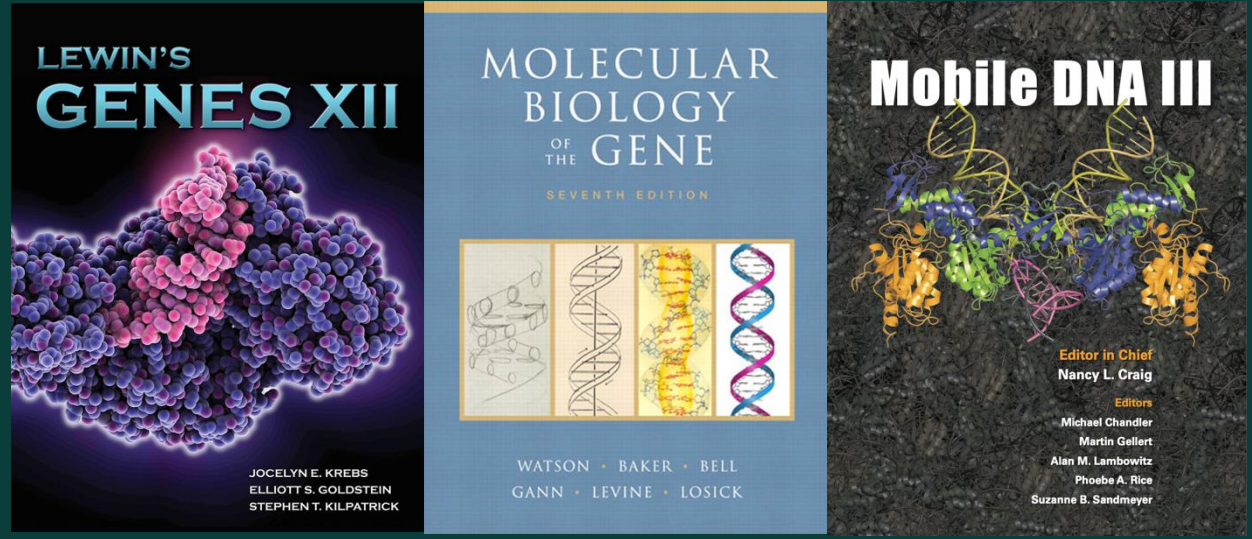


Copying, Moving & Rearranging the Genome

Replication · Extrachromosomal Elements · Recombination · Transposons & Retroviruses

Sources: Lewin's Genes XII (2018) · Watson, Molecular Biology of the Gene 7e (2013) · Mobile DNA III (Craig, Chandler, Gellert & Lambowitz eds., ASM Press, 2015)

Genes, Genomas y Proteomas

Four Machines That Act on DNA

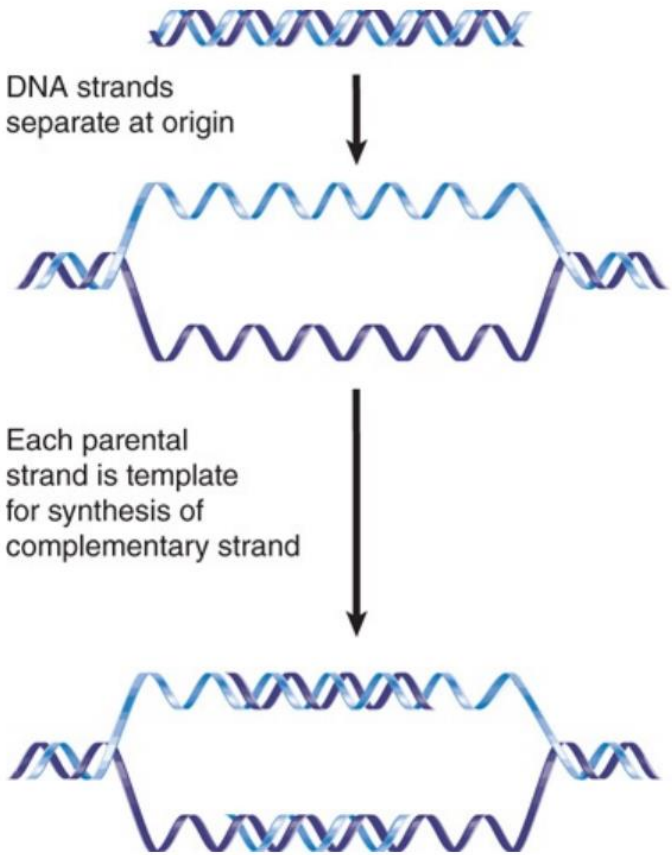
- **Replication —** copy the whole genome faithfully, and only once per cell cycle.
- **Extrachromosomal elements —** replicons that live beside the chromosome (plasmids) and can move between cells.
- **Recombination —** break and rejoin DNA to repair breaks, reshuffle alleles, or move defined segments.
- **Transposons & retroviruses —** DNA (or an RNA copy) that moves itself to new locations.
- **The unifying idea:** each is DNA breakage and synthesis carried out by a dedicated protein machine — the same chemistry, different jobs.

PART 1 · REPLICATION

Copying the Genome

The Replicon: Where Copying Begins

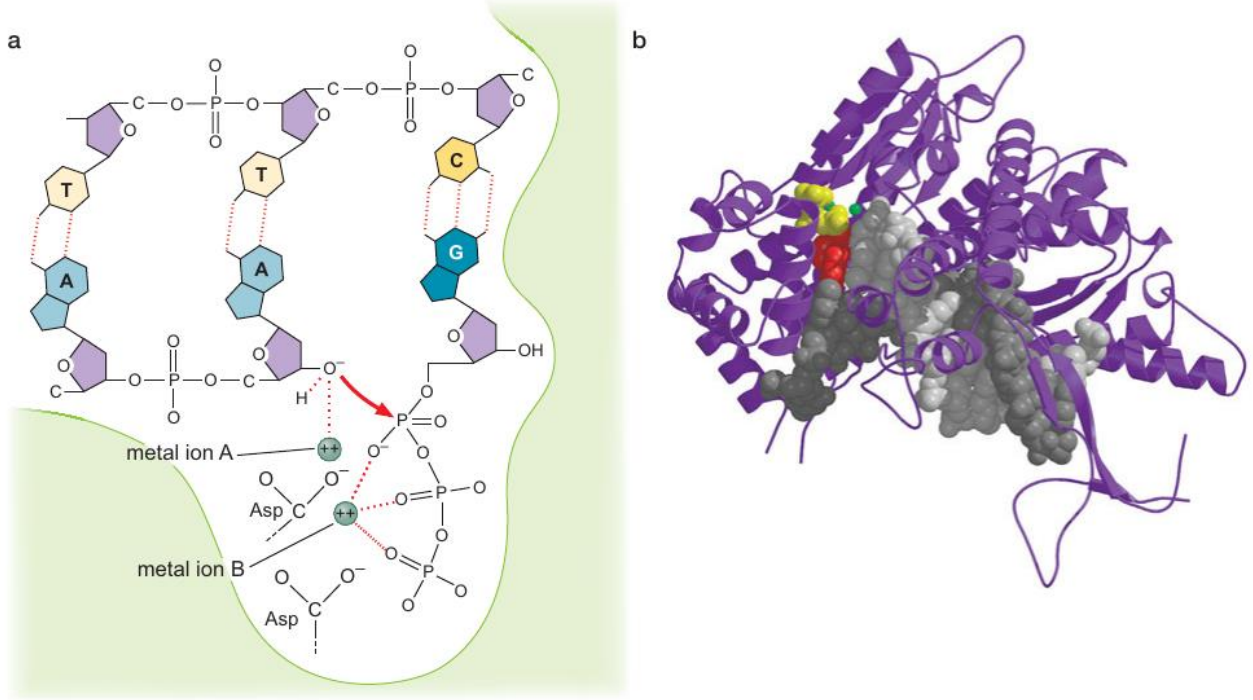
- **Semiconservative:** each parental strand is the template for a new one (proved by Meselson–Stahl).
- **The replicon** is the unit of replication, controlled by a cis-acting origin.
- **Bacteria:** one origin, oriC. DnaA-ATP melts an AT-rich region and loads the helicase.
- **Eukaryotes:** many origins, each licensed once per cycle (ORC → Cdc6/Cdt1 → MCM helicase).
- **Bidirectional:** two forks move outward from each fired origin.



