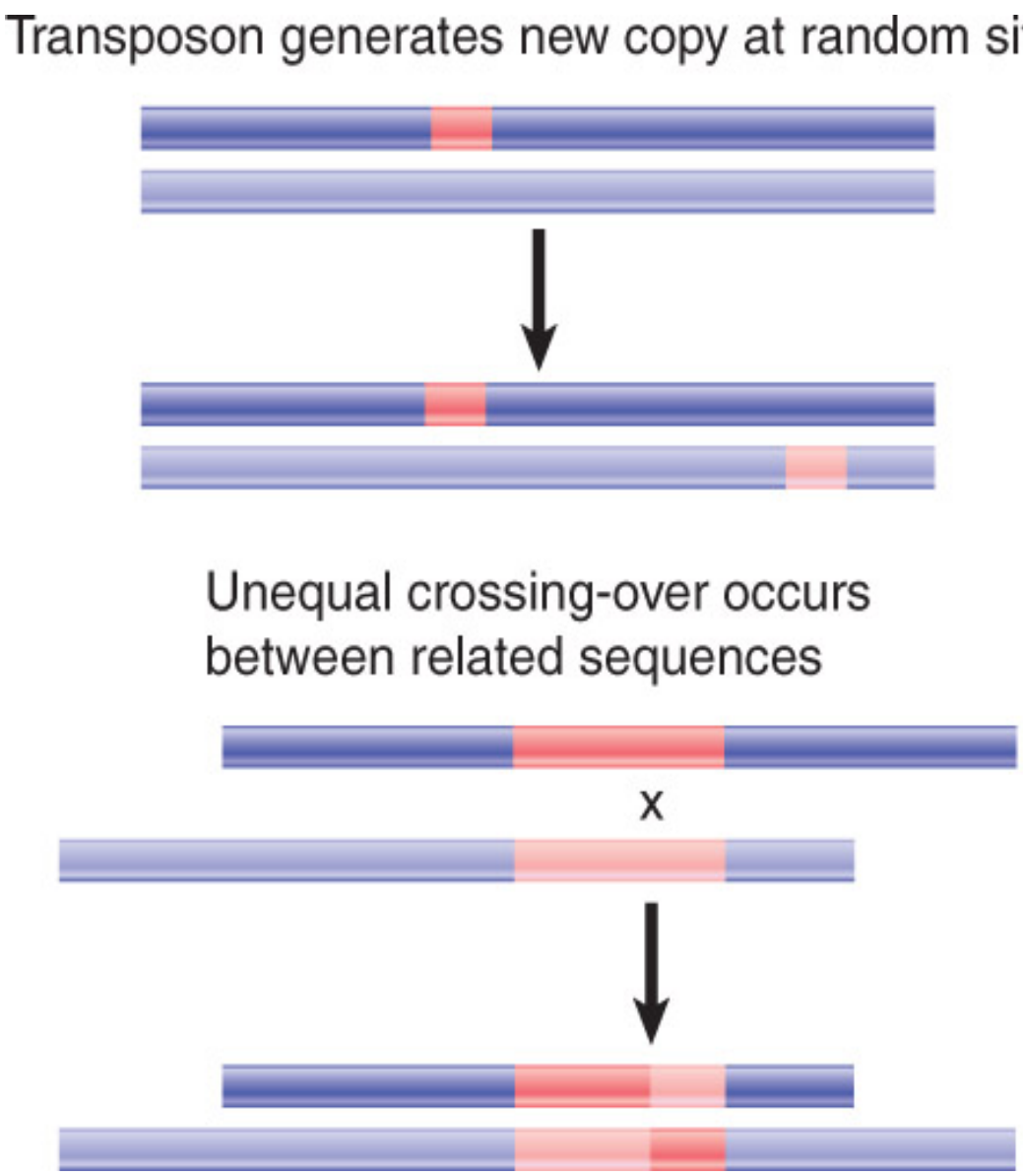


Tn10 (Bender and Kleckner, 1986 Cell 45, 801-815) vs. **Tn*3*** (Heffron, 1983)


λ
Chromosome
λ
Chromosome
Score for Lac phenotype on Xgal plates
λ::Tn10(lacZ+)
λ::Tn10(lacZ-)
Colonies are either Xgal+ or Xgal- because only the information from one strand of the transposon is transferred to the host chromosome
λ::Tn10(heteroduplex) Transposition of lacZ+ strand
λ::Tn10(heteroduplex) Transposition of lacZ- strand
λ
Chromosome
λ
Chromosome
Score for Lac phenotype on Xgal plates
λ::Tn10(lacZ+)
λ::Tn10(lacZ-)
λ::Tn10(heteroduplex) Both lacZ+ strand and lacZ- strand inherited
λ::Tn10(heteroduplex) Both lacZ+ strand and lacZ- strand inherited
Colonies arising from transposition of the heteroduplex transposons are half Xgal+ and half Xgal- due to segregration of the daughter strands after replication of the host chromosome

Integrations into Prophage





- Transposons move directly from one site in the genome to another.
- Transposons fall into two general classes: those that are mobilized directly through DNA and those that are mobilized in an RNA dependent manner.
- Transposable elements can promote rearrangements of the genome directly or indirectly
  - By causing deletions or inversions or leading to the movement of host sequences to a new location
  - By serving as substrates for cellular recombination systems

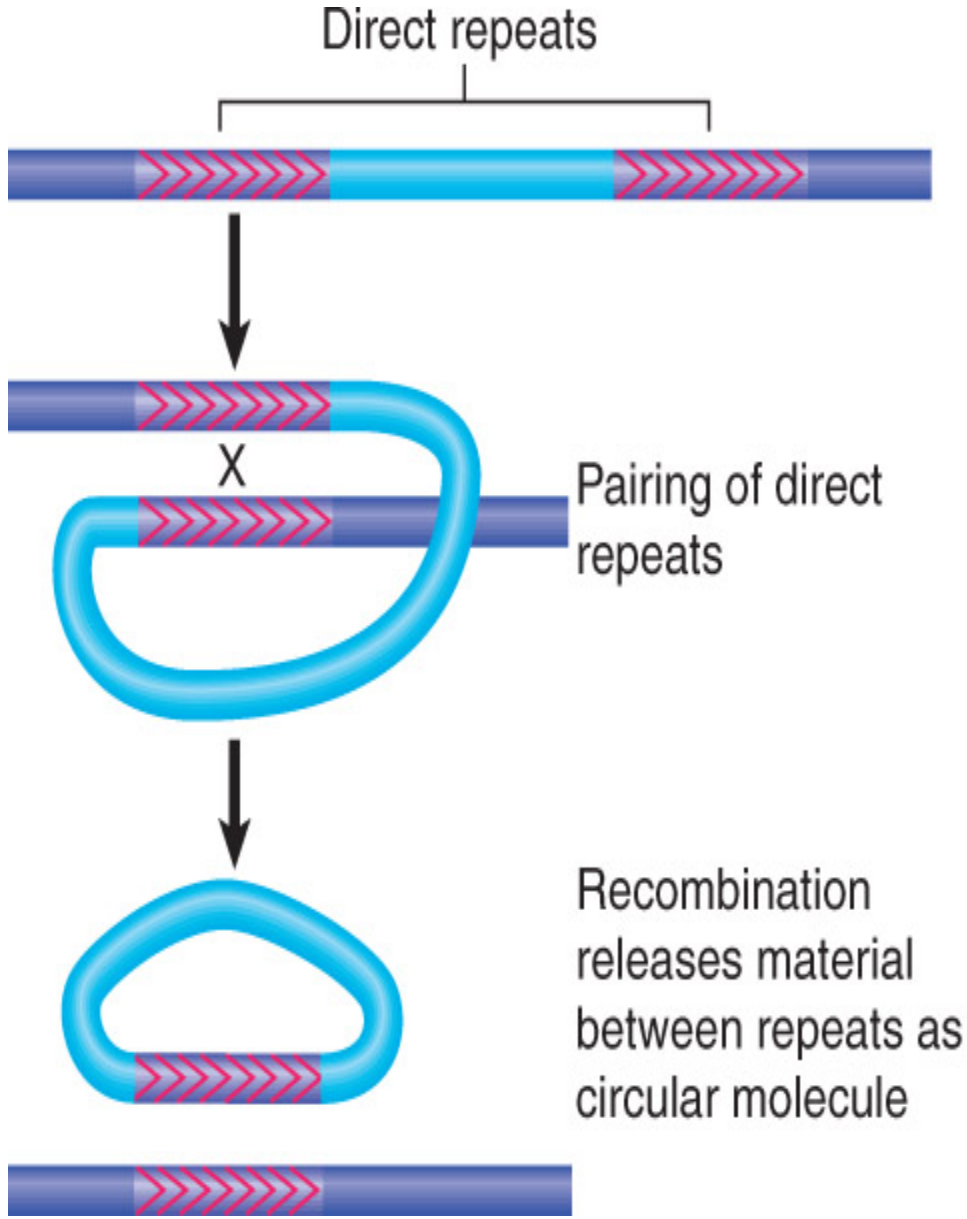
Transposons Cause Rearrangement of DNA

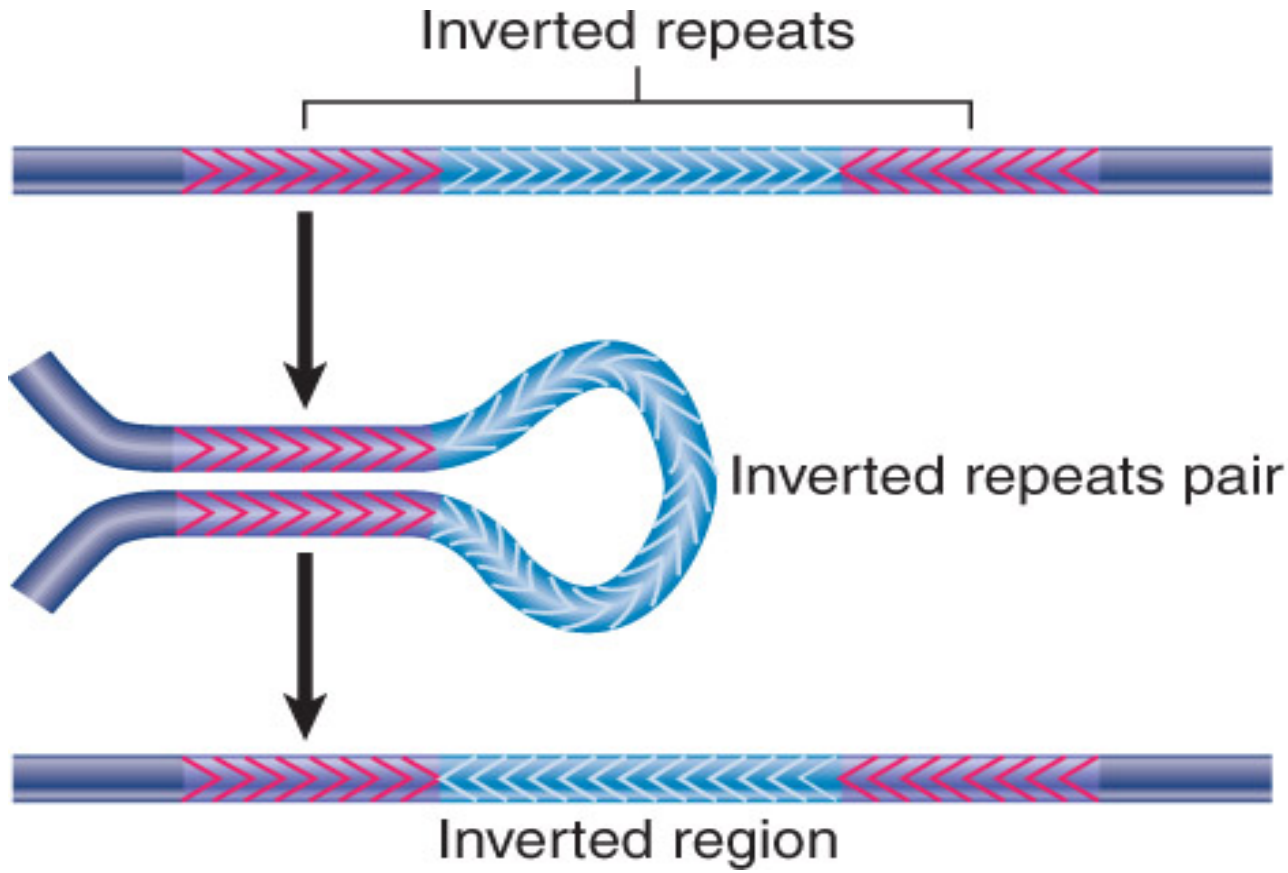
Homologous recombination between multiple copies of a transposon can cause subsequent rearrangements of host DNA.

Homologous recombination between the repeats of a transposon may lead to **precise** or **imprecise excision**.

Generally, recombination between any pair of direct repeats will **delete** the material between them.

The deletion of sequences adjacent to a transposon could therefore result from a two-stage process – transposition followed by recombination.





Reciprocal recombination between inverted repeats inverts the region between them.

- Maize provided one of the earliest models of transposition in eukayotes, which was discovered because of the effects of chromosome breaks that were generated by transposition of "**controlling elements**". The breaks generate on one chromosome that has a centromere, a broken end, and one **acentric fragment**.

- In maize, transposable elements often insert near genes that have visible but non-lethal effects on the phenotype.

- Maize displays clonal development which means that the occurrence and timing of a transposition event can be visualized as the resulting progeny following a "break" give rise to a **sector** – or patch of cells made up of a single altered cell and its progeny.

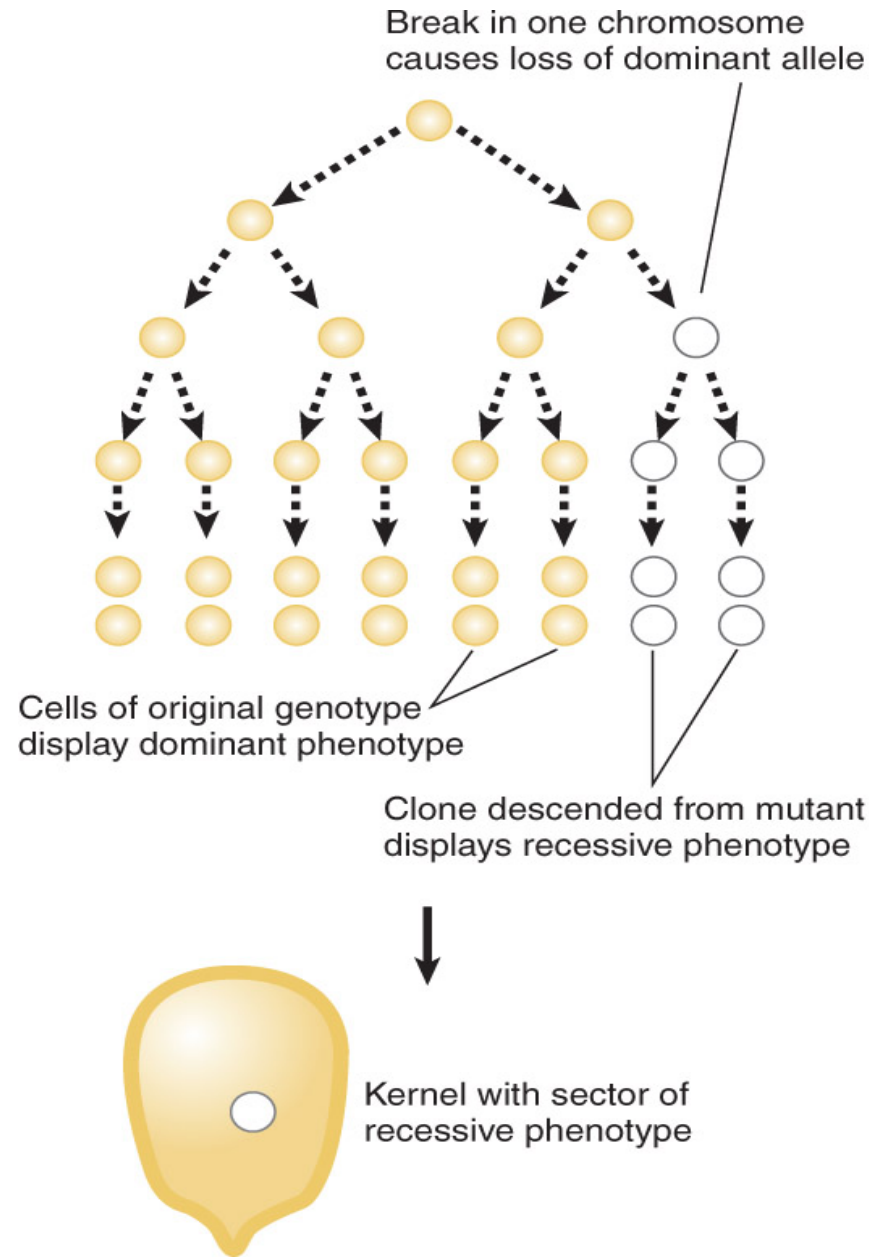
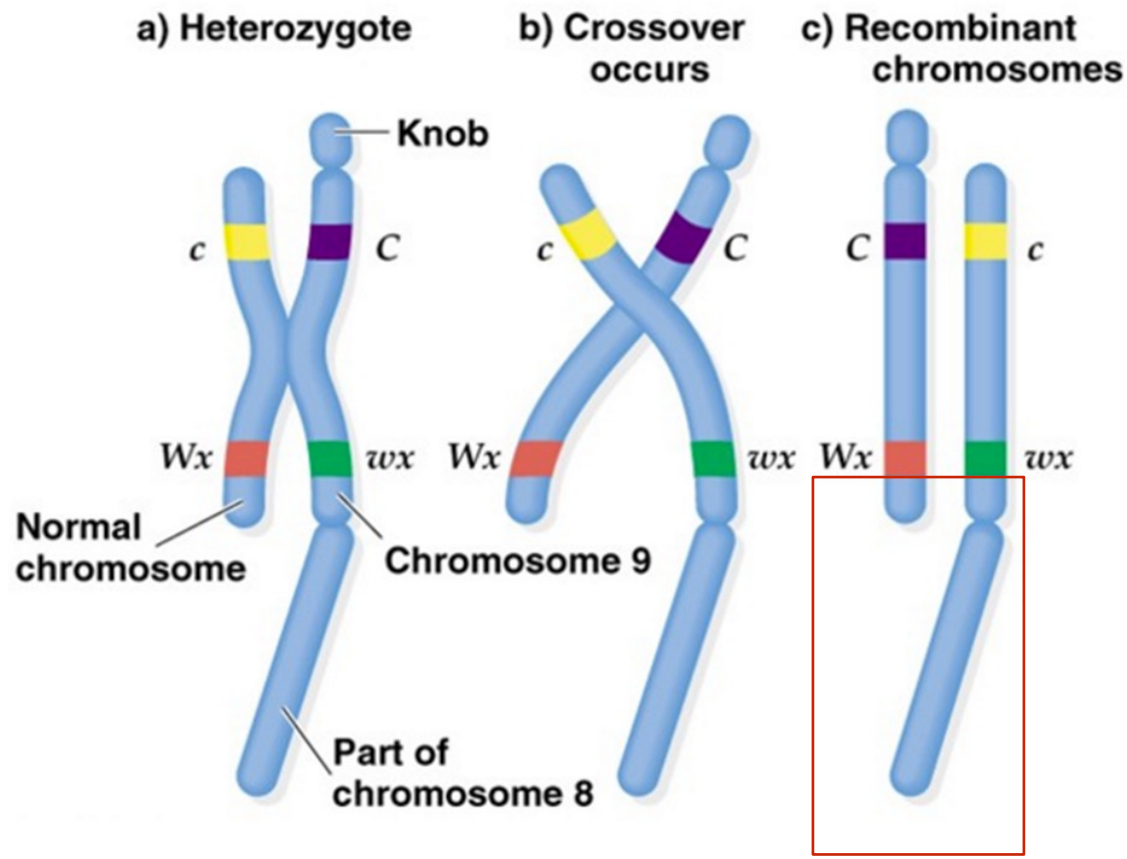
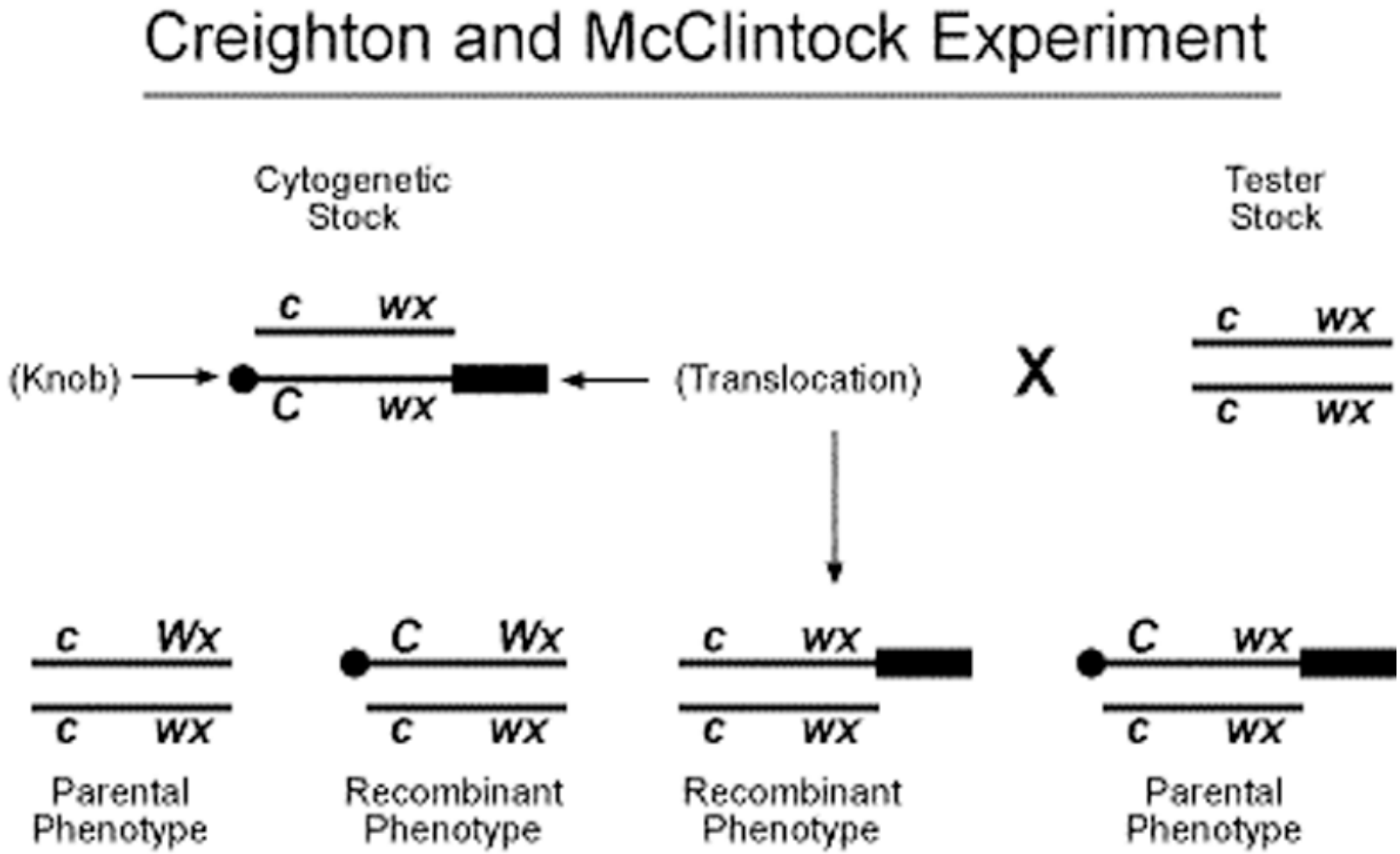


Figure 1: McClintock and her favorite research subject at Cold Spring Harbor.
image © Courtesy of the Barbara McClintock Papers, American Philosophical Society.

