COURSE BSc (BIOTECHNOLOGY) 2nd year

PAPER CODE: BBT-203

PAPER TITLE: MOLECULAR BIOLOGY

By:

DR. ANNIKA SINGH

DEPARTMENT OF BIOTECHNOLOGY

INSTITUTE OF BIOSCIENCE AND BIOTECHNOLOGY

Chhatrapati Shahu Ji Maharaj University, Kanpur
MOLECULAR BIOLOGY
Dr ANNIKA SINGH
DEPARTMENT OF
BIOTECHNOLOGY

MOLECULAR BIOLOGY

BBT-203
Paper Third
Molecular Biology **(Max. marks: 100)**

Unit I: Molecular basis of life: Structure and function of DNA and RNA; Organization of bacterial and eukaryotic chromosomes; Histone and DNA; Heterochromatin and Euchromatin; DNA reassociation kinetics (Cot curve analysis); DNase I hypersensitive regions; DNA methylation and Imprinting.

Unit II: Structure of DNA - A-,B-, Z- and triplex DNA: DNA replication in both prokaryotes

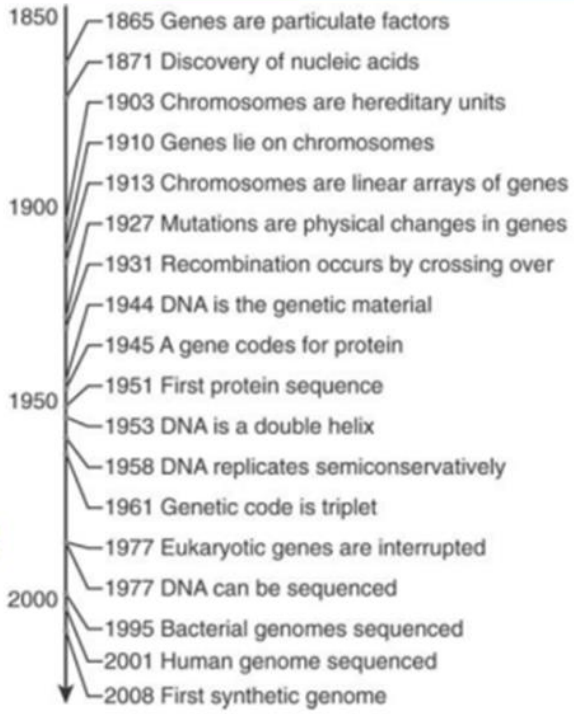
Unit III: Insertion elements and transposons: Transposable elements in *Drosophila* and maize; Organization of genetic material: split genes, overlapping genes, pseudo genes, cryptic genes.

Organization of genetic material: split genes, overlapping genes, pseudo genes, cryptic genes.

Unit IV: Prokaryotic Transcription- Transcription unit; Properties of Promoters- Constitutive and Inducible promoters; prokaryotic gene expression (lac, his, trp operon, catabolite repression); Initiation; Attenuation; Termination; Eukaryotic transcription and regulation; RNA polymerase (RNA polymerase I, II, III); Eukaryotic promoters, trans-acting factors and enhancers; General Transcription factors; TATA binding proteins (TBP) and TBP associated factors (TAF).

Unit V: Translation machinery; Ribosomes; Composition and assembly; Universal genetic code; Degeneracy of codons; Termination codons; Wobble hypothesis; Mechanism of initiation, elongation and termination; Co- and post-translational modifications; Genetic code in mitochondria.

A BRIEF HISTORY OF GENETICS



Chhatrapati Shahu Ji Maharaj University, Kanpur
MOLECULAR BIOLOGY
Dr ANNIKA SINGH
DEPARTMENT OF
BIOTECHNOLOGY

MOLECULAR BIOLOGY

GENE is the functional unit of heredity. Each gene is a sequence within the genome that functions by giving rise to a discrete product, which can be a polypeptide or an RNA.

ALLELES a gene can exist in alternative forms, called **alleles**.

CHROMOSOMES carry the genes.

Each chromosome consists of a linear array of genes, and each gene resides at a particular location on the chromosome, called as a genetic **locus**.

The alleles of a gene are the different forms that are found at its locus. **GENOME** is the complete set of genes of an organism. it is defined by the complete DNA sequence, **TRANSCRIPTOME** is the complete set of genes expressed under particular conditions. It is defined in terms of the set of RNA molecules present in a single cell type, a more complex assembly of cells, or a complete organism.

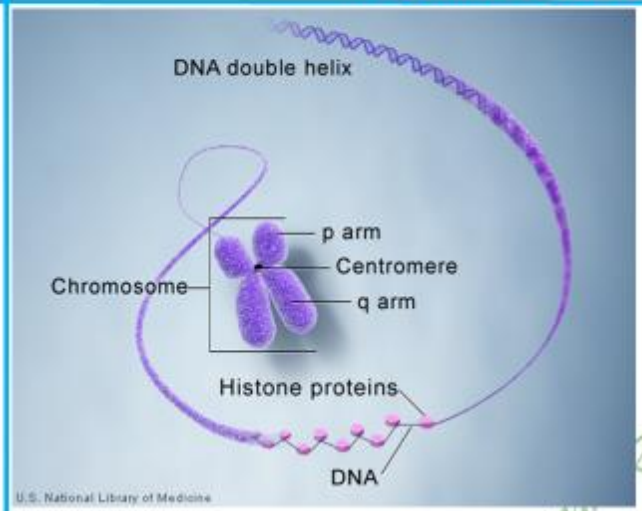
PROTEOME is the complete set of polypeptides encoded by the whole genome or produced in any particular cell or tissue. The maximum number of polypeptide-encoding genes in the genome can be identified directly by characterizing open reading frames (ORFs).