Lewin's Genes XII — Fig 10.1 — an origin is where replication initiates; parental strands separate and each templates a complementary new strand, giving two daughter duplexes.

The Chemistry of Adding a Base

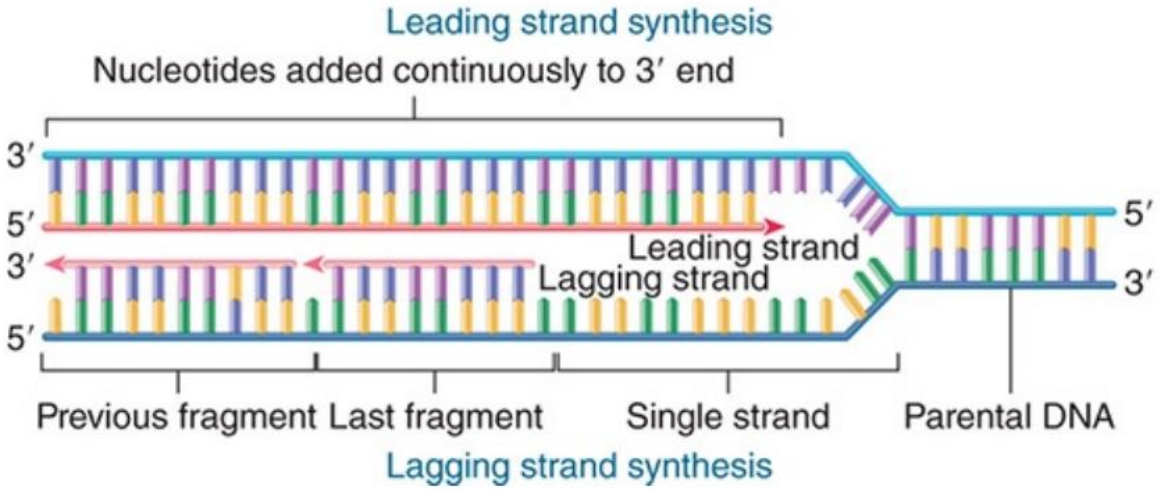
- **Substrates:** a primer:template junction with a free 3'-OH, plus the four dNTPs.
- **Two-metal-ion mechanism:** metal A activates the primer 3'-OH to attack the incoming dNTP's α-phosphate; metal B stabilizes the pyrophosphate leaving group.
- **Directionality:** synthesis is strictly 5'→3' — a chemical constraint, and the reason the lagging strand must exist.
- **Driving force:** release and hydrolysis of pyrophosphate makes each addition effectively irreversible.
- **Proofreading:** a separate 3'→5' exonuclease excises a misinserted base before synthesis continues — ~100–1000× more accurate.



Watson, Molecular Biology of the Gene — Fig 9-6 — two active-site metal ions catalyze nucleotide addition: metal A activates the primer 3'-OH to attack the incoming dNTP, and metal B stabilizes the pyrophosphate leaving group.

The Replication Fork Is Asymmetric

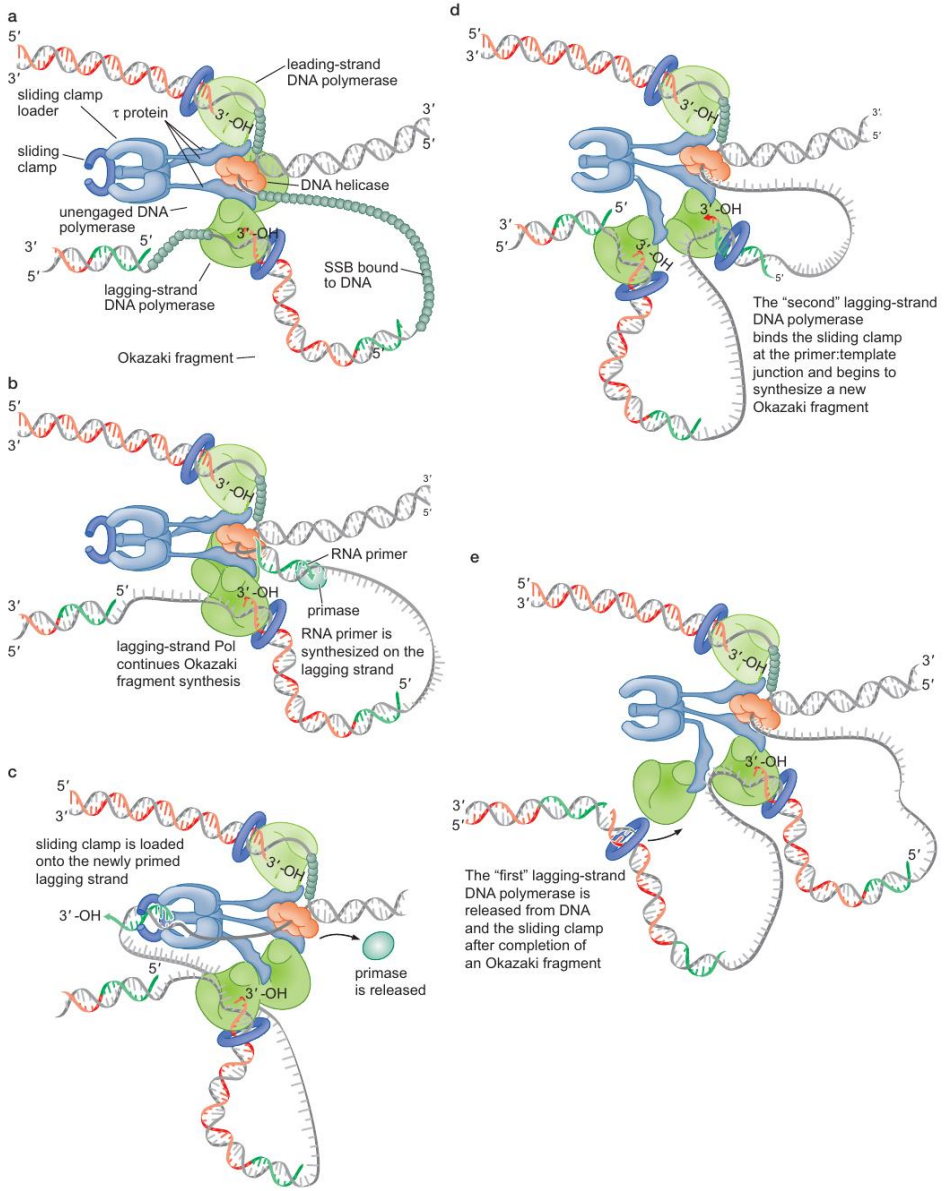
- **The problem:** polymerases only extend a 3'-OH, but the two strands are antiparallel.
- **Leading strand:** synthesized continuously, following the fork.
- **Lagging strand:** made discontinuously as Okazaki fragments, each started by an RNA primer.
- **Support enzymes:** helicase unwinds the duplex; SSB coats the exposed single strands; primase lays the primers.
- **Finishing:** primers are removed and gaps filled (Pol I), then ligase seals each nick.



Lewin's Genes XII — Fig 11.9 — the leading strand is made continuously while the lagging strand is built discontinuously as short Okazaki fragments.