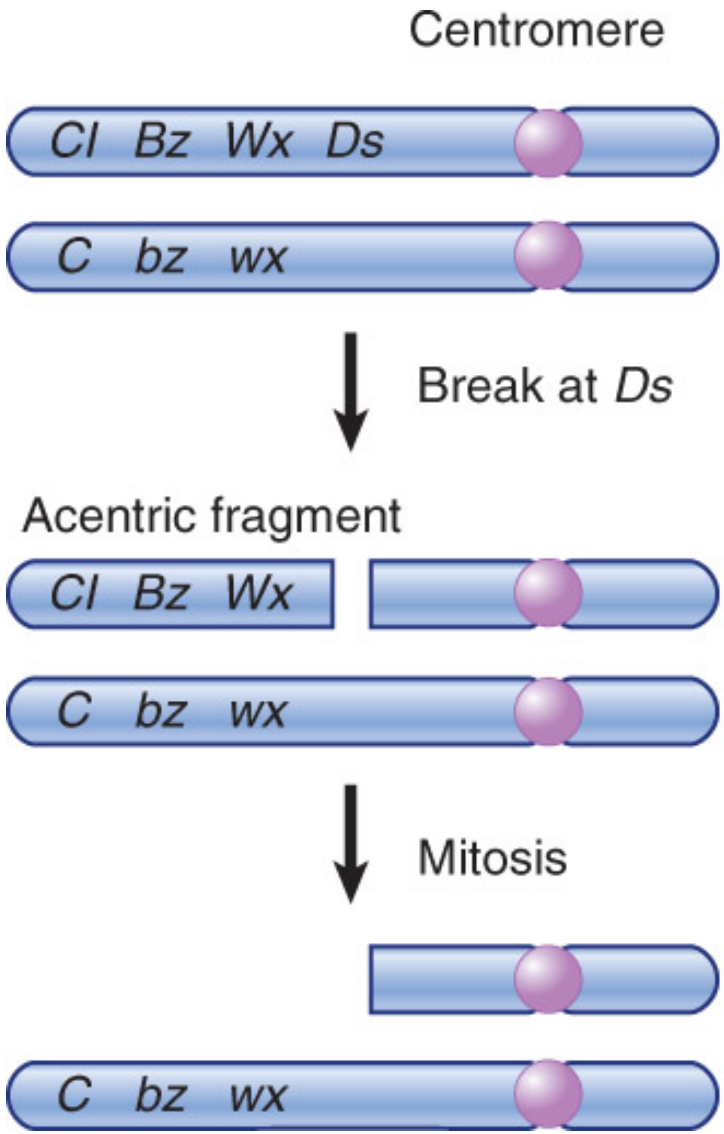


Creighton, H., and McClintock, B. 1931
A correlation of cytological and genetical crossing-over in Zea mays.
PNAS 17:492–497.

Figure 1: McClintock and her favorite research subject at Cold Spring Harbor.

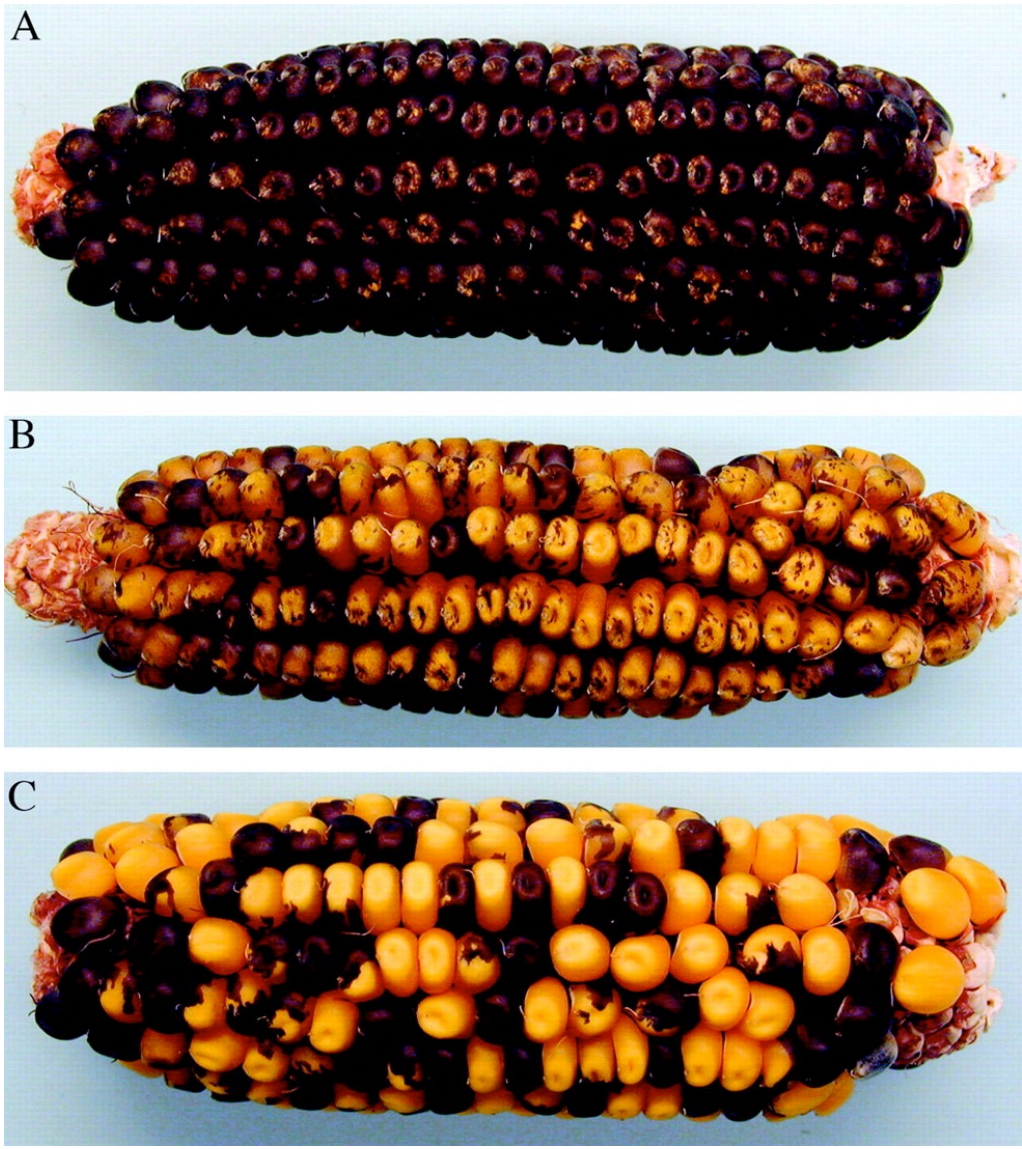


Creighton, H., and McClintock, B. 1931
A correlation of cytological and genetical crossing-over in Zea mays.
PNAS 17:492–497.



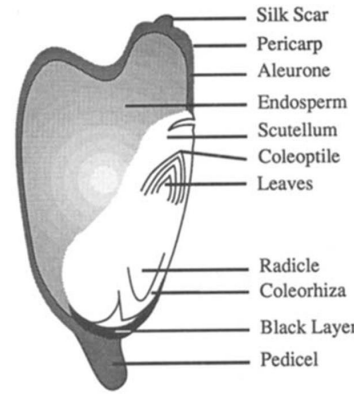
A break at a “controlling element”often caused the loss of an **acentric** fragment on this chromosome; if the fragment carried the dominant markers of a heterozygote, its loss would change the phenotype.

Ac-im contributes to negative dosage.



Ac-Immobilized, a Stable Source of *Activator* Transposase That Mediates Sporophytic and Gametophytic Excision of *Dissociation* Elements in Maize

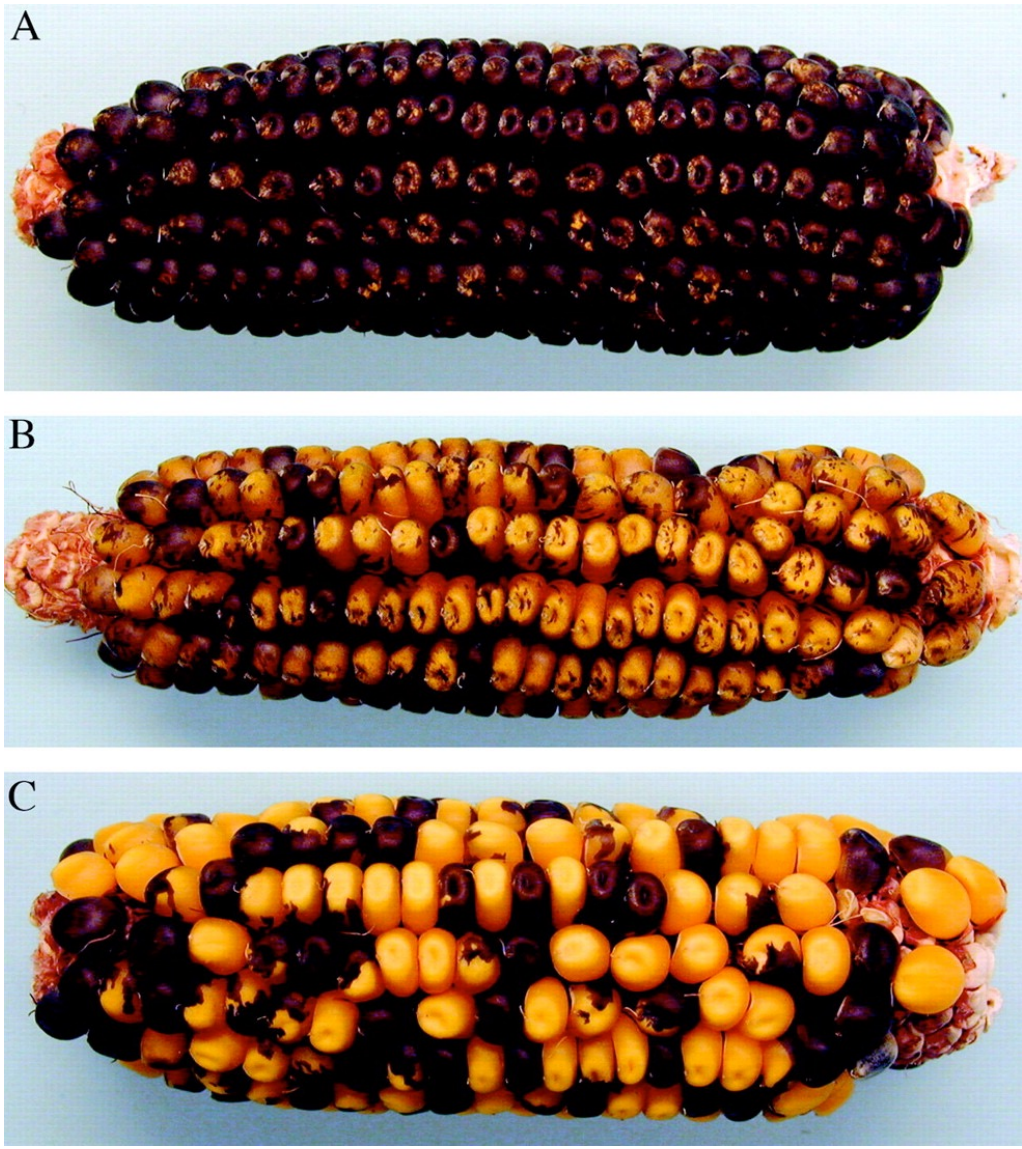
FIGURE 1.—
Ac-im contributes to negative dosage. Aleurone variegation patterns of ears carrying (A) one, (B) two, or (C) three copies of ***Ac-im*** in the triploid endosperm. Kernels with fully colored aleurone (clearly visible in B and C) carry revertant *R1-sc* alleles resulting from excisions of the *Ds* at *r1.*

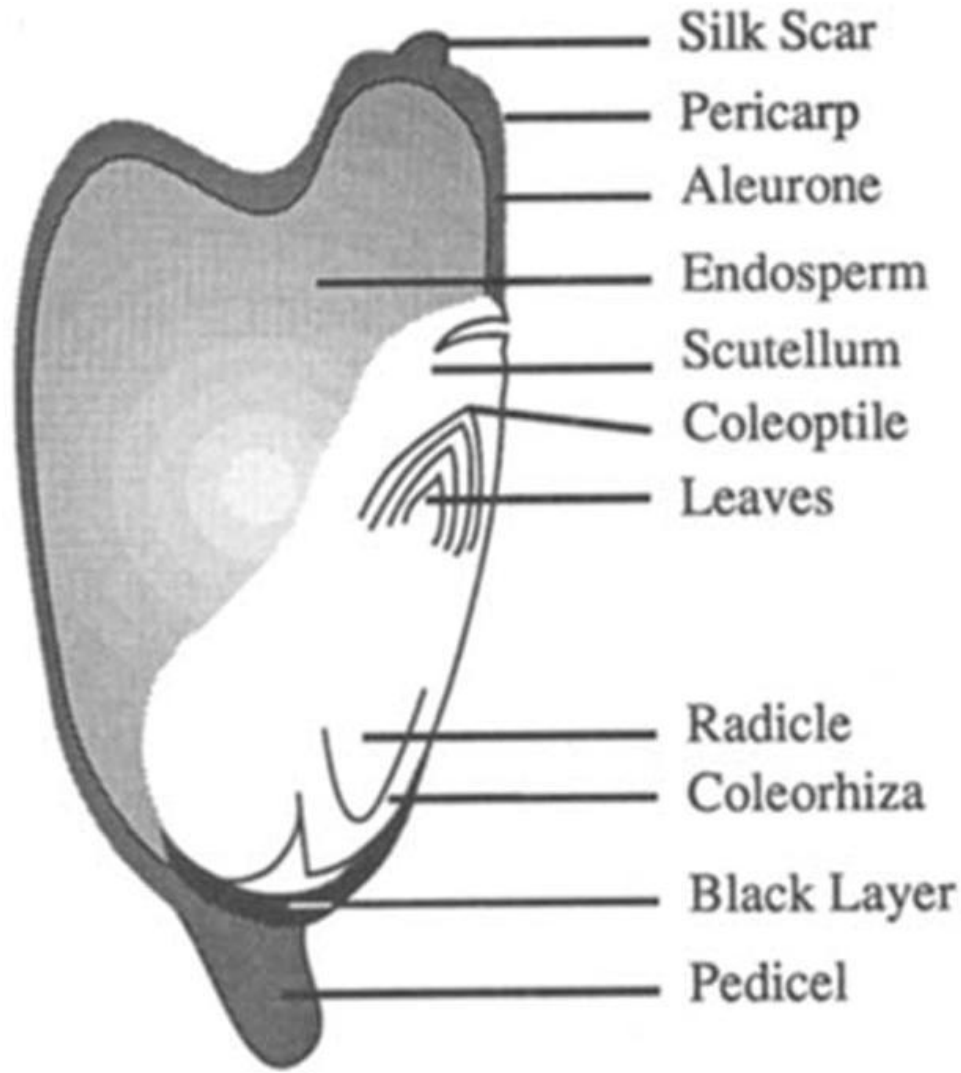


Conrad L J , Brutnell T P Genetics 2005;171:1999-2012

GENETICS

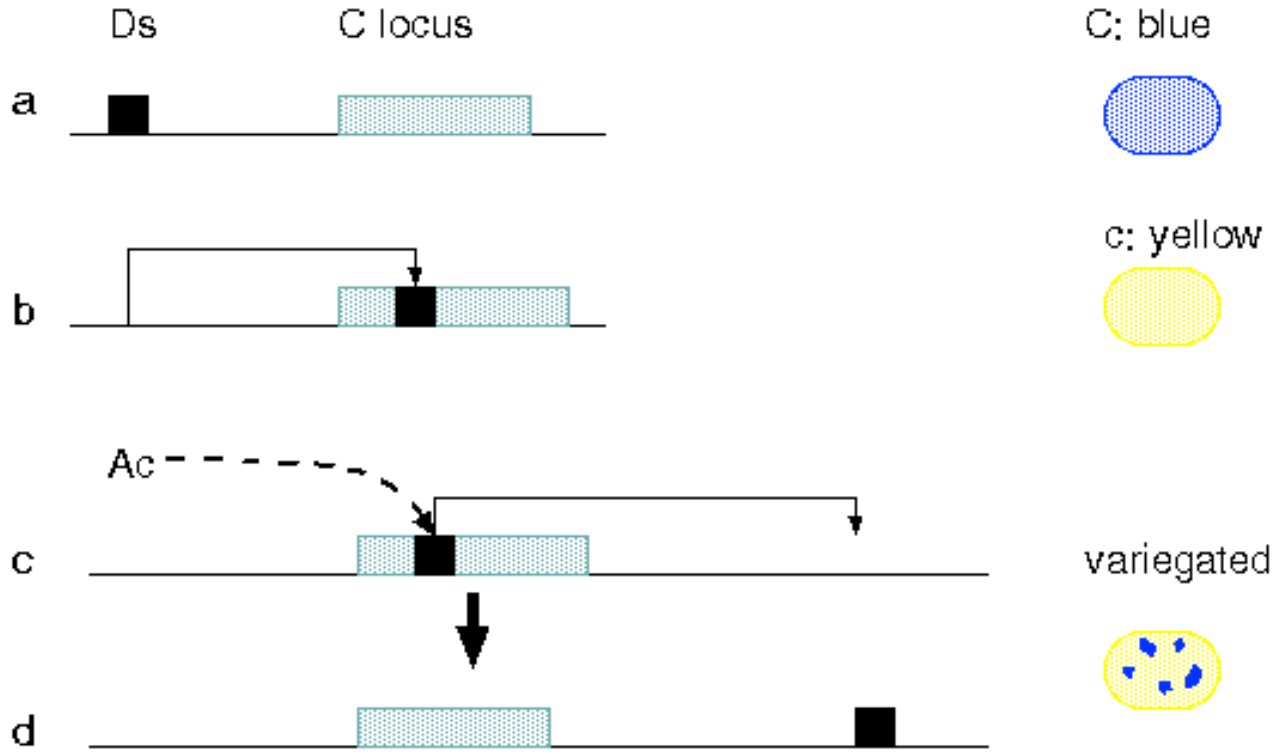
Ac-im contributes to negative dosage.





Conrad L J , Brutnell T P Genetics 2005;171:1999-2012

GENETICS



The C locus produces an **anthocyanin pigment**, resulting in a coloured **aleurone** ie. **wild-type.**

Insertion of a transposon into C blocks pigment production, resulting in the **colorless (c) phenotype.**

During seed development, excision of the Ds element can result in patches of cells (each progeny of a single cell) expressing a colored phenotype. This is referred to as **variegation.**

If the **Ac locus** moves to a position adjacent to **Ds**, where it is now able to promotes the movement of **Ds** away from the **C locus** during the development of the kernel -and the locus reverts to normal expression.

The larger the number of **Ac factors** present, the greater expression of variegation in the tissue.

Larger coloured patches are formed when **Ds** is expressed **early** in development;
smaller patches when **Ds** is expressed **late**.

Somatic variegation:

Kernel phenotypes of parental Rsc-24, the colorless r-g allele and Ds insertional mutants in the presence of Ac. Ac-st is a change of state derivative of Ac, which produces a much more sparsely variegated phenotype for Ds mutable alleles than does the standard Ac element

Genetics 135: 189-203 (September, 1993)

Transposons move from one place to another within the genetic material of Indian corn. When the *transposon* moves into the pigment gene, it stops pigments production and the Indian corn kernels appears white. When the *transposon* moves out, pigment is produced again. So, this variable in and out movement of *transposons* determines the colors of the Indian corn mosaic.

Genetics. 2012 Jul; 191(3): 747–756. PMCID: PMC3389971

doi: 10.1534/genetics.112.139642

A Hyperactive Transposase of the Maize Transposable Element *Activator* (*Ac*)

Katina Lazarow,[*] My-Linh Du,[*] Ruth Weimer,[†,1] and Reinhard Kunze[*,2]

Author information ▶ Article notes ▶ Copyright and License information ▶

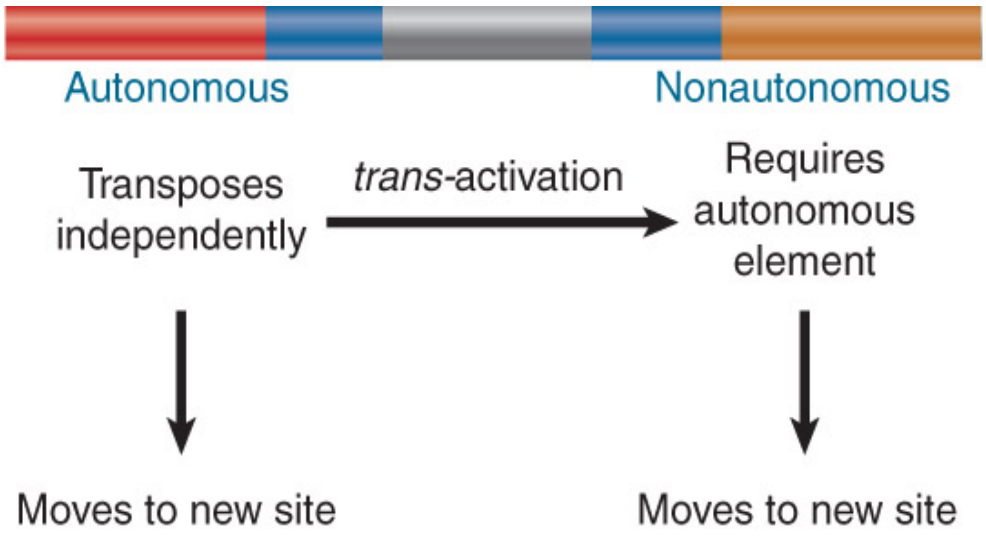
Abstract

Go to:

Activator/Dissociation (*Ac/Ds*) transposable elements from maize are widely used as insertional mutagenesis and gene isolation tools in plants and more recently also in medaka and zebrafish. They are particularly valuable for plant species that are transformation-recalcitrant and have long generation cycles or large genomes with low gene densities. *Ac/Ds* transposition frequencies vary widely, however, and in some species they are too low for large-scale mutagenesis. We discovered a hyperactive *Ac* transposase derivative, *Ac*TPase$_{4x}$, that catalyzes in the yeast *Saccharomyces cerevisiae* 100-fold more frequent *Ds* excisions than the wild-type transposase, whereas the reintegration frequency of excised *Ds* elements is unchanged (57%). Comparable to the wild-type transposase in plants, *Ac*TPase$_{4x}$ catalyzes *Ds* insertion preferentially into coding regions and to genetically linked sites, but the mutant protein apparently has lost the weak bias of the wild-type protein for insertion sites with elevated guanine–cytosine content and nonrandom protein-DNA twist. *Ac*TPase$_{4x}$ exhibits hyperactivity also in *Arabidopsis thaliana* where it effects a more than sixfold increase in *Ds* excision relative to wild-type *Ac*TPase and thus may be useful to facilitate *Ac/Ds*-based insertion mutagenesis approaches.

Keywords: *Activator* (*Ac*), *Dissociation* (*Ds*), DDE, hyperactive transposase, *hAT*

Maize elements form families of transposons that can be **autonomous** or **nonautonomous**.