MOLECULAR BIOLOGY

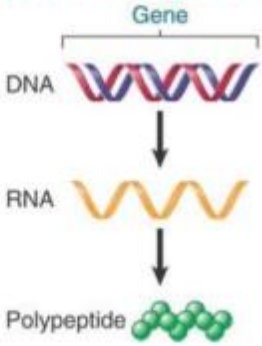
- The genetic material is seen in the form of the **NUCLEOID** in bacteria
- The genetic material is seen in as the mass of **CHROMATIN** in eukaryotic nuclei at interphase (between divisions), or as condensed chromosomes during mitosis
- DNA molecule is packaged into thread-like structures **(chromatin)** called chromosomes.
- Each chromosome is made up of DNA tightly coiled many times around proteins called histones
- Chromosomes are visible under a microscope during cell division only.
- Each chromosome has a constriction point called the centromere, which divides the chromosome into two sections, or "arms."
- The short arm of the chromosome is labeled the "p arm."
- The long arm of the chromosome is labeled the "q arm."
- The location of the centromere on each chromosome gives the chromosome its characteristic shape, and can be used to help describe the location of specific genes.

Chhatrapati Shahu Ji Maharaj University, Kanpur
MOLECULAR BIOLOGY
Dr ANNIKA SINGH
DEPARTMENT OF
BIOTECHNOLOGY

MOLECULAR BIOLOGY



- **During interphase, the chromatin is less densely packed than in the mitotic chromosome, called euchromatin, is relatively dispersed and occupies most of the nucleoplasm.**



- Some regions of chromatin are very densely packed, called **heterochromatin**, is typically found at centromeres, telomeres and highly repetitive sequences.

TRANSPOSONS

Transposons Are Discrete Sequences In The Genome That Are Mobile

- A **transposable element** or **TE or transposon**, or **jumping gene;** is a DNA sequence that can change its position within a genome
- Transposition often results in duplication of the same genetic material.
- **Barbara McClintock's discovery of them earned her a Nobel Prize in 1983.**
- Transposable elements make up a large fraction of the genome and are responsible for much of the mass of DNA in a eukaryotic cell.
- Generally TEs are selfish genetic elements, many are important in genome function and evolution.
- Transposition sometimes creating or reversing mutations and altering the cell's genetic identity and genome size.
- They can be a major source of mutations in the genome

Chhatrapati Shahu Ji Maharaj University, Kanpur
MOLECULAR BIOLOGY
Dr ANNIKA SINGH
DEPARTMENT OF
BIOTECHNOLOGY

TRANSPOSONS

- Transposition does not rely on any relationship between the sequences at the donor and recipient sites.
- Transposons are restricted to moving themselves, and sometimes additional sequences, to new sites elsewhere within the same genome
- Transposon content in eukaryotes varies over a wide range, from 4% in yeast to 70% or more in some amphibians and plants.
- *Zea mays* maize) transposable elements make up 85% of the genome.
- Transposon is an independent entity that resides in the genome
- Transposons are also very useful to researchers as a means to alter DNA inside a living organism.
- The simplest bacterial transposons are called **insertion sequence (IS)** elements e.g. **IS1, IS4** etc.

TRANSPOSONS

- Transposition does not rely on any relationship between the sequences at the donor and recipient sites.
- Transposons are restricted to moving themselves, and sometimes additional sequences, to new sites elsewhere within the same genome
- Transposon content in eukaryotes varies over a wide range, from 4% in yeast to 70% or more in some amphibians and plants.
- *Zea mays* maize) transposable elements make up 85% of the genome.
- Transposon is an independent entity that resides in the genome
- Transposons are also very useful to researchers as a means to alter DNA inside a living organism.
- The simplest bacterial transposons are called **insertion sequence (IS)** elements e.g. **IS1, IS4** etc.

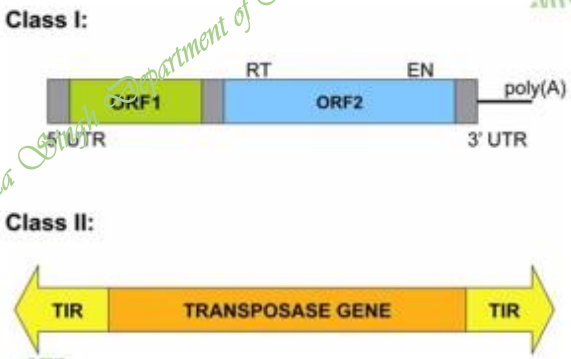
Chhatrapati Shahu Ji Maharaj University, Kanpur
MOLECULAR BIOLOGY
Dr ANNIKA SINGH
DEPARTMENT OF
BIOTECHNOLOGY

TRANSPOSONS

Transposons fall into two general classes:

- **Class I elements or Retroelements :** those whose source of mobility is the ability to make DNA copies of their RNA transcripts, which are then integrated at new sites in the genome .
- **Class II elements or DNA-type elements :** those that are able to directly manipulate DNA so as to propagate themselves within the genome
- **Class II elements** encode the protein transposase, which they require for insertion and excision, and some of these TEs also encode other proteins.
- Each transposon carries gene(s) that encode the enzyme activities required for its own transposition
- It also require ancillary products of the genome in which it resides (such as DNA polymerase or DNA gyrase).
- Transposition that involves an obligatory intermediate of RNA is primarily confined to eukaryotes and use some form of reverse transcriptase

TRANSPOSONS



Chhatrapati Shahu Ji Maharaj University, Kanpur
MOLECULAR BIOLOGY
Dr ANNIKA SINGH
DEPARTMENT OF
BIOTECHNOLOGY

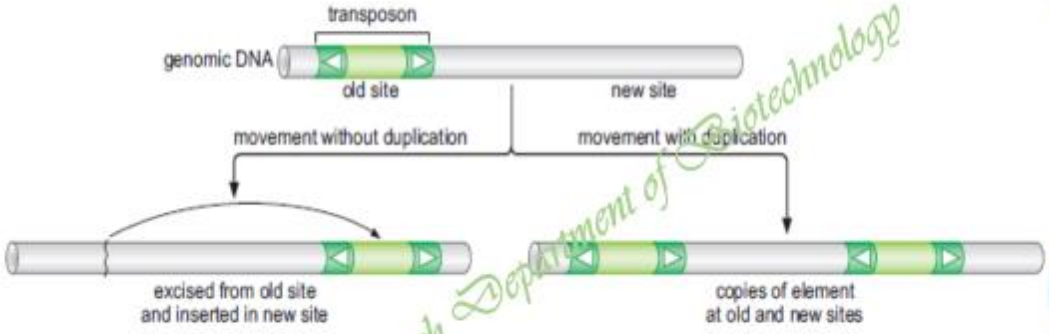


FIGURE 12-16 Transposition of a mobile genetic element to a new site in the host DNA. Recombination, in some cases, involves excision of the transposon from the old DNA location (left). In other cases, one copy of the transposon stays at the old location, and another copy is inserted into the new DNA site (right).

