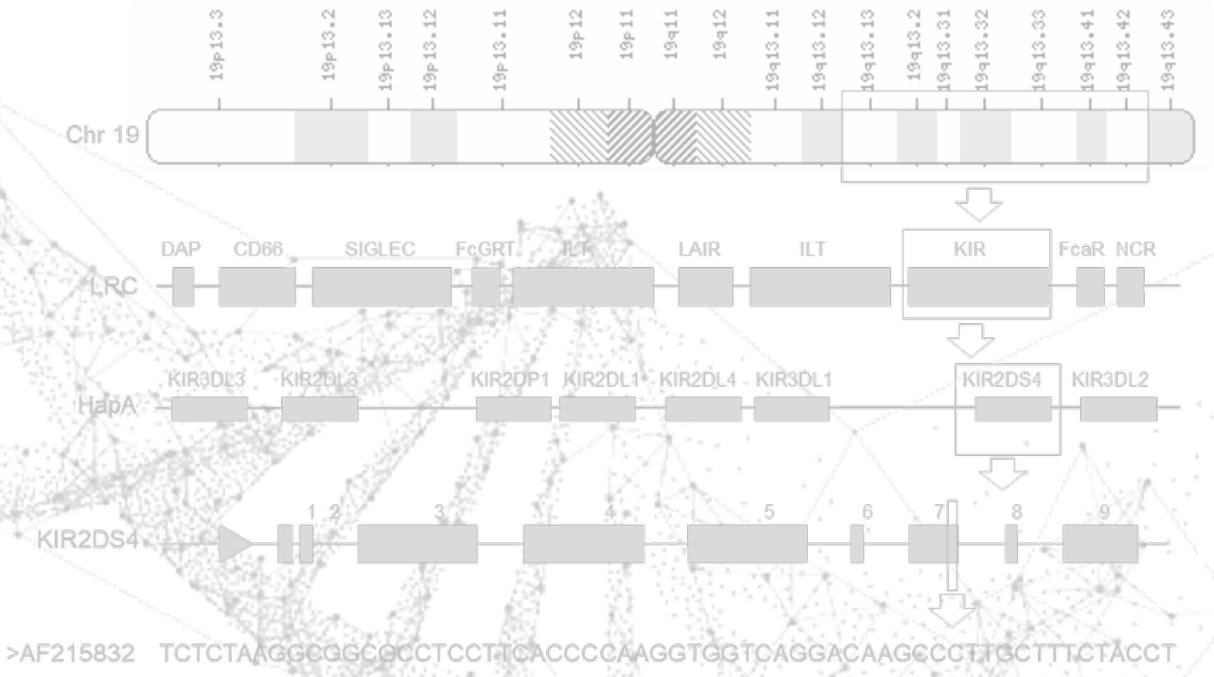


Transposable elements

San Luis Potosi State University (UASLP) Mexico
Molecular Biology Course, Faculty of Medicine post-graduate program

Dr. Christian A. García-Sepúlveda
Viral & Human Genomics BSL-3 Laboratory

Last updated October 09, 2025 v2

Evolution of genomes

Genomes evolve both by acquiring new sequences or by rearranging existing sequences.

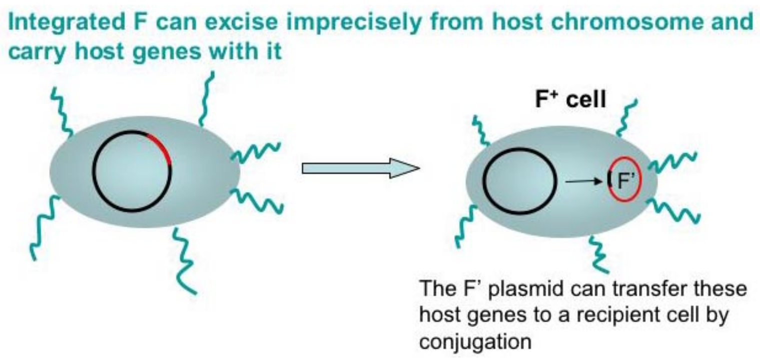
Acquisition of new sequences:

- Plasmids
- Virus (especially integrating and retrotranscriptive viruses.

Rearrangements sponsored by internal genomic events.

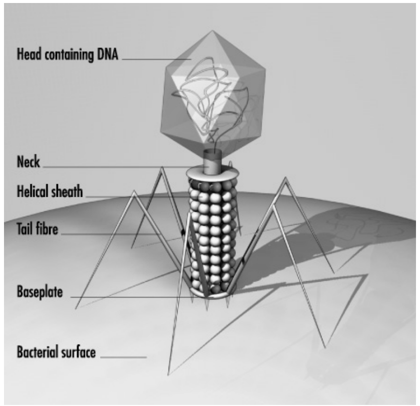
- Unequal recombination (non-recirpocal).
- Duplication or rearrangement of loci (Clusters & repeats).
- Duplication of sequences within a genome gives rise to further duplication.
- Somatic recombination.
- Chromosomal crossing-over.

Acquisition of new sequences



Plasmids allow bacterial conjugation.

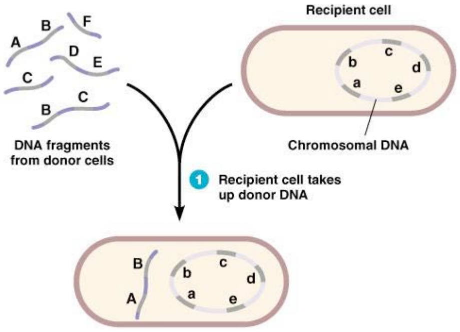
- Extrachromosomal elements move information horizontally by mediating the transfer of short lengths of genetic material.



Phages spread by infection.

- Both plasmids and phages occasionally transfer host genes along with their own replicon.

Transformation



- Direct transfer of DNA occurs between some bacteria by means of transformation.

Rearrangement of existing sequences

Another major cause of variation is provided by transposable elements or transposons:

These are discrete sequences in the genome that are mobile & able to transport themselves to other locations within the genome.

Found in both eukaryotes & prokaryotes.

Selfish DNA with the sole purpose of autoreplication.

Relationship of the transposon to the genome resembles that of a parasite with its host.

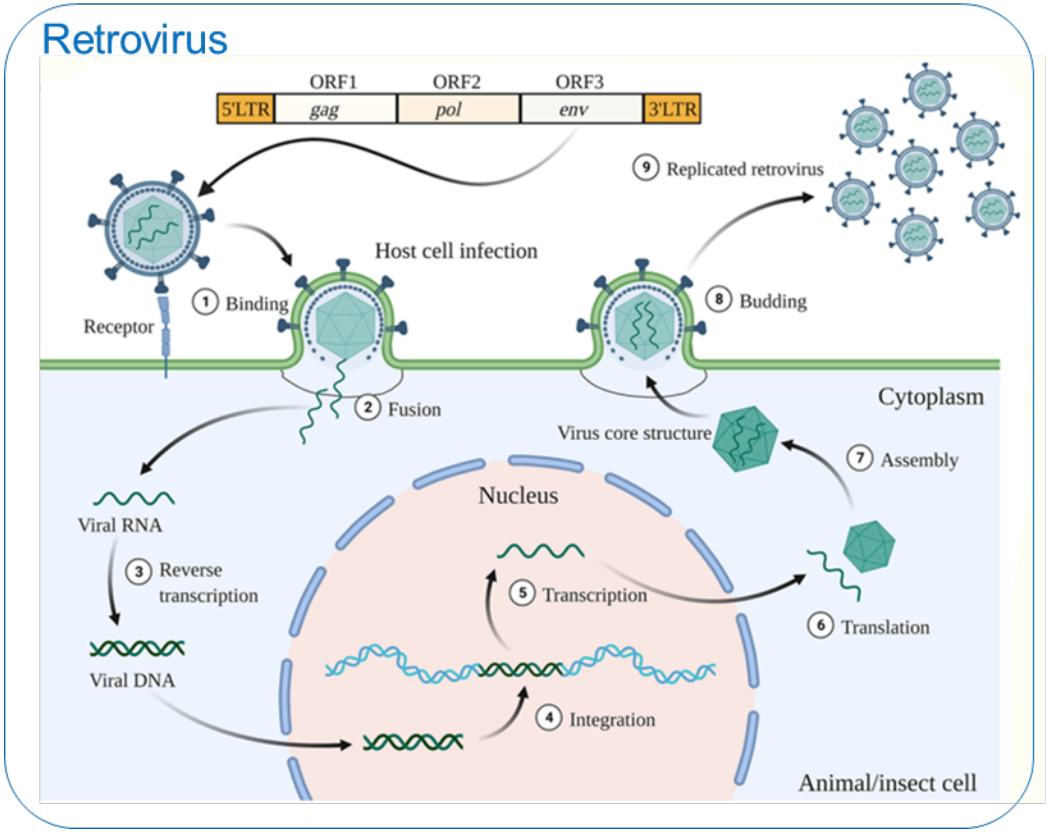
The propagation of an element by transposition is balanced by the harm done if a transposition event inactivates a necessary gene.

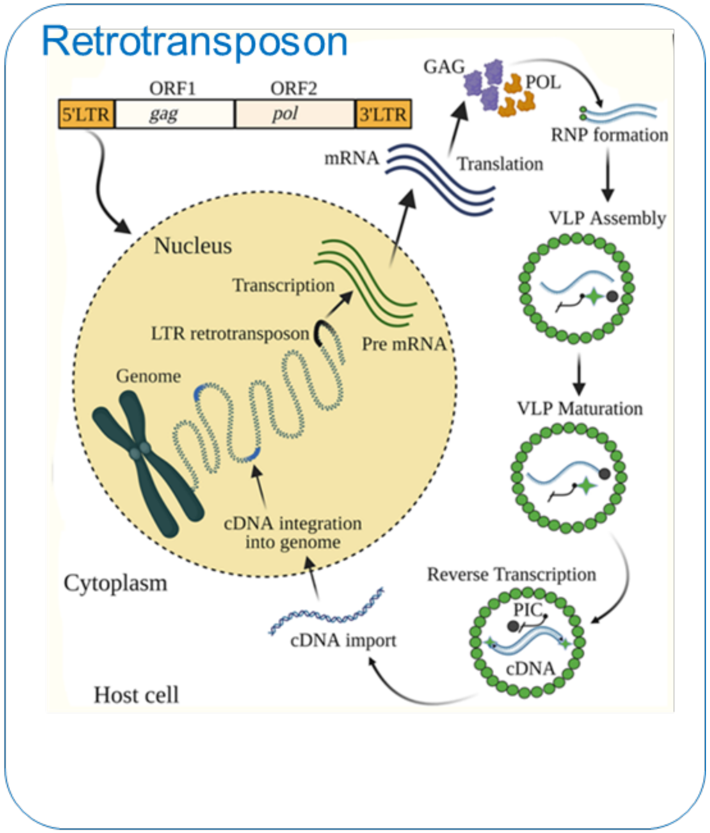
Any transposition event conferring a selective advantage will lead to preferential survival of the genome harboring the transposon !

Transposons are not virus

Transposons do not utilize an independent form (such as virus or plasmid DNA).

Move directly from one site of the genome to another.





Ramakrishnan M, et al. Plant Cell Rep. 2023 Jan

Transposons

Unlike other processes involved in genome restructuring, transposition does not rely on homology between donor and recipient sites.

Sometimes transfer contiguous host sequences to new sites elsewhere within the same genome as they move.

They are an internal counterpart to the vectors that can transport sequences from one genome to another.

Can promote rearrangements of the genome, directly or indirectly:

- Directly: The transposition event itself may cause deletions or inversions or lead to the movement of a host sequence to a new location.

- Indirectly: Transposon sequences serve as substrates for cellular recombination systems by functioning as "portable regions of homology";

Two copies of a transposon at different locations may provide sites for reciprocal recombination resulting in deletions, insertions, inversions, or translocations.

Clases of mobile elements or transposons

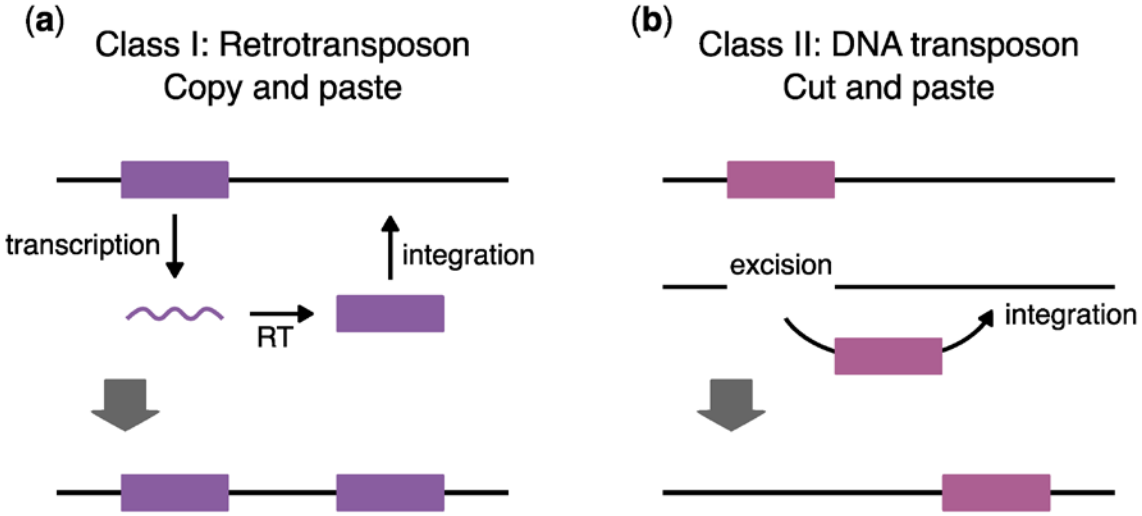
They may provide the major source of mutations in the genome!