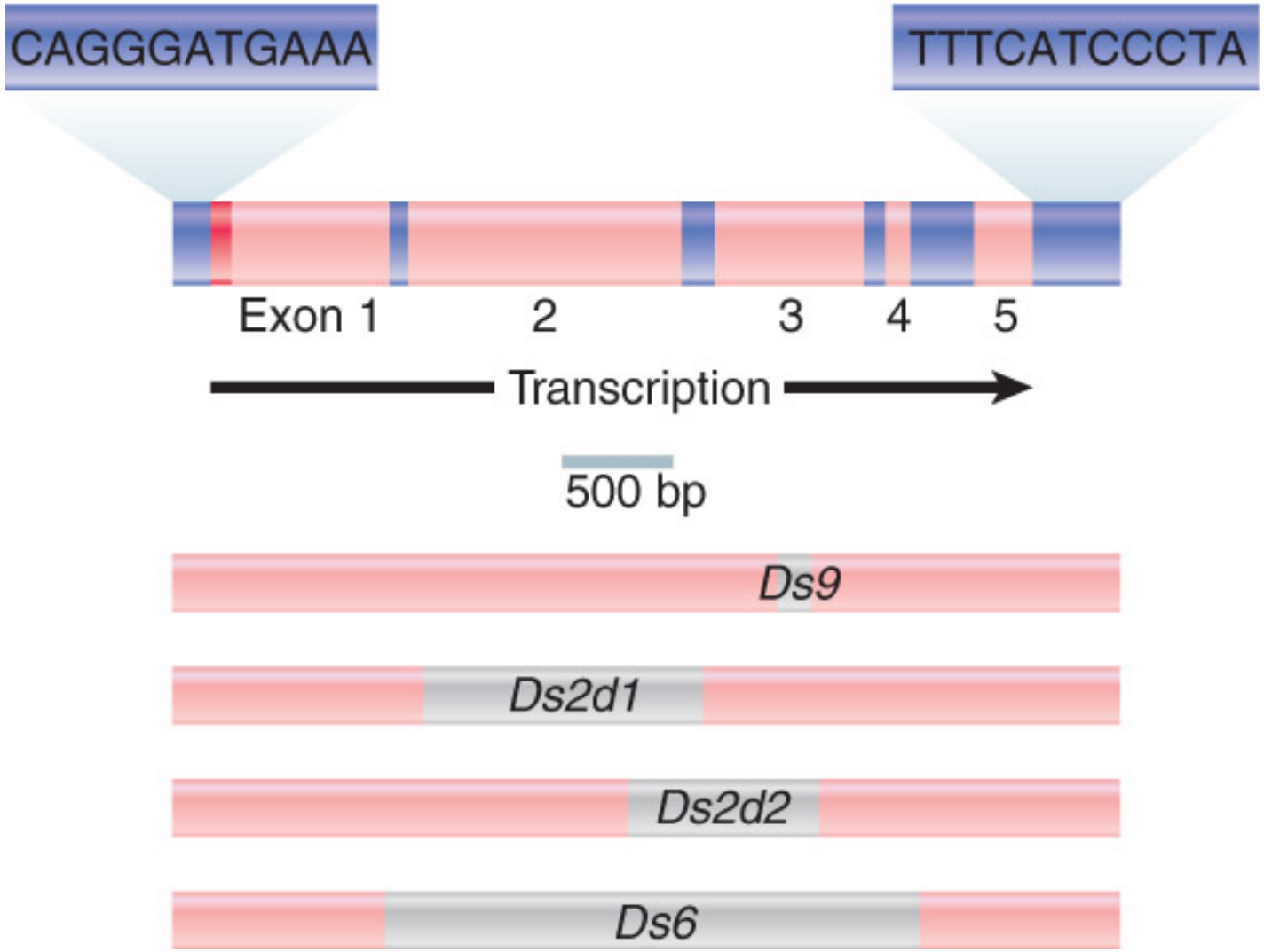


Maize transposon families	
Ac (activator) *Mp* (modulator)	*Ds* (dissociation)
Spm (suppressor-mutator) *En* (enhancer)	*dSpm* (defective *Spm*) *I* (inhibitor)
Dt (Dotted)	*rD+* (receptor of dotted)
MuDR (mutator)	*Mu*

***Ac* element –** Activator element; an **autonomous** transposable element in maize.

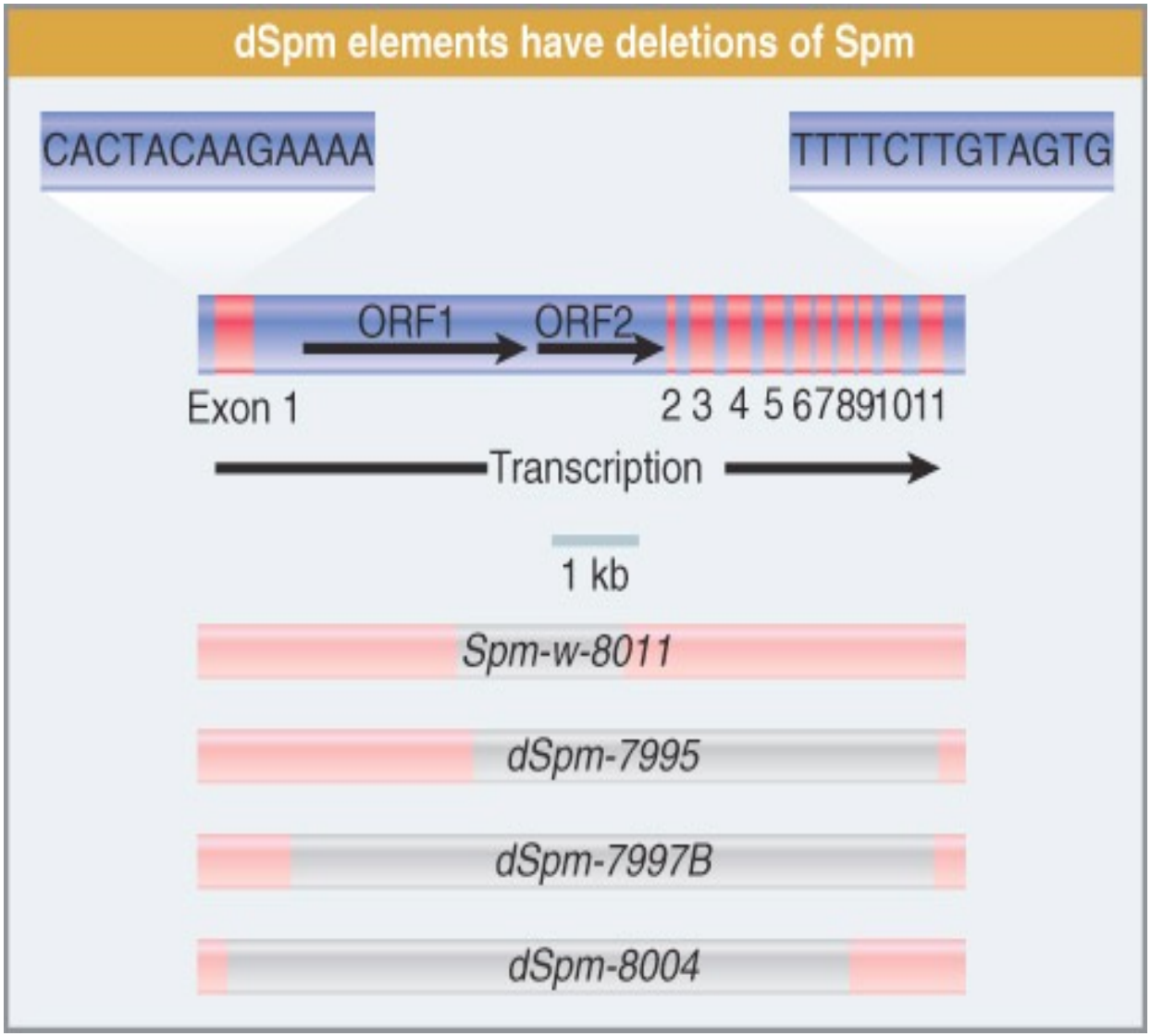
***Ds* element** – Dissociation element; a **nonautonomous** transposable element in maize, related to the autonomous Activator (*Ac*) element.

- **Autonomous** elements encode the transposase and have the ability to excise and transpose; insertion of such an element creates an unstable allele while the loss of the element converts a mutable allele into a stable one.
- **Nonautonomous** elements are relatively stable and do not transpose as they cannot catalyze transposition, but they can transpose when an autonomous element provides the necessary proteins. They are derived from autonomous elements by loss of the **transacting functions** needed for transposition.
- **Autonomous** elements are subject to changes of phase of their activity, which are heritable but relatively unstable alterations in their properties. In the **Ac** and **Mu** types of elements, these phases are regulated by **methylation state of DNA.**



Ds elements in maize arise by deletions of **Ac**

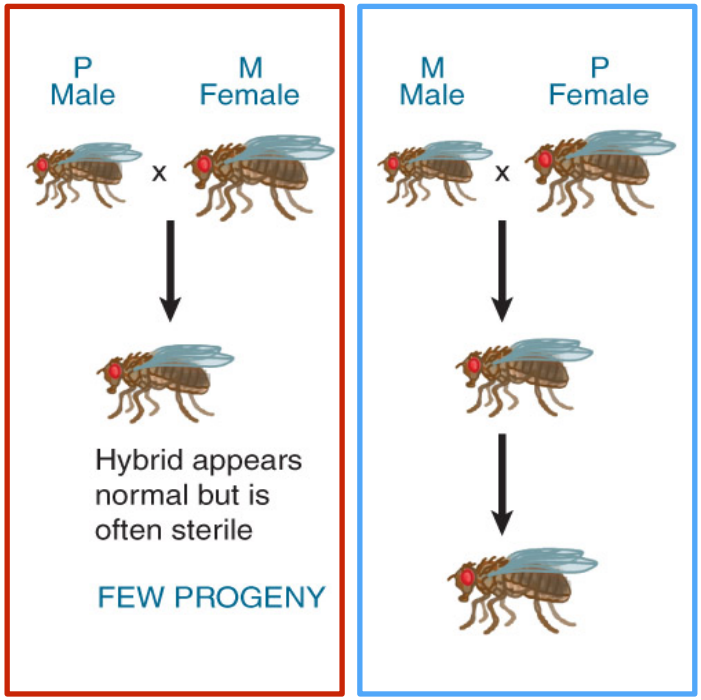
Transposons related to the **Spm family** of elements are found in a variety of plants. They all share nearly identical inverted terminal repeats and generate **3 bp duplications** of target DNA upon transposition. They are also known as the **CACTA group of transposons**.



- The **Spm element** includes the ***tnpA*** gene product (which is required for excision but may not be sufficient for transposition) and the ***tnpB*** product thought to bind to the 13 bp inverted repeats to cleave the termini for transposition

- **Spm** insertions can control the expression of a gene at the site of insertion.

- ***P* elements** and **hybrid dysgenesis** (hybrid deficiencies). P elements are found in 30-40 copies of certain *Drosophila* genomes. They come in various sizes and modifications, but all contain a **31 bp** perfect, terminal **inverted** repeats. Essentially there are two types of *Drosophila* populations **M** and **P** types (i.e. those that **haven't [M]** and those that **have [P]**).
- P elements have a canonical structure containing **31 bp terminal** inverted repeats and 11 bp internal inverted repeats located at around the transposase.
- Cross any **Male P** with a **Female M** and .....all heck breaks loose -genetically speaking…
- **Hybrid dysgenesis, –** The inability of certain strains of *D. melanogaster* to interbreed, because the hybrids are sterile (although otherwise they may be phenotypically normal).
- **P elements** this was first determined when "accidentally" pure bread *Drosophila* strains were mated with *Drosophila* from the wild.
- The resulting, “anomolous genetics” was hard to appreciate, at first.

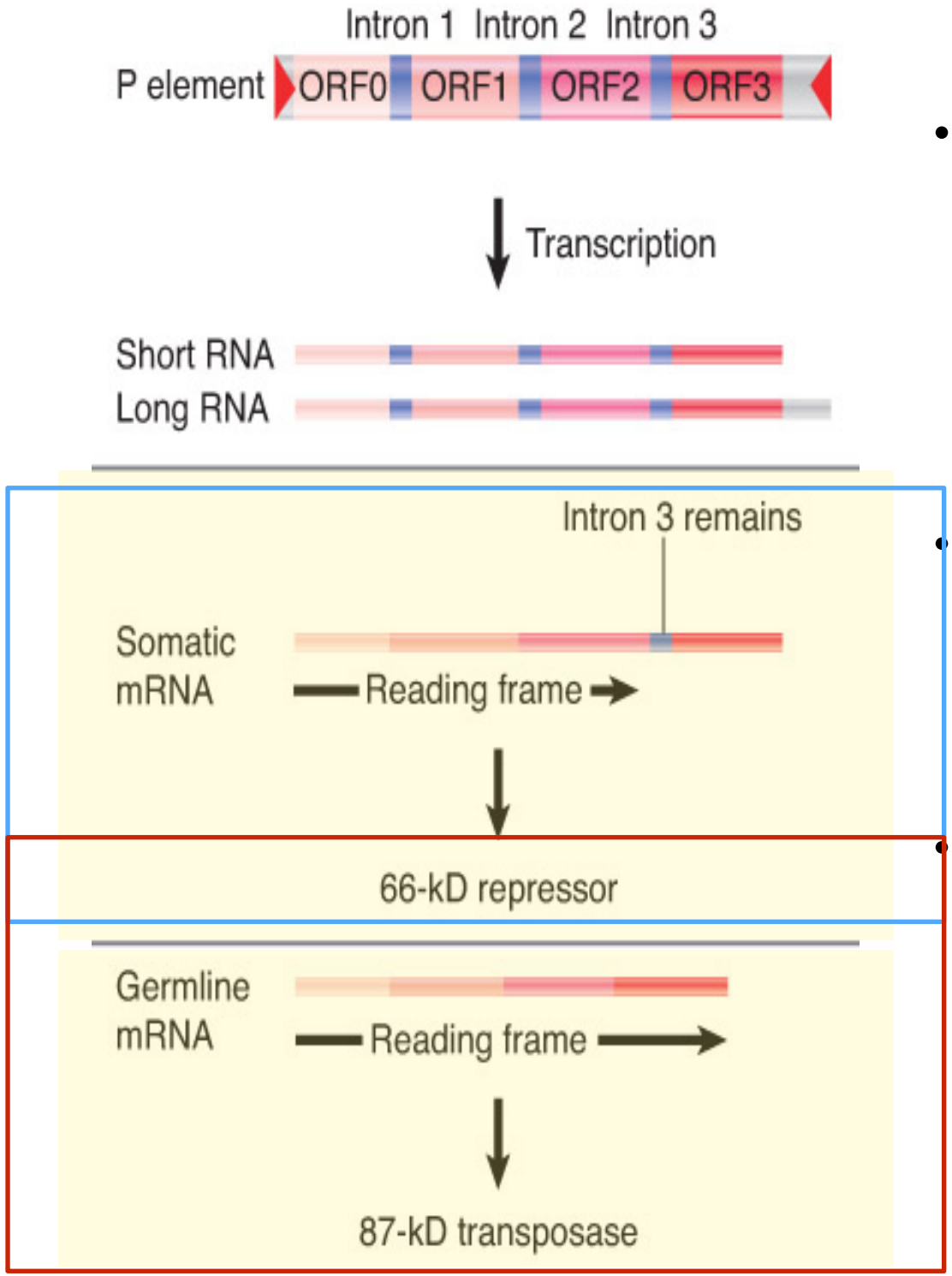


P elements are transposons that are carried in P strains of *Drosophila melanogaster.*

Transposition is activated when a **P male** is crossed with an **M female.**

Yet, **M male** is crossed with **P female**... nothing anomolous happens

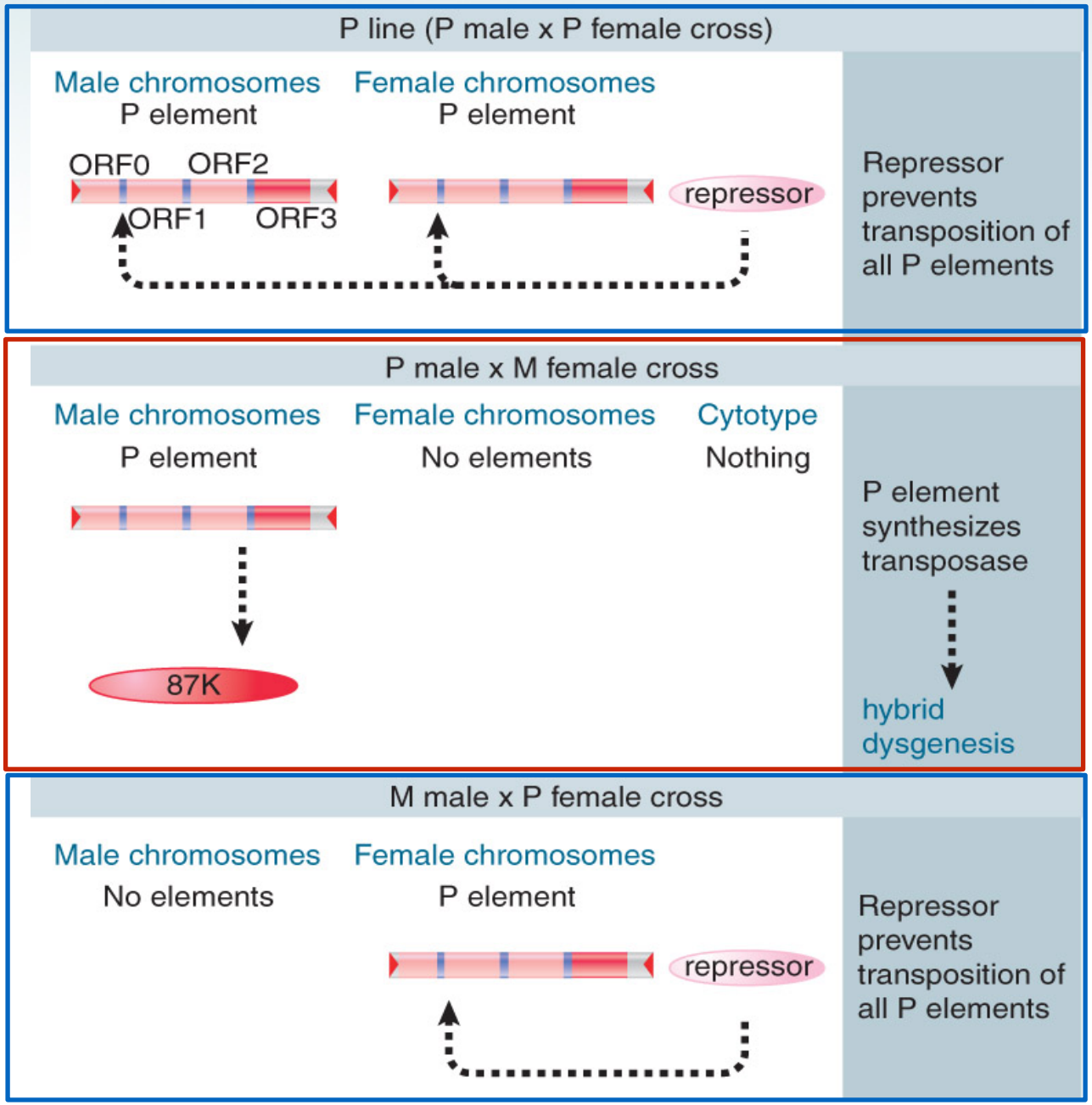
Any one chromosome of the P male can induce the **hybrid dysgenesis** effect.



The **P element** has four exons.

- A P strain carries 30 to 50 copies of a transposable element known as a P element, which becomes activated when a **P male** is crossed with a "naive" **M female.**
- The Tn insertion inactivates genes in which they are located, and often cause chromosomal breaks…. leading to sterility.
- This activation occurs because a tissue-specific splicing event, that occurs ONLY in the **germ-line cells,** and gives rise to the removal of intron 3 and generates an **87 kD protein,** which functions as a **transposase.**

P Elements Are Activated in the Germline



- The P element also produces a repressor of transposition, which is inherited maternally in the cytoplasm.
- The presence of the repressor explains why **M male** × **P female** crosses remain fertile.
- Hybrid dysgenesis is determined by the interactions between P elements in the genome and repressors in the cytotype.

Proc Natl Acad Sci U S A. 2011 Sep 20; 108(38): 15948–15953. PMCID: PMC3179094

Published online 2011 Sep 6. doi: 10.1073/pnas.1112960108

Genetics

Drosophila P elements preferentially transpose to replication origins

Allan C. Spradling,[a,b,1] Hugo J. Bellen,[a,c] and Roger A. Hoskins[d]

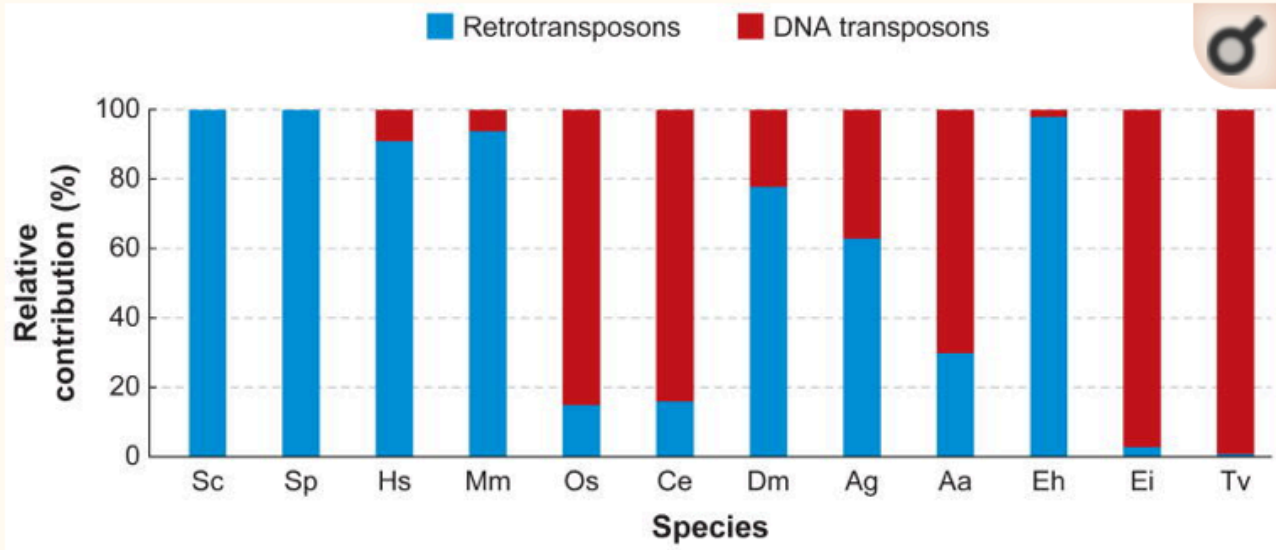
Author information ► Copyright and License information ►

ABSTRACT

Go to:

The *P* transposable element recently invaded wild *Drosophila melanogaster* strains worldwide. A single introduced copy can multiply and spread throughout the fly genome in just a few generations, even though its cut-and-paste transposition mechanism does not inherently increase copy number. *P* element insertions preferentially target the promoters of a subset of genes, but why these sites are hotspots remains unknown. We show that *P* elements selectively target sites that in tissue-culture cells bind origin recognition complex proteins and function as replication origins. The association of origin recognition complex-binding sites with selected promoters and their absence near clustered differentiation genes may dictate *P* element site specificity. Inserting at unfired replication origins during S phase may allow *P* elements to be both repaired and reduplicated, thereby increasing element copy number. The advantage transposons gain by moving from replicated to unreplicated genomic regions may contribute to the association of heterochromatin with late-replicating genomic regions.