TRANSPOSONS

- Elements that involves an obligatory intermediate of RNA, are called long terminal repeat (LTR) retrotransposons or simply **retrotransposons**.
- Members of a second class of elements that also use reverse transcriptase but lack LTRs, and that employ a distinct mode of transposition, are referred to as *non-LTR retrotransposons*, or simply **retroposons**.
- Transposition can be classified as either **"AUTONOMOUS"** or **"NON-AUTONOMOUS"** in both Class I and Class II TEs.
- Autonomous TEs can move by themselves, whereas non-autonomous TEs require the presence of another TE to move.
- This is often because dependent TEs lack transposase or reverse transcriptase.
- Activator element (*Ac*) is an example of an autonomous TE, and dissociation elements (*Ds*) is an example of a non-autonomous TE. Without *Ac*, *Ds* is not able to transpose.

Chhatrapati Shahu Ji Maharaj University, Kanpur
MOLECULAR BIOLOGY
Dr ANNIKA SINGH
DEPARTMENT OF
BIOTECHNOLOGY

TRANSPOSONS

Transposable Elements Of All Kinds Can Promote Rearrangements Of The Genome Directly Or Indirectly:

- The transposition event itself may cause deletions or inversions or lead to the movement of a host sequence to a new location.
- Transposons 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 aberrant reciprocal recombination.
- Such exchanges result in deletions, insertions, inversions, or translocations.

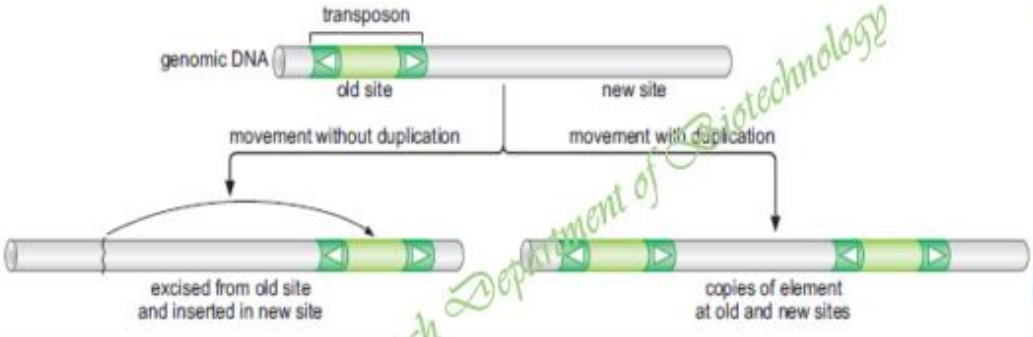


FIGURE 12-16 **Transposition of a mobile genetic element to a new site in the host DNA.** Recombination, in some cases, involves excision of the transposon from the old DNA location (left). In other cases, one copy of the transposon stays at the old location, and another copy is inserted into the new DNA site (right).

Chhatrapati Shahu Ji Maharaj University, Kanpur
MOLECULAR BIOLOGY
Dr ANNIKA SINGH
DEPARTMENT OF
BIOTECHNOLOGY

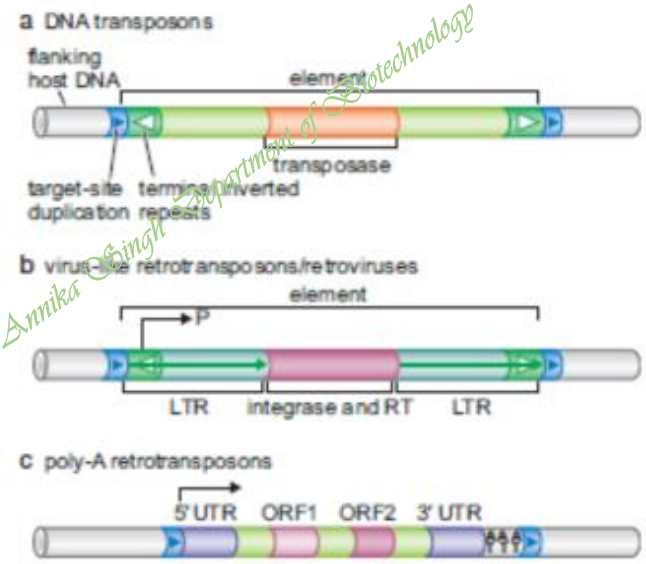
Transposons can be divided into the following three families on the basis of their overall organization and mechanism of transposition:

1. DNA transposons.
2. Retrotransposons

Virus-like retrotransposons. This class includes the retroviruses. These elements are also called long terminal repeat (LTR) retrotransposons.

Poly-A retrotransposons. These elements are also called non-viral retrotransposons.

- DNAtransposons remain as DNA throughout a cycle of recombination. They move using mechanisms that involve the cleavage and rejoining of DNA strands, and in this way, they are similar to elements that move by conservative site-specific recombination.
- Both types of retrotransposons move to a new DNA location using a transient RNA intermediate.



