

TRANSPOSABLE ELEMENTS

Structure, Mechanisms of Transposition, and Biological Significance in Prokaryotes and Eukaryotes

1. Introduction

Transposable elements (TEs), or transposons, are mobile segments of DNA capable of moving from one location in a genome to another — a process called **transposition**. They were first discovered in the 1940s-50s by **Barbara McClintock**, who studied unstable pigmentation patterns in maize kernels and correctly deduced the existence of mobile "controlling elements" (the Ac/Ds system) long before DNA sequencing existed — work for which she was awarded the Nobel Prize in 1983.

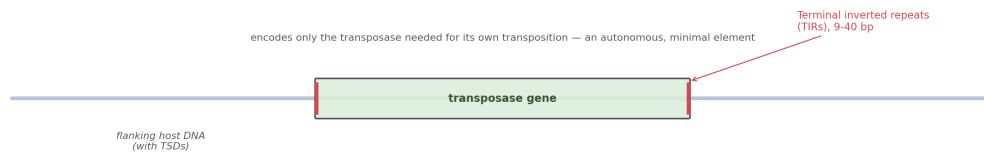
- Transposable elements are found in virtually every organism examined, from bacteria to humans, and often make up a very large fraction of genome size.
 - Remarkably, TEs and their remnants make up roughly **45% of the human genome**, far more than the ~1.5% that is protein-coding sequence.
 - TEs are broadly grouped by their mechanism of movement into **Class I retrotransposons** (which move via an RNA intermediate — "copy-and-paste") and **Class II DNA transposons** (which move directly as DNA — "cut-and-paste").
 - Because transposition can disrupt genes, alter gene regulation, and drive chromosome rearrangements, TEs are often described as "**molecular parasites**" or "**genomic parasites**" — yet they have also been repeatedly co-opted by their hosts and are now recognised as a major force in genome evolution.
-

2. Classification and Structural Organisation

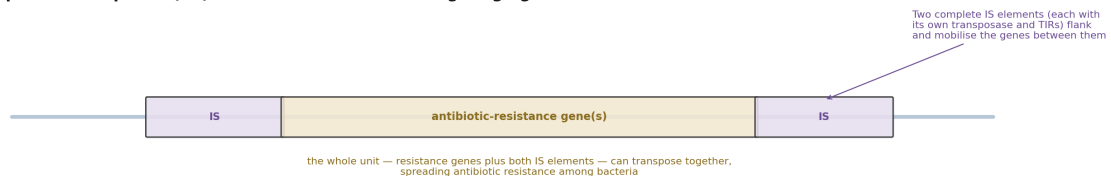
Although individual transposable elements are extremely diverse, they fall into a small number of structural classes, each with a distinctive arrangement of coding and non-coding sequence that reflects its particular transposition mechanism.

Structural Anatomy of the Major Transposable-Element Classes

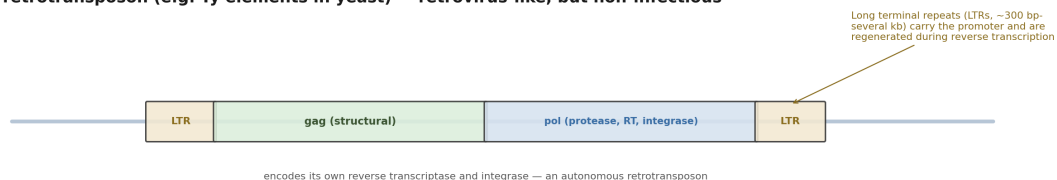
1. Insertion sequence (IS element) — the simplest prokaryotic transposon



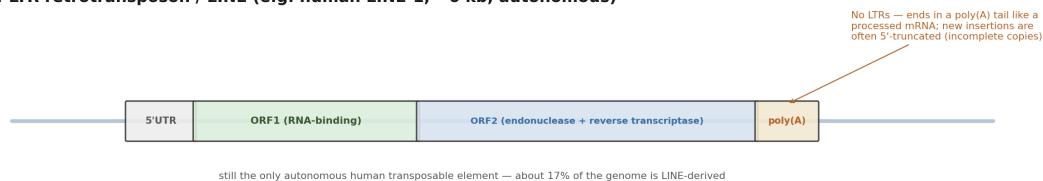
2. Composite transposon (Tn) — two IS elements flanking cargo genes