Keywords: genome evolution, cell cycle, DNA replication, pre-replication complex



Open in a separate window

Figure 2

The relative amount of retrotransposons and DNA transposons in diverse eukaryotic genomes. The graph shows the contribution of DNA transposons and retrotransposons in percentage relative to the total number of transposable elements in each species. The data were compiled from papers reporting draft genome sequences (references available upon request) and from the Repeatmasker output tables available at the UCSC Genome Browser (http://genome.ucsc.edu) or from the following sources: *E. histolytica* and *E. invadens*: (159); *T. vaginalis*: E. Pritham, unpublished data. Species abbreviations: Sc: *Saccharomyces cerevisiae*; Sp: *Schizosaccharomyces pombe*; Hs: *Homo sapiens*; Mm: *Mus musculus*; Os: *Oryza sativa*; Ce: *Caenorhabditis elegans*; Dm: *Drosophila melanogaster*; Ag: *Anopheles gambiae*, malaria mosquito; Aa: *Aedes aegypti*, yellow fever mosquito; Eh: *Entamoeba histolytica*; Ei: *Entamoeba invadens*; Tv: *Trichomonas vaginalis*.

Retroelements:

Retrons, **retrotransposons** (*copia* elements in *Drosophila*, and DIRSI in slime mold)

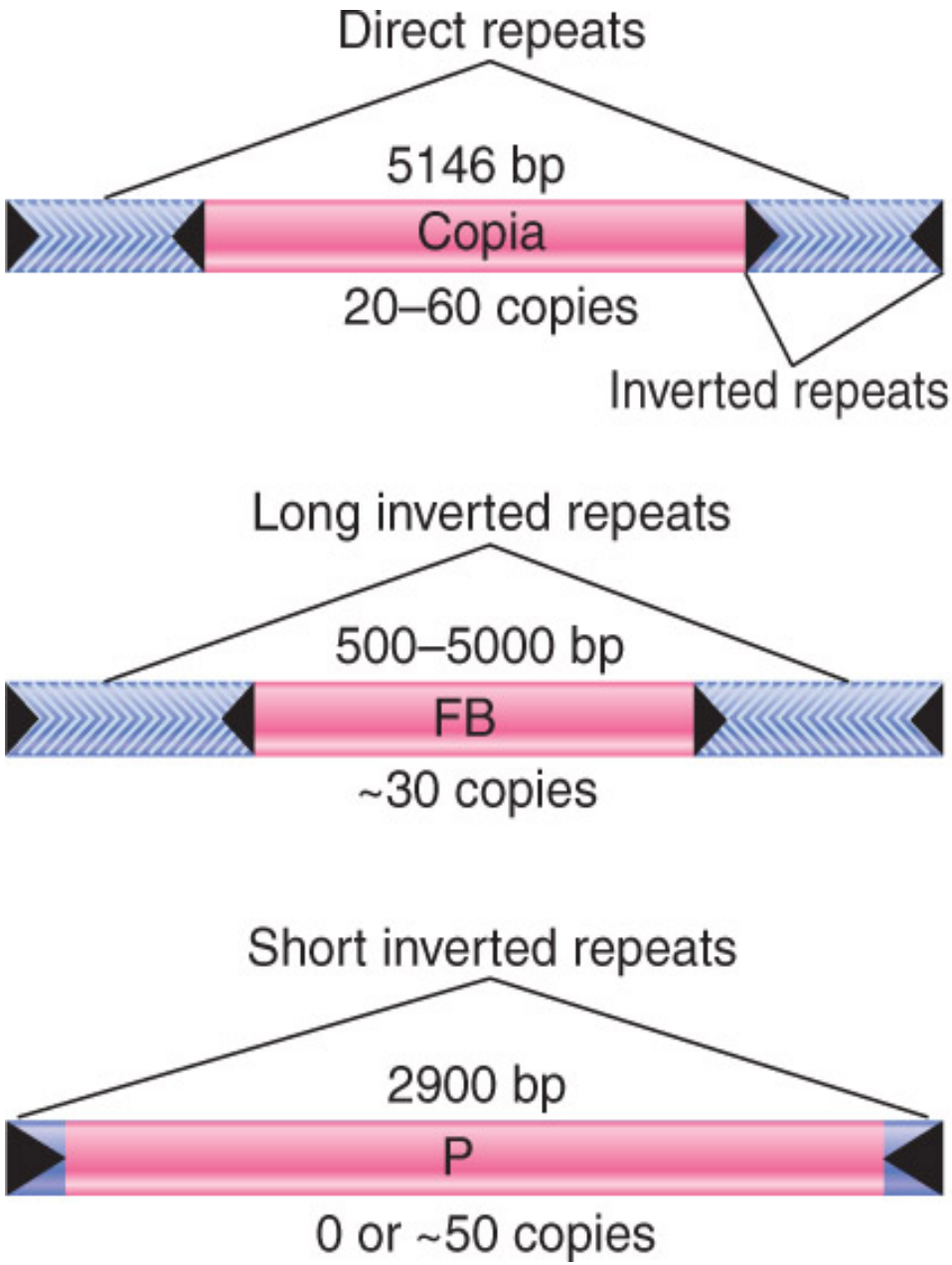
and

retroviruses, (HIV, Feline leukaemia retrovirus).

All contain **Reverse transcriptase** in their sequence, which they use to replicate themselves (consequently all possess a high mutation rate:

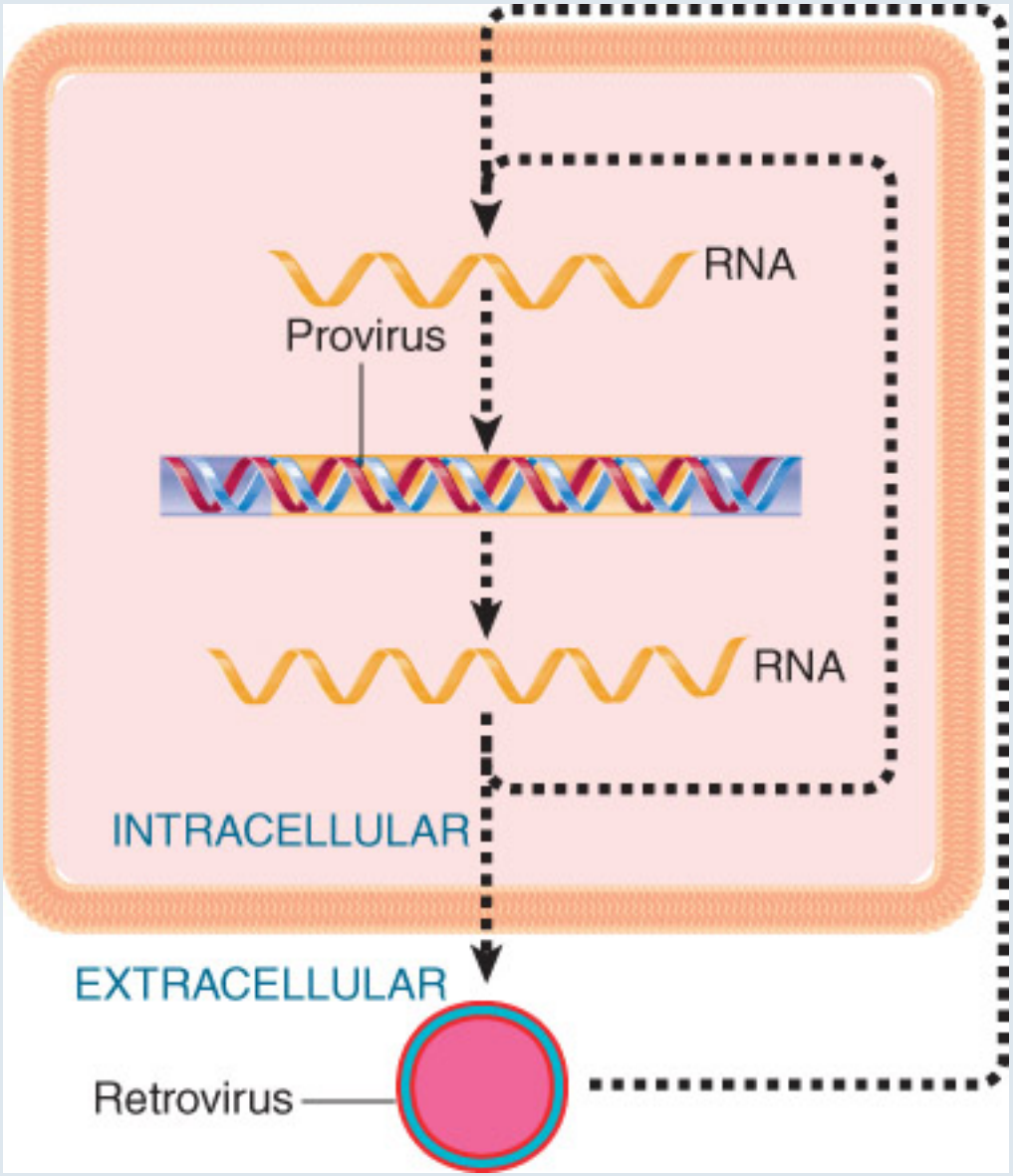
eg HIV, $1x10^{-4}$ (1 in 10,000 base pairs replicated incorporate a mutation) vs. 1×10^{-8} for *most human genes*.

Retroviruses and Retrotransposons

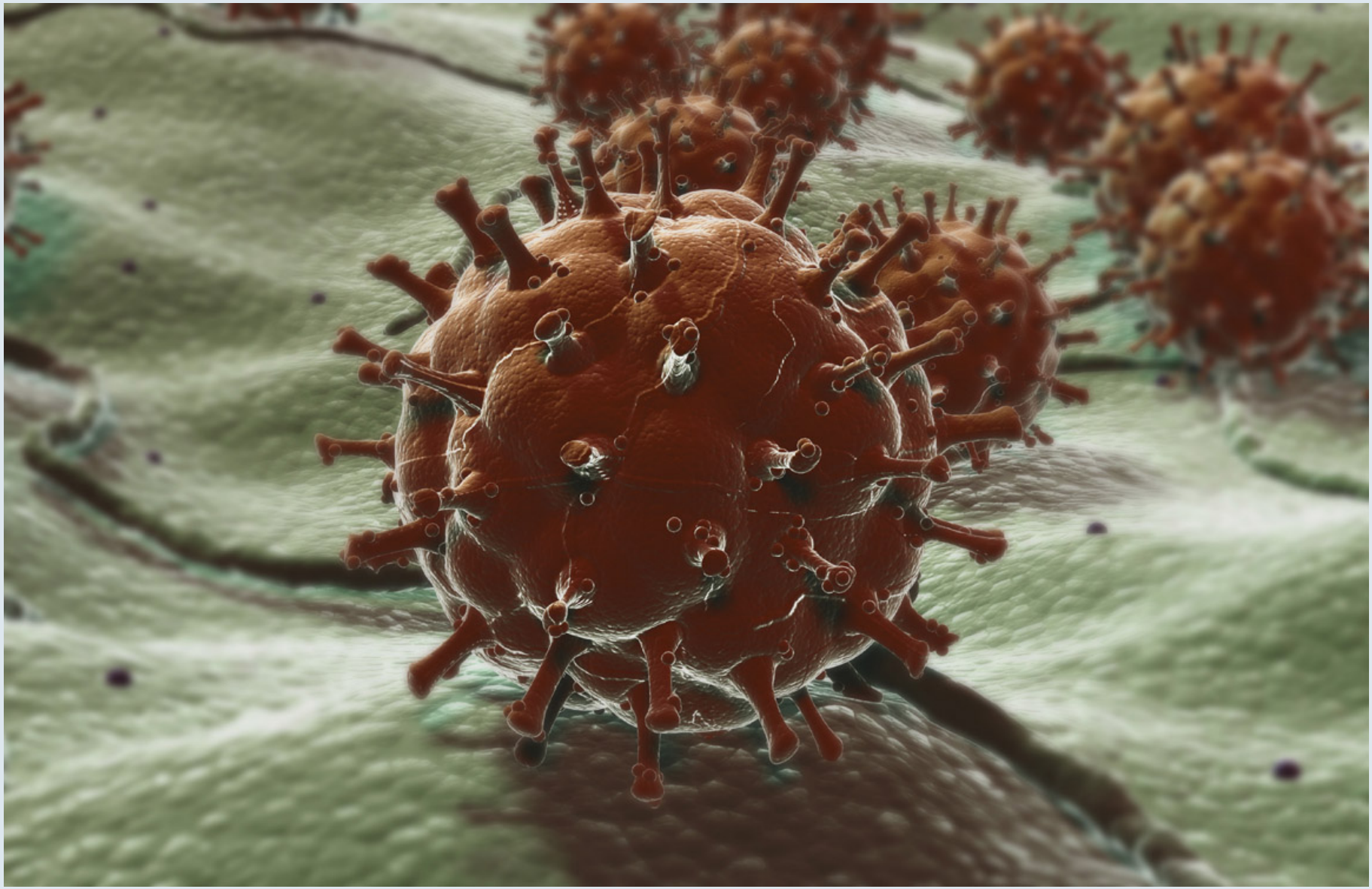


- **Copia elements represent a family of retrotransposons. The number of copia elements depends on the strain of fly and is typically between 20 and 60 copies.**

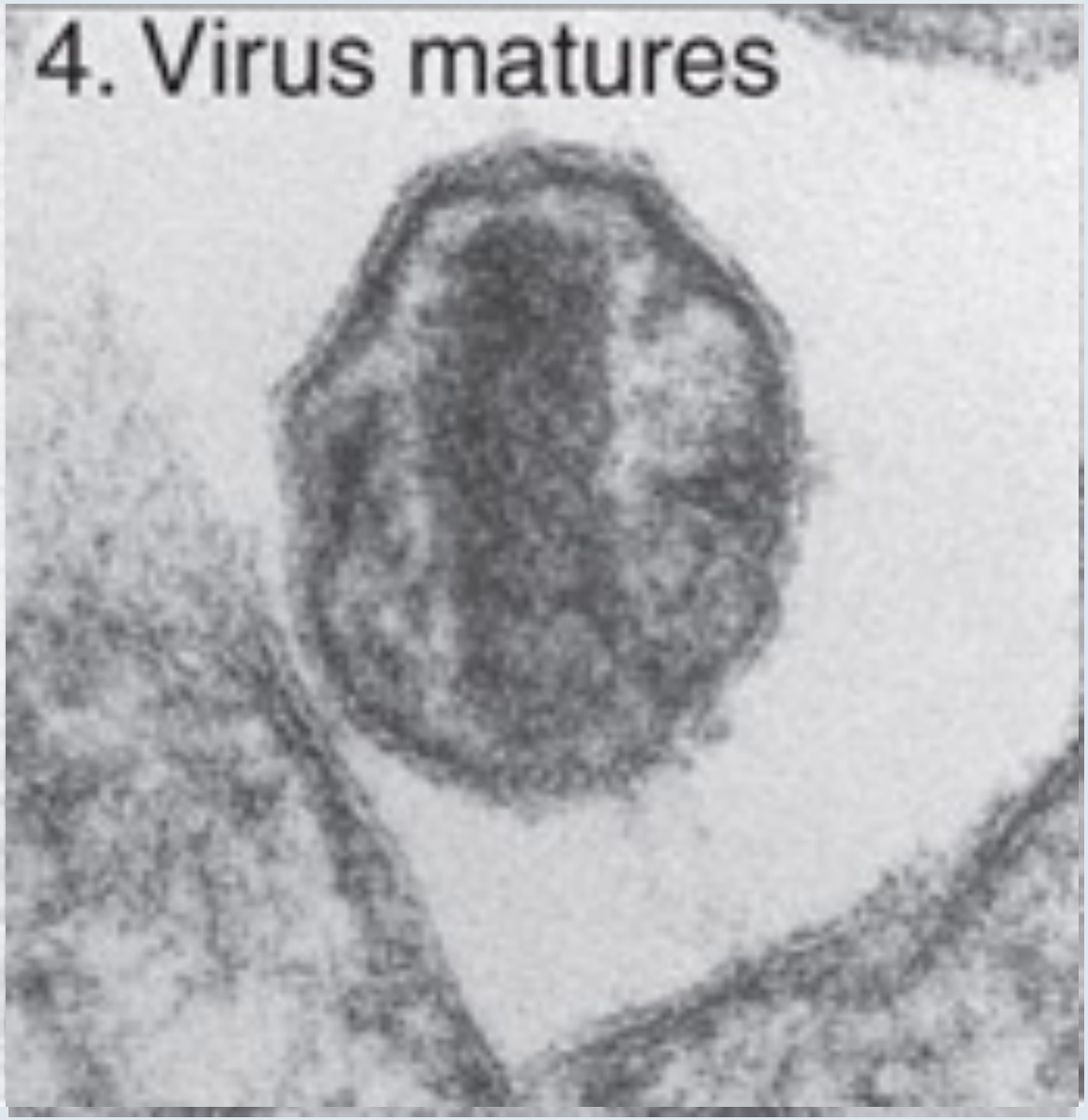
  - The element is ~5000 bp long, with identical direct terminal repeats of 276 bp.
  - A direct repeat of 5 bp of target DNA is generated at the site of insertion.
  - There is very little divergence between individual family members; variants usually contain small deletions.
  - Transcripts of copia are found as abundant poly(A+) mRNAs representing both full-length and part-length transcripts.



The reproductive cycles of retroviruses and retrotransposons.

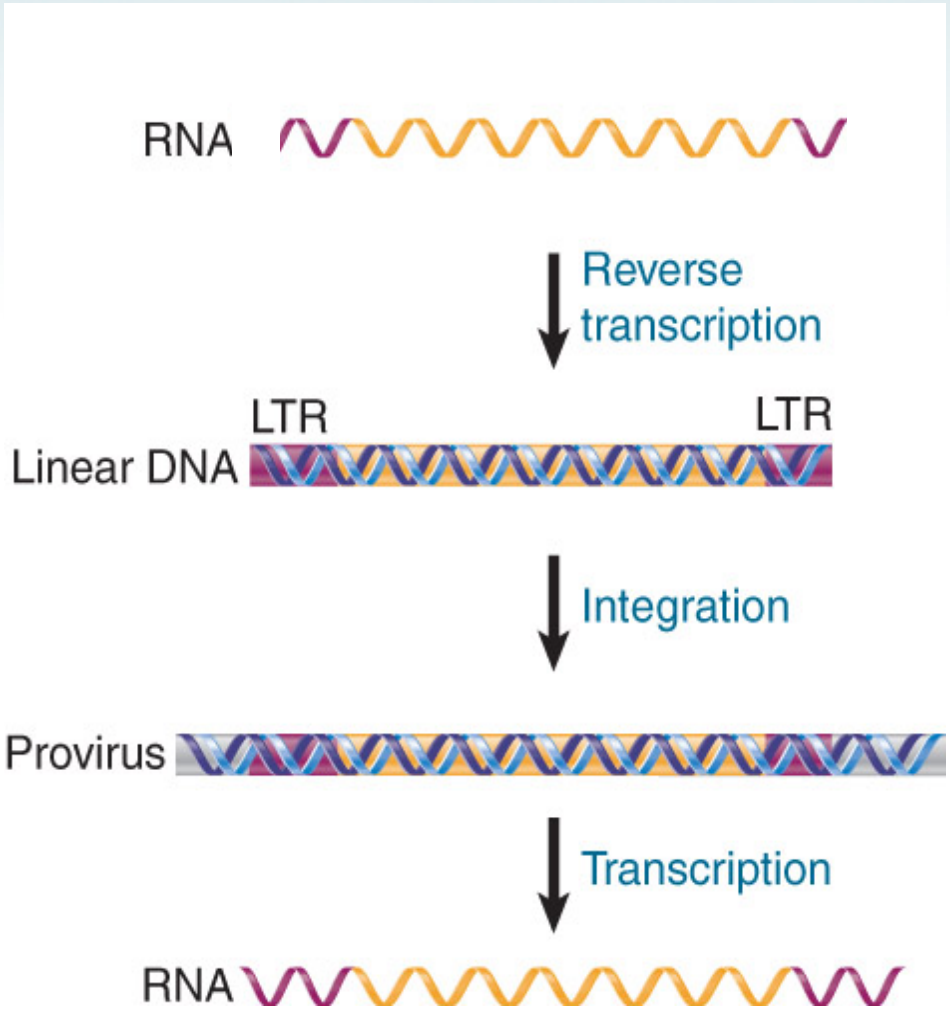


Artist's rendering of an electron micrograph showing retroviruses.



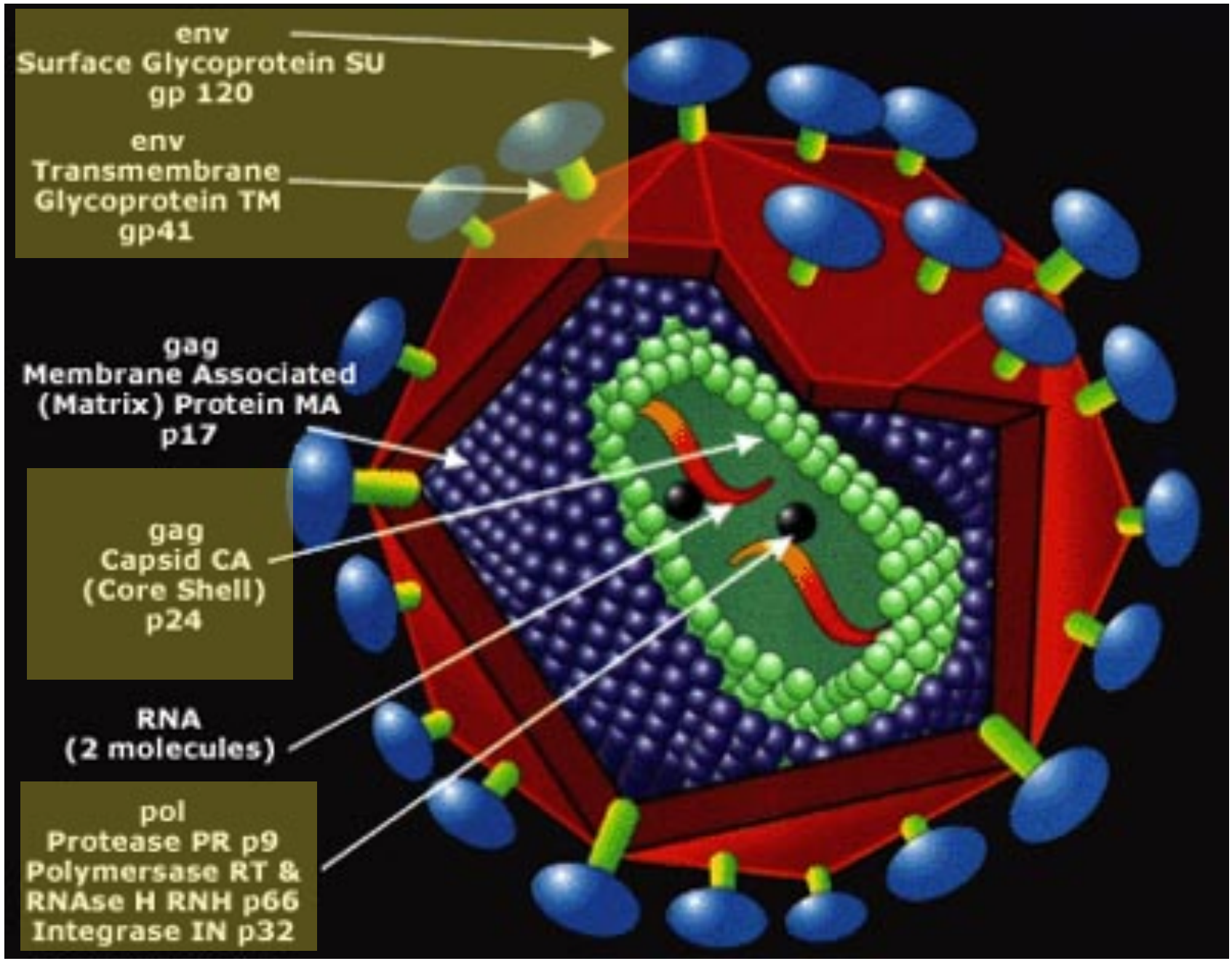
Retroviruses (HIV) bud from the plasma membrane of an infected cell.