One Machine Copies Both Strands

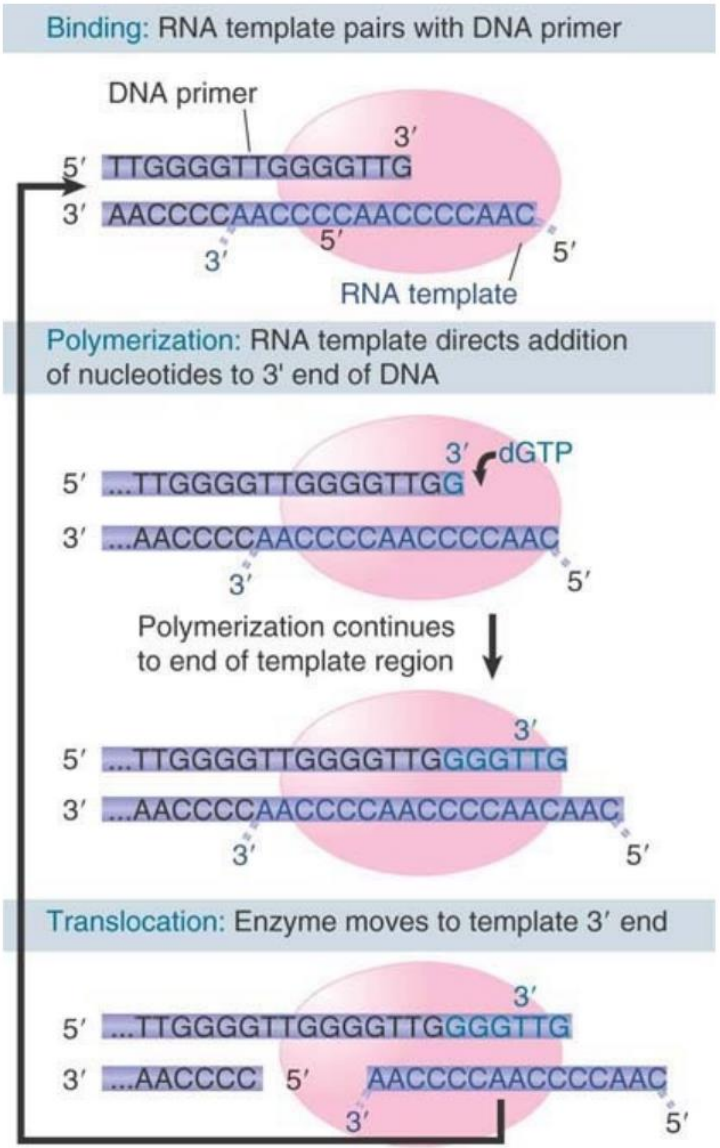
- **Pol III holoenzyme:** two (or three) polymerase cores are coupled at a single fork.
- **Sliding clamp (β / PCNA):** a protein ring that encircles DNA and tethers the polymerase → high processivity.
- **Clamp loader:** uses ATP to open the clamp and place it on each new primer.
- **Trombone model:** the lagging-strand template loops out so both cores move in the same direction.
- **Fidelity:** base selection + 3'→5' proofreading exonuclease → ~1 error per 10^7–10^9 bases.



Watson, Molecular Biology of the Gene — Fig 9-23 — the 'trombone' model: the two replisome polymerases stay coupled at the fork while the lagging-strand template loops out, so both strands extend toward fork movement.

Solving the Ends: Telomerase

- **End-replication problem:** the lagging strand can't be completed at a linear end, so chromosomes would shorten each round.
- **Telomeres:** tandem G-rich repeats (TTAGGG in humans) with a 3' single-strand overhang.
- **Telomerase:** a reverse transcriptase that carries its own RNA template to extend the G-strand.
- **Consequence:** without telomerase telomeres erode → senescence; reactivation is a hallmark of many cancers.
- **Protection:** a t-loop plus the shelterin complex hide the end so it isn't read as a break.



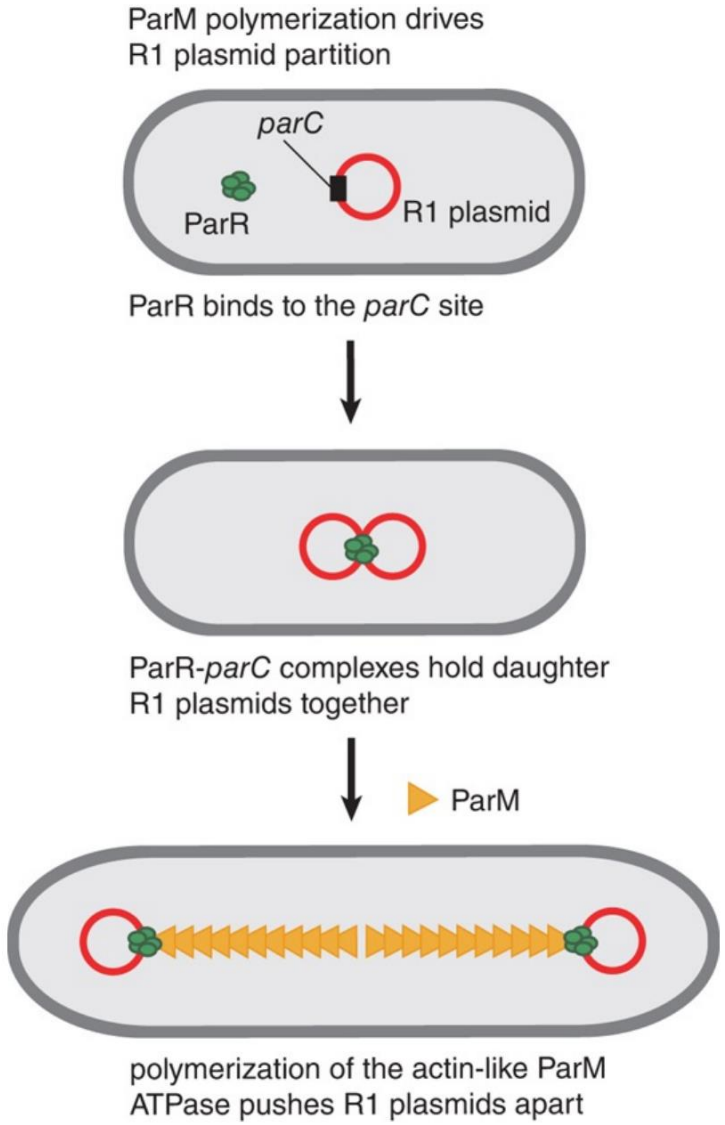
Lewin's Genes XII — Fig 7.33 — telomerase base-pairs its RNA template to the single-stranded end and adds G/T repeats one base at a time to extend the telomere.