Chhatrapati Shahu Ji Maharaj University, Kanpur
MOLECULAR BIOLOGY
Dr ANNIKA SINGH
DEPARTMENT OF
BIOTECHNOLOGY

DNA TRANSPOSONS

- DNA Transposons Carry a Transposase Gene, Flanked by Recombination Sites
- DNA transposons carry both DNA sequences that function as recombination sites and genes encoding proteins that participate in recombination.
- The recombination sites are at the two ends of the element and are organized as inverted-repeat sequences.
- These terminal inverted repeats vary in length from 25 bp to a few hundred base pairs, are not exact sequence repeats, and carry the recombinase recognition sequences.
- The recombinases responsible for transposition are usually called transposases (integrases).
- They may carry a few additional genes, sometimes encoding proteins that regulate transposition or provide a function useful to the element or its host cell.
- For example, many bacterial DNA transposons carry genes encoding proteins that promote resistance to one or more antibiotic(s).
- The presence of the transposon therefore causes the host cell to be resistant to that antibiotic.

Dr. Annika Singh Department of Biotechnology

VIRUS-LIKE RETROTRANSPOSONS

- Virus-like retrotransposons and retroviruses also carry inverted terminal
- repeat sequences that are the sites of recombinase binding and action
- The terminal inverted repeats are embedded within longer repeated sequences; these sequences are organized on the two ends of the element as direct repeats and are called long terminal repeats (LTRs).
- Virus-like retrotransposons encode two proteins needed for their mobility: integrase (the
- transposase) and reverse transcriptase.
- Reverse transcriptase is needed for transposition because an RNA intermediate is required for the transposition reaction.
- Because these elements convert RNA into DNA, they are known as "retro" elements.
- The distinction between virus-like retrotransposons and retroviruses is that the genome of a retrovirus is packaged into a viral particle, escapes its host cell, and infects a newcell. In contrast, the retrotransposons can move only to new DNA sites within a cell but can never leave that cell.
- Like the DNA transposons, these elements are flanked by short target-site duplications that are generated during recombination.

Chhatrapati Shahu Ji Maharaj University, Kanpur
MOLECULAR BIOLOGY
Dr ANNIKA SINGH
DEPARTMENT OF
BIOTECHNOLOGY

POLY-A RETROTRANSPOSONS

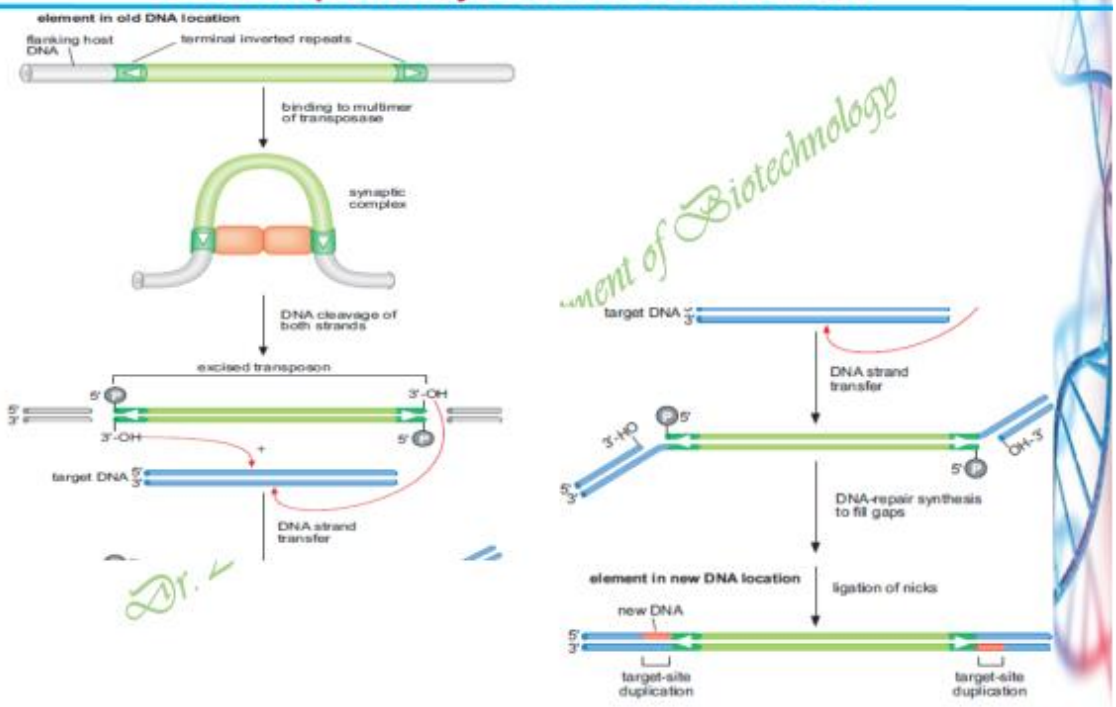
- The poly-A retrotransposons do not have the terminal inverted repeats present in the other transposon classes.
- Instead, the two ends of the element have distinct sequences
- One end is called the 50-UTR, whereas the other end has a region called the 30-UTR followed by a stretch of A:T base pairs called the poly-A sequence.
- These elements are also flanked by short target-site duplications.
- Retrotransposons carry two genes, known as ORF1 and ORF2.
- ORF1 encodes an RNA-binding protein.
- ORF2 encodes a protein with both reverse transcriptase activity and an endonuclease activity.
- This protein, although distinct from the transposases and integrases encoded by the other classes of elements, has essential roles during recombination.
- poly-A retrotransposons exist commonly in both autonomous and nonautonomous forms. truncated elements that do not have a complete 50-UTR sequence and have lost their
- ability to transpose.