3. LTR retrotransposon (e.g. Ty elements in yeast) — retrovirus-like, but non-infectious



4. Non-LTR retrotransposon / LINE (e.g. human LINE-1, ~6 kb, autonomous)



5. SINE (e.g. human Alu element, ~300 bp, non-autonomous)

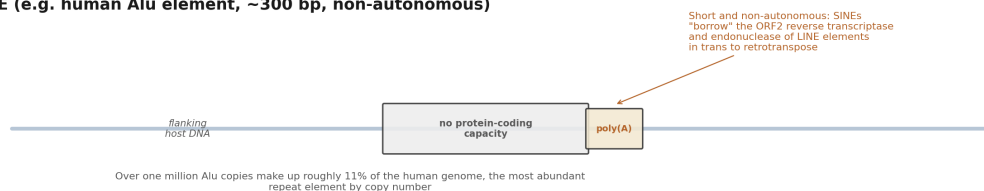


Fig. 1 — Structural anatomy of the five major classes of transposable element, from the simplest prokaryotic insertion sequence to the human Alu SINE.

- **Autonomous elements** (IS elements, LTR retrotransposons, LINEs) encode all the proteins they need for their own transposition.
- **Non-autonomous elements** (SINEs such as Alu, and many defective DNA transposons) lack functional coding sequence and must "borrow" the transposition machinery encoded by an autonomous partner element.

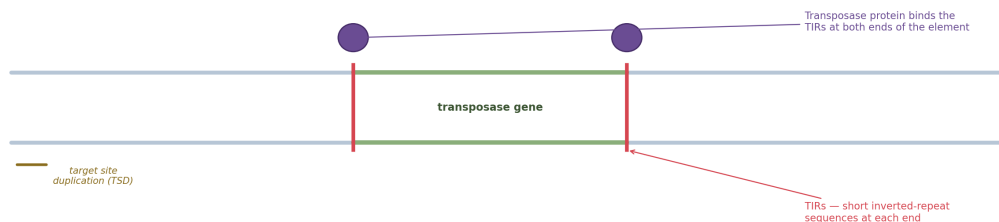
- Nearly all DNA transposons and retrotransposons are flanked by short **target site duplications (TSDs)** — direct repeats of host sequence created at the moment of insertion — which serve as a molecular signature of a past transposition event even after the element itself has decayed or been deleted.

3. Class II: DNA Transposons ("Cut-and-Paste")

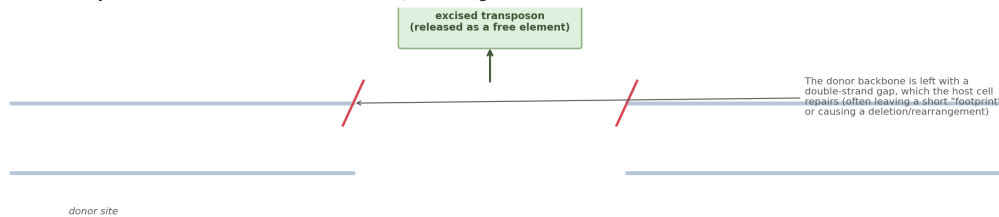
DNA transposons move directly as DNA, without an RNA intermediate, using an enzyme called **transposase** that recognises **terminal inverted repeats (TIRs)** at each end of the element.