PART 2 · EXTRACHROMOSOMAL ELEMENTS

Replicons Beside the Chromosome

Plasmids: Independent Replicons

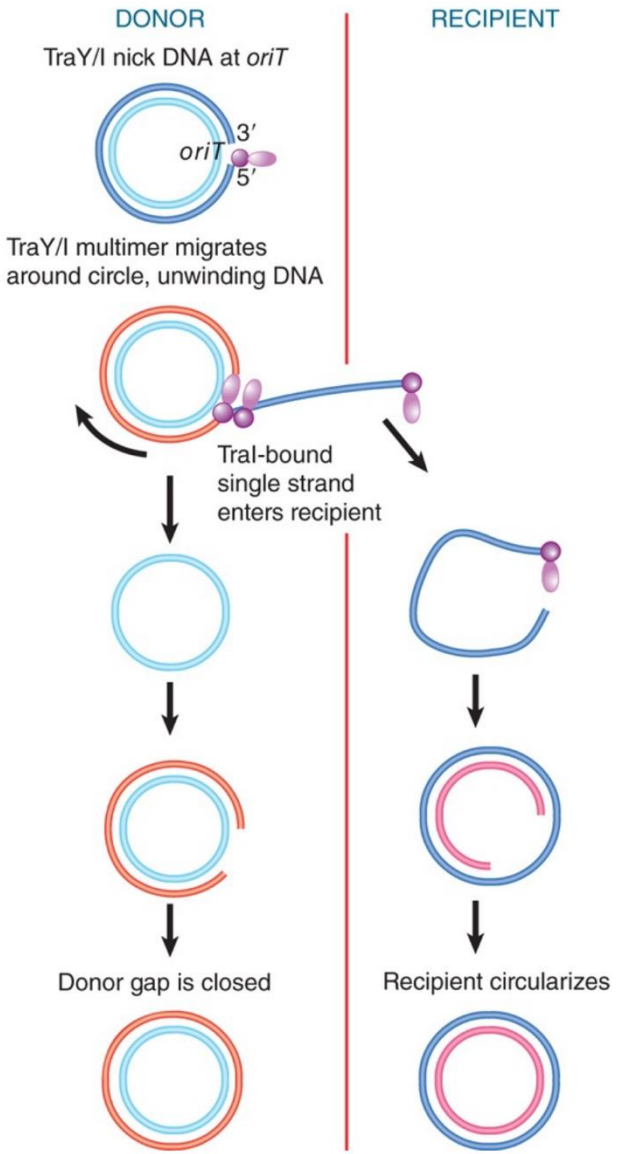
- **Plasmids** are autonomous circular replicons carrying non-essential but useful genes (resistance, virulence, metabolism).
- **Copy number** is set by the plasmid's own control system (iterons or antisense RNA).
- **Incompatibility:** two plasmids that share a control system can't be stably co-inherited.
- **Partition:** low-copy plasmids are actively segregated — ParM filaments push copies to opposite poles.
- **Addiction:** toxin–antidote systems kill any daughter cell that loses the plasmid.



Lewin's Genes XII — Fig 12.14 — low-copy R1 plasmid partition: ParR binds the parC site and the actin-like ParM ATPase polymerizes between the two copies, pushing them to opposite poles.

Conjugation & Rolling-Circle Transfer

- **Conjugation:** the F plasmid transfers a copy of itself to a recipient through a pilus.
- **Rolling circle:** transfer replication nicks the plasmid at oriT; a single strand, 5' end first, crosses into the recipient.
- **Result:** each cell rebuilds the complementary strand, so both become F^+.
- **Hfr:** an F integrated into the chromosome drags host genes along — the historical basis of gene mapping.
- **Beyond F:** rolling-circle replication also copies many phages and is exploited to amplify DNA in vitro.



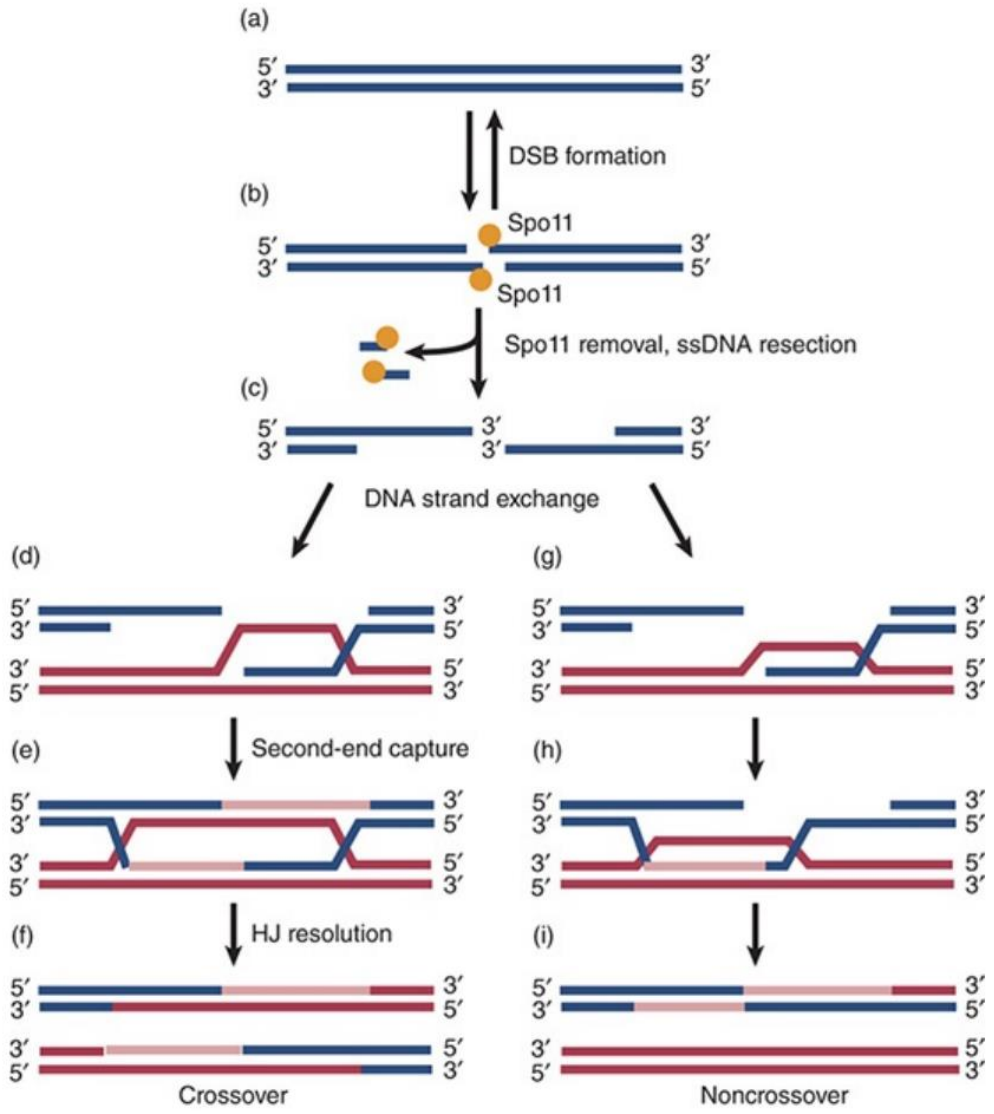
Lewin's Genes XII — Fig 12.11 — the F plasmid is nicked at oriT, and a single strand (5' end led by TraI) is transferred to the recipient; each cell rebuilds the complement.