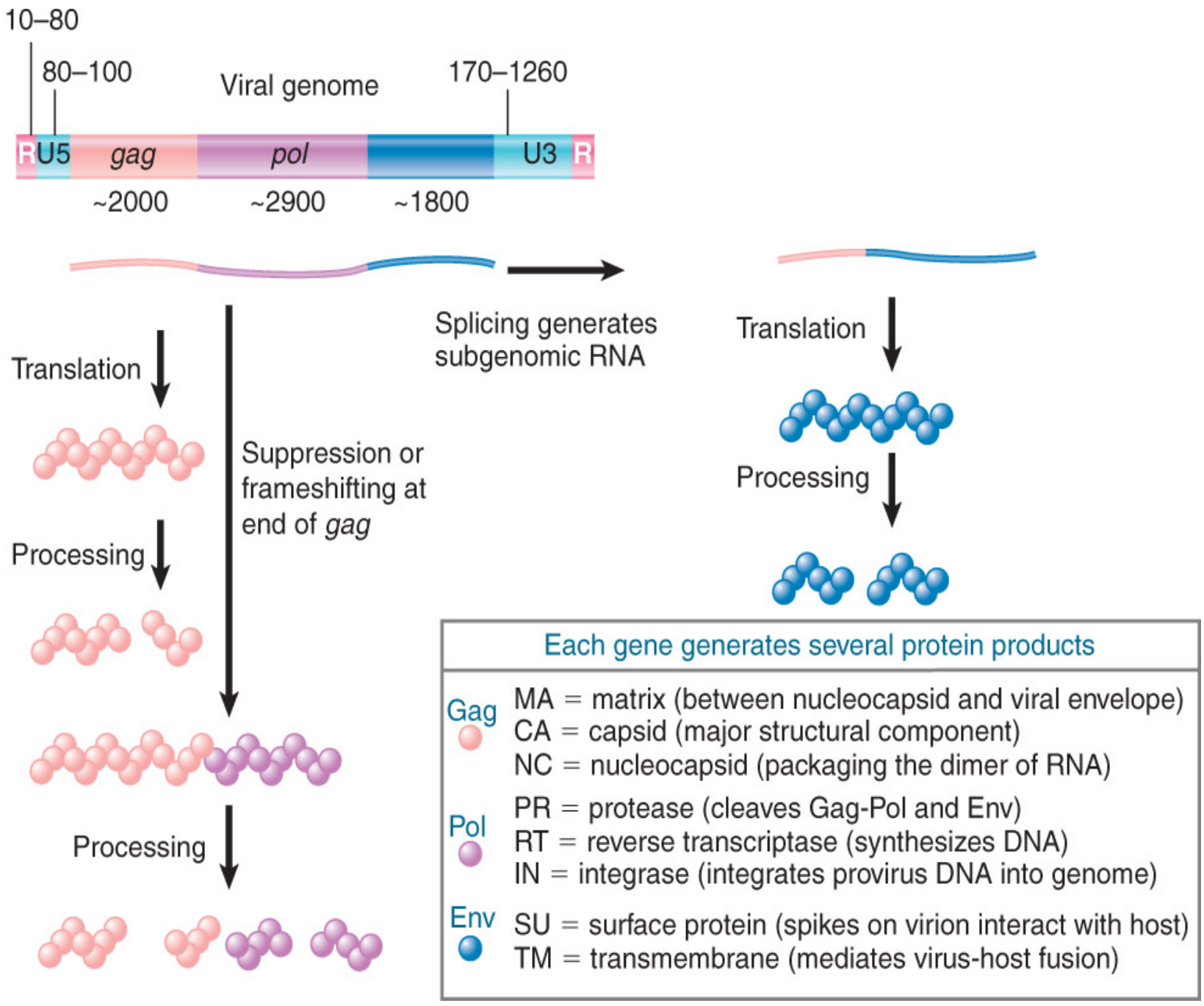
Photos courtesy of Matthew A. Gonda, Ph.D., Partner at Power Ten Medical Ventures I

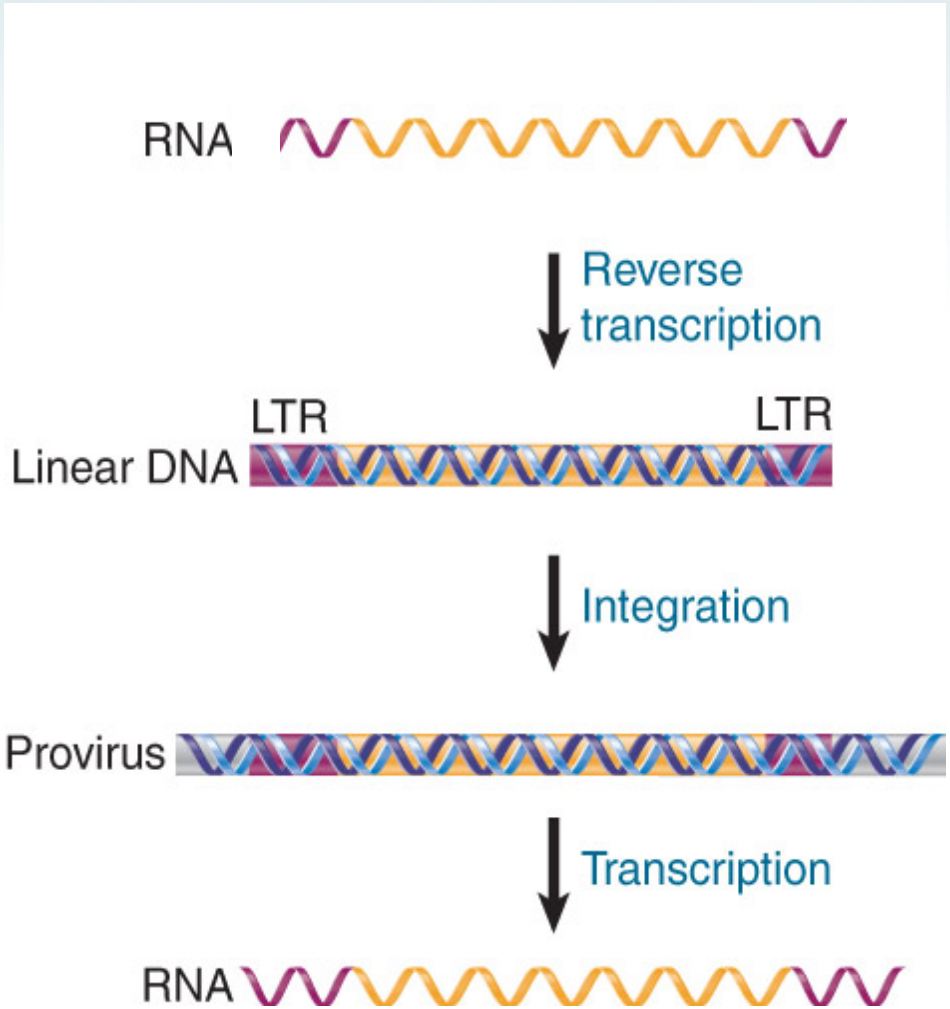


- A retrovirus normally has two copies of its single-stranded RNA genome packaged in the viral particle.
- An integrated **provirus** is a double-stranded DNA sequence.
- A retrovirus generates a provirus by reverse transcription of the retroviral genome.

RNA --> DNA --> RNA

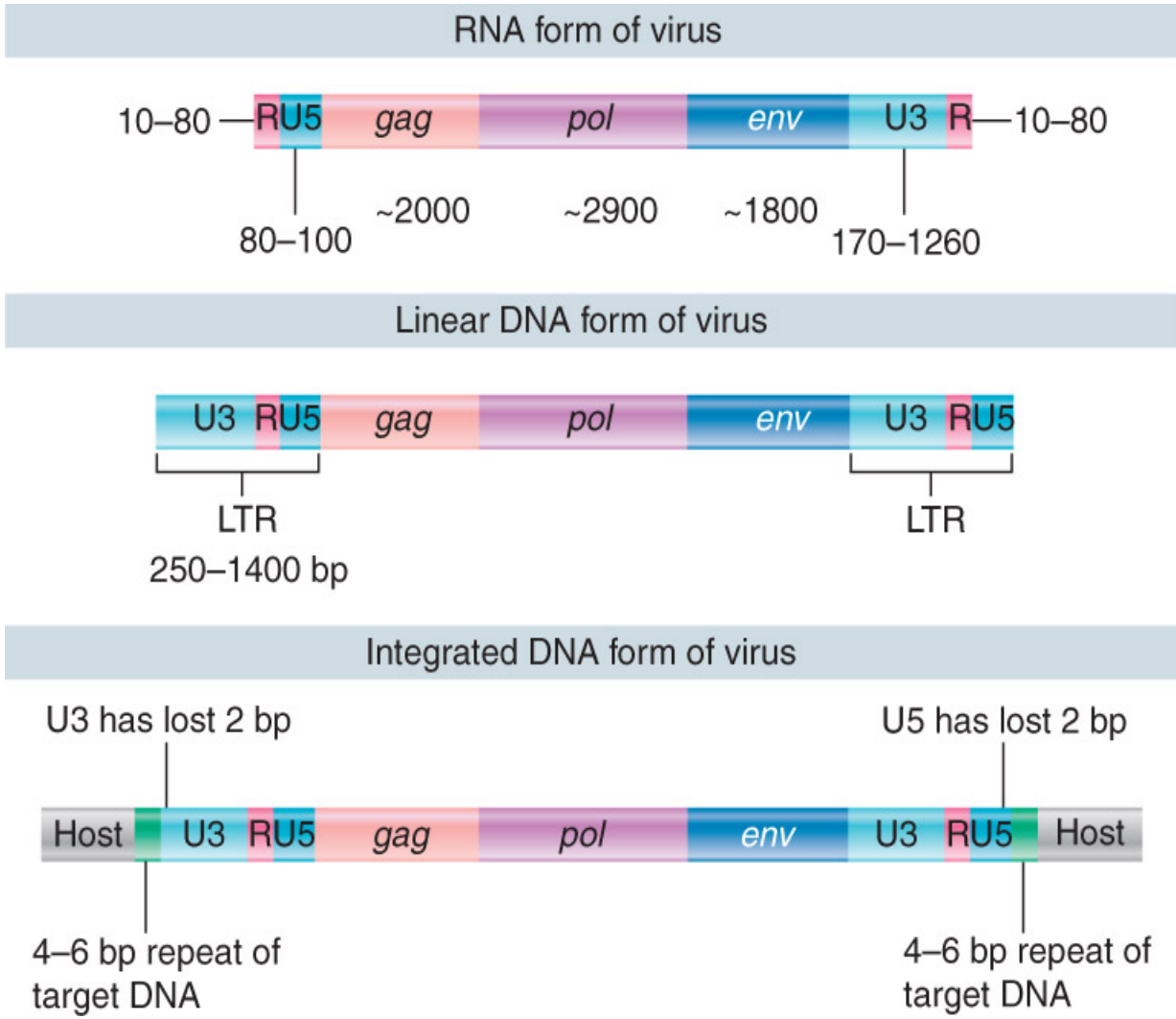

env
Surface Glycoprotein SU
gp 120
env
Transmembrane
Glycoprotein TM
gp41
gag
Membrane Associated
(Matrix) Protein MA
p17
gag
Capsid CA
(Core Shell)
p24
RNA
(2 molecules)
pol
Protease PR p9
Polymersase RT &
RNAse H RNH p66
Integrase IN p32


10–80
80–100
Viral genome
170–1260
R
U5
gag
pol
U3
R
~2000
~2900
~1800
Splicing generates subgenomic RNA
Translation
Processing
Suppression or frameshifting at end of gag
Translation
Processing
Processing
Each gene generates several protein products
Gag
MA = matrix (between nucleocapsid and viral envelope)
CA = capsid (major structural component)
NC = nucleocapsid (packaging the dimer of RNA)
Pol
PR = protease (cleaves Gag-Pol and Env)
RT = reverse transcriptase (synthesizes DNA)
IN = integrase (integrates provirus DNA into genome)
Env
SU = surface protein (spikes on virion interact with host)
TM = transmembrane (mediates virus-host fusion)

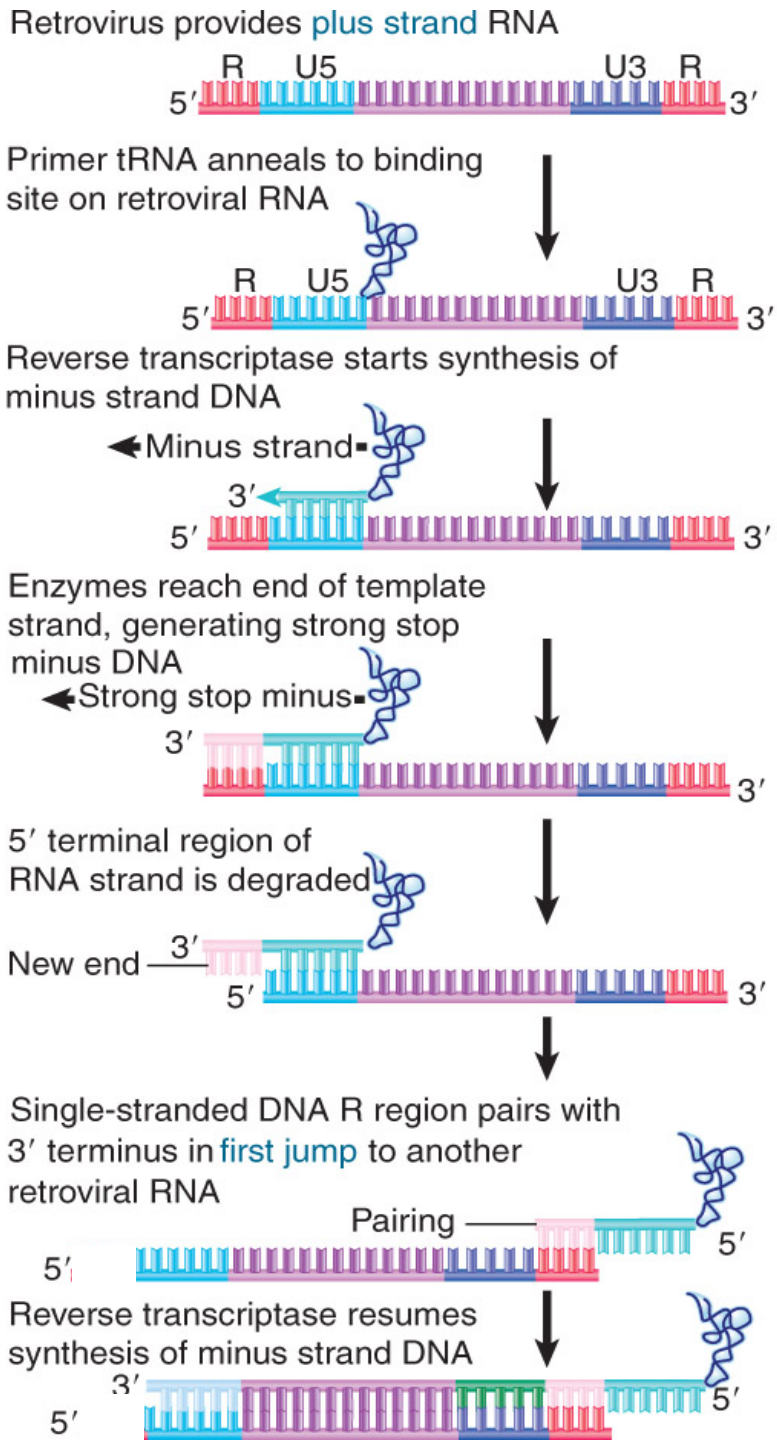


- A retrovirus normally has two copies of its single-stranded RNA genome packaged in the viral particle.
- An integrated **provirus** is a double-stranded DNA sequence.
- A retrovirus generates a provirus by reverse transcription of the retroviral genome.

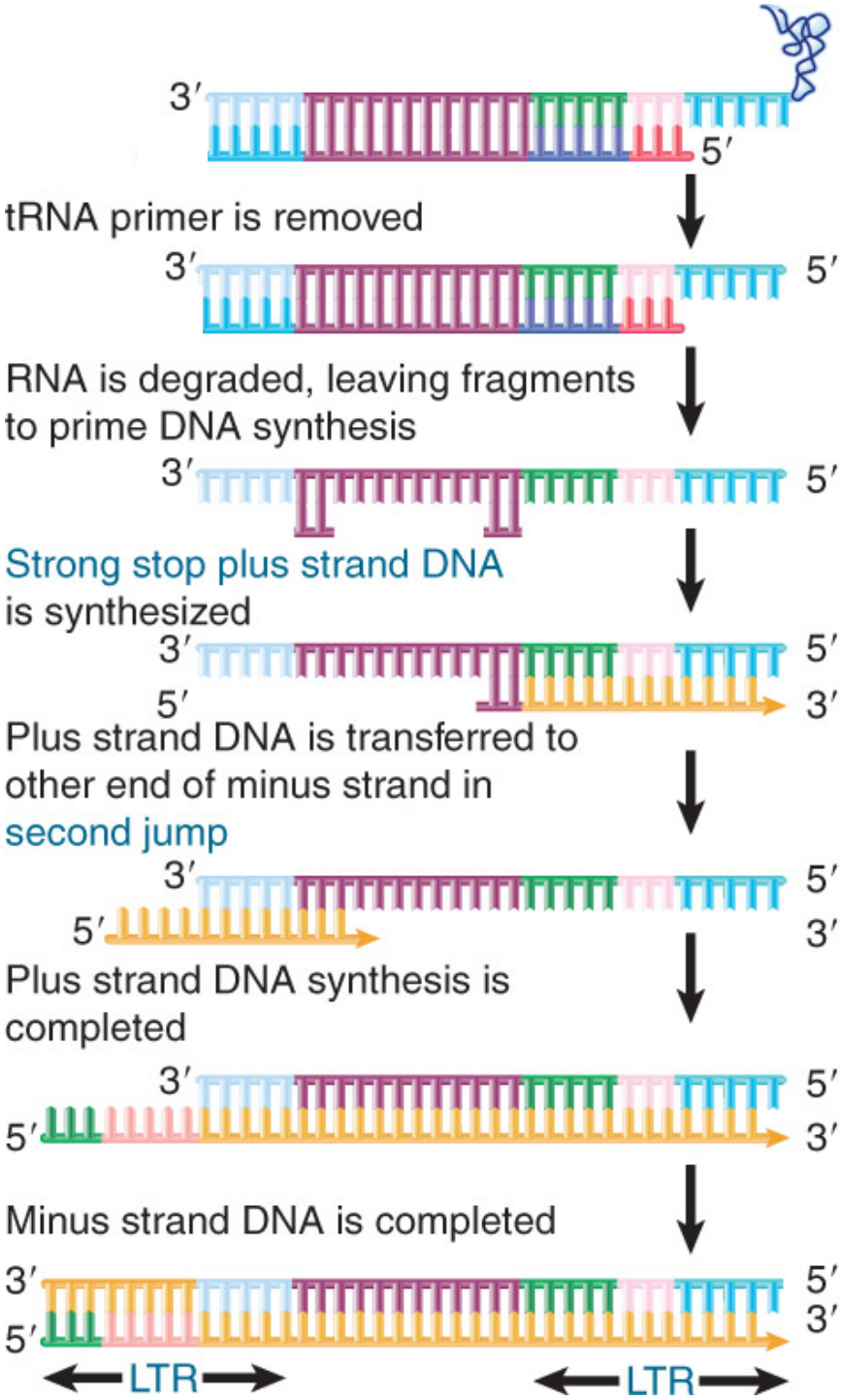
RNA --> DNA --> RNA



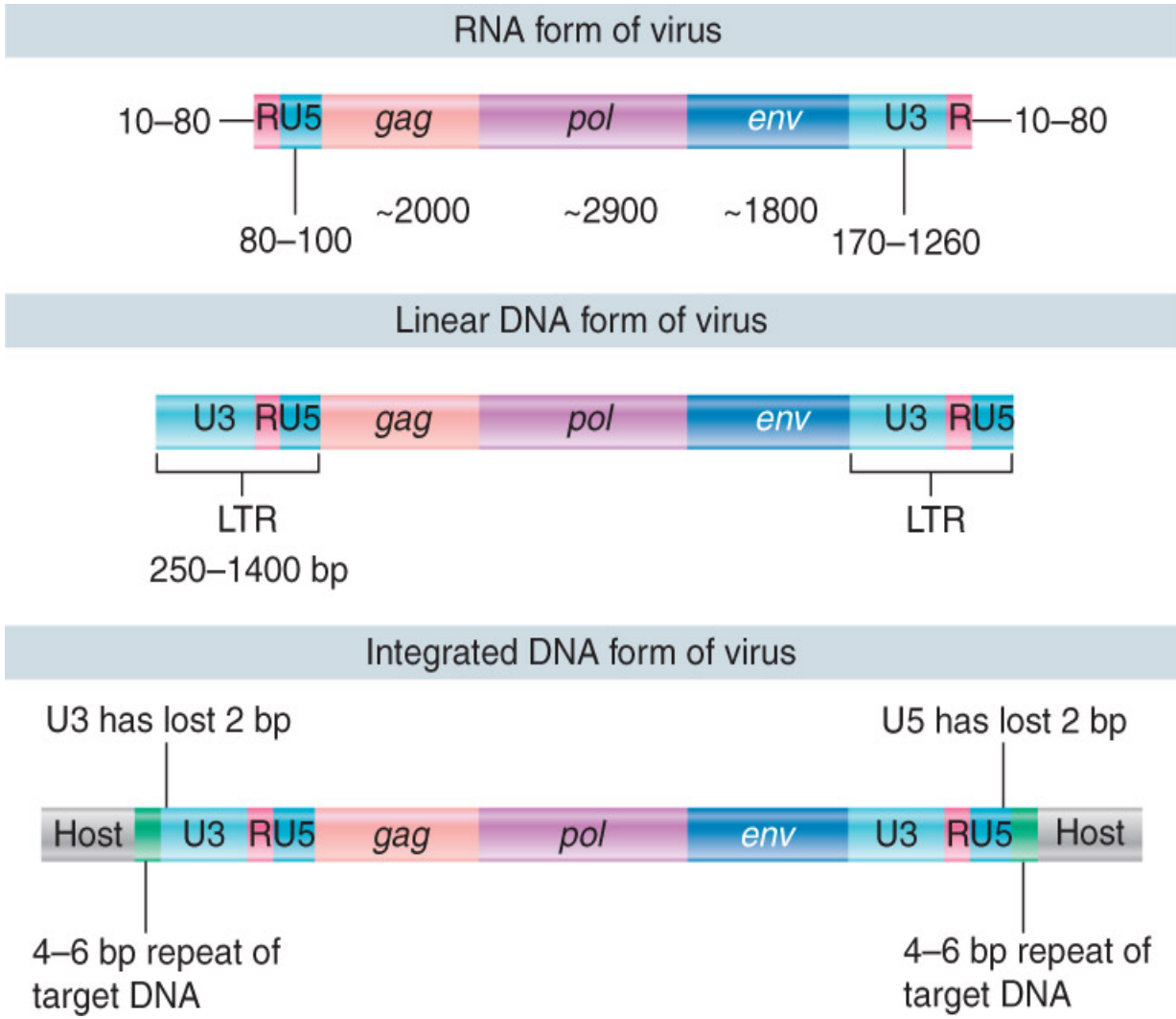
- Retroviral genomes exist as RNA and DNA sequences
- A short sequence (R) is repeated at each end of the viral RNA.
  - The 5′ and 3′ ends are R-U5 and U3-R, respectively.



- Reverse transcriptase starts synthesis when a tRNA primer binds to a site 100 to 200 bases from the 5′ end.
- When the enzyme reaches the end, the 5′-terminal bases of RNA are degraded.
  - This exposes the 3′ end of the DNA product.
- The exposed 3′ end base pairs with the 3′ terminus of another RNA genome.
- Synthesis continues, generating a product in which the 5′ and 3′ regions are repeated.
  - This gives each end the structure U3-R-U5.