Cut-and-Paste Transposition (Class II DNA Transposon)

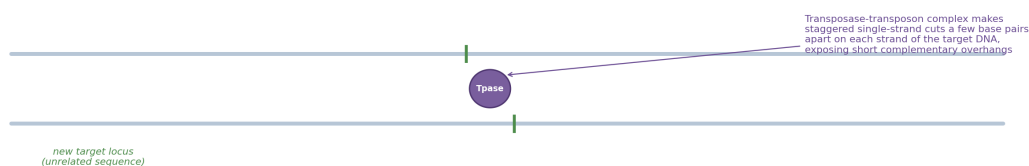
1. Transposase recognises terminal inverted repeats (TIRs) flanking the element in donor DNA



2. Excision — transposase cuts both strands at each TIR, releasing the element from the donor site



3. Target capture — transposase makes a staggered cut at a new target site elsewhere in the genome



4. Insertion — the element ligates into the staggered ends; gap repair duplicates the target sequence (new TSDs)

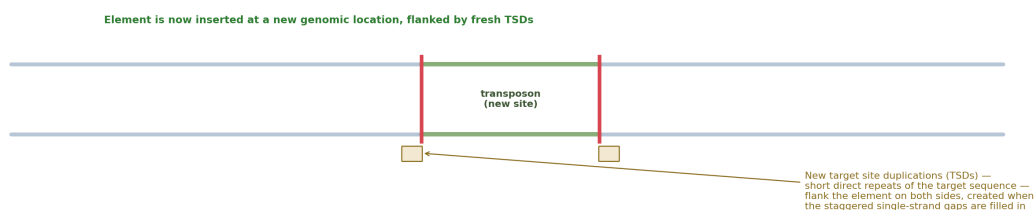


Fig. 2 — Cut-and-paste transposition: transposase excises the element from the donor site and inserts it at a new target site, generating new target site duplications.

- Transposase binds the TIRs and cuts both DNA strands at each end, **excising** the element from the donor locus; the donor backbone is then rejoined by host repair, sometimes leaving mutations or a small "footprint".
- The same (or an associated) transposase makes a **staggered cut** at a new target site elsewhere in the genome, and the excised element is ligated into the offset ends.
- Host repair enzymes fill in the resulting short single-stranded gaps, which duplicates a few base pairs of target sequence on each side of the new insertion — the target site duplication.
- Because the element is physically removed from the donor site, cut-and-paste transposition does not by itself increase the number of copies of the element (unless transposition occurs during S phase, when a repaired donor site can be restored from a sister chromatid, effectively duplicating the element).

(a) Prokaryotic DNA Transposons

- **Insertion sequences (IS elements)** are the simplest bacterial transposons: just a transposase gene flanked by TIRs, typically 700-2500 bp.
- **Composite transposons (Tn elements)** arise when two IS elements flank one or more additional genes — very often **antibiotic-resistance genes** — allowing the whole cassette to move as a unit. This is a major mechanism by which antibiotic resistance spreads among bacterial populations and between plasmids and chromosomes.

(b) Eukaryotic DNA Transposons

- The **Ac/Ds system** in maize — McClintock's original discovery — consists of an autonomous **Activator (Ac)** element encoding transposase, and non-autonomous **Dissociation (Ds)** elements that can only move when Ac is also present, producing the variegated kernel colours she studied.
- **P elements** in *Drosophila* are DNA transposons that spread rapidly through wild fly populations during the 20th century and are now standard tools for making transgenic flies in the laboratory.
- Eukaryotic genomes generally carry many more inactive, mutated "fossil" DNA transposons than active ones, since most have lost transposase function over evolutionary time.

4. Class I: Retrotransposons ("Copy-and-Paste")

Retrotransposons move via an **RNA intermediate**: the element is transcribed into RNA, which is then reverse-transcribed back into DNA by a **reverse transcriptase** and inserted at a new site. Because the original copy is transcribed rather than excised, retrotransposition is inherently **replicative** — the element increases in copy number with every transposition event.