PART 3 · RECOMBINATION

Breaking and Rejoining DNA

Homologous Recombination Repairs Breaks

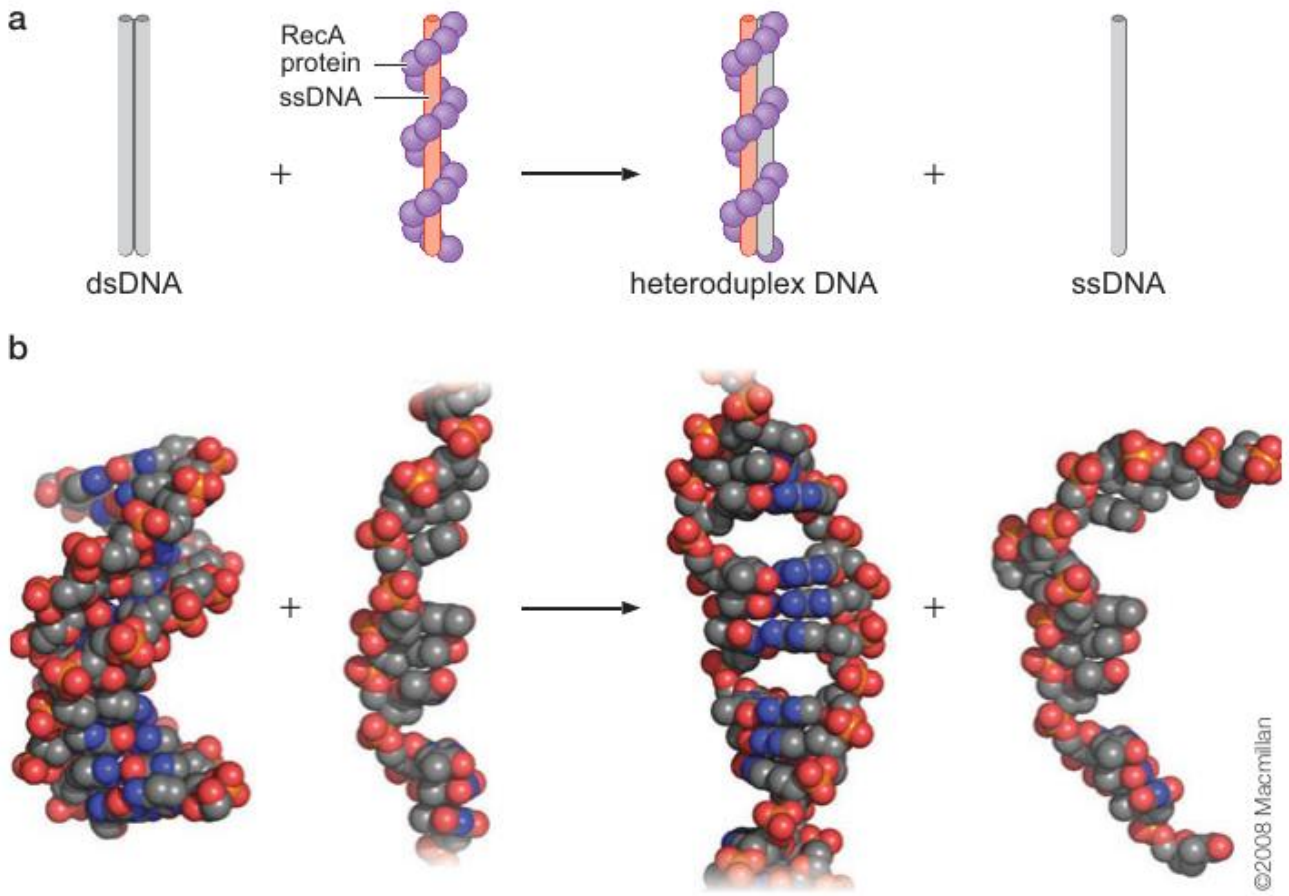
- **Purpose:** repair a double-strand break using an intact homologous DNA as template — error-free.
- **DSBR model:** break → 5' resection → a 3' tail invades the homolog (D-loop) → repair synthesis.
- **Double Holliday junction:** second-end capture forms two junctions; how they're cut sets crossover vs non-crossover.
- **SDSA:** the invading strand is displaced and re-anneals → non-crossover, no junction.
- **In meiosis:** Spo11 makes breaks on purpose → crossovers that shuffle alleles.



Lewin's Genes XII — Fig 13.15 — Spo11 makes a double-strand break; ends are resected to 3' tails, one invades the homolog, and repair builds a double Holliday junction that is resolved.

The Strand-Exchange Machine

- **RecA / Rad51:** coat ssDNA into a filament that searches the genome for homology and invades it.
- **RecBCD:** resects a broken end until it reaches a Chi site, then loads RecA onto the 3' tail.
- **Resolution:** RuvAB drives branch migration; the RuvC resolvase cuts the Holliday junction.
- **Core step:** homology search + strand exchange is the committed, ATP-driven heart of recombination.
- **In the lab:** the same strand-invasion logic powers CRISPR homology-directed repair.

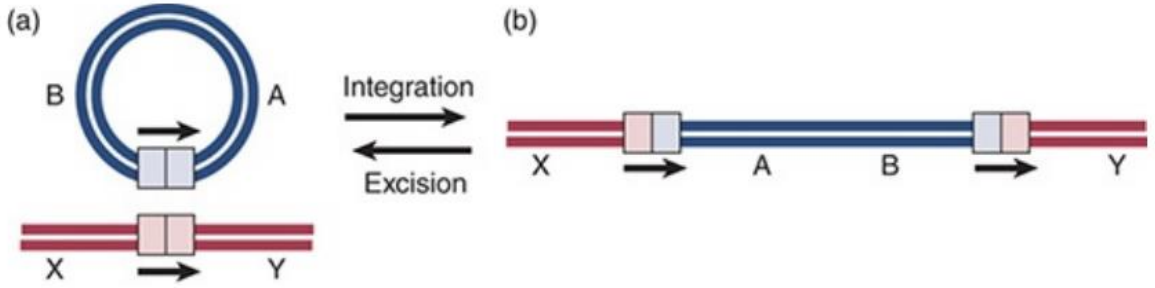


Watson, Molecular Biology of the Gene — Fig 11-11 — RecA-promoted strand exchange: a RecA-coated single strand pairs with a homologous duplex and swaps strands to form heteroduplex DNA.

Site-Specific Recombination: Precise Cut-and-Paste

Lewin 13 · Watson 12

- **Recombinases** act only at specific short sites — no long homology and no new DNA synthesis.
- **Outcome by orientation:** sites in the same vs opposite orientation give excision/insertion vs inversion.
- **Two chemistries:** tyrosine recombinases (λ Int, Cre) and serine recombinases differ in how they cut and swap strands.
- **Phage λ:** Int integrates the genome at attB × attP; the reverse reaction excises the prophage.
- **Toolbox:** Cre/lox and FLP/FRT are everyday genome-engineering systems built on this.



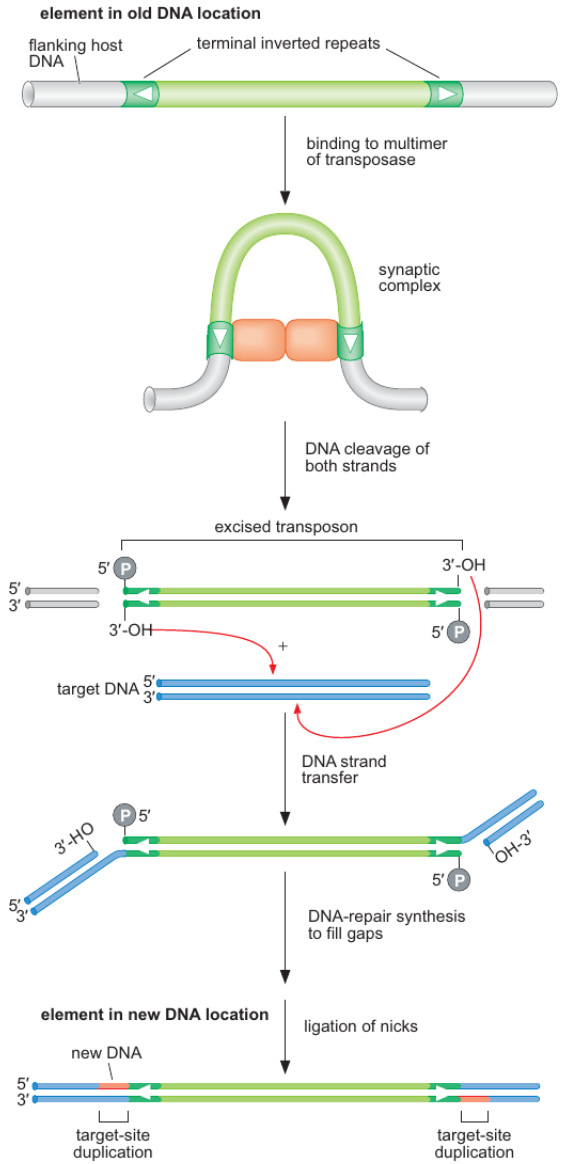
Lewin's Genes XII — Fig 13.2 — site-specific recombination at defined sites; integration inserts one DNA into another and the reverse reaction excises it, set by site orientation.