DNA Transposition by a Cut-and-Paste Mechanism

- DNA transposons, virus-like retrotransposons, and retroviruses all use a similar mechanism of recombination to insert their DNA into a new site.
- In cut-and-paste transposition mechanism the movement of a DNA transposon involves the excision of the transposon from its initial location in the host DNA, followed by integration of this excised transposon into a new DNA site.
- To initiate recombination, the transposase binds to the terminal inverted repeats at the end of the transposon.
- Once the transposase recognizes these sequences, it brings the two ends of the transposon DNA together to generate a stable protein–DNA complex.
- This complex is called the synaptic complex or transpososome.
- It contains a multimer of transposase—usually two or four subunits—and the two DNA ends.
- It also protects the DNA ends from cellular enzymes during recombination.
- The next step is the excision of the transposon DNA from its original location in the genome.

Chhatrapati Shahu Ji Maharaj University, Kanpur
MOLECULAR BIOLOGY
Dr ANNIKA SINGH
DEPARTMENT OF
BIOTECHNOLOGY

DNA Transposition by a Cut-and-Paste Mechanism

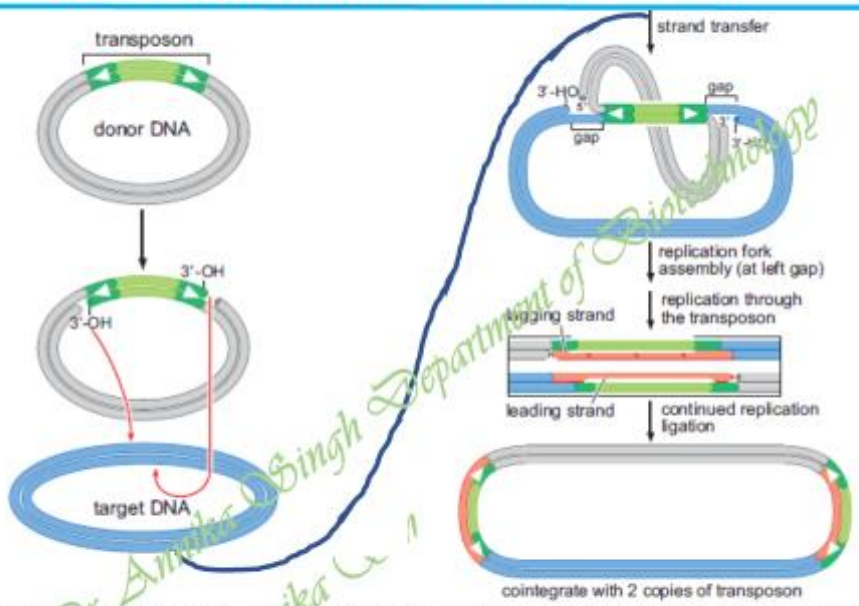


- To achieve this, the transposase subunits within the transpososome first cleave one DNA strand at each end of the transposon, exactly at the junction between the transposon DNA and the host sequence in which it is inserted (a region called the flanking host DNA).
- The transposase cleaves the DNA such that the transposon sequence terminates with free 3'-OH groups at each end of the element's DNA.
- Different transposons use different mechanisms to cleave the "second" DNA strands (those strands that terminate with 5' ends at the transposon host DNA junction.
- After excision of the transposon, the 3'-OH ends of the transposon DNA— the ends first liberated by the transposase—attack the DNA phosphodiester bonds at the site of the new insertion. This DNA segment is called the target DNA.
- As a result of this attack, the transposon DNA is covalently joined to the DNA at the target site.
- During each DNA-joining reaction, a nick is also introduced into the target DNA
- This DNA-joining reaction occurs by a one-step transesterification reaction that is called DNA strand transfer.

Chhatrapati Shahu Ji Maharaj University, Kanpur
MOLECULAR BIOLOGY
Dr ANNIKA SINGH
DEPARTMENT OF
BIOTECHNOLOGY

- The transpososome ensures that the two ends of the transposon DNA attack the two DNA strands of the same target site together.
- The sites of attack on the two strands are usually separated by a few nucleotides (e.g., 2-, 5-, and 9-nucleotide spacings are common). This distance is fixed for each type of transposon and gives rise to the short target-site duplications that flank
- transposed copies of the element (as explained in the next section). Once DNA strand transfer is complete, the job of the transpososome is also complete.
- The remaining recombination steps are performed by cellular DNA repair protein.

DNA Transposition by a Replicative Mechanism



Mechanism for replicative transposition. The transpososome introduces a single-strand nick at each of the ends of the transposon DNA. This cleavage generates a 30-OH group at each end. These OH groups then attack the target DNA and become joined to the target by DNA strand transfer. Note that at each end of the transposon, only one strand is transferred into the target at this point, resulting in the formation of a doubly branched DNA structure. The replication apparatus assembles at one of these "forks" (the left one in the figure). Replication continues through the transposon sequence. The resulting product, called a cointegrate, has the two starting circular DNA molecules joined by two copies of the transposon. The single stranded DNA gaps in the branched intermediate give rise to the target-site duplications.

Chhatrapati Shahu Ji Maharaj University, Kanpur
MOLECULAR BIOLOGY
Dr ANNIKA SINGH
DEPARTMENT OF
BIOTECHNOLOGY

Mechanism for replicative transposition.

The transpososome introduces a single-strand nick at each of the ends of the transposon DNA.

This cleavage generates a 30-OH group at each end.

These OH groups then attack the target DNA and become joined to the target by DNA strand transfer.

at each end of the transposon, only one strand is transferred into the target at this point, resulting in the formation of a doubly branched DNA structure.