Retrotransposition: LTR vs Non-LTR Retrotransposons

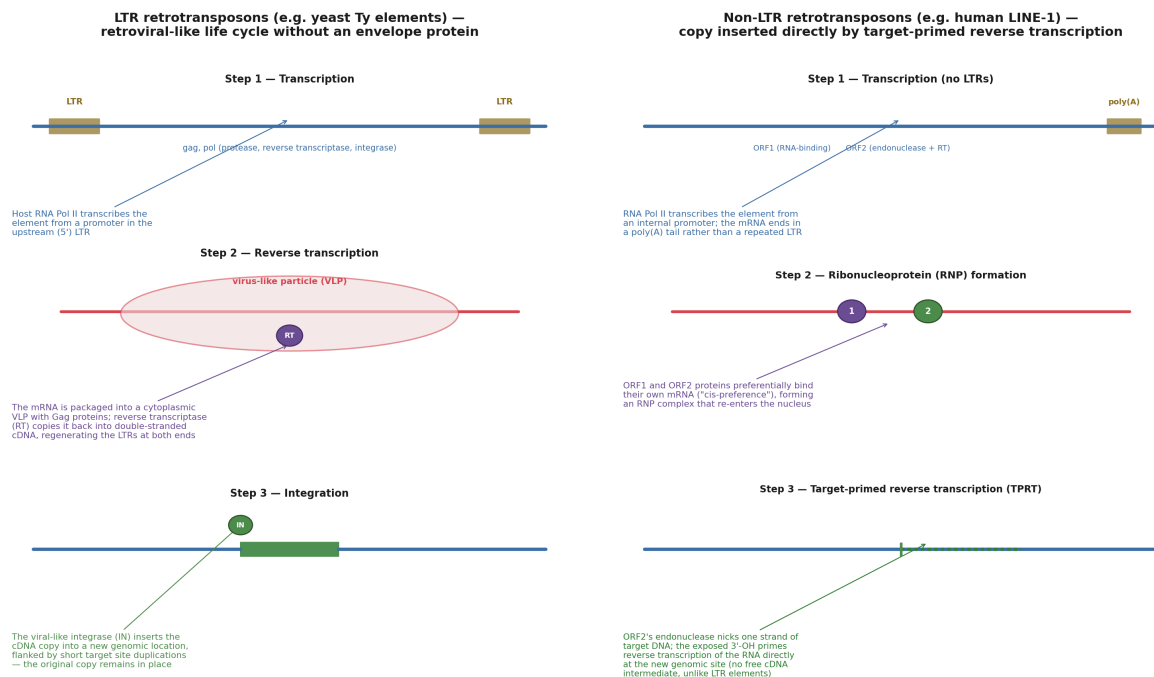


Fig. 3 — LTR retrotransposons replicate through a retrovirus-like virus-like particle, while non-LTR elements such as LINE-1 insert directly by target-primed reverse transcription (TPRT).

(a) LTR Retrotransposons

- Flanked by **long terminal repeats (LTRs)** that carry the promoter and are regenerated at both ends during reverse transcription, closely resembling the structure of an integrated retrovirus.
- Encode *gag* (structural/capsid-like proteins) and *pol* (protease, reverse transcriptase, and integrase) genes; the RNA transcript assembles with Gag proteins into a cytoplasmic **virus-like particle (VLP)** in which reverse transcription occurs.
- The resulting double-stranded cDNA is inserted elsewhere in the genome by the element's own integrase, exactly as a retrovirus integrates its provirus — but LTR retrotransposons (e.g. yeast **Ty elements**) lack an envelope gene and so cannot leave the cell as infectious particles.

(b) Non-LTR Retrotransposons

- Lack LTRs and instead end in a poly(A) tail, resembling a processed mRNA; movement occurs by **target-primed reverse transcription (TPRT)**, in which an endonuclease nicks the target DNA and the exposed 3'-OH primes reverse transcription directly at the new site, without a free cDNA intermediate.
- LINEs (long interspersed nuclear elements)**, such as human **LINE-1 (L1)**, are autonomous, encoding ORF1 (an RNA-binding protein) and ORF2 (endonuclease and reverse transcriptase); L1 is the only currently active autonomous transposable element in humans and makes up about 17% of the genome.

- **SINEs (short interspersed nuclear elements)**, such as the human **Alu element** (~300 bp), are short and non-autonomous, encoding no protein of their own; they retrotranspose only by hijacking the ORF2 machinery of LINEs in *trans*. Alu elements are the single most abundant repetitive sequence in the human genome, present in over one million copies (~11% of total genome sequence).