DNA transposons

- Exist as sequences of DNA coding for proteins that are able directly to manipulate DNA so as to propagate themselves within the genome.

RNA transposons

- They are related to retroviruses and move as a consequence of their ability to make DNA copies of their RNA transcripts, the DNA copies then become integrated at new sites in the genome.



Insertion sequences

Transposable elements were first identified in the form of spontaneous insertions in bacterial operons.

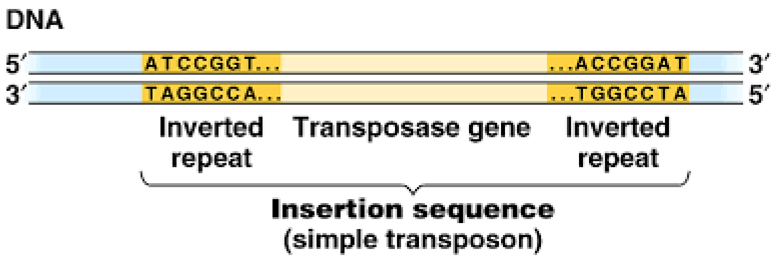
Such an insertion prevents transcription and/or translation of the gene in which it is inserted.

The first transposons that were discovered were simple and called insertion sequences (IS).

Each type is given the prefix IS, followed by a number that identifies the type. The original classes were IS1-4, more classes have been discovered since.

Insertion into a particular site described with a double colon:

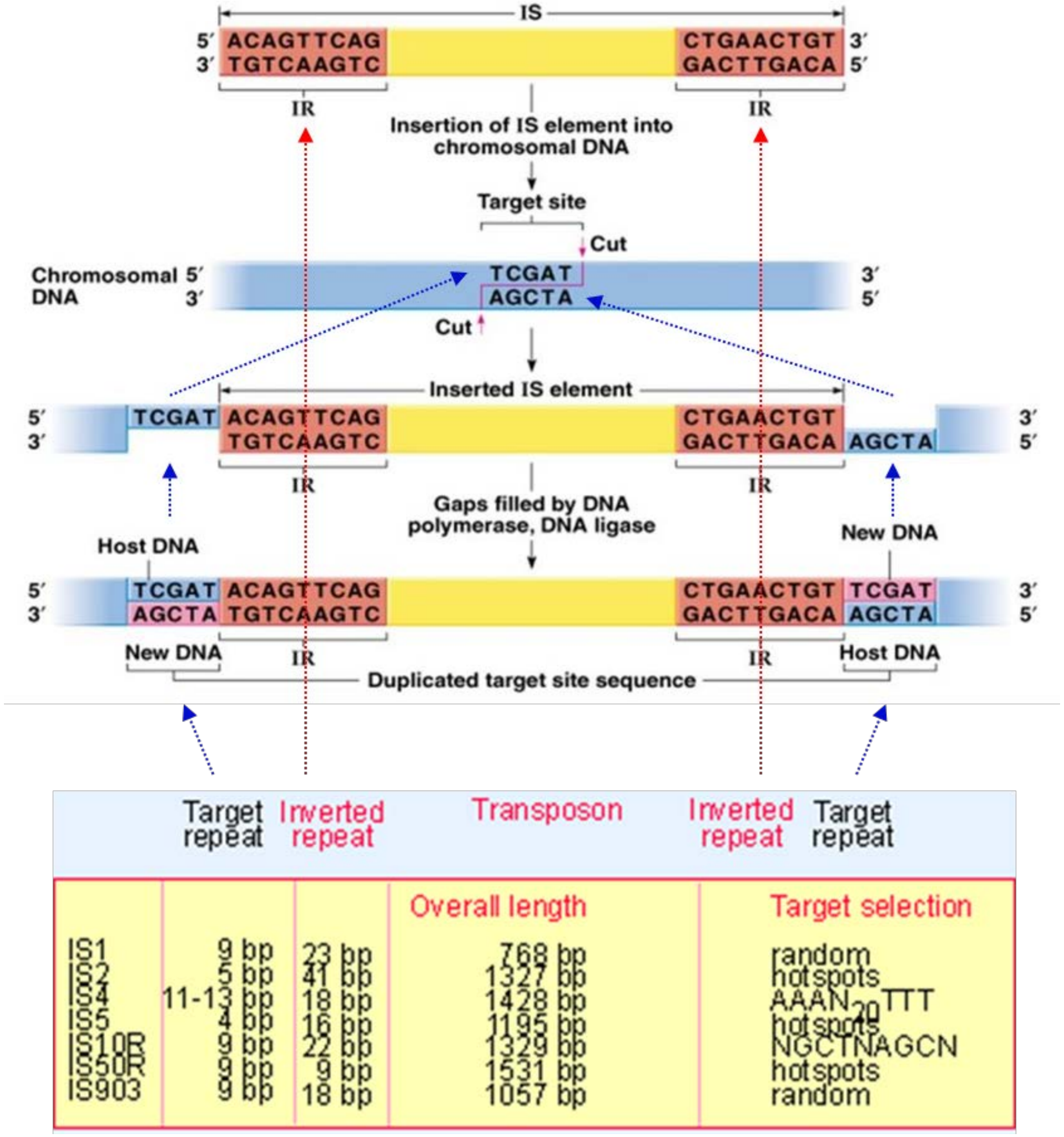
l::IS1 = An IS1 element inserted into phage lambda.



Insertion sequences (IS)

IS elements are normal constituents of bacterial chromosomes & plasmids.

- A standard strain of E. coli contains <10 copies of the more common IS elements.
- The IS elements code only for the proteins needed to sponsor its own transposition (Transposase).
- Each IS element is different in sequence, but there are common organizational features.



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

Insertion sequence organizational features

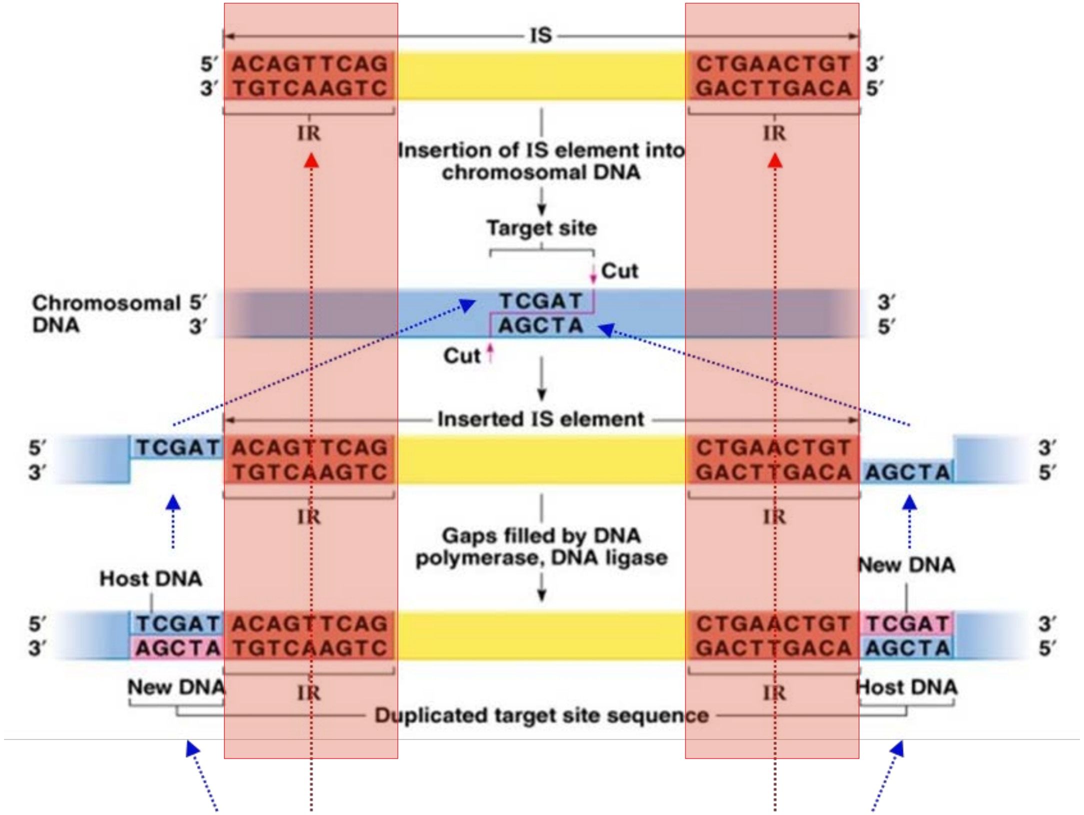
An IS element ends in short inverted terminal repeats which are not identical but closely related.

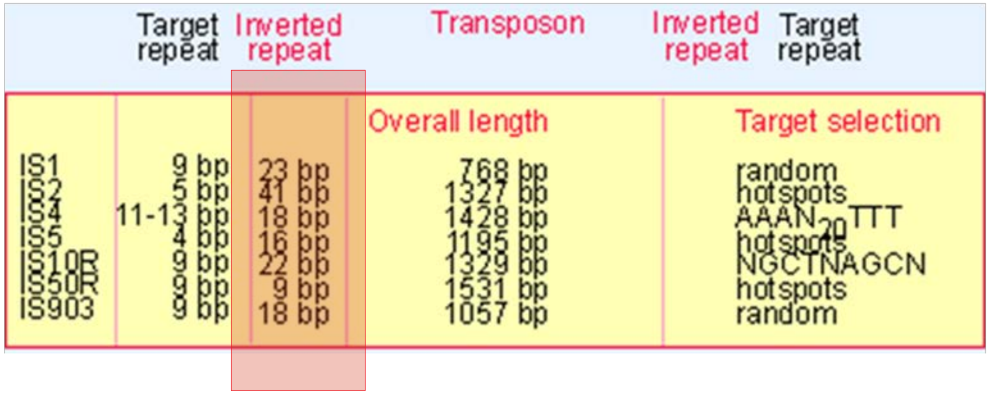
Inverted repeats recognized by TRANSPOSASE.

Ensure that the same sequence is encountered proceeding toward the element from any direction.

Inverted repeat recognition is common to transposition events sponsored by all transposons.

Cis-acting mutations of these ends prevent transposition.





Target repeat | Inverted repeat | Transposon | Inverted repeat | Target repeat

			Overall length	Target selection
IS1	9 bp	23 bp	768 bp	random
IS2	5 bp	41 bp	1327 bp	hotspots
IS4	11-13 bp	18 bp	1428 bp	AAAN$_{20}$TTT
IS5	4 bp	16 bp	1195 bp	hotspots
IS10R	9 bp	22 bp	1329 bp	NGCTNAGCN
IS50R	9 bp	9 bp	1531 bp	hotspots
IS903	9 bp	18 bp	1057 bp	random

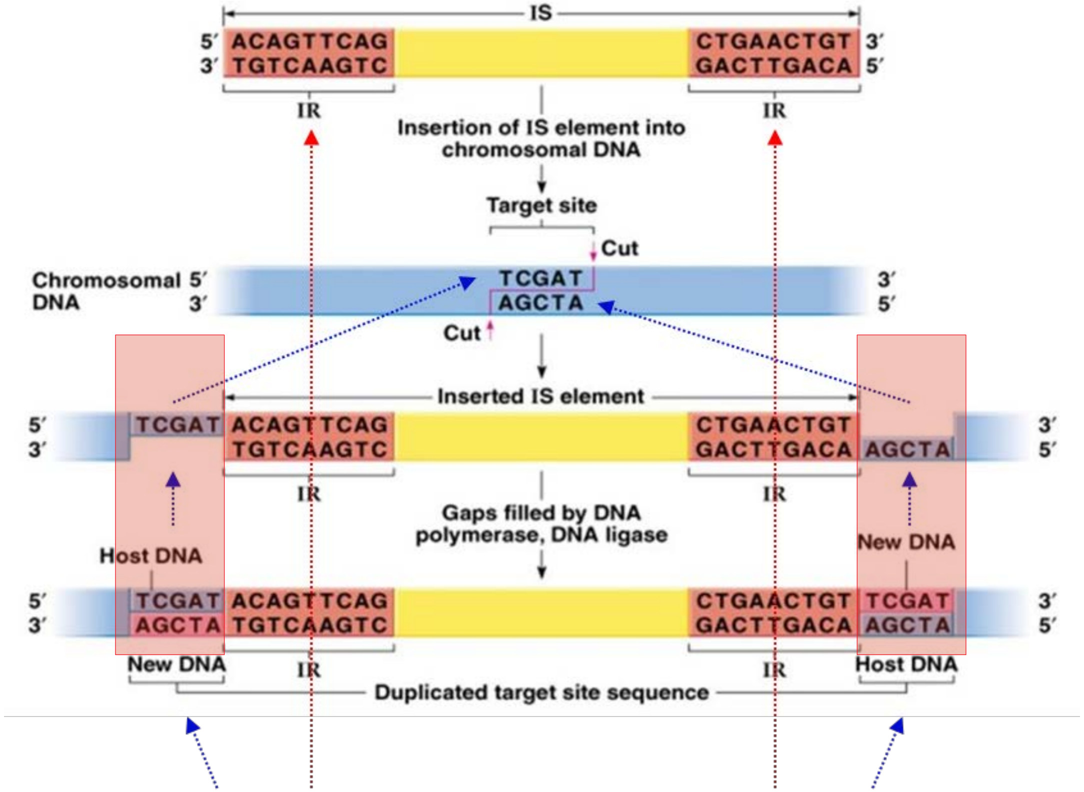
Insertion sequence organizational features