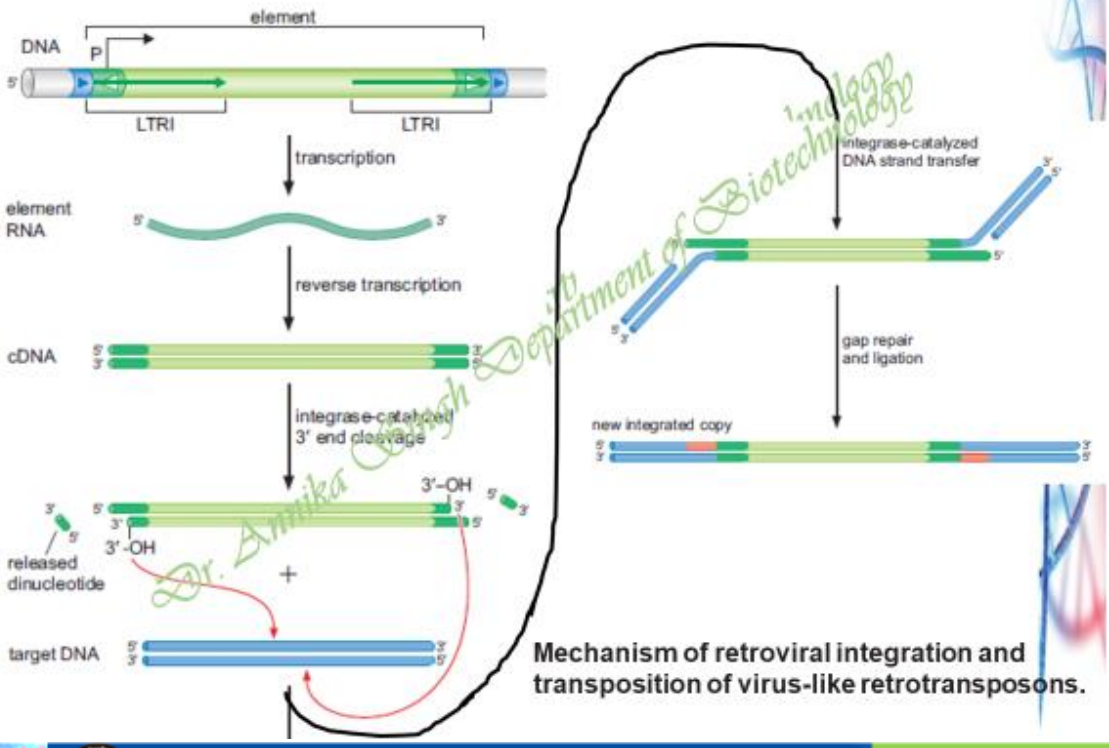
The replication apparatus assembles at one of these "forks"

Replication continues through the transposon sequence.

The resulting product, called a cointegrate, has the two starting circular DNA molecules joined by two copies of the transposon.

The singlestranded DNA gaps in the branched intermediate give rise to the target-site duplications.

Mechanism of transposition of virus-like retrotransposons

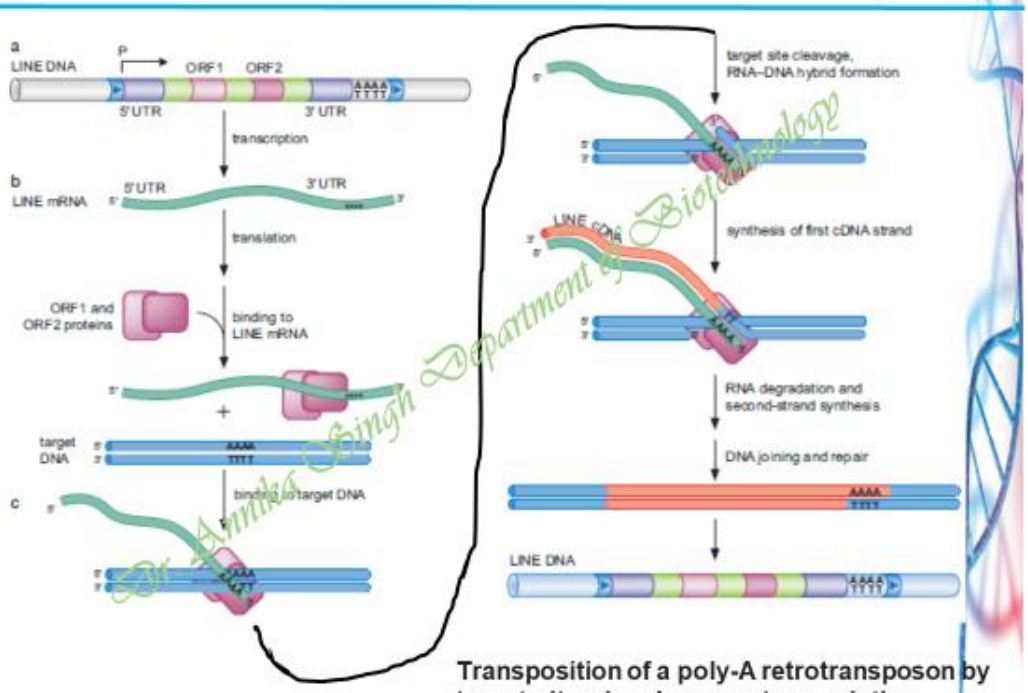


Mechanism of retroviral integration and transposition of virus-like retrotransposons.

Chhatrapati Shahu Ji Maharaj University, Kanpur
MOLECULAR BIOLOGY
Dr ANNIKA SINGH
DEPARTMENT OF
BIOTECHNOLOGY

Mechanism of retroviral integration and transposition of virus-like retrotransposons

- Virus-like retrotransposons and retroviruses insert into new sites in the
- genome of the host cell, using the same steps of DNA cleavage and DNA
- strand transfer as in the DNA transposons.
- recombination for these retroelements involves an RNA intermediate.
- transposition starts with transcription of the retrotransposon DNA sequence into RNA by a cellular RNA polymerase.
- Transcription initiates at a promoter sequence within one of the LTRs and continues across the element to generate a nearly full-length RNA copy of the element's DNA .
- The RNA is then reverse-transcribed to generate a double-stranded DNA molecule, cDNA
- the cDNA is then recognized by the integrase protein for recombination with a new target DNA site.
- Integrase assembles on the ends of this cDNA and then cleaves a few nucleotides off the 3′ end of each strand.
- This cleavage reaction is identical to the DNA cleavage step of DNA transposition.
- Integrase then catalyzes the insertion of these cleaved 3′ ends into a DNA target site in the host-cell genome using the DNA strand transfer reaction.
- Host-cell DNA-repair proteins fill the gaps at the target site generated during DNA strand transfer to complete recombination.
- This gap-repair reaction generates the target-site duplications.