- Similar strand switching events occur when reverse transcriptase uses the DNA product to generate a complementary strand.
- Strand switching is an example of the **copy choice** mechanism of recombination.



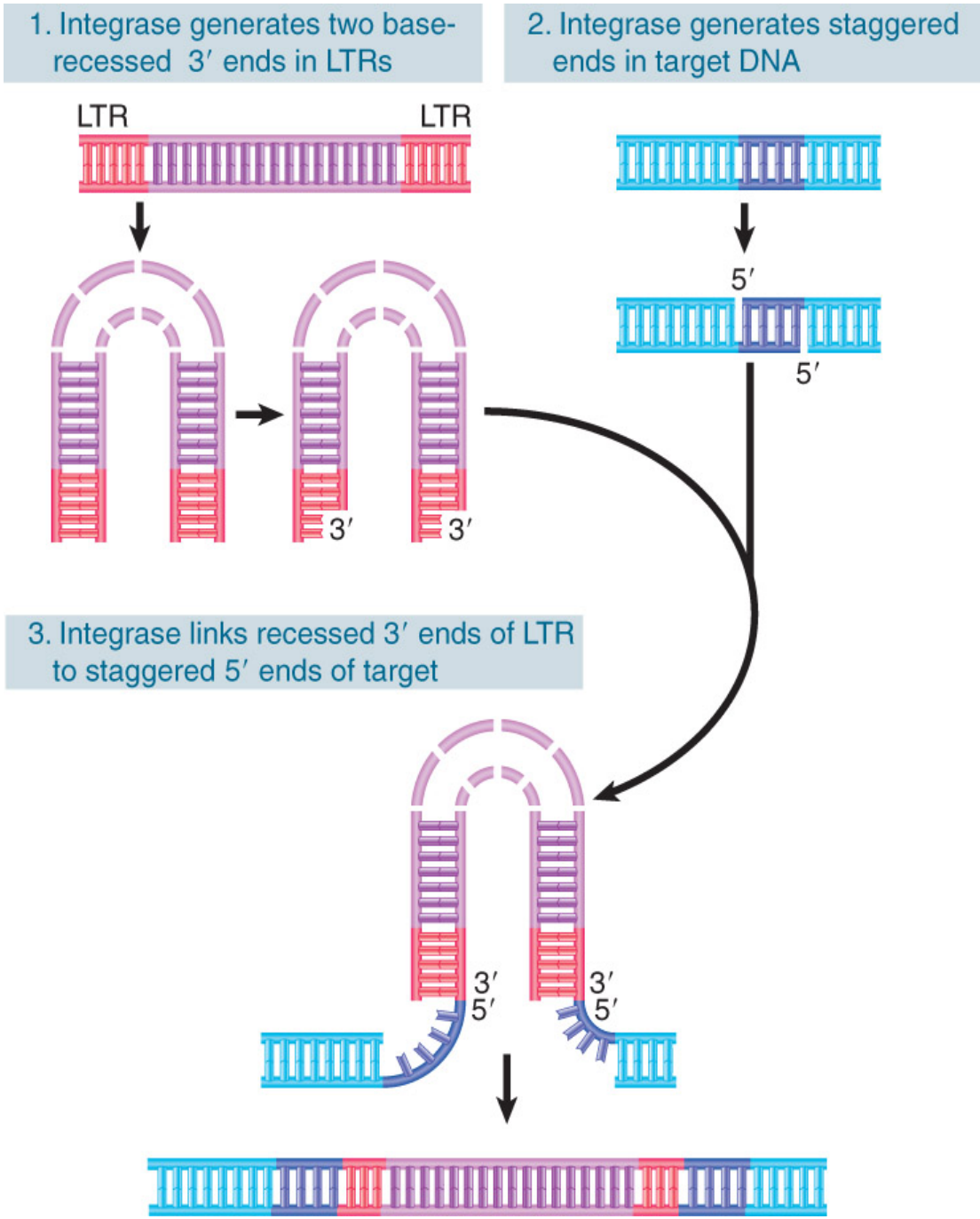
- Retroviral genomes exist as RNA and DNA sequences
- A short sequence (R) is repeated at each end of the viral RNA.
  - The 5′ and 3′ ends are R-U5 and U3-R, respectively.

Viral DNA Integrates into the Chromosome

- The organization of proviral DNA in a chromosome is the same as a transposon, with the provirus flanked by short direct repeats of a sequence at the target site.

- Linear DNA is inserted directly into the host chromosome by the retroviral **integrase** enzyme (derived from proteolysis of the **pol gene** product).

- Two base pairs of DNA are lost from each end of the retroviral sequence during the integration reaction.

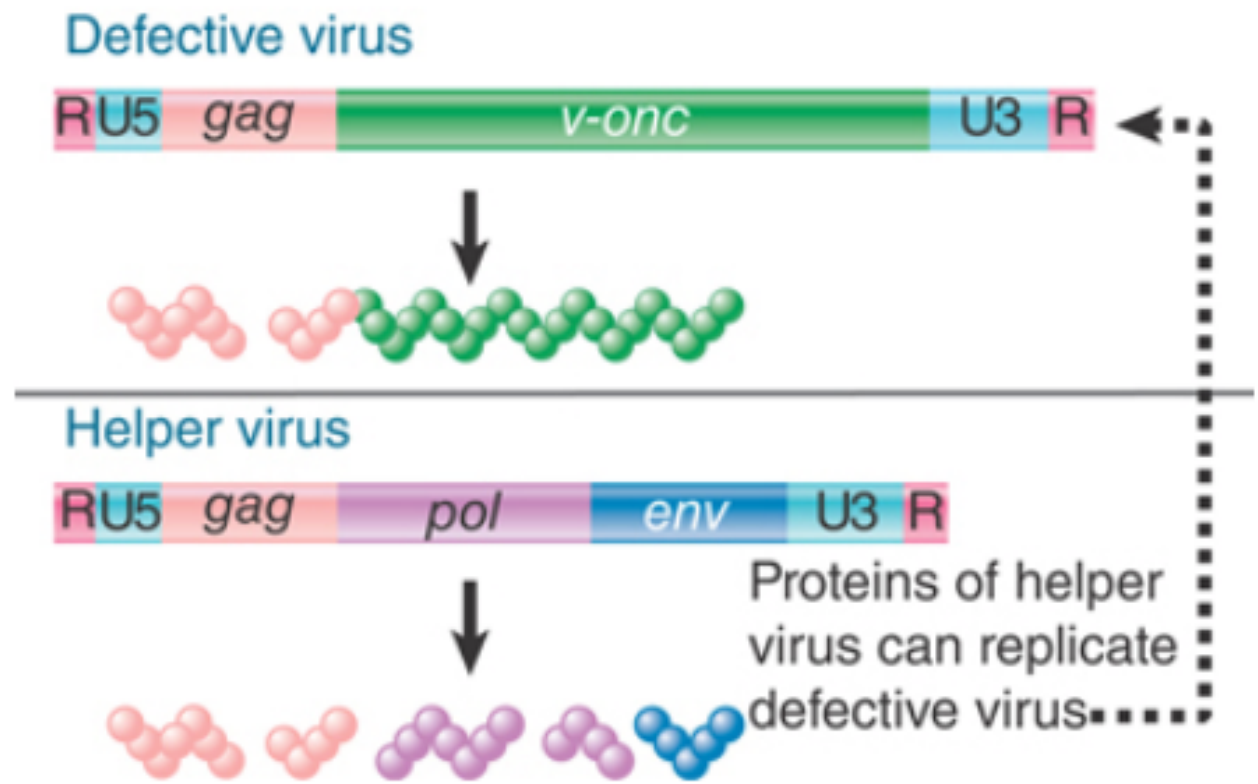


The organization of proviral DNA in a chromosome is the same as a transposon, with the provirus flanked by short direct repeats of a sequence at the target site.

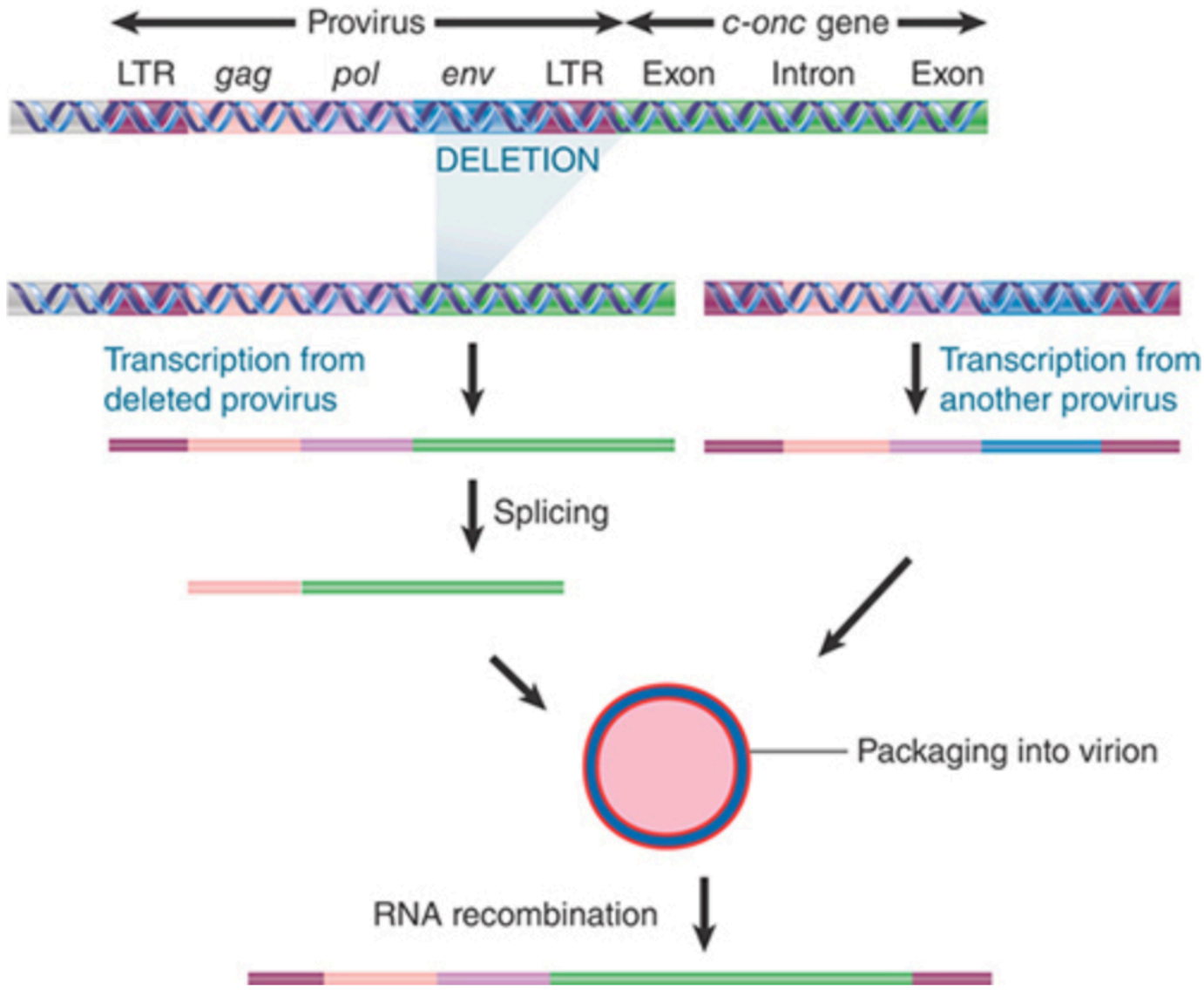
Linear DNA is inserted directly into the host chromosome by the retroviral integrase enzyme.

- Two base pairs of DNA are lost from each end of the retroviral sequence during the integration reaction.

- **Integrase** is the only viral protein required for the integration reaction.



Replication-defective transforming viruses have a cellular sequence substituted for part of the viral sequence. The defective virus may replicate with the assistance of a helper virus that carries the wild-type functions

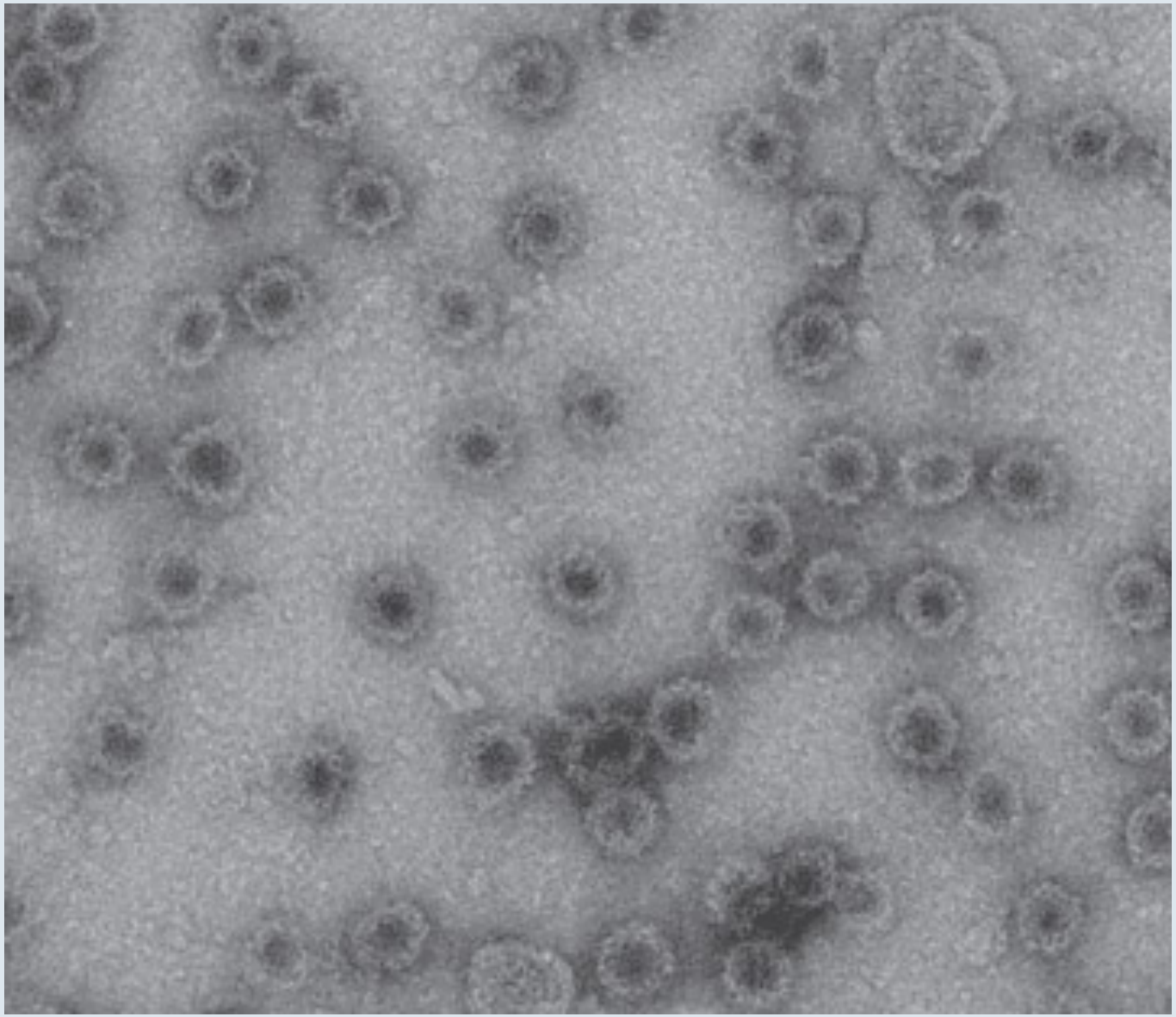


Replication-defective viruses may be generated through integration and deletion of a viral genome to generate a fused viral–cellular transcript that is packaged with a normal RNA genome. **Nonhomologous** recombination is necessary to generate the replication-defective transforming genome.

- **Retrotransposons** of the viral superfamily are transposons that mobilize via an RNA that does not form an infectious particle.
- Some **retrotransposons** directly resemble **retroviruses** in their use of LTRs. Others do not, and have no LTRs.

	LTR retrotransposons	non-LTR retroposons	SINES
Common types	Ty (*S. cerevisiae*) copia (*D.melanogaster*)	L1 (human) B1, B2 ID, B4 (mouse)	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

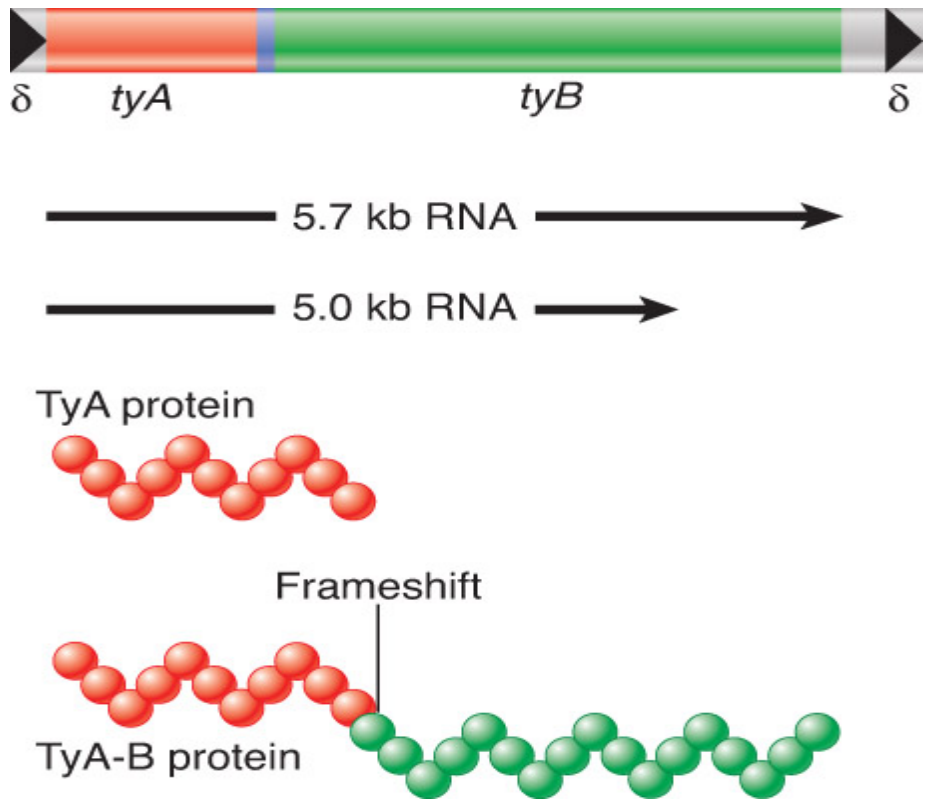
- Despite having an **RT activity**, LINES lack the **LTRs** of the viral superfamily and use a unique mechanism to prime the reverse transcription **rxn.**
- The non-viral superfamily may have originated from RNA sequences;
- **SINES** are derived from RNA polymerase III transcripts.



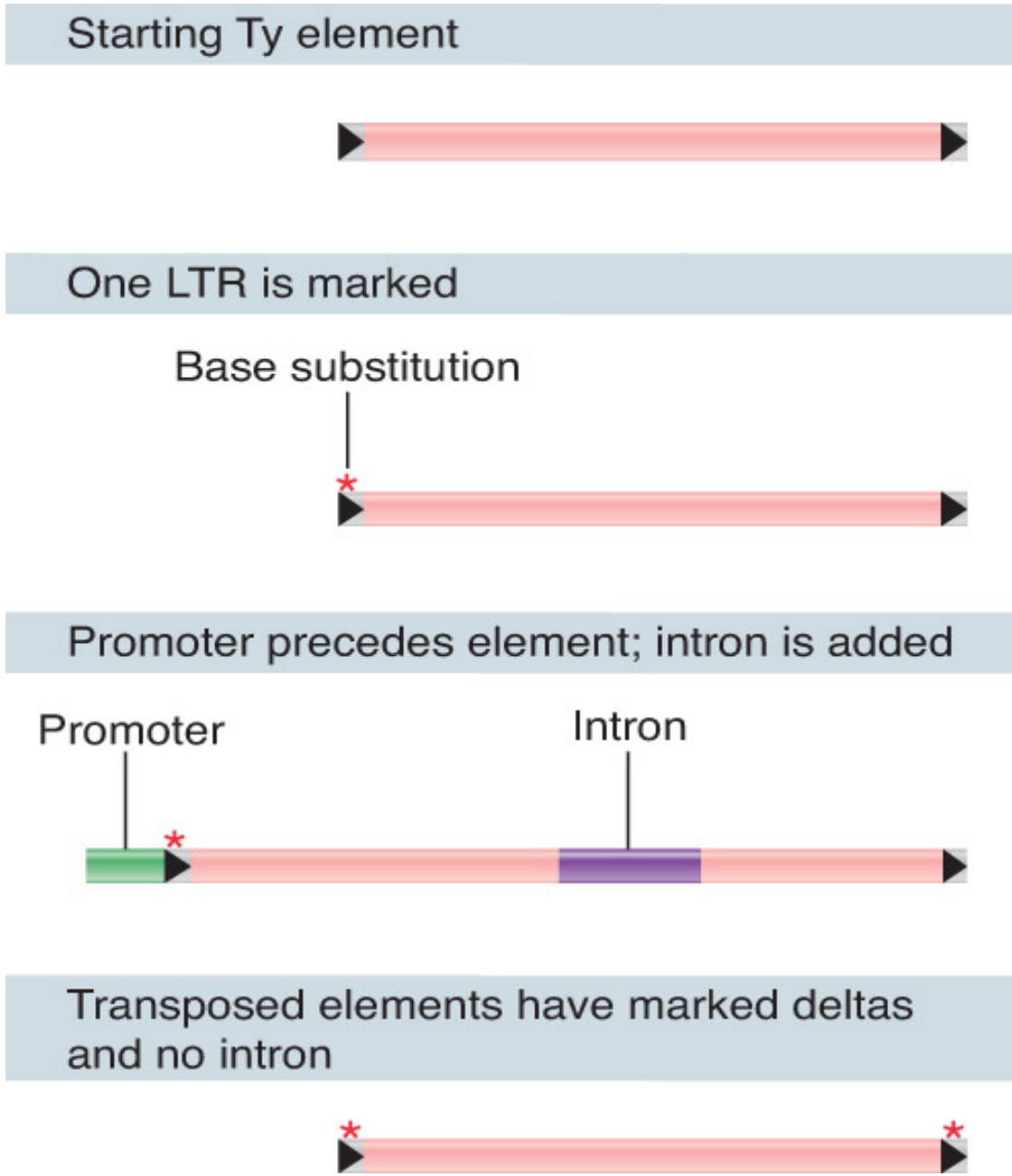
Ty elements in yeast generate virus-like particles.

Reprinted from J. Mol. Biol., vol. 292, H. A. AL-Khayat, et al., Yeast Ty retrotransposons..., pp. 65-73. Copyright 1999, with permission from Elsevier [http://www.sciencedirect.com/science/journal/00222836]. Photo courtesy of Dr. Hind A. AL-Khayat, Imperial College London, United Kingdom.

Ty elements in Yeast (~35 copies per genome) represent a third type of transpososition al insertion....**Retrotransposition**.