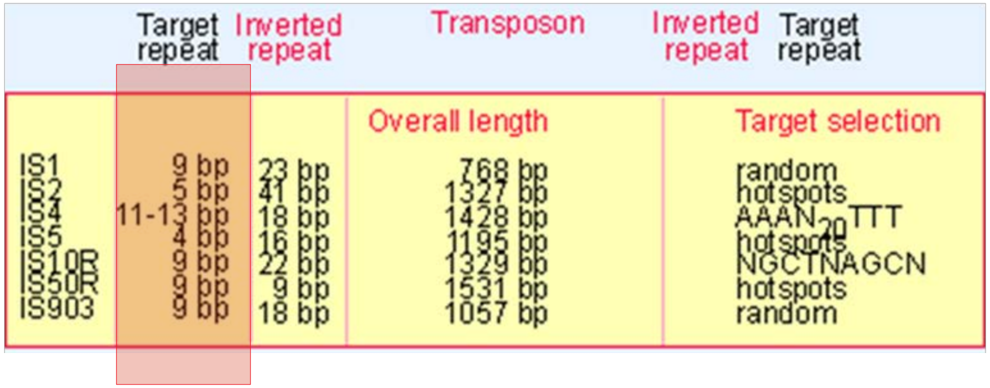
When an IS element transposes, a sequence of host DNA at the site of insertion is duplicated: DIRECT REPEATS.

IS DNA is always flanked by very **short direct repeats** with the same direction.

Pre-transposition genomic sequences exhibit only one of these “repeats” (i.e.: TCGAT).

Post-transposition sequence will have this sequence duplicated and flanking the transposon sequence.





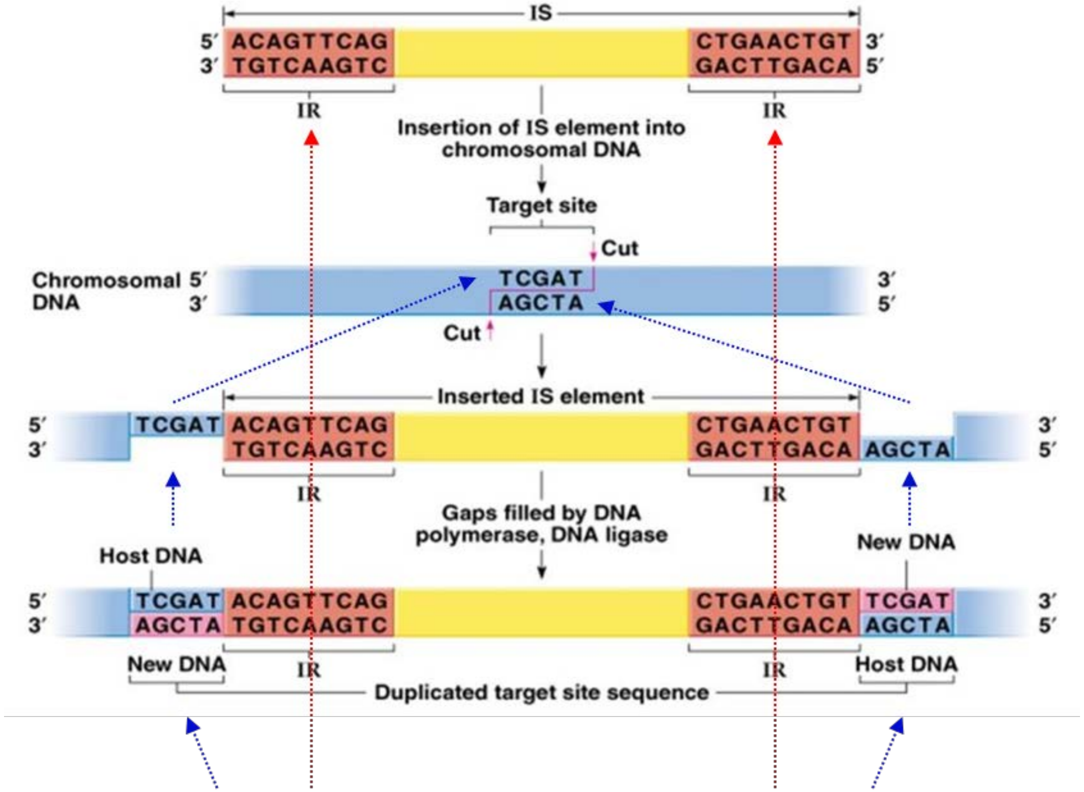
	Target repeat	Inverted repeat	Transposon Overall length	Target selection
IS1	9 bp	23 bp	768 bp	random
IS2	5 bp	41 bp	1327 bp	hotspots
IS4	11-13 bp	18 bp	1428 bp	AAAN$_{20}$TTT
IS5	4 bp	16 bp	1195 bp	hotspots
IS10R	9 bp	22 bp	1329 bp	NGCTNAGCN
IS50R	9 bp	9 bp	1531 bp	hotspots
IS903	9 bp	18 bp	1057 bp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

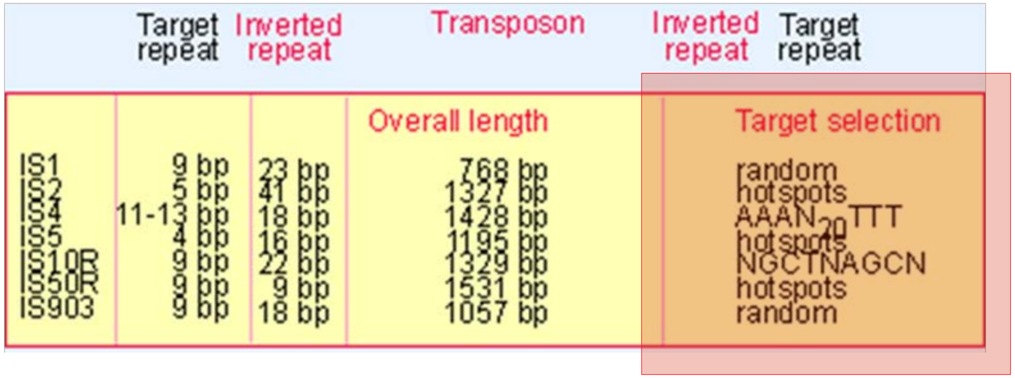
Insertion sequence organizational features

IS display a characteristic structure in which its ends possess inverted terminal repeats while the adjacent ends of the flanking host DNA possess short direct repeats.

This type of organization is taken to be diagnostic and suggest that the sequence originated in a transposition event.

IS elements insert at a variety of sites within host DNA, some show preference for particular hotspots

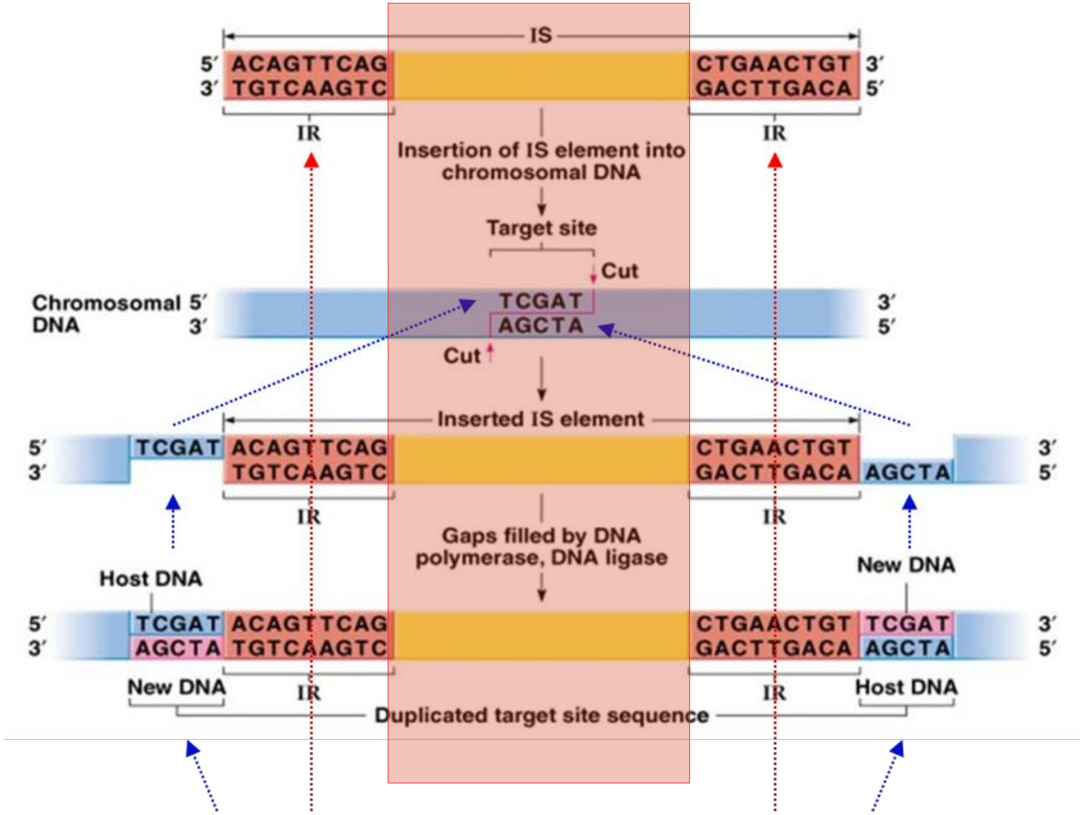


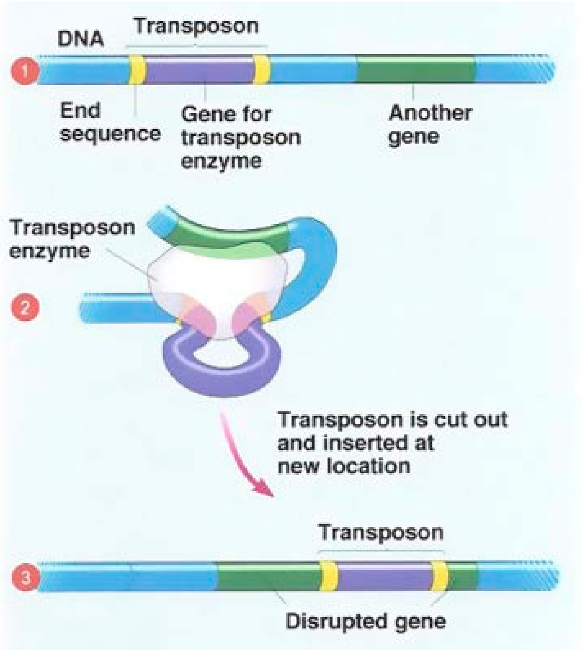


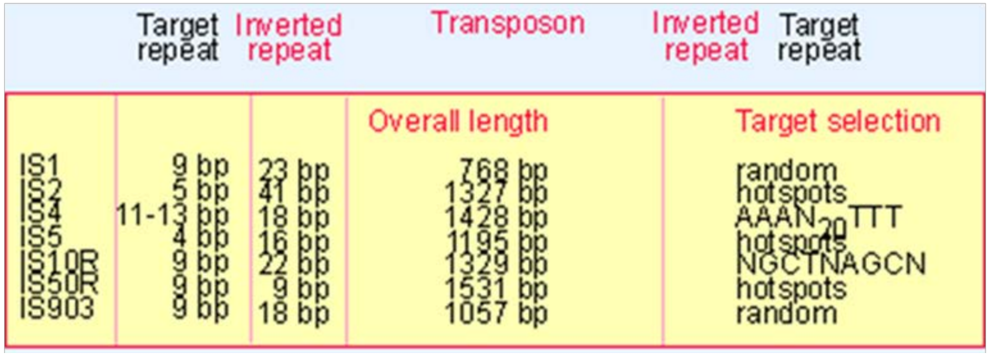
	Target repeat	Inverted repeat	Transposon	Inverted repeat Target repeat
			Overall length	Target selection
IS1	9 bp	23 bp	768 bp	random
IS2	5 bp	41 bp	1327 bp	hotspots
IS4	11-13 bp	18 bp	1428 bp	$AAAN_{20}TTT$
IS5	4 bp	16 bp	1195 bp	hotspots
IS10R	9 bp	22 bp	1329 bp	NGCTNAGCN
IS50R	9 bp	9 bp	1531 bp	hotspots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

Insertion sequence organizational features

All IS elements (except IS1) contain a single long coding region for transposase starting after the inverted repeat at one end and terminating before or within the inverted repeat at the other end.