Transposition of a poly-A retrotransposon by target-siteprimed reverse transcription.

Chhatrapati Shahu Ji Maharaj University, Kanpur
MOLECULAR BIOLOGY
Dr ANNIKA SINGH
DEPARTMENT OF
BIOTECHNOLOGY

Poly-A Retrotransposons Transposition by a "Reverse Splicing" Mechanism

- The poly-A retrotransposons (e.g., human LINE elements) move using an RNA intermediate
- This mechanism is called target-site-primed reverse transcription
- The first step is transcription of the DNA of an integrated element by a cellular RNA polymerase.
- This newly synthesized RNA is exported to the cytoplasm and translated to generate the ORF1 and ORF2 proteins.
- These proteins remain associated with the RNA that encoded them
- The protein–RNA complex then re-enters the nucleus and associates with the cellular DNA
- The ORF2 protein has both a DNA endonuclease activity and a reverse transcriptase activity.
- The endonuclease initiates the integration reaction by introducing a nick in the chromosomal DNA.T-rich sequences are preferred cleavage sites.
- The presence of these Ts at the cleavage site permits the DNA to base pair with the poly-A tail sequence of the element RNA.
- The 3' -OH DNA end generated by the nicking reaction then serves as the primer for reverse transcription of the element RNA .
- The ORF2 protein also catalyzes this DNA synthesis.
- The remaining steps of transposition, include synthesis of the second cDNA strand, repair of DNA gaps at the insertion site, and ligation to seal the DNA strands.

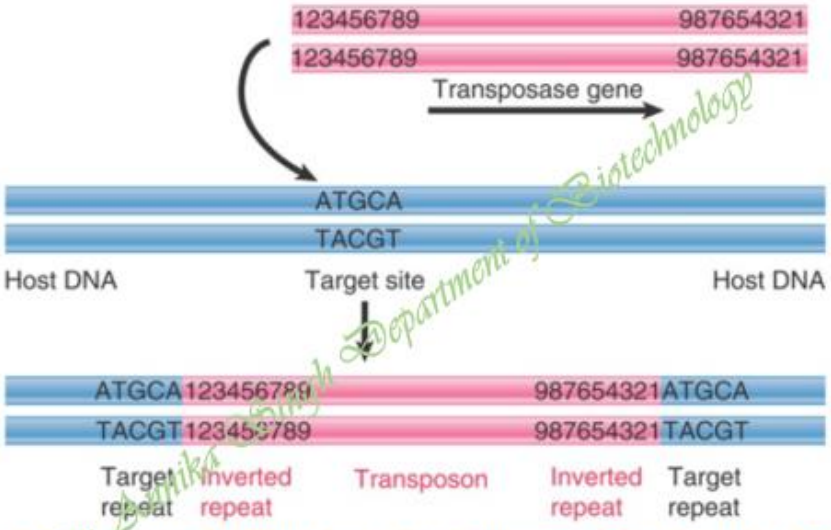
Type	Structural Features	Mechanism of Movement	Examples
DNA-mediated transposition			
Bacterial replicative transposons	Terminal inverted repeats that flank antibiotic-resistance and transposase genes	Copying of element DNA accompanying each round of insertion into a new target site	Tn*3*, γδ, phage Mu
Bacterial cut-and-paste transposons	Terminal inverted repeats that flank antibiotic-resistance and transposase genes	Excision of DNA from old target site and insertion into new site	Tn*5*, Tn*10*, Tn*7*, IS*911*, Tn*917*
Eukaryotic transposons	Inverted repeats that flank coding region with introns	Excision of DNA from old target site and insertion into new site	P-elements (*Drosophila*), *hAT* family elements, Tc*1*/*Mariner* elements
RNA-mediated transposition			
Virus-like retrotransposons	~250- to 600-bp direct terminal repeats (LTRs) flanking genes for reverse transcriptase, integrase, and retrovirus-like Gag protein	Transcription into RNA from promoter in 5′ LTR by RNA polymerase II followed by reverse transcription and insertion at target site	Ty elements (yeast), *Copia* elements (*Drosophila*)
Poly-A retrotransposons	3′-A-T-rich sequence and 5′-UTR flank genes encoding an RNA-binding protein and reverse transcriptase	Transcription into RNA from internal promoter; target- primed reverse transcription initiated by endonuclease cleavage	F and G elements (*Drosophila*), LINE and SINE elements (mammals), *Alu* sequences (humans)

Chhatrapati Shahu Ji Maharaj University, Kanpur
MOLECULAR BIOLOGY
Dr ANNIKA SINGH
DEPARTMENT OF
BIOTECHNOLOGY

INSERTION SEQUENCE (IS)

- The simplest bacterial transposons are called **insertion sequence (IS)** elements
- Each type is given the prefix "IS," followed by a number that identifies the type.
- The IS elements are normal constituents of bacterial chromosomes and plasmids.
- A standard strain of *Escherichia coli* contains several (fewer than 10) copies of any one of the more common IS elements.
- To describe an insertion into a particular site, a double colon is used; thus λ::IS1 describes an IS1 element inserted into phage lambda.
- Most IS elements insert at a variety of sites within host DNA, show varying degrees of preference for particular hotspots.
- The IS elements are autonomous units, each of which encodes only the proteins needed to sponsor its own transposition.
- Each IS element is different in sequence, but share some common features in organization.

INSERTION SEQUENCE (IS)



IS elements have inverted terminal repeats and generate direct repeats of flanking DNA at the target site. In this example, the target is a 5-bp sequence. The ends of the transposon consist of inverted repeats of 9 bp, where the numbers 1 through 9 indicate a sequence of base pairs.