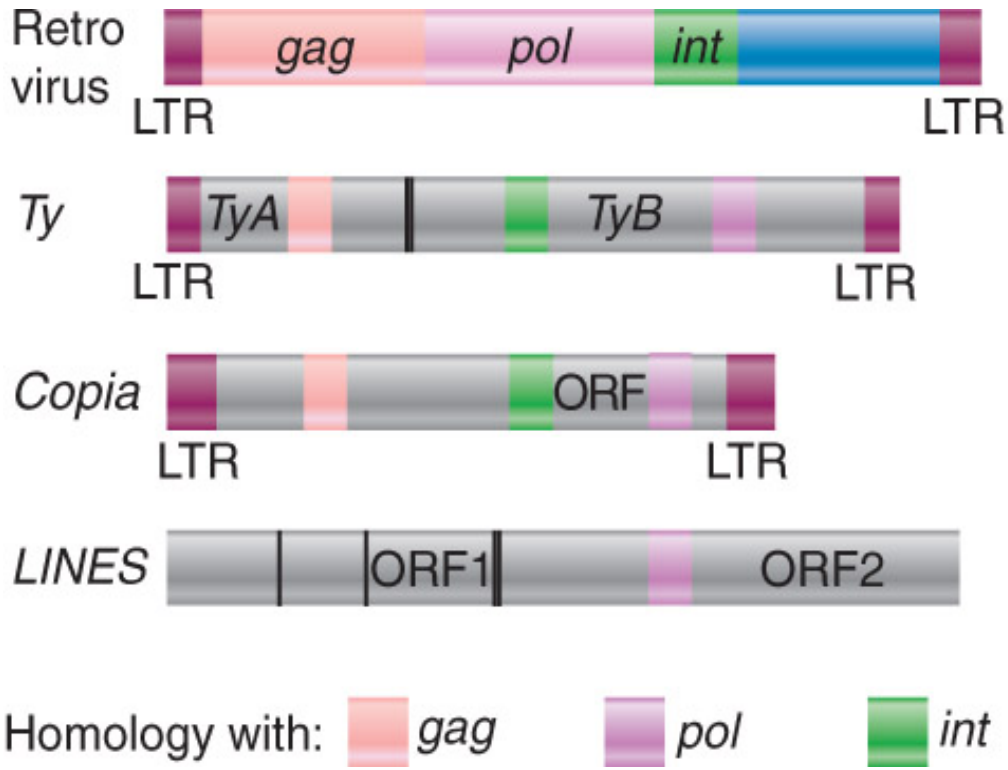
Ty elements ~5.1kbp in size and encode for a 300 bp **tandem repeats** (δ's), which can be seen scattered around Yeast genomes. They cause **5 bp repeats** in target site and transpose through an **RNA intermediate**!!! How can we know this?



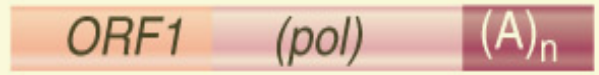
Ty transposes through a spliced RNA form

- Although **retroelements** that lack LTRs, also transpose *via* reverse transcriptase, they employ a distinct method of integration and are phylogenetically distinct from both **retroviruses** and **LTR retrotransposons**.

- Other elements can be found that were generated by an RNA-mediated transposition event, but they do not themselves code for enzymes that can catalyze transposition.



- **Retroelements** constitute almost half (48%) of the human genome.

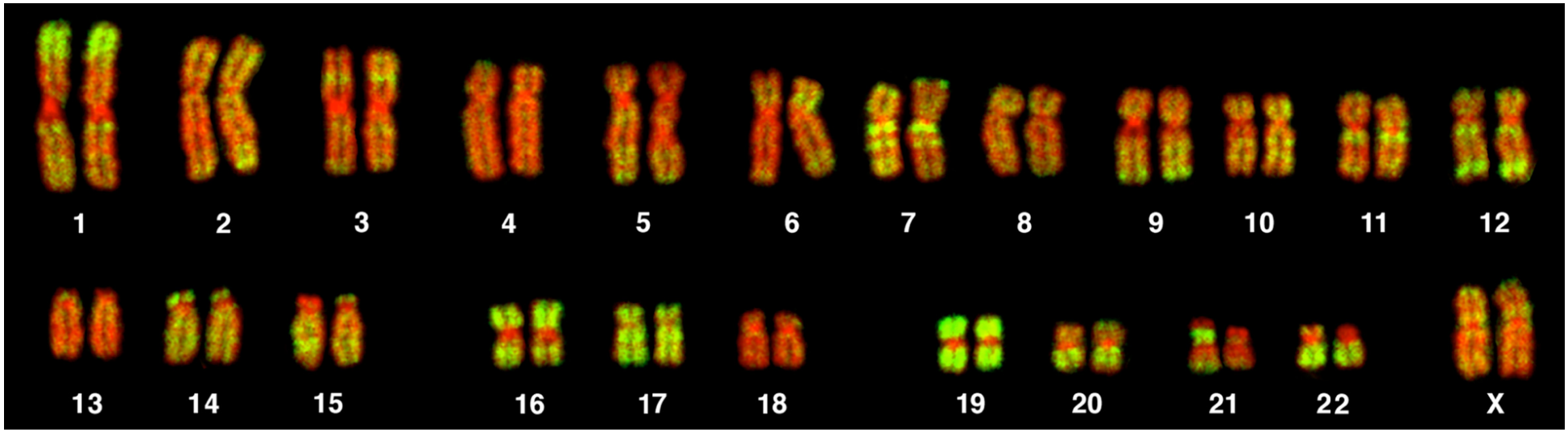
Element	Organization	Length (Kb)	Human genome	
			Number	Fraction
Retrovirus/LTR retrotransposon		1–11	450,000	8%
LINES (autonomous), e.g., L1		6–8	850,000	17%
SINES (nonautonomous), e.g., Alu		<0.3	1,500,000	15%
DNA transposon		2–3	300,000	3%

- LINES and SINES comprise a major part of the animal genome. They were originally defined by the existence of a large number of relatively short sequences that were related to one another.

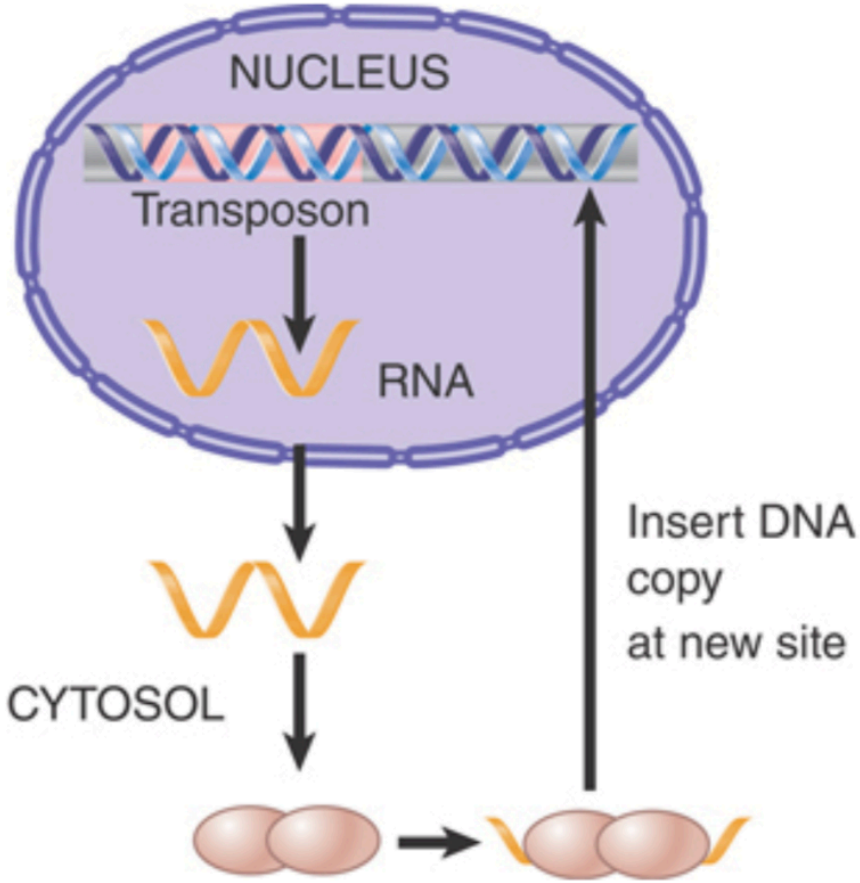
- They are described as interspersed nuclear elements because of their common occurrence and widespread distribution. **L1** = active human LINES; **ALU** = active human SINES

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%

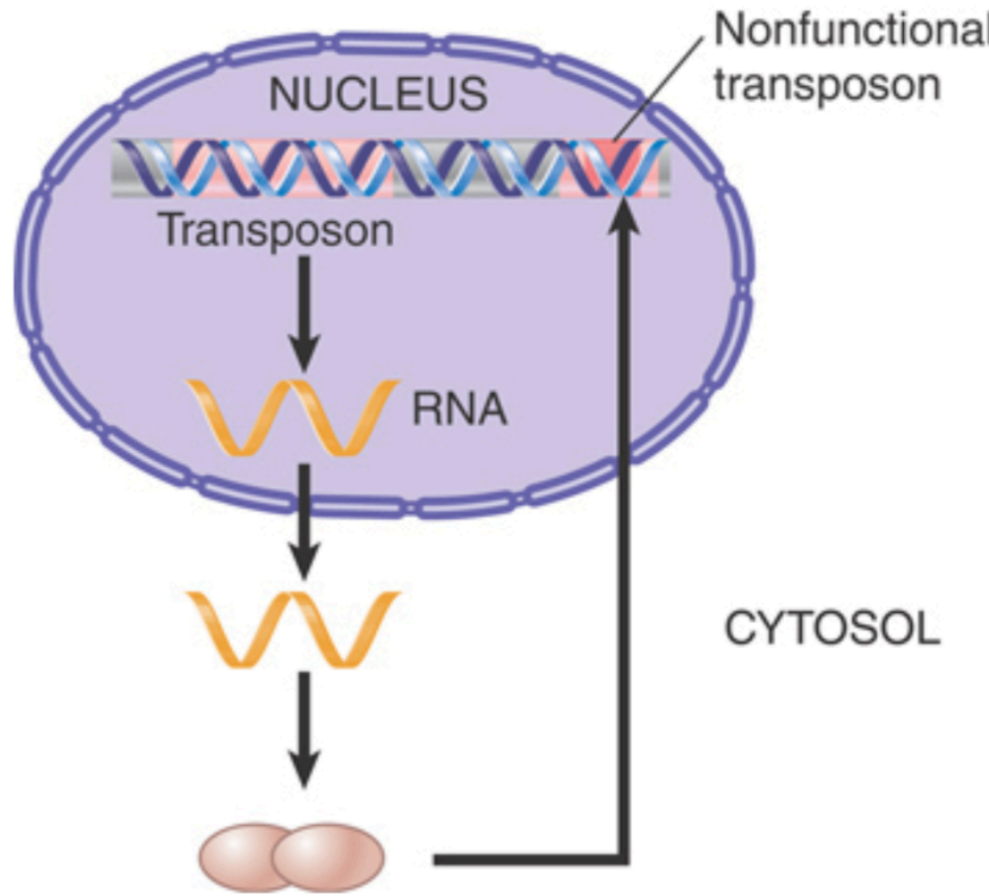
- **short-interspersed elements (SINEs)** – A major class of short (<500 bp) **nonautonomous** retrotransposons that occupy ~13 -15% of the human genome.
    - **Alu element** – One of a set of dispersed, related sequences, each ~300 bp long, in the human genome (members of the SINE family).



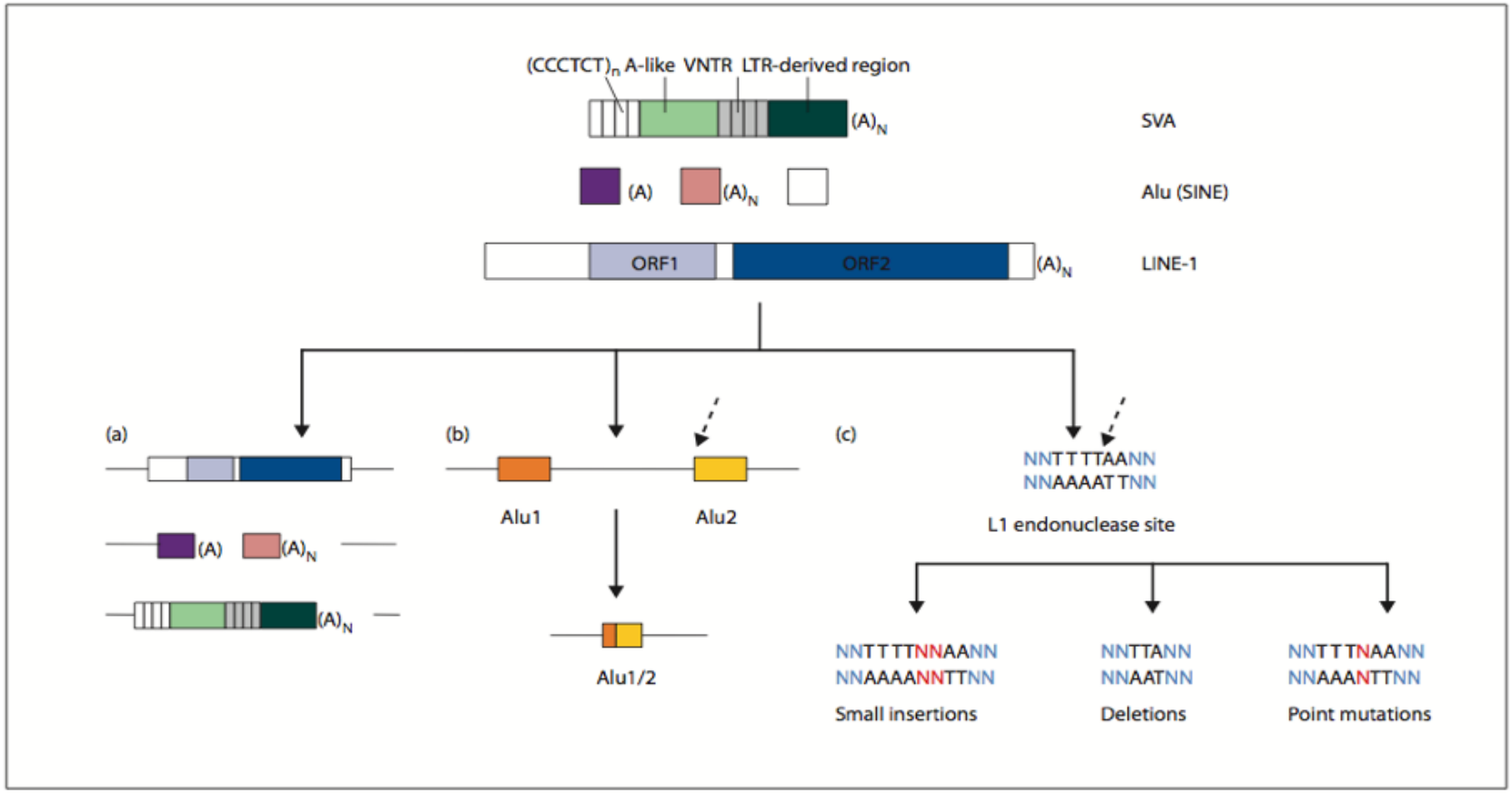
Karyotype from a female human lymphocyte (46, XX). Chromosomes were hybridized with a probe for Alu elements (green) and counterstained with TOPRO-3 (red). Alu elements were used as a marker for chromosomes and chromosome bands rich in genes.



A LINE is transcribed into an RNA that is translated into proteins that assemble into a complex with the RNA. The complex translocates to the nucleus, where it inserts a DNA copy into the genome



A transposon is transcribed into an RNA that is translated into proteins that move independently to the nucleus, where they act on any pair of inverted repeats with the same sequence as the original transposon.



L1 expression leads to different types of DNA damage.

Schematic structures of an **SVA element** (labeled SVA), showing the CCCTCT repeat, the **Alu derived (Alike) region**, the variable number tandem repeat (VNTR) region, and the long terminal repeat (LTR)derived region; an Alu element (labeled Alu (SINE)), showing left (purple) and right (pink) halves separated by the Arich region (A) and the variable length Atail ((A)n) followed by the 3' region (white), which has a variable length and sequence; and an L1 element (labeled LINE1), showing open reading frame (ORF)1 (light blue) and ORF2 (dark blue) and the 5' untranslated region, interORF region and 3' untranslated region (white).

(a) The typical insertion of these elements into the genome, which can lead to insertional mutagenesis. In breast cancer **BRCA1** and **BRCA2** are also known to be disrupted by TE insertion

(b) Dispersed repetitive elements such as **Alu elements** can undergo **non-allelic homologous recombination**, which can cause a deletion (shown) or duplication (not shown). The dashed arrow indicates the potential site of DNA damage by an L1 endonuclease that may help initiate these recombination events.

(c) Potential outcomes of the repair of the **L1 induced double strand breaks** (**DSBs**). The **L1** recognition site is in black; surrounding sequence is in blue; inserted nucleotides are in red. The associated changes are typical of what might be seen with repair of the DSB by **non homologous end joining (NHEJ) mechanisms**. It is also possible that the sites are simply re-ligated with no mutation occurring, or alternatively, these sites may cause recombination, as shown in **(b).**

Belancio *et al.: Genome Medicine* 2009, **1:97**

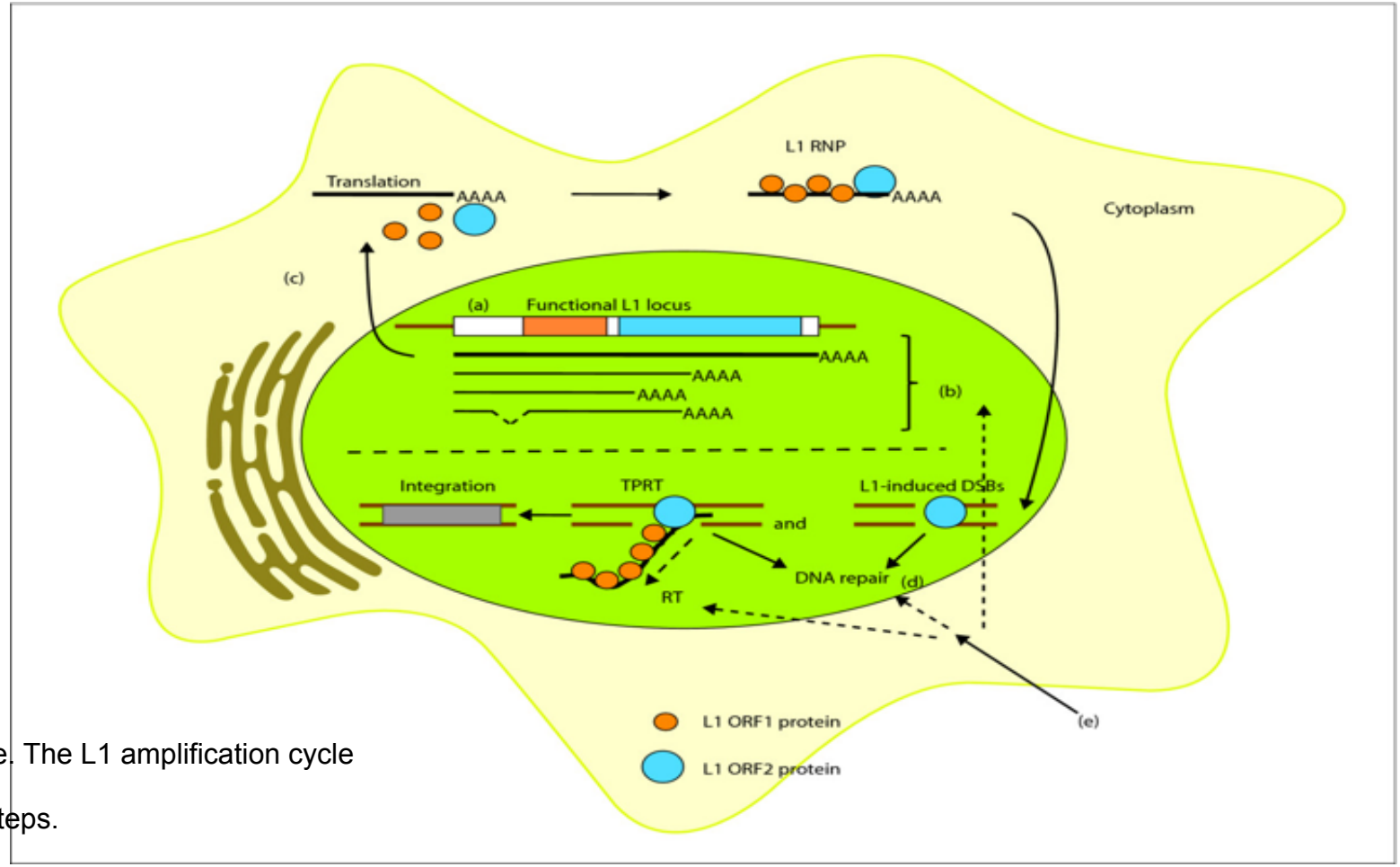


Figure 2

Modulators of the **L1** lifecycle. The L1 amplification cycle

can be divided into several steps.

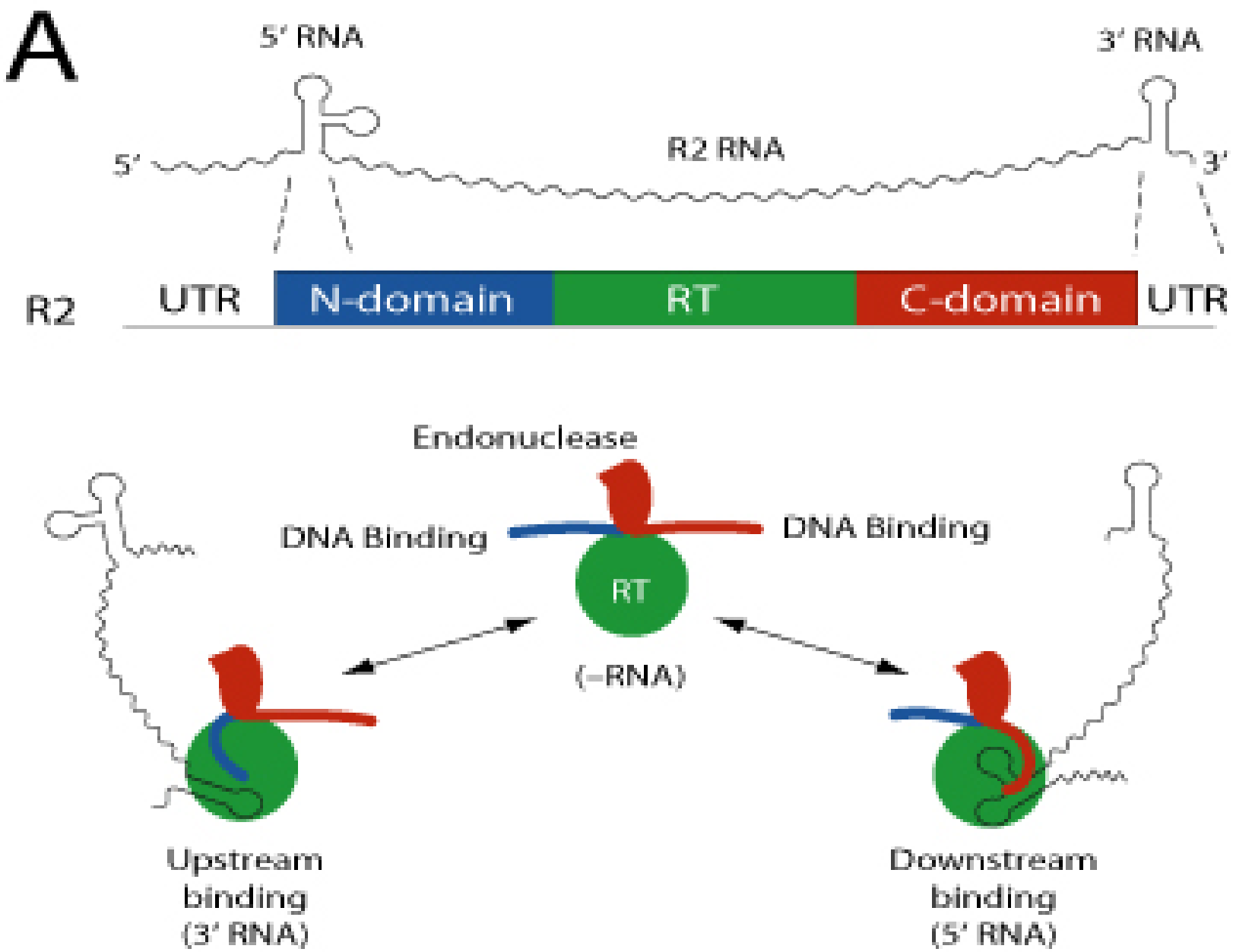
(a) Transcription. L1 amplification initiates with transcription, and regulation of L1 at this step can be modified by epigenetic modifications, **DNA methylation, and recruitment of transcription factors**.