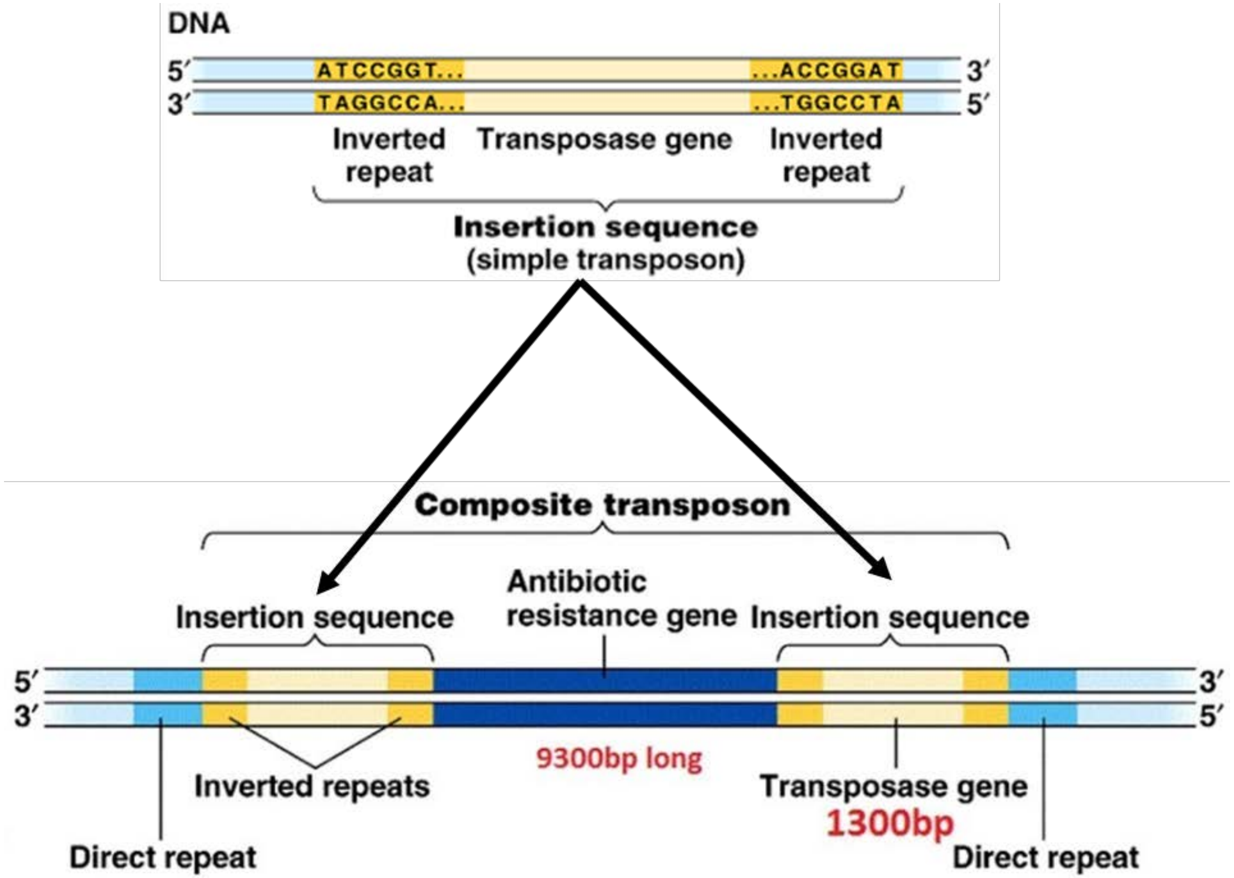


	Target repeat	Inverted repeat	Transposon	Inverted repeat	Target repeat
			Overall length		Target selection
IS1	9 bp	23 bp	768 bp		random
IS2	5 bp	41 bp	1327 bp		hotspots
IS4	11-13 bp	18 bp	1428 bp		AAAN$_{20}$TTT
IS5	4 bp	16 bp	1195 bp		hotspots
IS10R	9 bp	22 bp	1329 bp		NGCTNAGCN
IS50R	9 bp	9 bp	1531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

[Transposase Chapter in Nucleic Acids and Molecular Genetics, Science Direct](#)

Evolution of composite transposons

What would happen if an IS transposed near the original position... or if two IS sequences were separated by genomic DNA?



[Transposase Chapter in Nucleic Acids and Molecular Genetics, Science Direct](#)

Composite transposons (Tn)

Code for more than proteins involved in transposition.

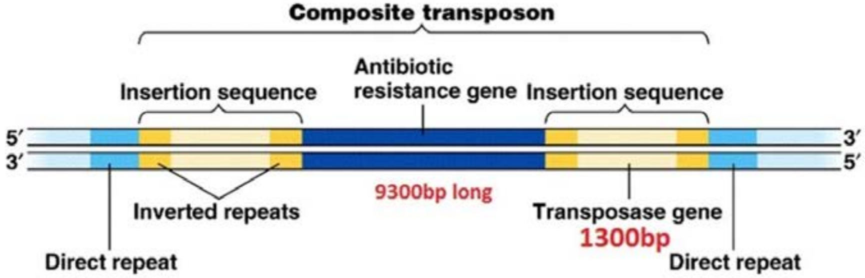
A central genomic “core” flanked by two IS.

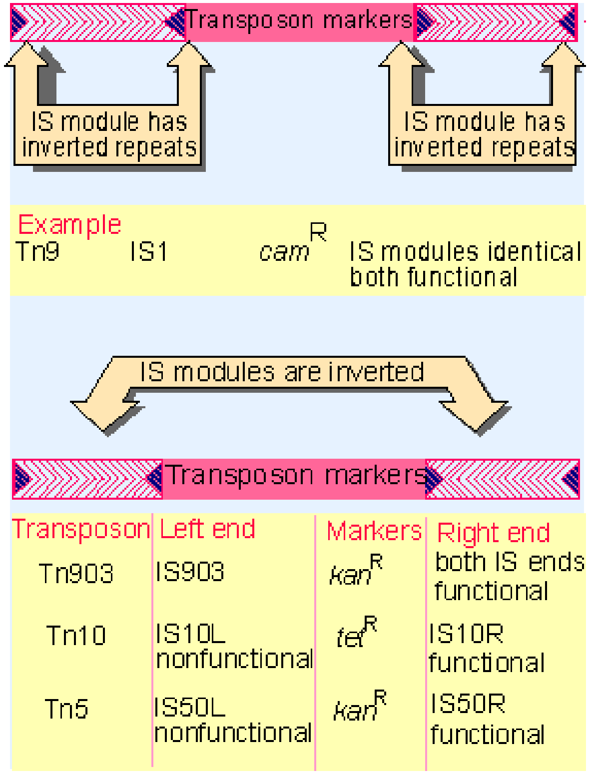
A functional IS module can transpose either itself or the entire transposon.

Either identical module of a composite transposon can sponsor movement (IS10L or IS10R).

In transposons with different modules transposition might depend entirely or principally on one of the modules (Tn10 or Tn5).

What is responsible for transposing a composite transposon instead of just the individual module?



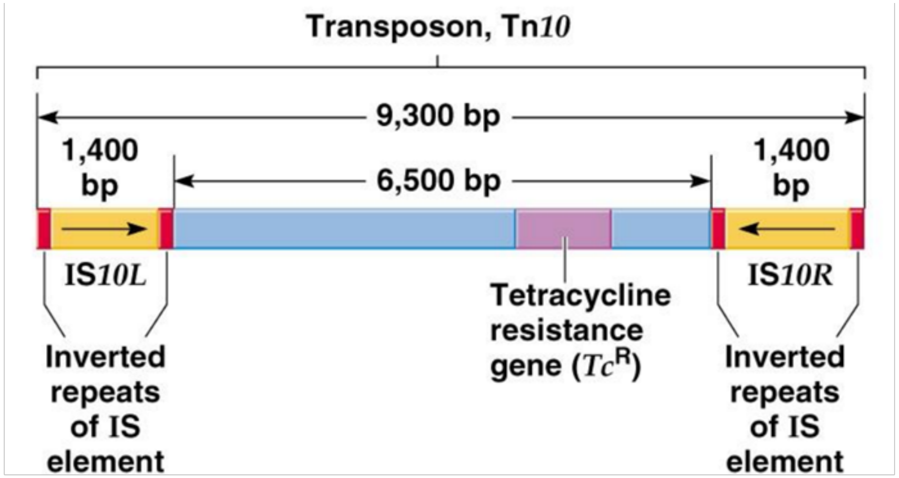


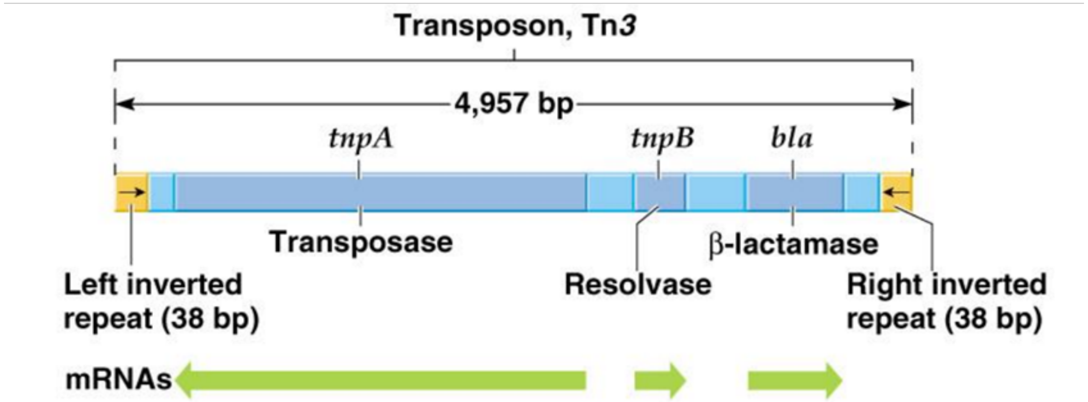
Example
Tn9 IS1 cam^R IS modules identical both functional

Transposon	Left end	Markers	Right end
Tn903	IS903	kan^R	both IS ends functional
Tn10	IS10L nonfunctional	tet^R	IS10R functional
Tn5	IS50L nonfunctional	kan^R	IS50R functional

Composite transposons (Tn)

Natural selection, ensures a composite transposon is copied together as a whole rather than the individual IS.





Composite transposons (Tn)

IS vs Tn equally feasible & useful from a “selfish point of view”.

Selective pressure.

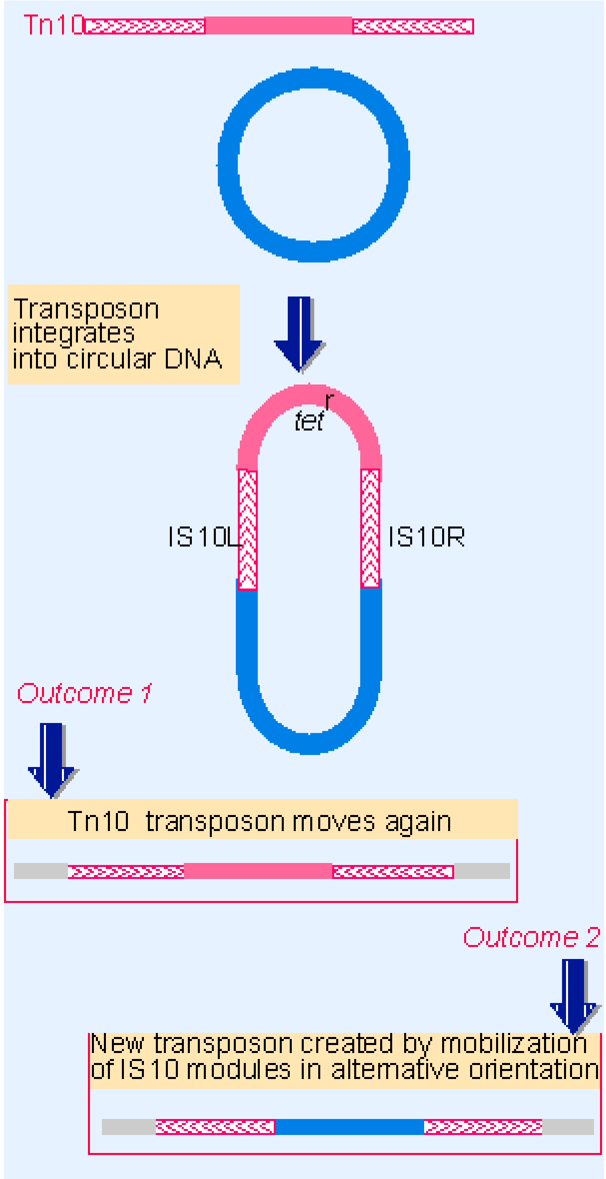
Two IS elements can transpose any sequence residing between them just as well as themselves.

Exemplified by transposons in bacteria where the two modules flank an antibiotic resistance gene.

Selection for the trait(s) carried in the central region.

An IS10 module mobilizes an order of magnitude more frequently than Tn10.

But Tn10 is held together by selection for tetR; so that under selective conditions, the relative frequency of intact Tn10 transposition is higher.