Feature	DNA transposons (Class II)	Retrotransposons (Class I)
Intermediate	None — moves directly as DNA	RNA intermediate, reverse-transcribed back to DNA
Effect on copy number	Unchanged (cut-and-paste); the element relocates	Increases with each event (copy-and-paste)
Key enzyme	Transposase	Reverse transcriptase (plus integrase or endonuclease)
Flanking repeats	Terminal inverted repeats (TIRs)	LTRs (LTR type) or poly(A) tail (non-LTR type)
Examples	Bacterial IS elements/Tn; maize Ac/Ds; Drosophila P elements	Yeast Ty elements (LTR); human LINE-1, Alu/SINEs (non-LTR)
Share of human genome	~3%, almost all inactive	~42%, dominated by LINEs and SINEs

5. Biological Significance

(a) Mutagenesis and Disease

- An insertion within or near a gene can disrupt its coding sequence, splicing, or regulatory elements, causing loss-of-function mutations; TE insertion has been documented as the direct cause of individual cases of **haemophilia A and B**, **neurofibromatosis type 1**, and **certain cancers** in humans, most often via new LINE-1 or Alu insertions.
- Because retrotransposons carry sequence homology to other copies scattered throughout the genome, they can promote **non-allelic homologous recombination** between mismatched repeat copies, producing deletions, duplications, and other chromosomal rearrangements.

(b) Genome Evolution

- TE insertions are a major source of new genetic variation and structural diversity between individuals and species, contributing substantially to genome size differences across eukaryotes.
- Transposons have repeatedly been "**domesticated**" or **exapted** by their hosts for new cellular functions — well-known examples include the RAG1/RAG2 recombinase genes of the vertebrate immune system (V(D)J recombination), which are thought to have evolved from an ancient DNA transposon, and the placental gene *syncytin*, derived from a retroviral/LTR retrotransposon envelope gene.
- TE-derived sequences frequently supply new promoters, enhancers, alternative splice sites, and even entirely new exons ("exonisation"), making them an important source of regulatory and coding innovation over evolutionary time.

Barbara McClintock's legacy

McClintock's maize experiments in the 1940s-50s showed that genes are not always fixed in position — a radical idea at the time that was largely ignored for decades until molecular techniques in the 1970s-80s confirmed transposable elements in bacteria, yeast, and other organisms, finally earning her the 1983 Nobel Prize in Physiology or Medicine.

6. Host Defence and Regulation of Transposition

Because unrestrained transposition threatens genome integrity, host cells have evolved several layers of defence to keep transposable elements silenced, particularly in the germline (where new insertions can be inherited).

- **DNA methylation** of cytosines within TE sequences, and associated repressive **histone modifications**, package most TE copies into transcriptionally silent heterochromatin.
- The **piRNA (PIWI-interacting RNA) pathway** is a germline-specific small-RNA silencing system that recognises TE transcripts by sequence complementarity and directs both their degradation and the establishment of repressive chromatin at the corresponding genomic loci.
- In bacteria, DNA transposons are comparatively less tightly restrained but transposition is generally kept infrequent by the low activity and copy number of transposase genes, plus target-site immunity mechanisms in some elements that prevent repeated insertion into an already-occupied site.

7. Quick Revision Summary

Aspect	Key point
Discovery	Barbara McClintock, maize Ac/Ds system, 1940s-50s (Nobel Prize 1983)
Two classes	Class I retrotransposons (RNA intermediate, copy-and-paste) vs Class II DNA transposons (cut-and-paste)
DNA transposon mechanism	Transposase excises element via TIRs, inserts at staggered cut, creating new TSDs
LTR retrotransposons	Retrovirus-like (gag, pol, LTRs); reverse-transcribed in a cytoplasmic VLP, then integrated
Non-LTR retrotransposons	LINE-1 (autonomous) and Alu/SINEs (non-autonomous); inserted by target-primed reverse transcription (TPRT)
Human genome share	~45% TE-derived overall; LINEs/SINEs (~42%) far outnumber DNA transposons (~3%)
Significance	Mutagenesis and disease, genome rearrangement, source of new regulatory/coding sequence, gene domestication (e.g. RAG1/2, syncytin)
Host silencing	DNA methylation, repressive chromatin, and the piRNA pathway in the germline

Compiled study note — molecular biology / genetics reference. Figures are original schematic diagrams prepared for this note and are not to scale.