PART 4 · MOBILE DNA

Sequences That Move Themselves

DNA Transposons: Cut-and-Paste vs Replicative

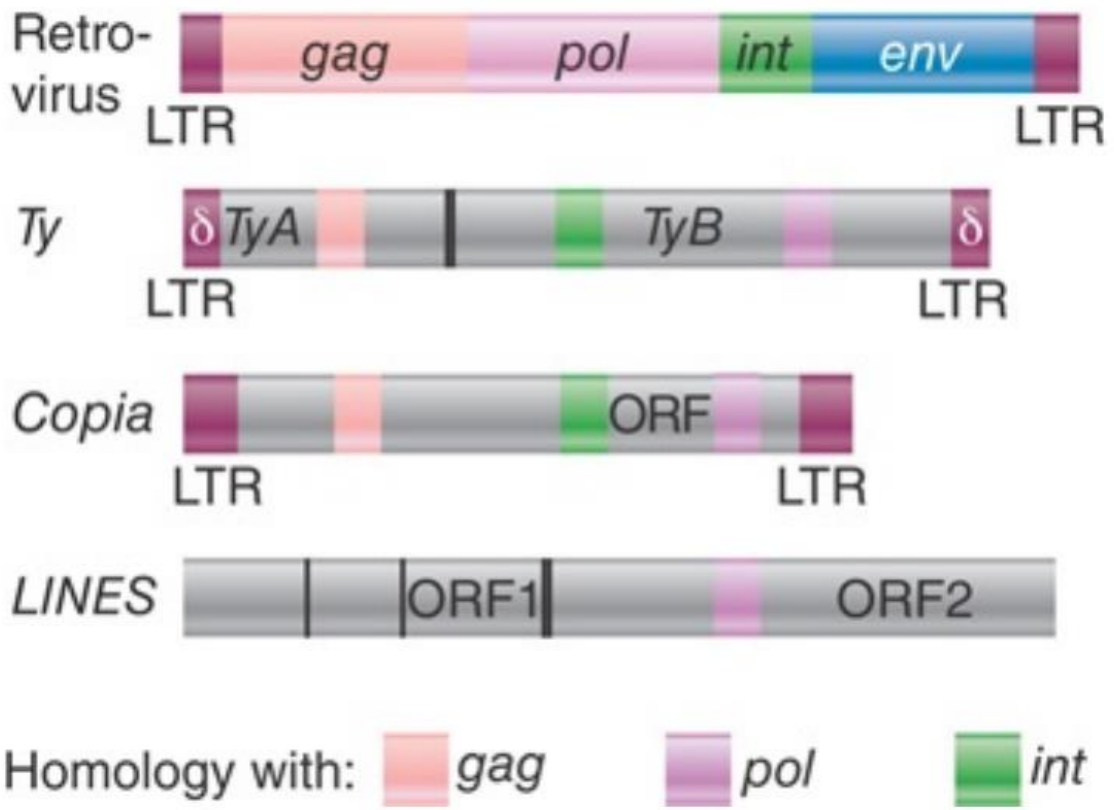
- **Transposons** are DNA segments that move to new sites, carried by a transposase they encode (bounded by inverted terminal repeats).
- **Shared signature:** transposase makes staggered cuts in the target; filling the gaps leaves target-site duplications.
- **Cut-and-paste:** the element is excised and reinserted elsewhere (nonreplicative) — Tn5/Tn10, mariner, P element.
- **Replicative:** a copy is made and inserts via a cointegrate that a resolvase then separates — Tn3, phage Mu.
- **One chemistry:** most transposases share a DDE active site running the same metal-dependent strand-transfer reaction — the unifying theme of Mobile DNA III.



Watson, Molecular Biology of the Gene — Fig 12-19 — cut-and-paste transposition: transposase excises the element and inserts it at a staggered cut in new target DNA, leaving target-site duplications.

Retroelements Move Through RNA

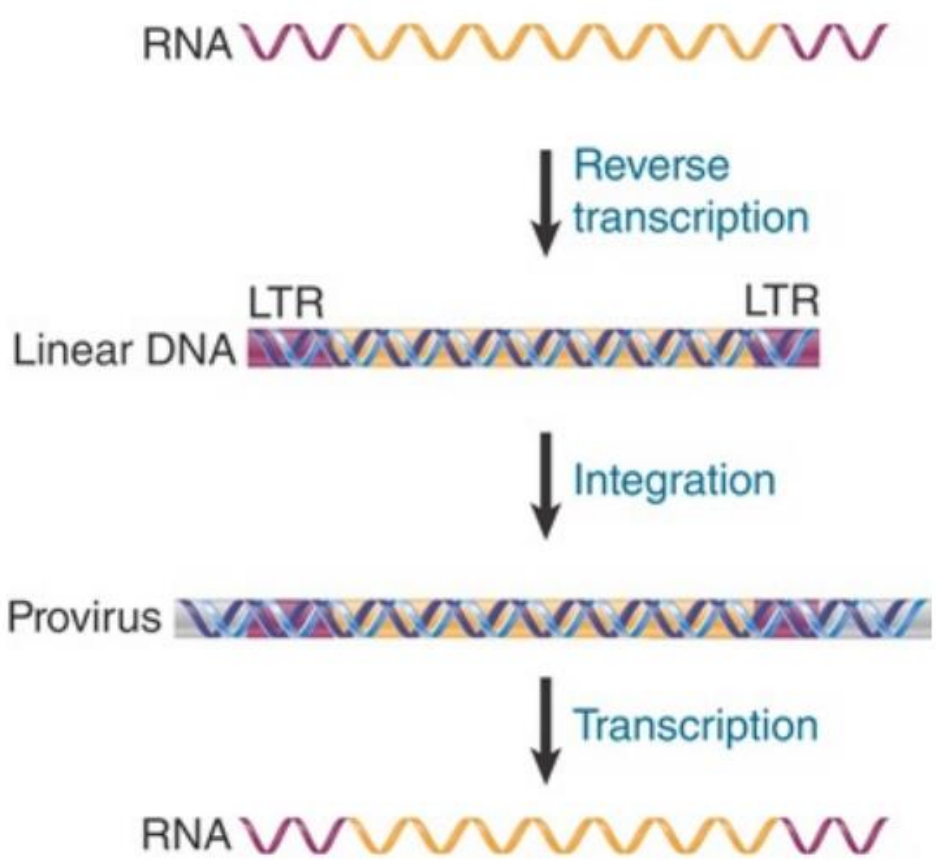
- **Retroelements** move via an RNA intermediate + reverse transcriptase — one-way, so they only expand genomes.
- **Two families:** LTR retrotransposons (integrase-based, like retroviruses) vs non-LTR elements (LINEs/SINEs) that insert by target-primed reverse transcription.
- **Scale:** ~45% of the human genome is retroelement-derived — the single biggest force shaping mammalian genome size.
- **Copy, don't move:** unlike DNA transposons, the original stays put and a new DNA copy is made each cycle.
- **Ahead:** we look at the non-LTR (TPRT) mechanism, then retroviruses, then the ancient group II introns they descend from.



Lewin's Genes XII — Fig 15.30 — retroelement organization: LTR retrotransposons (Ty, copia) and retroviruses carry gag/pol (± int, env) flanked by LTRs; non-LTR LINEs keep only ORF1/ORF2 (reverse transcriptase) and lack LTRs.