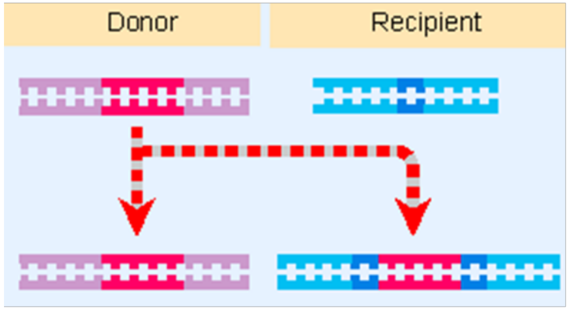
Transposon replication

Three different types of mechanism by which a transposon moves:

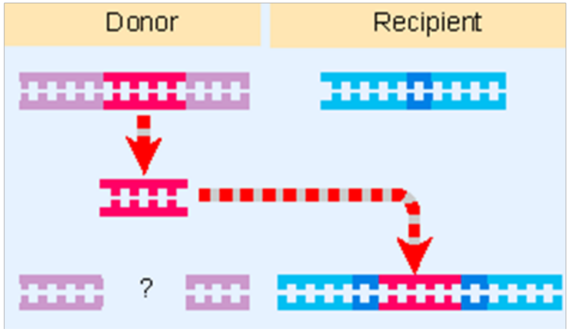
- Replicative Transposition
- Common Non-replicative Transposition
- Conservative Non-replicative Transposition

Episomal transposition

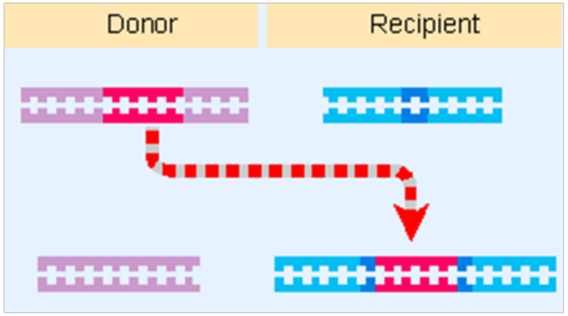
Replicative transposition



Common non-replicative transposition



Conservative non-replicative transposition

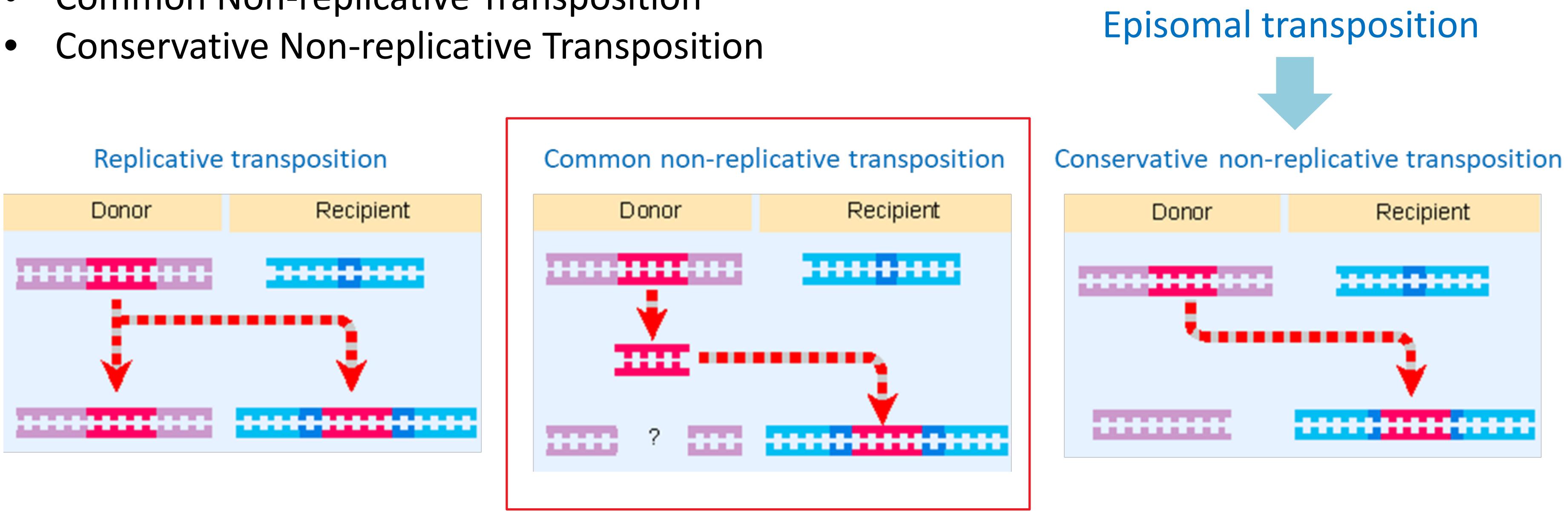


- Involves two types of enzymatic activity:
- Transposase that acts on the ends of the original transposon.
- Resolvase that acts on the duplicated copies.

Transposon replication

Three different types of mechanism by which a transposon moves:

- Replicative Transposition
- Common Non-replicative Transposition
- Conservative Non-replicative Transposition



Episomal transposition

Replicative transposition

Common non-replicative transposition

Conservative non-replicative transposition

- The transposing element moves as a physical entity from one site to another
- No copies and no change
- Requires only a Transposase
- Disregards double strand cleavage of genomic DNA from which it originated.
- Relies on host repair mechanisms to repair double strand breaks.

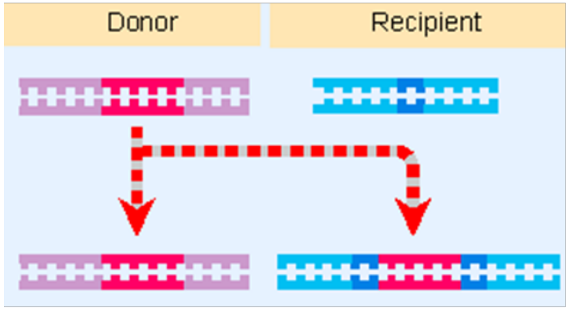
Transposon replication

Three different types of mechanism by which a transposon moves:

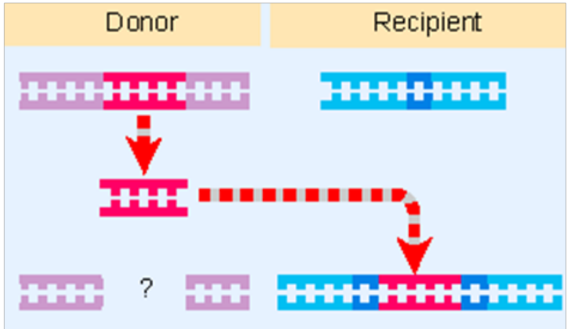
- Replicative Transposition
- Common Non-replicative Transposition
- Conservative Non-replicative Transposition

Episomal transposition

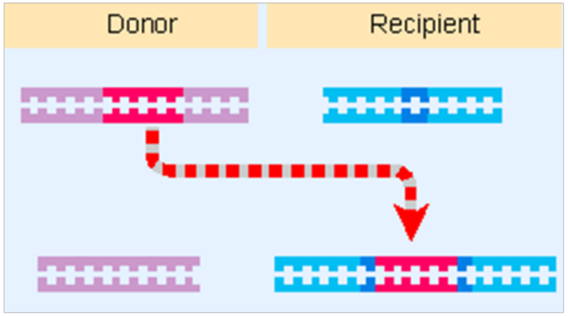
Replicative transposition



Common non-replicative transposition



Conservative non-replicative transposition



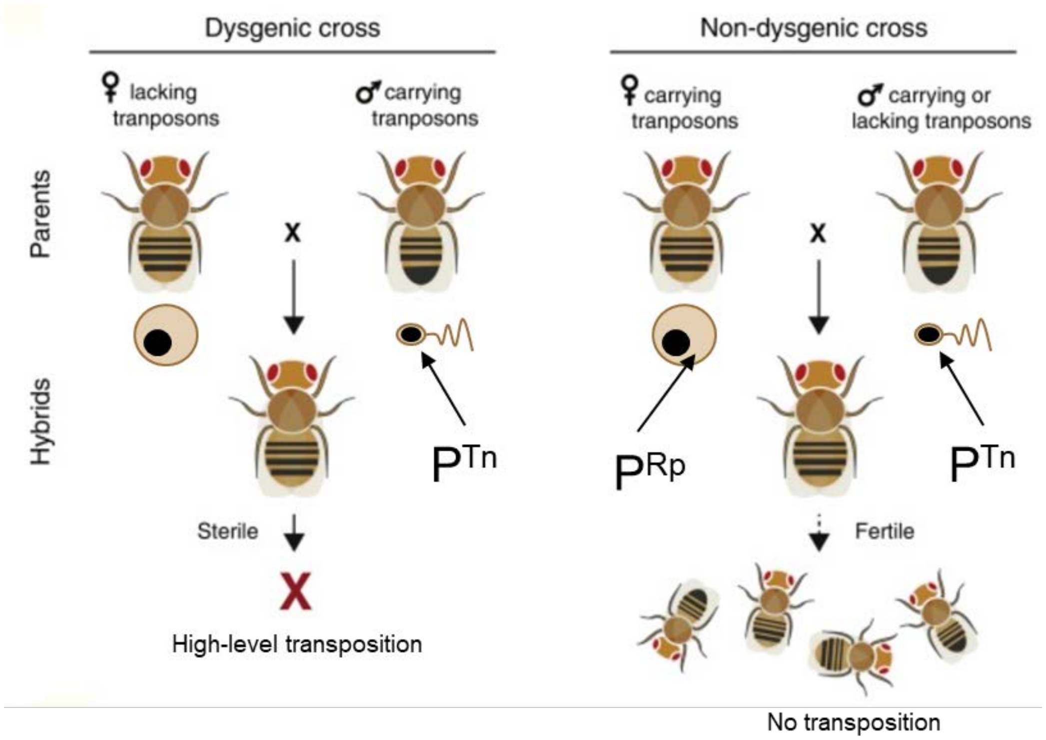
- The element is excised from the donor site and inserted into a new target site.
- Series of events in which every nucleotide bond is conserved and gap is healed.
- Transposon looking after the health of its host.
- Mechanism of lambda Phage integration.

Transposons & Hybrid Dysgenesis

Certain strains of D. melanogaster encounter difficulties in interbreeding.

When flies from two of these different strains are crossed, the progeny display "dysgenic traits".

Dysgeneic Traits are a series of defects including: mutations, chromosomal aberrations, distorted segregation at meiosis, and sterility.



Why is there a direction?

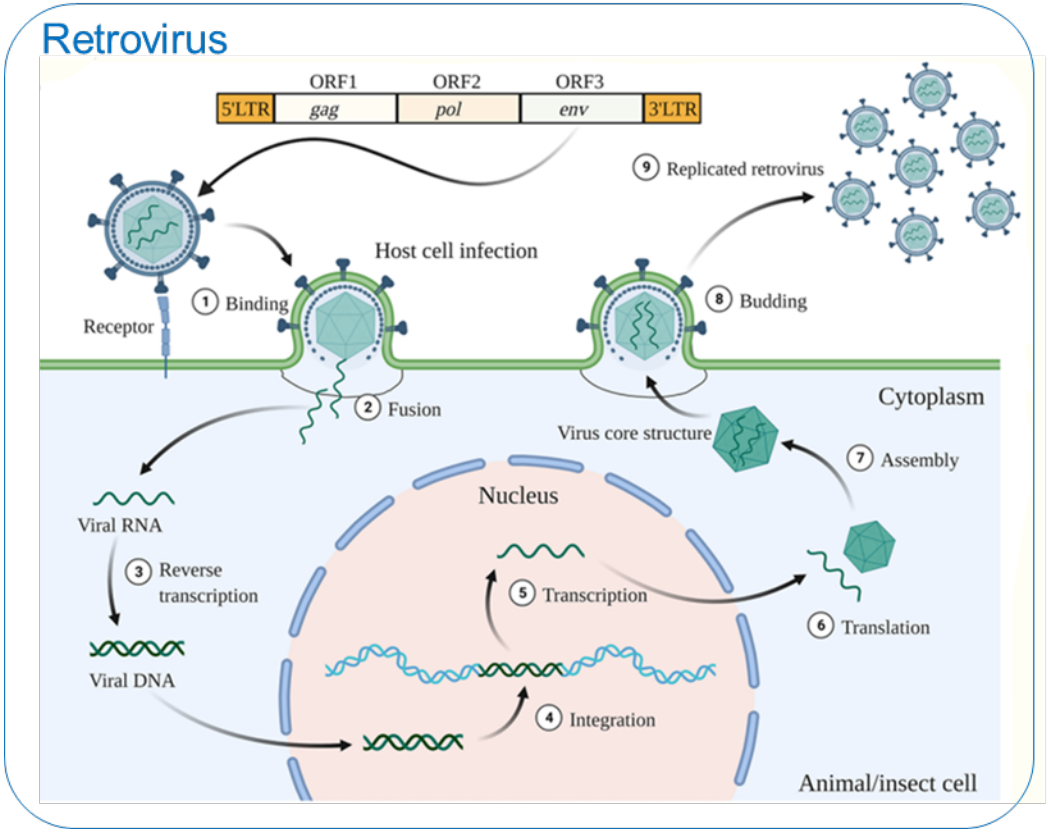
P♂ + M♀ = HD
M♂ + P♀ = Wt
P♂ + P♀ = Wt

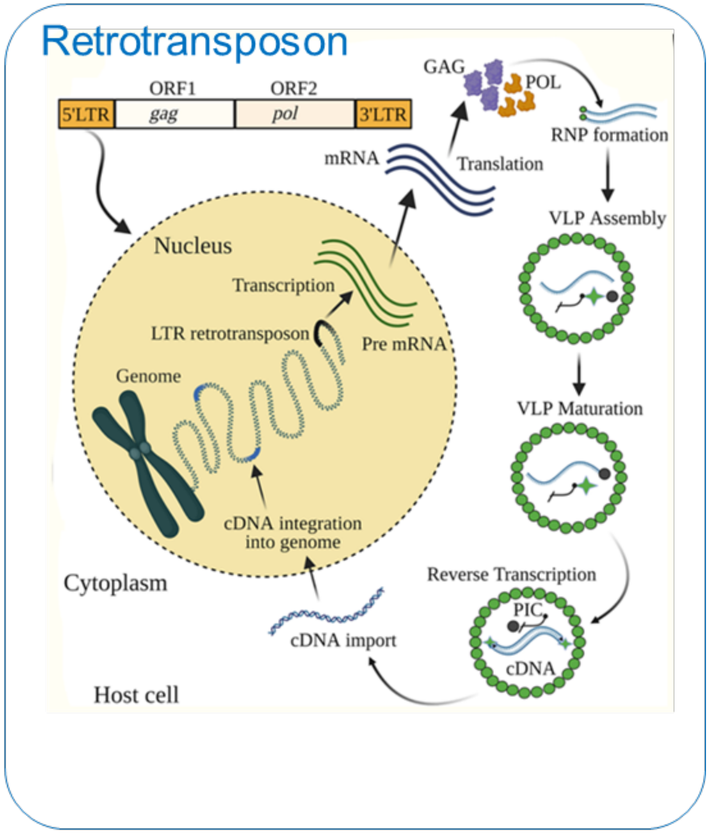
P♂♀ have 30-50 P transposons (P^{Tn})
Only P♀ have Repressors (P^{Rp})
M♂♀ do not.