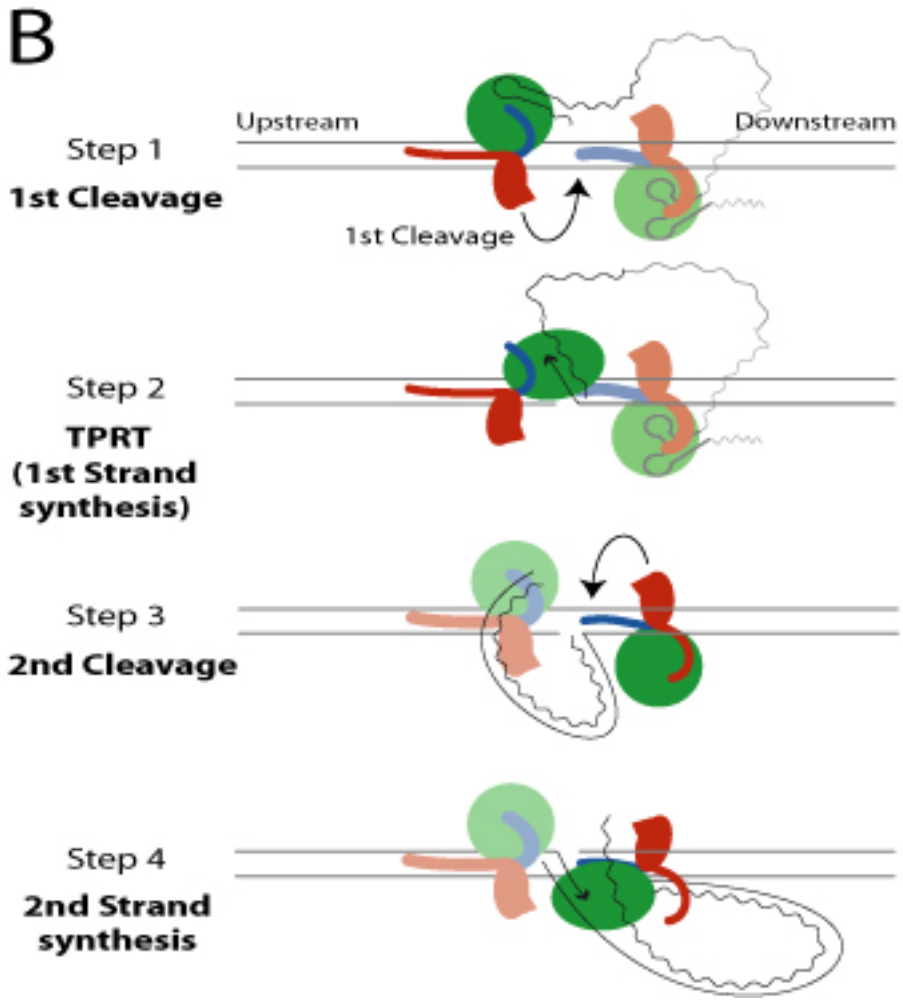
(b) Before leaving the nucleus, the number of retrocompetent full length **L1** transcripts can be reduced by RNA processing through premature polyadenylation and splicing.

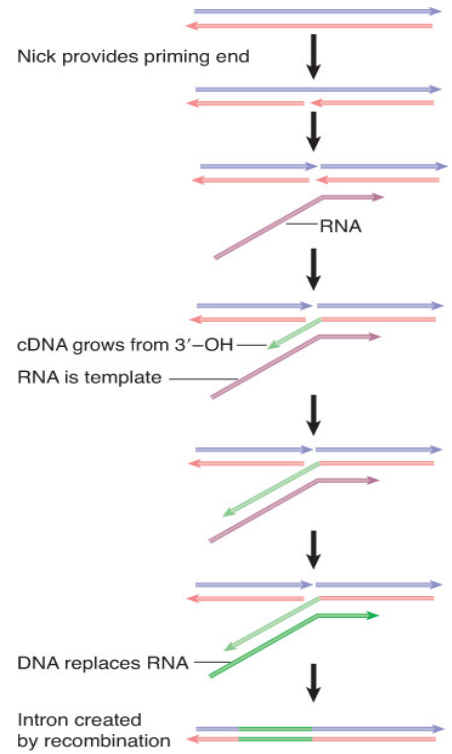
(c) Translation. Full length L1 enters the cytoplasm to be translated, producing **ORF1** and **ORF2** proteins for retrotransposition. The two proteins interact with the L1 transcript to form an **L1 ribonucleoprotein** particle (RNP). RNA interference can affect this step.

(d) Insertion of a new **L1** copy. The L1 RNP reaches the nucleus, where the DNA is cleaved by the L1 **ORF2** endonuclease activity. It is proposed that reverse transcription occurs through a process referred to as "**t**arget **p**rimed **r**everse **t**ranscription" (**TPRT**) [71]. The **L1 ORF2** reverse transcriptase activity generates the first strand of DNA. DNA repair proteins are likely to be involved in inhibiting the **L1** integration step.

(e) Effects of external stimuli. Ionizing radiation or heavy metals can affect L1 at multiple steps, such as transcriptional activation or altering DNA repair pathways.


A
5' RNA
3' RNA
5'
R2 RNA
3'
R2
UTR
N-domain
RT
C-domain
UTR
Endonuclease
DNA Binding
DNA Binding
RT
(–RNA)
Upstream
binding
(3' RNA)
Downstream
binding
(5' RNA)



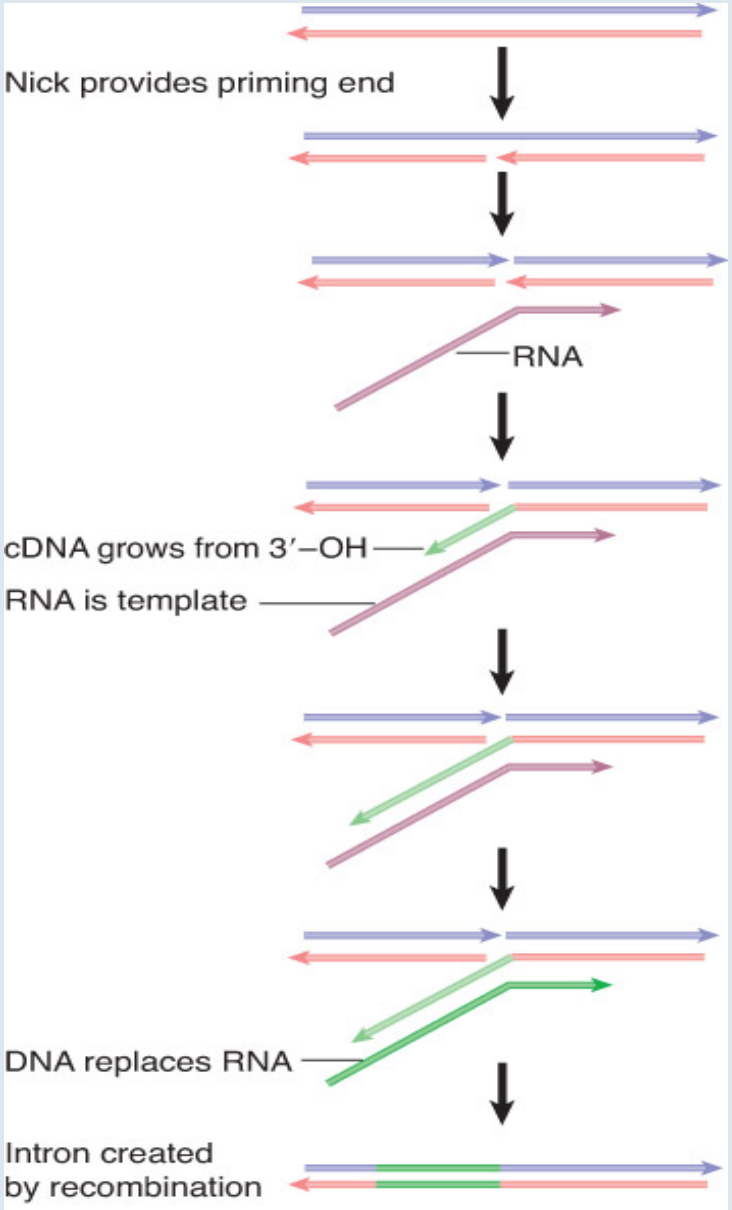


(Step 1) the **endonuclease** (red oval) from the upstream subunit is responsible for first strand cleavage.

(Step 2) The **RT** (green oval) of the upstream subunit catalyzes reverse transcription of the RNA template using the cleaved DNA target site as primer, a reaction we call **Target Primed Reverse Transcription**, **TPRT**.

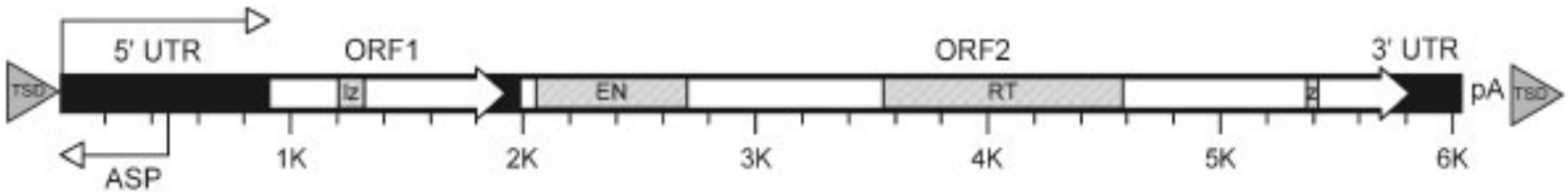
(Step 3) The downstream subunit cleaves the second DNA strand.

(Step 4) The downstream subunit provides the polymerase to perform **second strand TPRT** displacing the RNA template as it uses the first DNA strand as template.



Retrotransposition of non-LTR retroposons (such as L1) occurs by nicking the target to provide a primer for cDNA synthesis on an RNA template.

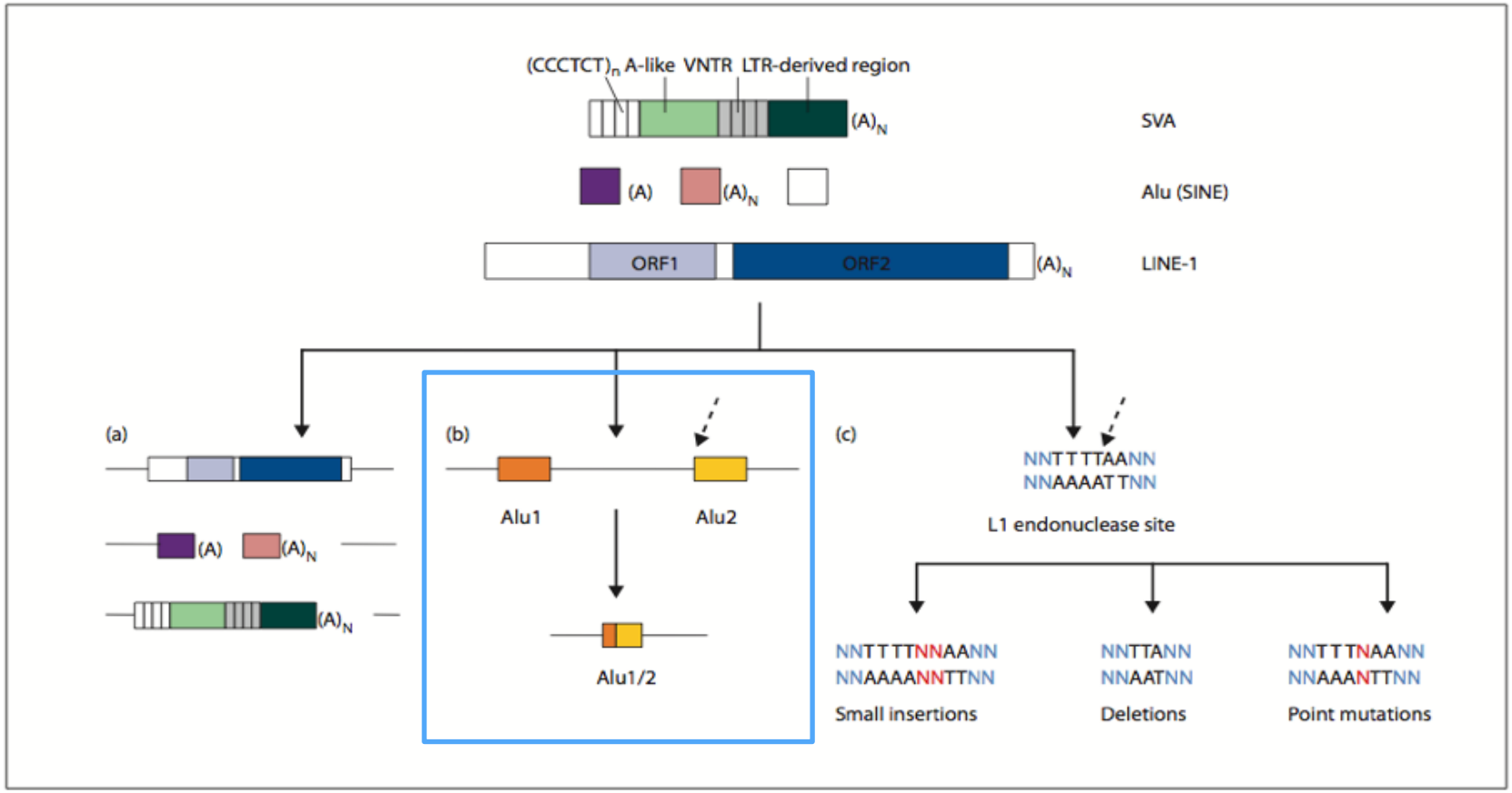
- **DNA transposons** and **LTR elements** are believed to be “extinct” in the human genome, but the average human carries approximately **80 – 100 potentially active L1 elements** in a diploid genome.
- **LINES** and **SINES** are NOT extinct
- The full length human **L1 retrotransposon** is 6kb and contains
  - a 910 bp 5’-UTR with bidirectional promoter activity
  - An **ORF1** region which encodes an **RN binding protein** with a **leucine zipper domain**
  - An **ORF2** region which encodes a 150 kDa protein with **endonuclease and reverse transcriptase activities**
  - A 3’-UTR which contains a functional **polyadenylation sequence**
- The L1 element is flanked by 2 to 20 bp target site duplications.



- The EN domain is thought to originate from a host endonuclease present in early eukaryotes. The element can move without a functional EN, but the endonuclease-independent integration is less efficient and occurs rarely.

- **L1** is active in BOTH the **germ line** and **somatic cells.**
- In mice and humans **L1** transcripts and **ORF1** protein have been detected in cells of the testes, placenta, cardiac tissue, and in neural precursor cells. **L1** retrotransposition has also been detected in primary culture of both human and rat cells.
- Frequency of transposition is thought to be determined by **methylation of CpG islands in** 5'-UTR -giving rise to repression of the **L1** 5'-UTR.

  Abherrant hypomethylation is believed to **increase** expression in some human tissues.
- Structurally, L1 appears to be an RNA **Pol III dependent transcript.**
  - Internal promoter similar to those associated with short noncoding RNAs produced by **RNA pol III**
- Yet other characteristics are more consistent with it being RNA Pol II dependent:
  - Long (6 kb) transcript
  - Encodes protein(s)
  - 3' polyadenylation



L1 expression leads to different types of DNA damage.

Schematic structures of an **SVA element** (labeled SVA), showing the CCCTCT repeat, the **Alu derived (Alike) region**, the variable number tandem repeat (VNTR) region, and the long terminal repeat (LTR)derived region; an Alu element (labeled Alu (SINE)), showing left (purple) and right (pink) halves separated by the Arich region (A) and the variable length Atail ((A)n) followed by the 3' region (white), which has a variable length and sequence; and an L1 element (labeled LINE1), showing open reading frame (ORF)1 (light blue) and ORF2 (dark blue) and the 5' untranslated region, interORF region and 3' untranslated region (white).

(a) The typical insertion of these elements into the genome, which can lead to insertional mutagenesis. In breast cancer **BRCA1** and **BRCA2** are also known to be disrupted by TE insertion

(b) Dispersed repetitive elements such as **Alu elements** can undergo **non-allelic homologous recombination**, which can cause a deletion (shown) or duplication (not shown). The dashed arrow indicates the potential site of DNA damage by an L1 endonuclease that may help initiate these recombination events.

(c) Potential outcomes of the repair of the **L1 induced double strand breaks** (**DSBs**). The **L1** recognition site is in black; surrounding sequence is in blue; inserted nucleotides are in red. The associated changes are typical of what might be seen with repair of the DSB by **non homologous end joining (NHEJ) mechanisms**. It is also possible that the sites are simply re-ligated with no mutation occurring, or alternatively, these sites may cause recombination, as shown in **(b).**

Belancio *et al.: Genome Medicine* 2009, **1:97**

2005: Callinan Pauline A; Wang Jianxin; Herke Scott W; Garber Randall K; Liang Ping; Batzer Mark A
Alu retrotransposition-mediated deletion.
Journal of molecular biology 2005;348(4):791-800

Alu Retrotransposition-mediated Deletion

Pauline A. Callinan[a, †], Jianxin Wang[b, †], Scott W. Herke[a, †], Randall K. Garber[a], Ping Liang[b] and Mark A. Batzer[a,]

[a]Department of Biological Sciences, Biological Computation and Visualization Center, Center for BioModular Multi-Scale Systems, Louisiana State University, 202 Life Sciences Building, Baton Rouge, LA 70803, USA

[b]Department of Cancer Genetics, Roswell Park Cancer Institute, Elm and Carlton Streets, Buffalo, NY 14263, USA

Received 11 January 2005; revised 17 February 2005; accepted 18 February 2005. Edited by J. Karn. Available online 17 March 2005.

Alu repeats contribute to genomic instability in primates *via* insertional and recombinational mutagenesis. Here, we report an analysis of *Alu* element-induced genomic instability through a novel mechanism termed retrotransposition-mediated deletion, and assess its impact on the integrity of primate genomes. For human and chimpanzee genomes, we find evidence of 33 retrotransposition-mediated deletion events that have eliminated approximately 9000 nucleotides of genomic DNA. Our data suggest that, during the course of primate evolution, *Alu* retrotransposition may have contributed to over 3000 deletion events, eliminating approximately 900 kb of DNA in the process. Potential mechanisms for the creation of *Alu* retrotransposition-mediated deletions include L1 endonuclease-dependent retrotransposition, L1 endonuclease-independent retrotransposition, internal priming on DNA breaks, and promiscuous target primed reverse transcription. A comprehensive analysis of the collateral effects by *Alu* mobilization on all primate genomes will require sequenced genomes from representatives of the entire order.

Keywords: short interspersed elements; target primed reverse transcription

Abbreviations used: SINE, short interspersed element; LINE, long interspersed element; ARD, *Alu* retrotransposition-mediated deletion; HuARD, human-specific ARD; TSD, target site duplication; pTPRT, promiscuous target primed reverse transcription

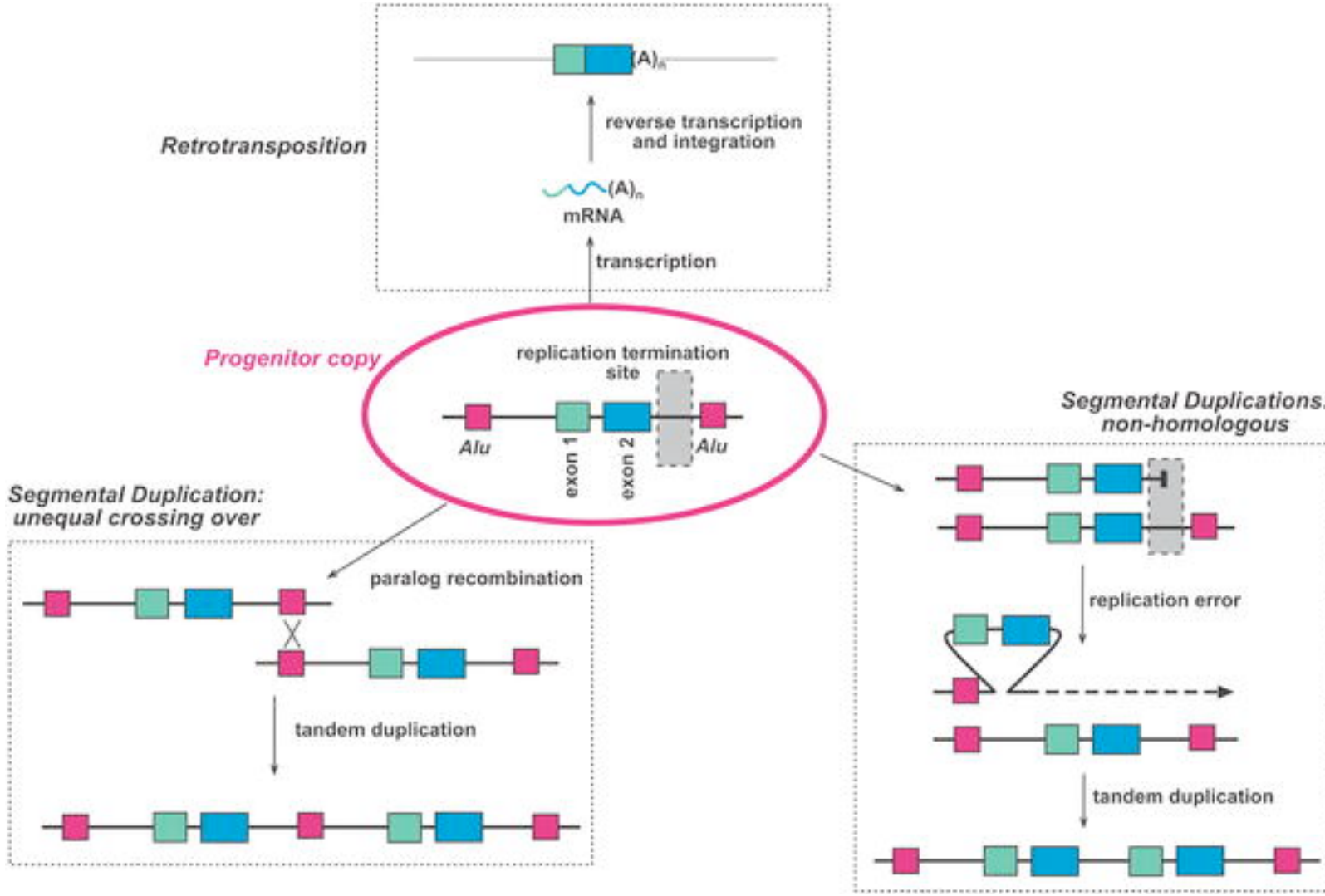
Article Outline

PRIMER OPEN ACCESS

Gene Duplication: The Genomic Trade in Spare Parts

Matthew Hurles

Citation: Hurles M (2004) Gene Duplication: The



A two-exon gene is flanked by two ***Alu* elements** and a neighbouring replication termination site. Recombination between the two *Alu* elements leads to a tandem duplication event, as does a replication error instigated by the replication termination site.

Retrotransposition of the mRNA of the gene leads to the random integration of an intron-less paralogue at a distinct genomic location.