Retroviruses: Mobile DNA That Went Infectious

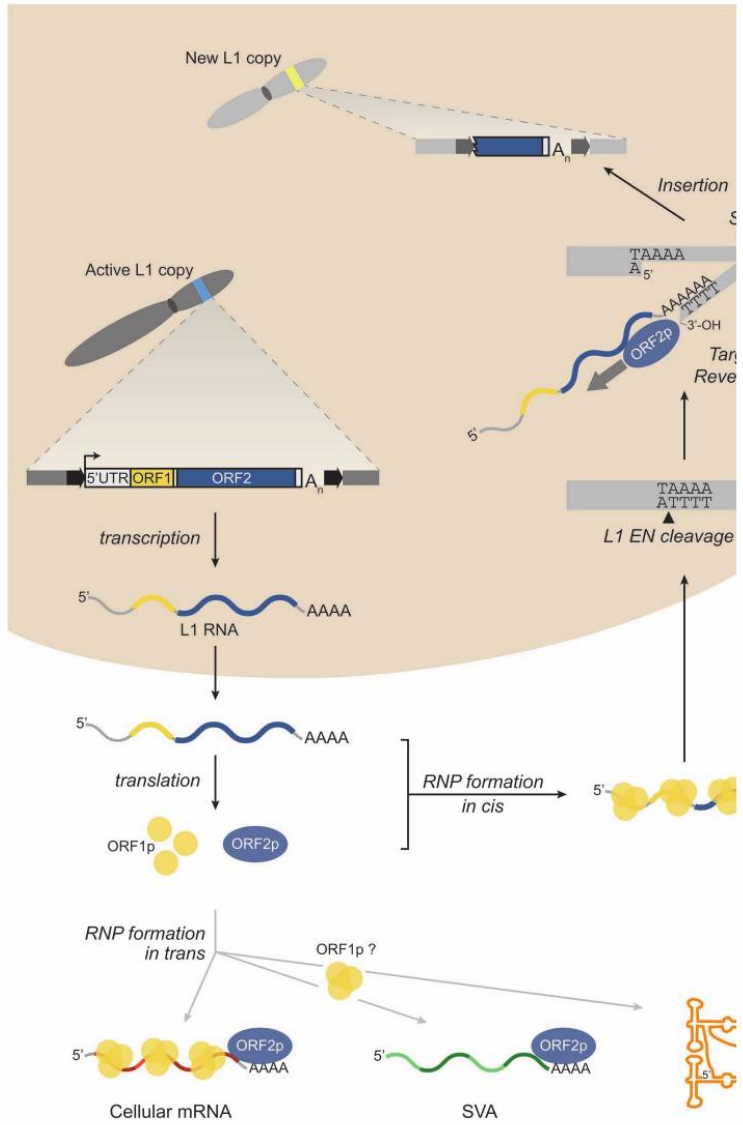
- **A retrovirus** is essentially an LTR retroelement that gained an envelope and became infectious.
- **Life cycle:** RNA genome → reverse transcriptase → double-stranded DNA → integrase inserts a provirus.
- **Expression:** host machinery transcribes the provirus into new genomes and gag–pol–env polyproteins.
- **Permanence:** integration is irreversible → the basis of HIV persistence and of retroviral gene-therapy vectors.
- **Evolution:** reverse transcriptase is error-prone → rapid variation and drug resistance.



Lewin's Genes XII — Fig 15.20 — the retroviral RNA genome is reverse-transcribed to dsDNA, integrated as a provirus, then transcribed to make new virus.

Non-LTR Retrotransposition: Copy-and-Paste by TPRT

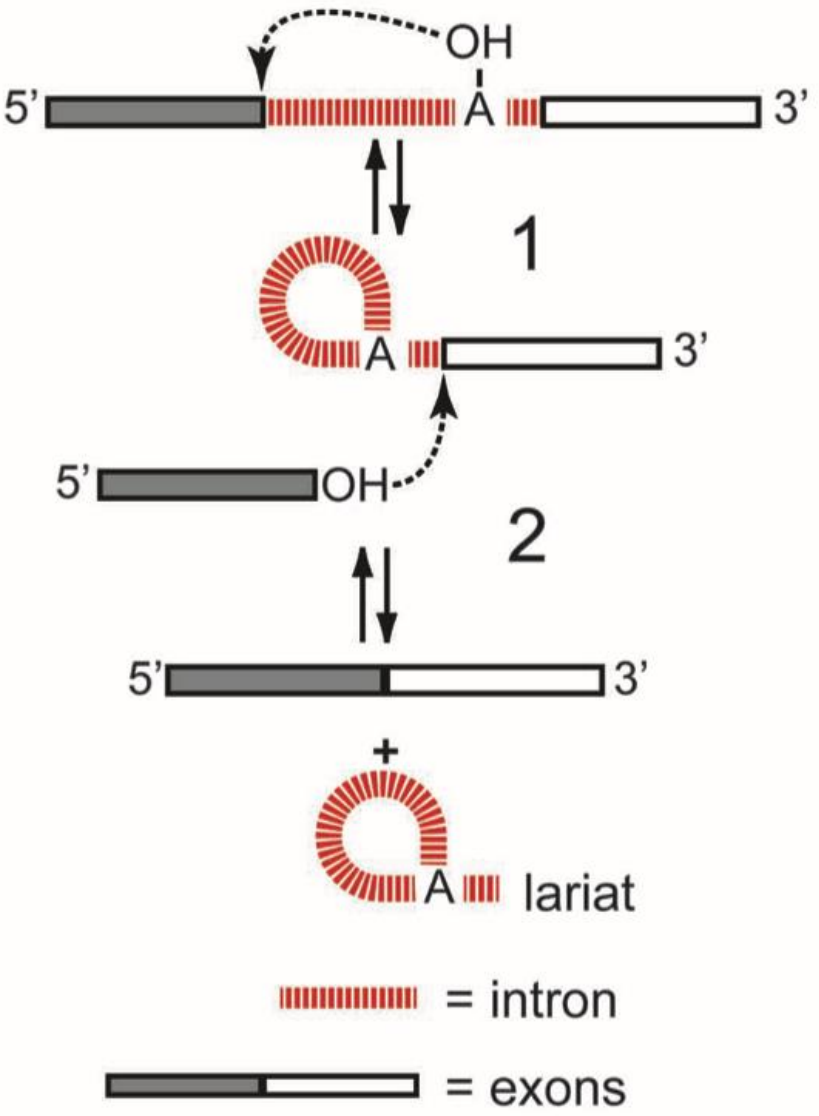
- **No LTRs, no integrase:** non-LTR elements copy themselves straight into a chromosomal nick — target-primed reverse transcription (TPRT).
- **The two-part engine:** L1 ORF2p is an endonuclease + reverse transcriptase; it nicks the target and uses the freed 3'-OH to prime cDNA synthesis directly on the element's own RNA.
- **cis-preference:** ORF1p/ORF2p bind the very mRNA that encoded them, so it is the active copies that move.
- **SINEs cheat:** Alu and SVA make no proteins — they hijack the L1 machinery in trans.
- **Genome impact:** insertions are often 5'-truncated with target-site duplications; ~a third of the human genome is L1/Alu-derived — a live source of mutation and structural variation.



Mobile DNA III (Craig et al., 2015) — Fig 3 (ch 51) — the LINE-1 cycle: ORF1p/ORF2p bind their own mRNA (cis-preference) to form an RNP; in the nucleus ORF2p nicks the target and primes reverse transcription of L1 RNA (TPRT), inserting a new, often 5'-truncated copy. Alu/SVA hijack the machinery in trans.