Retrotransposons are not retrovirusvirus (but related)

Do not utilize an extracellular independent form.

Move directly from one site of the genome to another.





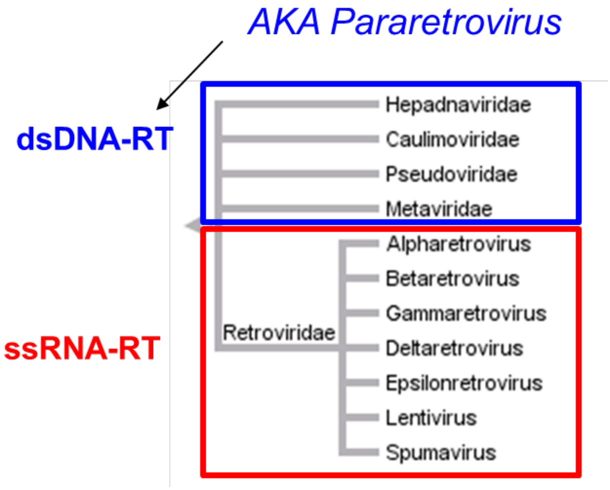
Ramakrishnan M, et al. Plant Cell Rep. 2023 Jan

Retroviridae family overview

Taxonomy & Distribution

Family contains 2 subfamilies and 11 genera

Widespread distribution across diverse vertebrate hosts



Virion Structure

Diameter: 80-100 nm

Lipid envelope displaying viral glycoproteins

Inner protein core

Core morphology varies by genus and distinguishes each

Replication Strategy

Reverse transcription: ssRNA (+) → dsDNA

Integration into host chromosomal DNA creates provirus

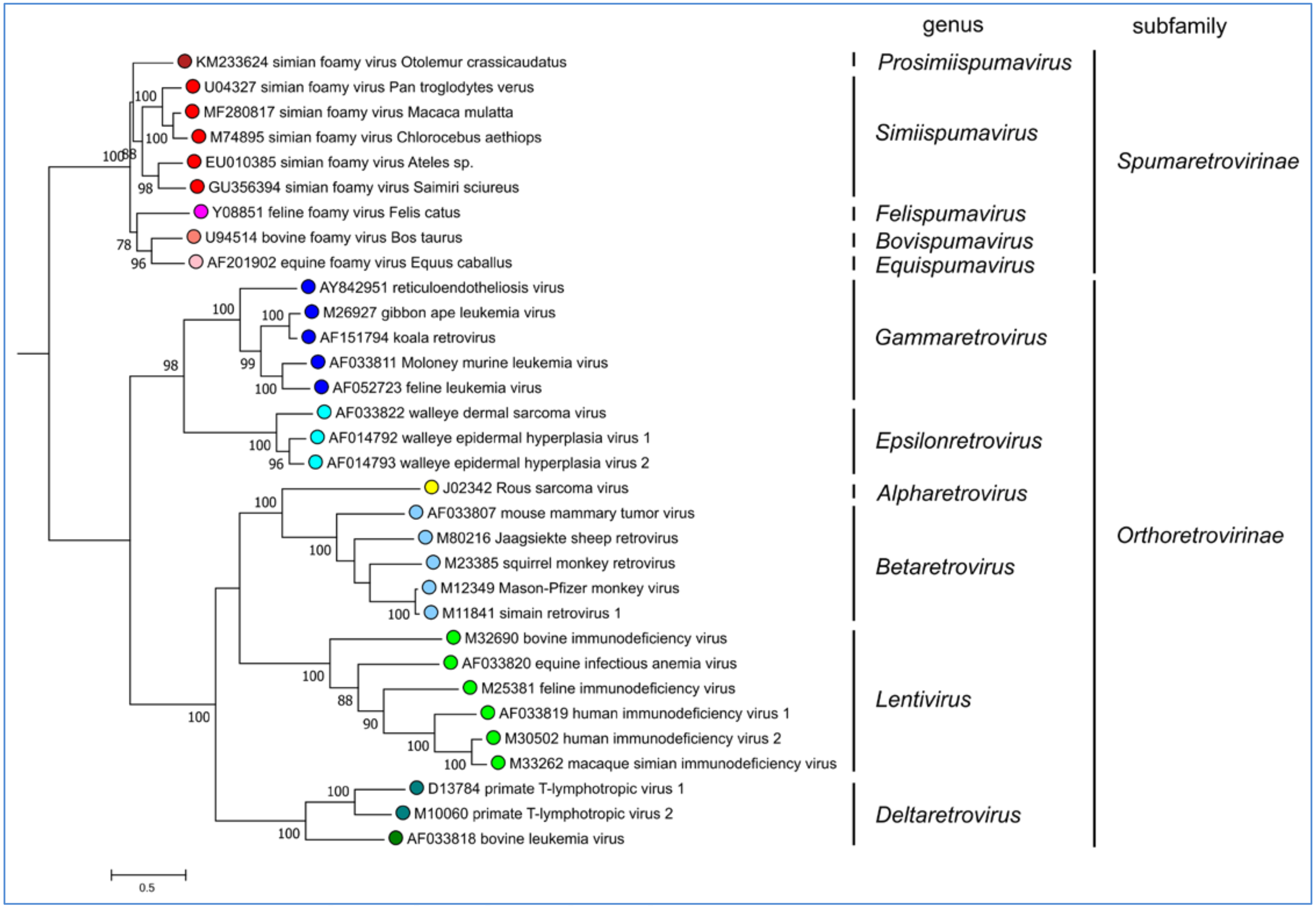
Evolutionary Significance

Germline integration produces heritable endogenous retroviruses (ERVs)

Most vertebrate genomes contain thousands of ERV loci

Retroviridae family taxonomy

Contains 2 subfamilies and 11 genera with widespread across diverse vertebrate hosts



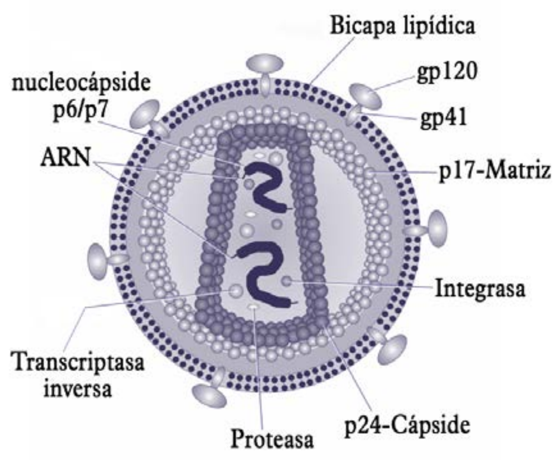
- *Avian leukosis virus*

Retroviridae family overview

Taxonomy & Distribution

Family contains 2 subfamilies and 11 genera

Widespread distribution across diverse vertebrate hosts



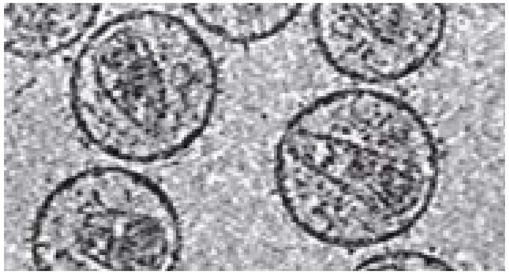
Virion Structure

Diameter: 80-100 nm

Lipid envelope displaying viral glycoproteins

Inner protein core

Core morphology varies by genus and distinguishes each



Replication Strategy

Reverse transcription: ssRNA (+) → dsDNA

Integration into host chromosomal DNA creates provirus

Evolutionary Significance

Germline integration produces heritable endogenous retroviruses (ERVs)

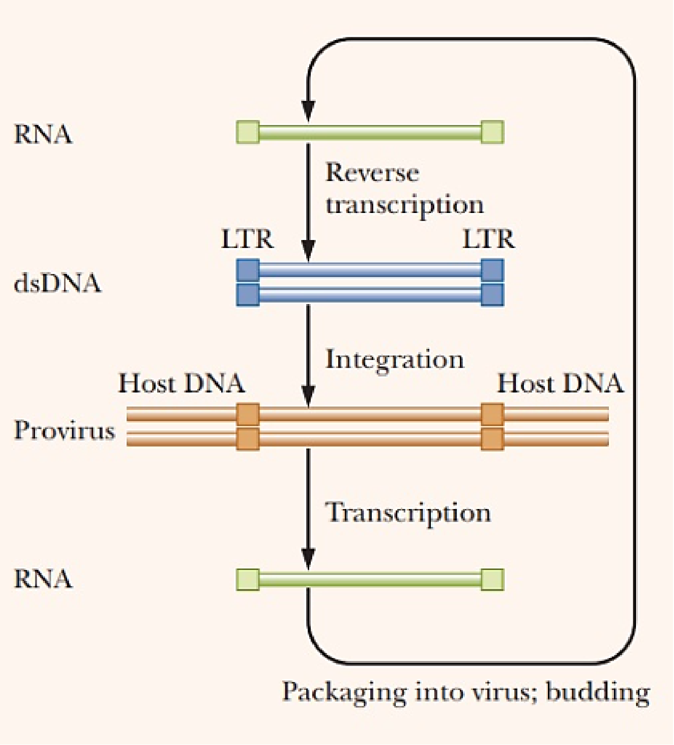
Most vertebrate genomes contain thousands of ERV loci

Retroviridae family overview

Taxonomy & Distribution

Family contains 2 subfamilies and 11 genera

Widespread distribution across diverse vertebrate hosts



Virion Structure

Diameter: 80-100 nm

Lipid envelope displaying viral glycoproteins

Inner protein core

Core morphology varies by genus and distinguishes each

Replication Strategy

Reverse transcription: ssRNA (+) → dsDNA

Integration into host chromosomal DNA creates provirus

Evolutionary Significance

Germline integration produces heritable endogenous retroviruses (ERVs)

Most vertebrate genomes contain thousands of ERV loci

Retroviridae family overview

Taxonomy & Distribution

Family contains 2 subfamilies and 11 genera

Widespread distribution across diverse vertebrate hosts

Virion Structure

Diameter: 80-100 nm

Lipid envelope displaying viral glycoproteins

Inner protein core

Core morphology varies by genus and distinguishes each

Replication Strategy

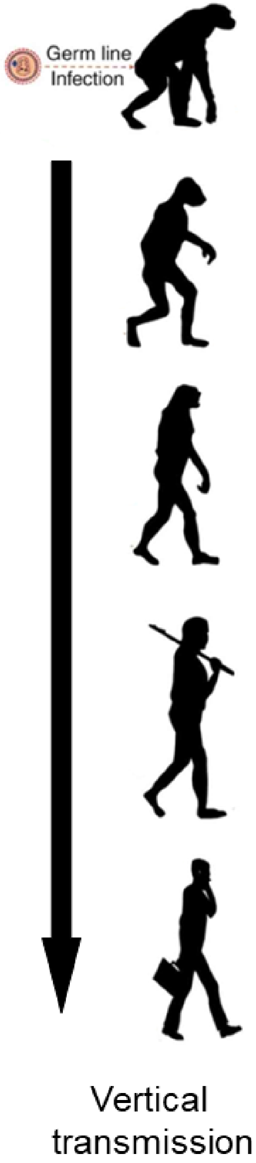
Reverse transcription: ssRNA (+) → dsDNA

Integration into host chromosomal DNA creates provirus

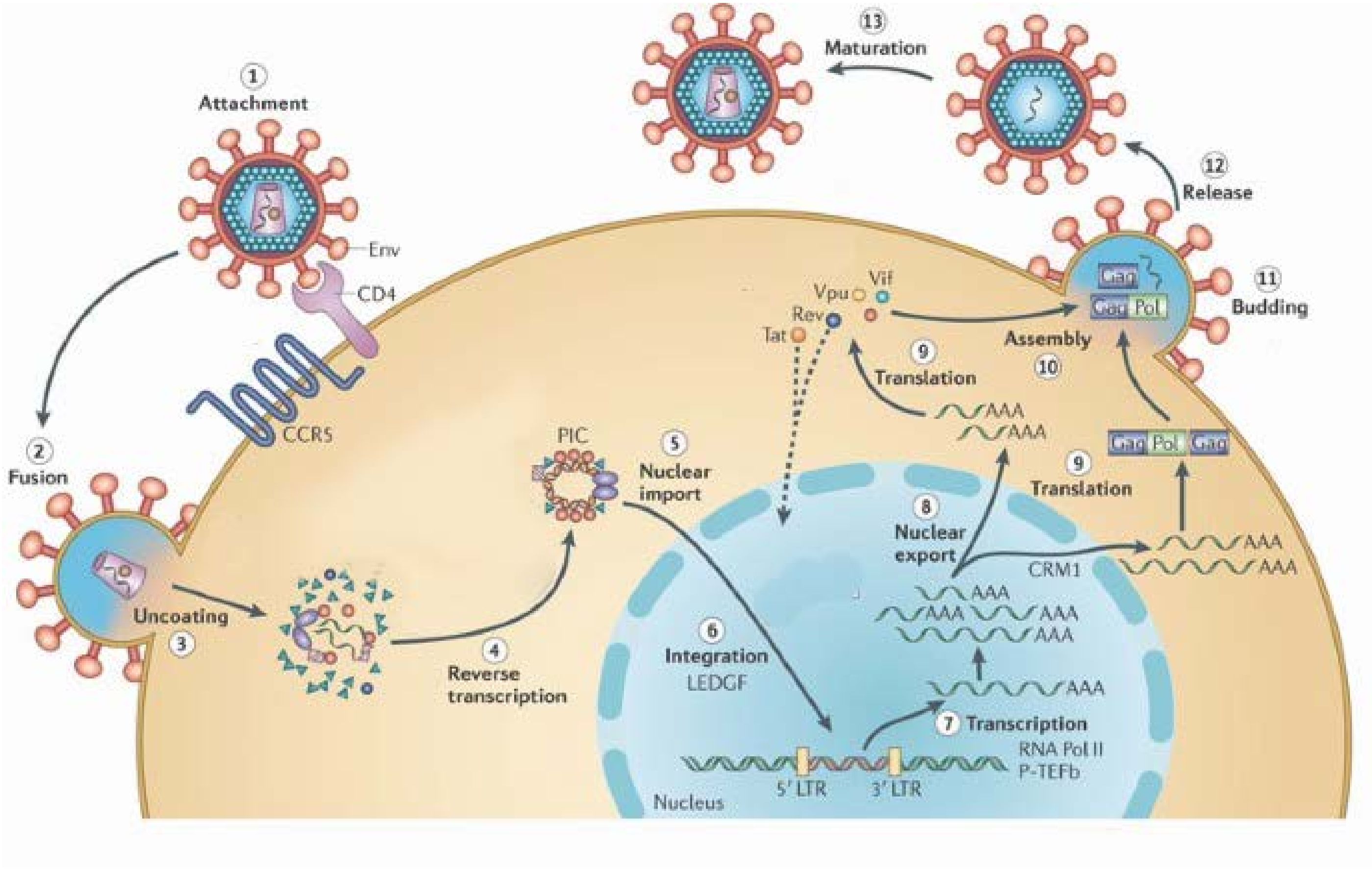
Evolutionary Significance

Germline integration produces heritable endogenous retroviruses (ERVs)

Most vertebrate genomes contain thousands of ERV loci



HIV life cycle



Engelman A, Cherepanov P. Nat Rev Microbiol. 2012 Mar

HIV life cycle (molecular detail)

- Retroviruses have ssRNA genomes.

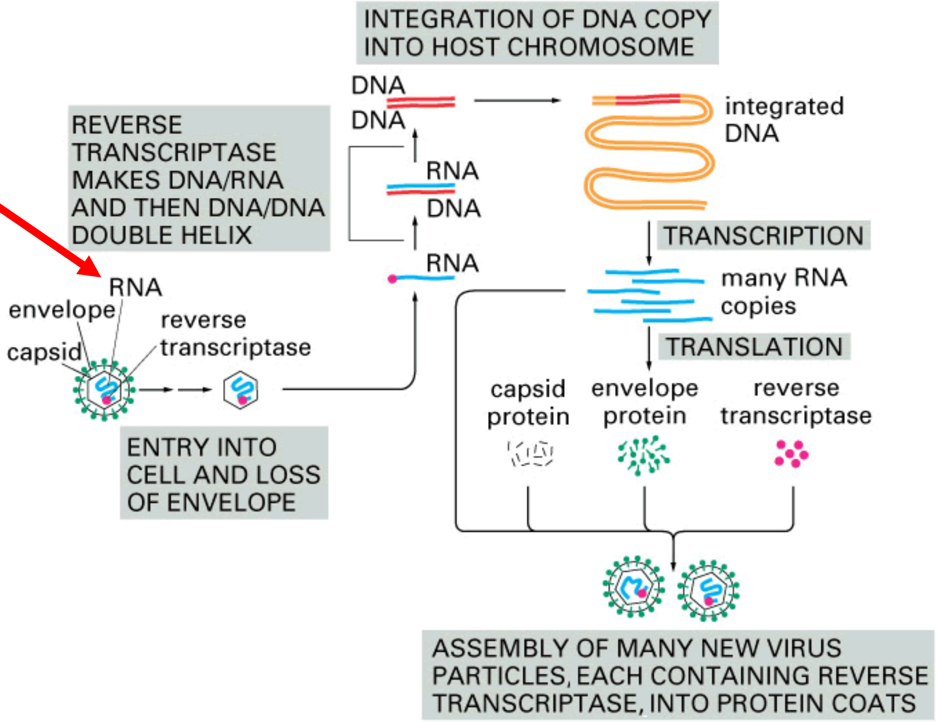


Figure 6-39 **Essential Cell Biology, 2/e. (© 2004 Garland Science)**

HIV life cycle (molecular detail)

- Retroviruses have ssRNA genomes.
- Genomes are replicated through a dsDNA intermediate.

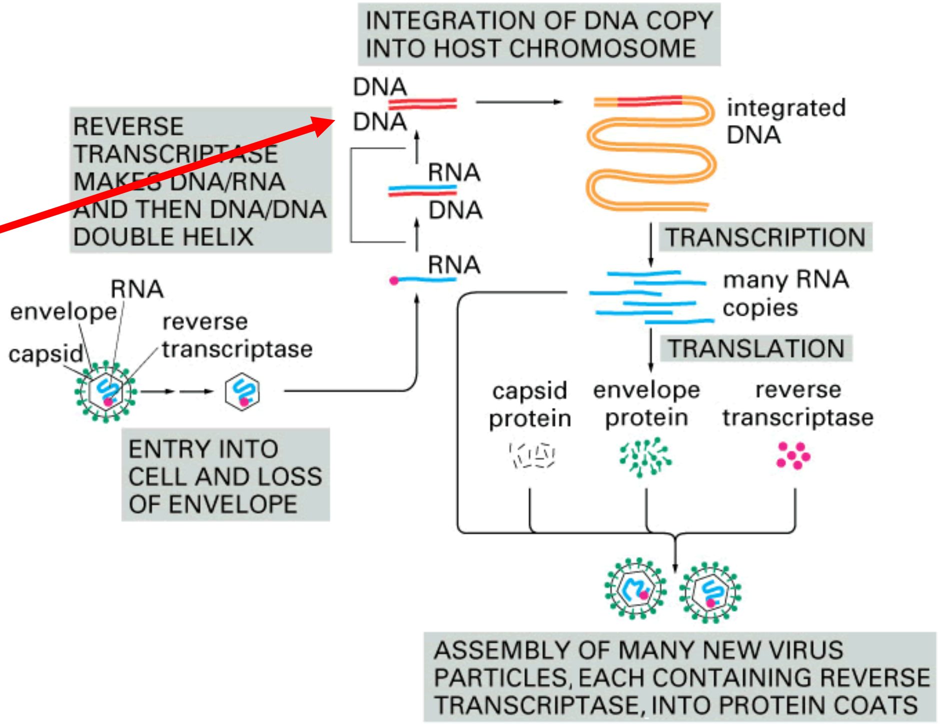


Figure 6-39 **Essential Cell Biology, 2/e. (© 2004 Garland Science)**

HIV life cycle (molecular detail)

- Retroviruses have ssRNA genomes.
- Genomes are replicated through a dsDNA intermediate.
- The life cycle involves an obligatory stage of integration in which the proviral dsDNA is inserted into the host genome by a transposition-like event that generates short direct repeats of target DNA.

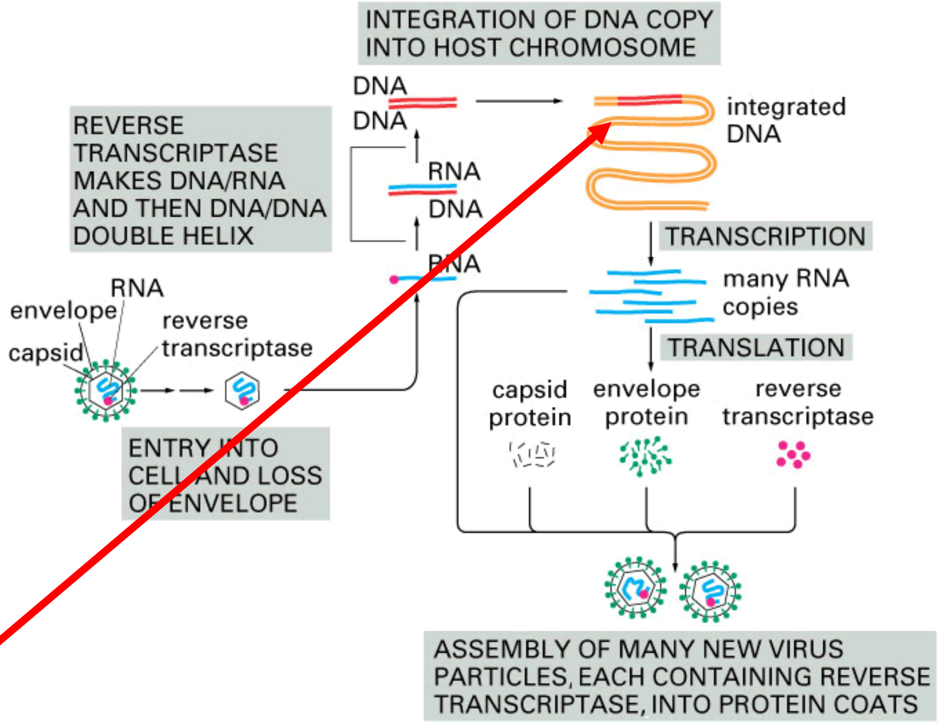


Figure 6-39 Essential Cell Biology, 2/e. (© 2004 Garland Science)

HIV life cycle (molecular detail)

- The enzyme responsible for generating the initial DNA copy of the RNA is reverse transcriptase.

- Converts the RNA into a linear (or circular) duplex of DNA in the cytoplasm of the infected cell.

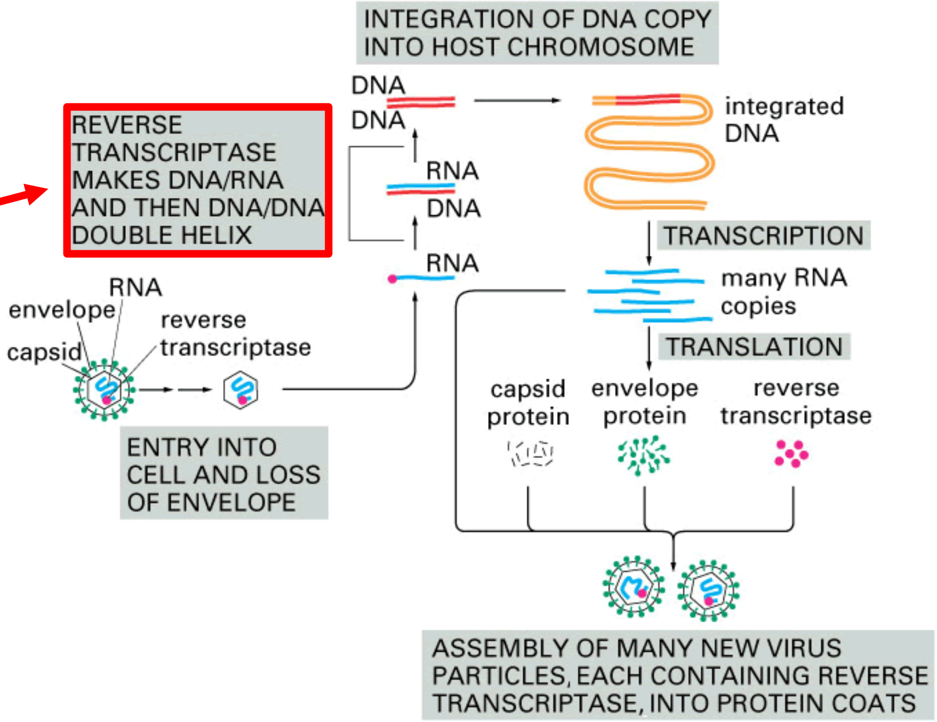


Figure 6-39 **Essential Cell Biology, 2/e. (© 2004 Garland Science)**

HIV life cycle (molecular detail)

- The enzyme responsible for generating the initial DNA copy of the RNA is reverse transcriptase.
- Converts the RNA into a linear (or circular) duplex of DNA in the cytoplasm of the infected cell.
- The enzyme responsible for integrating the DNA copy is integrase.