Group II Introns: Roots of the Eukaryotic Gene

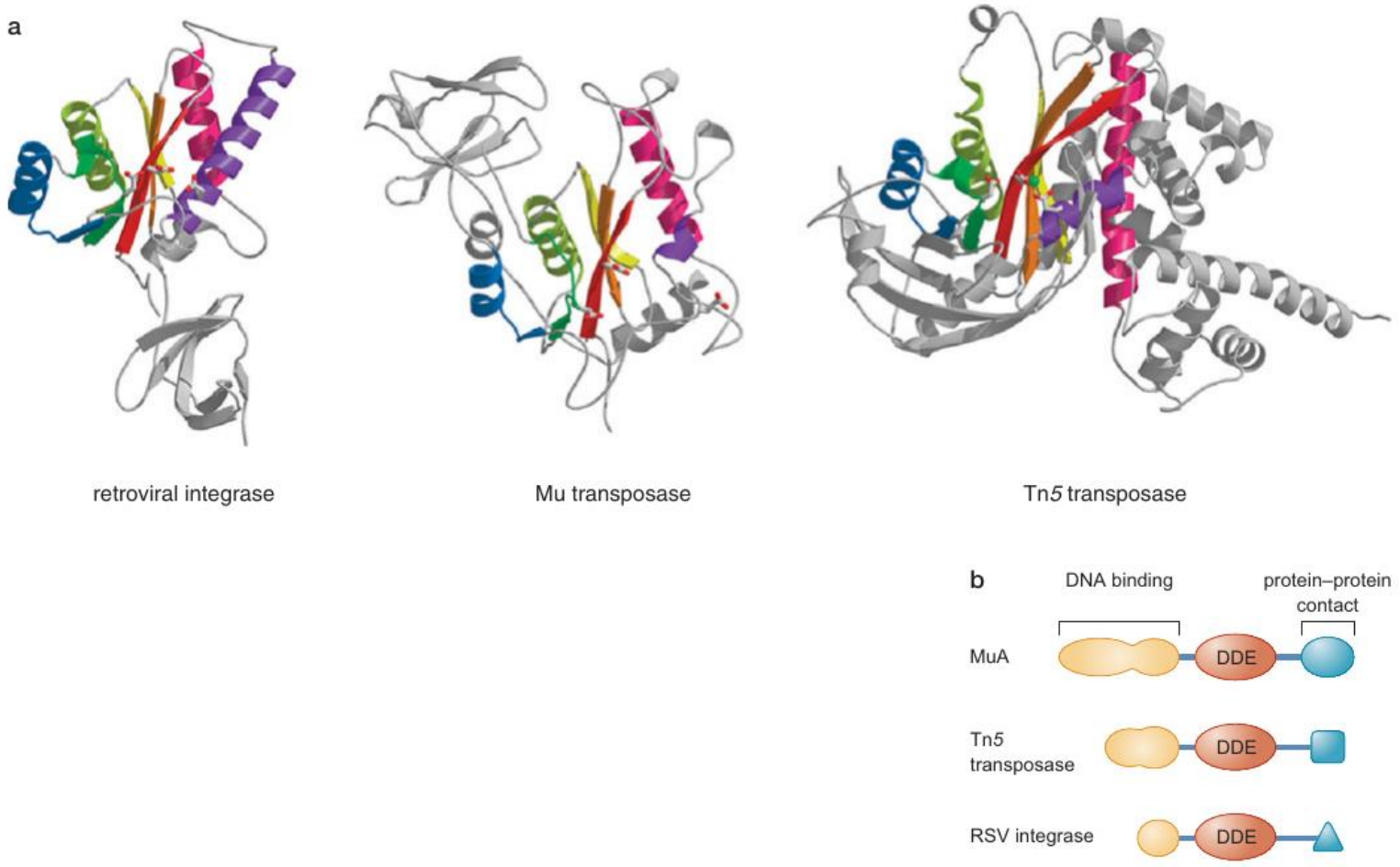
- **A ribozyme that moves:** bacterial group II introns are self-splicing RNAs that also encode a reverse transcriptase — retroelement and enzyme in one.
- **Splicing chemistry:** two transesterifications through a branch-point adenosine make a lariat — the exact chemistry the spliceosome uses.
- **Retrohoming:** run in reverse, the intron reverse-splices into a DNA target and its RT copies it in place — a TPRT-like reaction.
- **Evolutionary payoff:** group II introns are the likely ancestors of spliceosomal introns, the snRNAs, and non-LTR retroelements.
- **One thread:** mobile DNA didn't just clutter genomes — it may have built the interrupted eukaryotic gene.



Mobile DNA III (Craig et al., 2015) — Fig 1A (ch 52) — group II intron self-splicing: a branch-point adenosine attacks the 5' splice site (step 1) to form a lariat, then exon ligation (step 2) frees the intron. Run in reverse, the same chemistry reverse-splices the intron into a DNA target.

Transposases & Integrases: One Catalytic Core

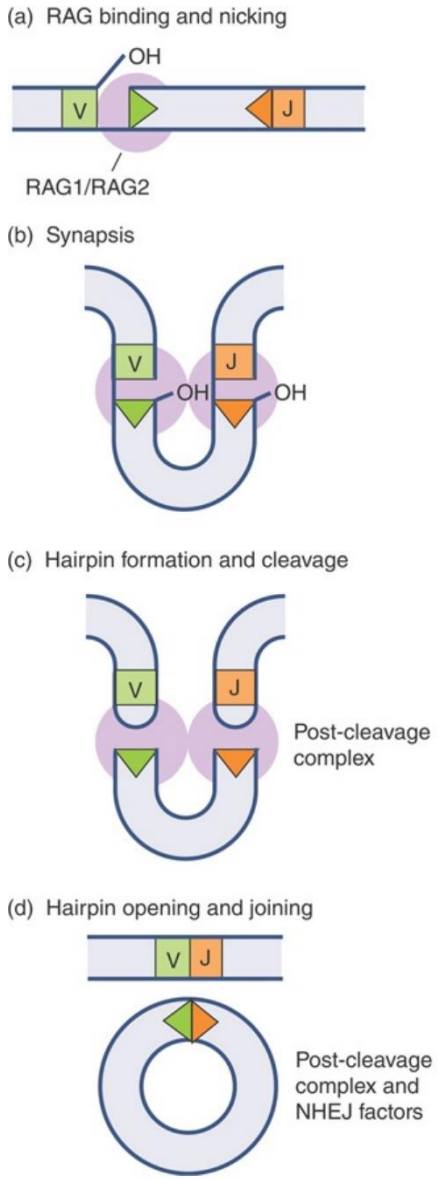
- **Shared chemistry:** DNA transposases and retroviral integrase adopt a common fold and use the same DDE active-site motif.
- **Same reaction:** a metal-dependent transesterification cuts DNA and joins the mobile-element end to a new target site.
- **Implication:** cut-and-paste transposition, replicative transposition, and retroviral integration are variations on one enzymatic theme.
- **Evolutionary link:** DNA transposons, LTR retrotransposons, and retroviruses are branches of a single mobile-DNA family.
- **Why it matters:** this shared active site is the target of HIV integrase inhibitors (e.g. raltegravir) used in the clinic.



Watson, Molecular Biology of the Gene — Fig 12-24 — transposase and retroviral-integrase catalytic domains share a common fold and a DDE active-site motif — marking them as mechanistic relatives.

Why It Matters: Genomes, Immunity & Tools

- **Genome shaping:** mobile DNA and recombination create mutations, duplications, and shuffled exons — raw material for evolution.
- **Immunity:** V(D)J recombination builds antibody diversity from gene segments — RAG1/2 (a domesticated transposase) cut at signal sequences.
- **Targeted variation:** diversity-generating retroelements use error-prone reverse transcription to hypervary a chosen gene — mobile DNA as a controlled mutagen.
- **Engineering:** tamed transposons (Sleeping Beauty, piggyBac) and recombinases (Cre/lox) deliver and edit genes.
- **One chemistry:** the same break–copy–move logic is both a driver of disease and a toolbox.



Lewin's Genes XII — Fig 16.14 — V(D)J recombination breaks and rejoins at recombination signal sequences (RSSs) to assemble a complete antigen-receptor gene from separate segments.

One Logic, Four Machines

- **Common core:** dedicated proteins break, copy, or join DNA with defined, controlled chemistry.
- **Division of labor:** replication copies faithfully and once per cycle; recombination repairs and reshuffles; transposons and retroelements relocate DNA.
- **Fidelity vs. change:** proofreading and origin licensing guard the genome, while recombination and mobile DNA generate the variation evolution needs.
- **Recurring players:** reverse transcriptase (telomerase, retroelements, retroviruses); strand invasion (HR, CRISPR-HDR); recombinases (λ, Cre, RAG).