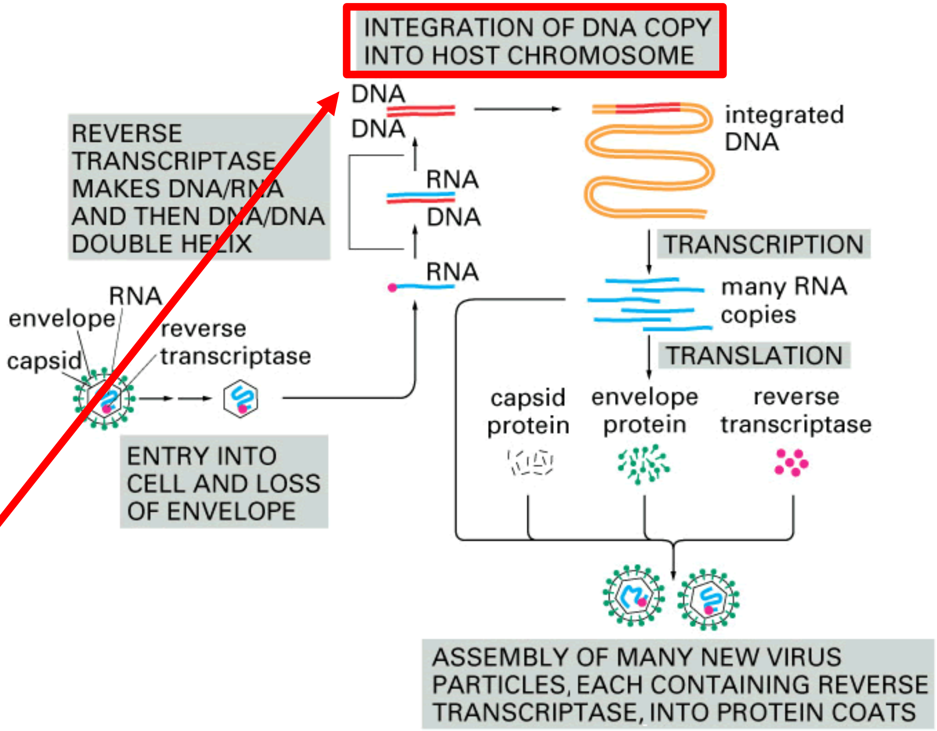


Figure 6-39 **Essential Cell Biology, 2/e. (© 2004 Garland Science)**

Endogenous retrovirus (ERVs)

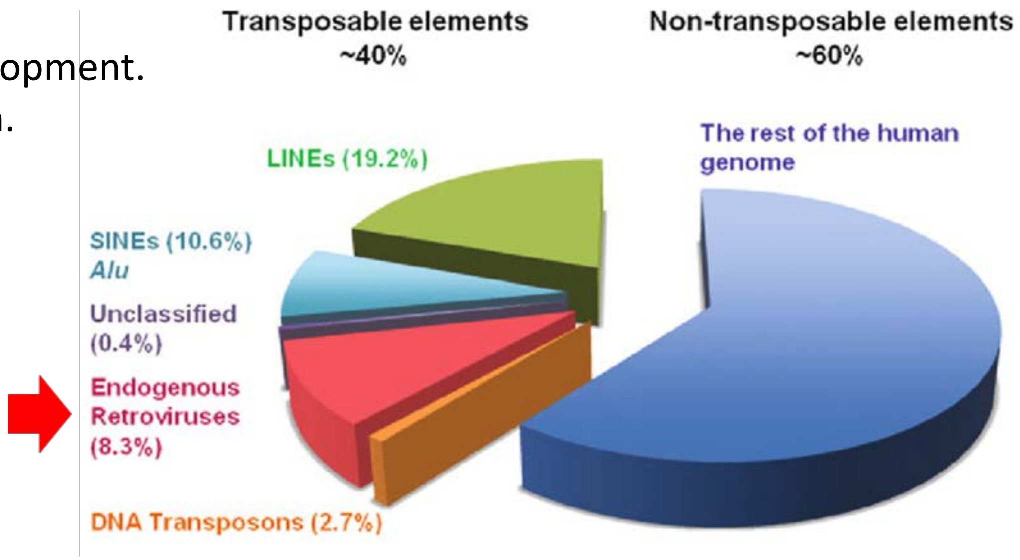
Retrovirus that have been incorporated to the genome are called ENDOGENOUS (ERV).

They are derived from ancient infections of germ cells in humans, mammals & other vertebrates.

ERVs make up 5-8% of the human genome (98,000 elements).

Most insertions have no known function (junk DNA) but some play important roles in host biology:

- Control of gene transcription.
- Control of cell fusion during placental development.
- Resistance to exogenous retroviral infection.
- Immunosuppresion



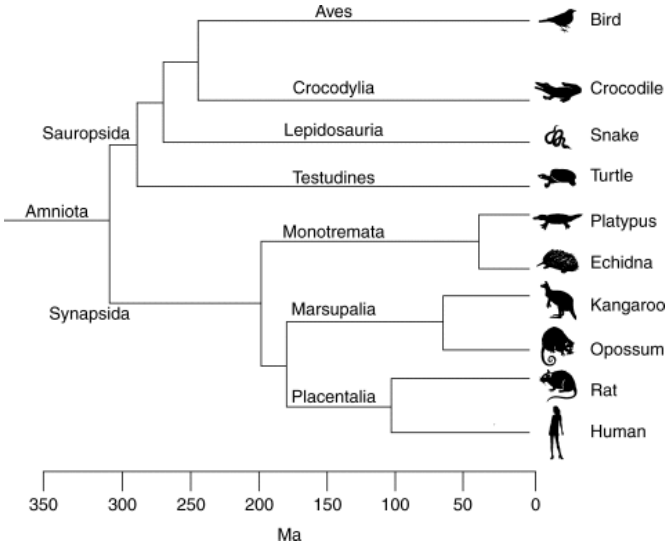
Endogenous retrovirus (ERVs)

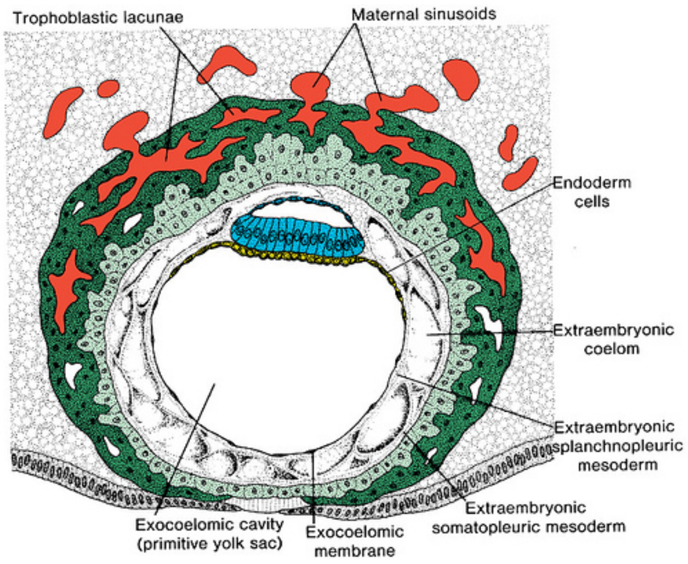
ERV's are similar to HIV (which causes AIDS in humans).

The immunodepressive action was the initial normal behavior of the virus, as in HIV.

The fusion proteins were a way to spread the infection to other cells by merging them with the infected one, as in HIV.

It is believed that the ancestors of modern vivipary mammals evolved after an infection by this virus, enabling the fetus to survive the immune system of the mother.





Endogenous retrovirus (ERVs)

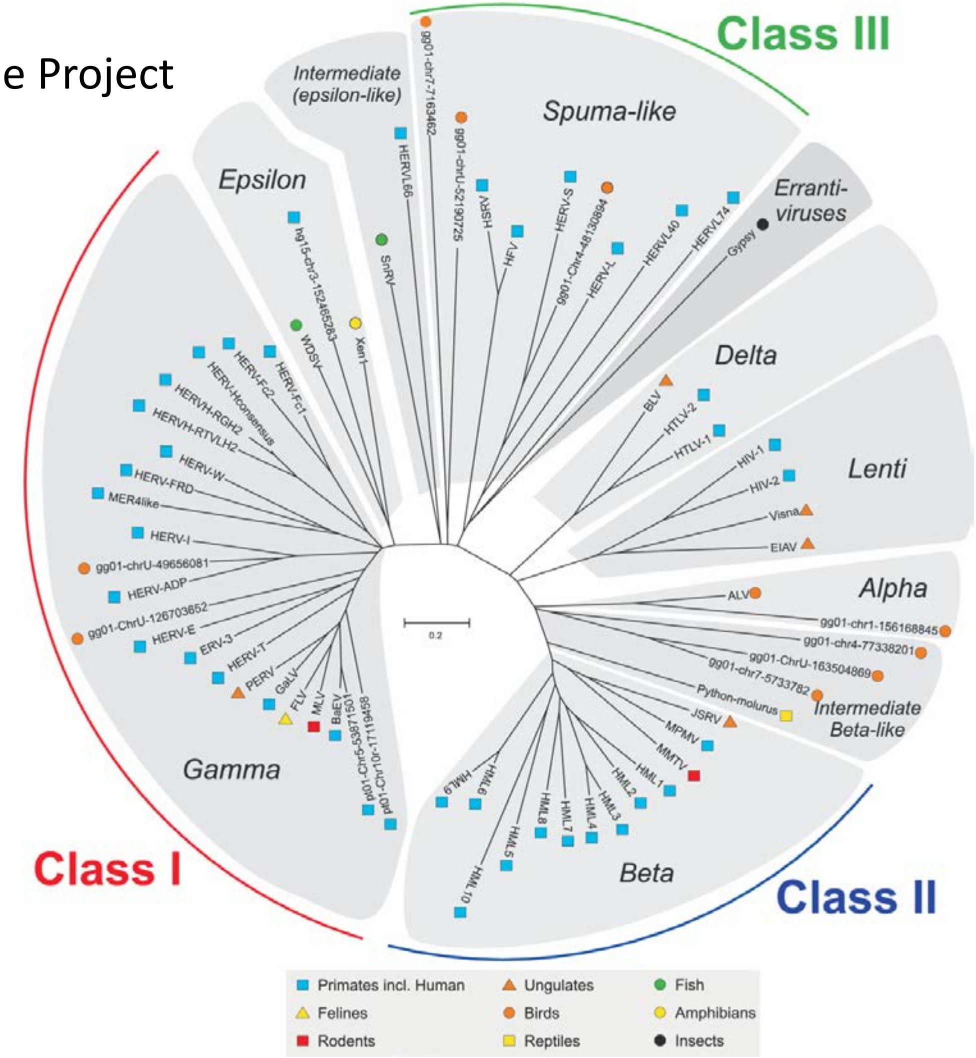
24 ERV families identified by Human Genome Project (HGP).

Broadly classified into 3 classes on the basis of relatedness to exogenous genera:

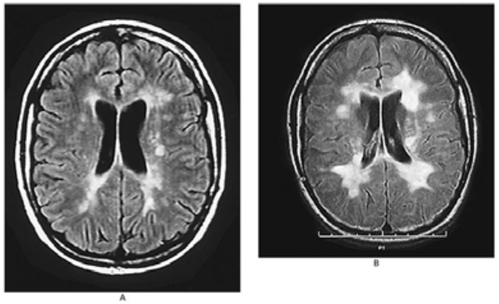
Class I are similar to the gammaretroviruses

Class II are similar to the betaretroviruses & alpharetroviruses

Class III are similar to the spumaviruses



Endogenous retrovirus (ERVs)



Human ERVs (hERVs) are suspected of involvement in some autoimmune diseases (multiple sclerosis).

Especially human endogenous retrovirus W known (MSRV).

Also a possible hERV involvement in the HELLP (Hemolytic anemia, Elevated Liver enzymes & Low Platelet count) syndrome & pre-eclampsia.

hERVs very likely associated with some types of schizophrenia.

Viral & Human Genomics Laboratory

CDC Commissioned Biosafety Level 3 (BSL-3) High Biocontainment Facility

Laboratorio de Genómica Viral y Humana

Instalaciones de Alta Contención Biológica Nivel de Bioseguridad 3 (BSL-3) CDC-certificadas

Facultad de Medicina UASLP
San Luis Potosí, México

www.genomica.uaslp.mx

GenomicaUASLP

GenomicaUASLP