Chhatrapati Shahu Ji Maharaj University, Kanpur
MOLECULAR BIOLOGY
Dr ANNIKA SINGH
DEPARTMENT OF BIOTECHNOLOGY

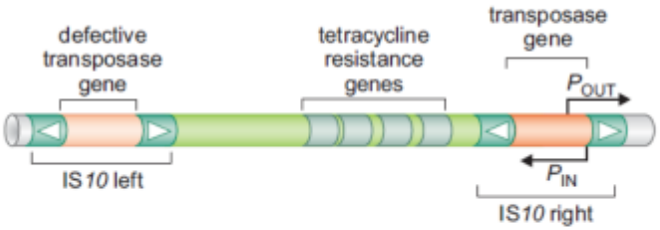
INSERTION SEQUENCE (IS)

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

- An IS element ends in short **inverted terminal repeats**; usually the two copies of the repeat are closely related rather than identical.
- When an IS element transposes, a sequence of host DNA at the site of insertion is duplicated.
- The nature of the duplication is revealed by comparing the sequence of the target site before and after an insertion has occurred.
- At the site of insertion the IS DNA is always flanked by very short **direct repeats**.

Tn10

- The bacterial transposon Tn10 is a well-characterized representative of the IS4 family, which also includes Tn5.
- Tn10 is a compact element of 9 kb and encodes a gene for its own transposase and genes imparting resistance to the antibiotic tetracycline.
- Tn10 transposes via the cut-and-paste mechanism, using the DNA hairpin strategy to cleave the non transferred strands
- Tn10 is organized into three functional modules. This organization is relatively common, and elements that have it are called composite transposons.



Chhatrapati Shahu Ji Maharaj University, Kanpur
MOLECULAR BIOLOGY
Dr ANNIKA SINGH
DEPARTMENT OF
BIOTECHNOLOGY

Maize *Ac*/*Ds* Family

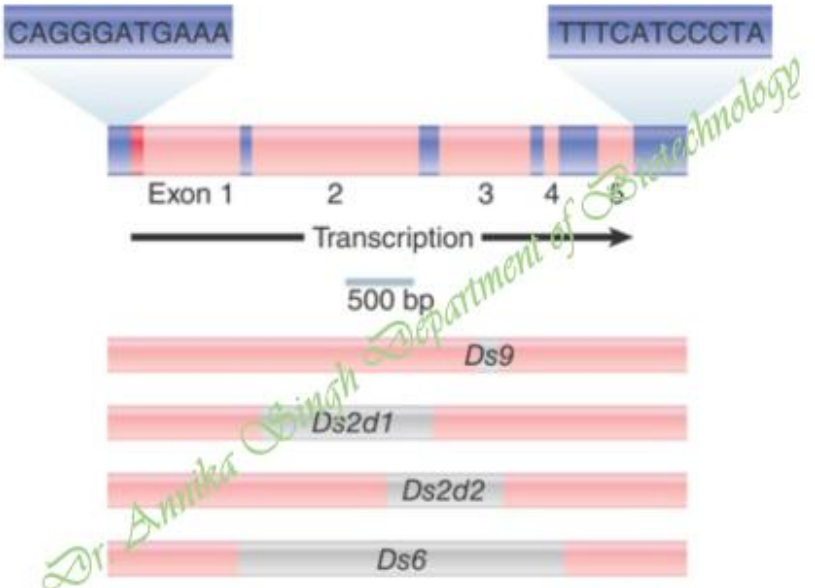
- **The *Ac*/*Ds* family**, was first discovered by Barbara McClintock in the 1940s (she received the Nobel Prize in 1983).
- Transposition of *Ac*/*Ds* occurs by a non replicative cut-and-paste mechanism

***Ac* (Activator) element**

- ***Ac* (Activator) element are** autonomous in nature
- Element is occupied by a single gene consisting of five exons codes for transposase.
- The element itself ends in inverted repeats of 11 bp, and a target sequence of 8 bp is duplicated at the site of insertion.

***Ds* (Dissociator) elements**

- Ds are nonautonomous elements and related to *Ac*.
- derived from autonomous elements by deletions (or other changes) that inactivate the *trans* acting transposase but leave the sites (including the termini) on which the transposase acts intact.
- They end in the same 11-bp inverted repeats.
- They are shorter than *Ac*, and the length of deletion varies.
- The element ***Ds9*** has a deletion of only 194 bp.
- The ***Ds6* element** retains a length of only 2 kb, representing 1 kb from each end of *Ac*.

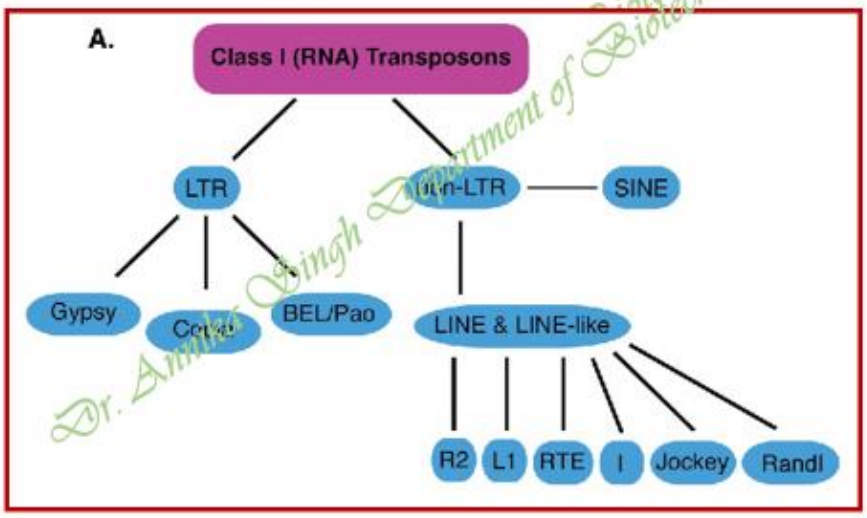


The *Ac* element has five exons (pink) that encode a transposase; *Ds* elements have internal deletions (gray).

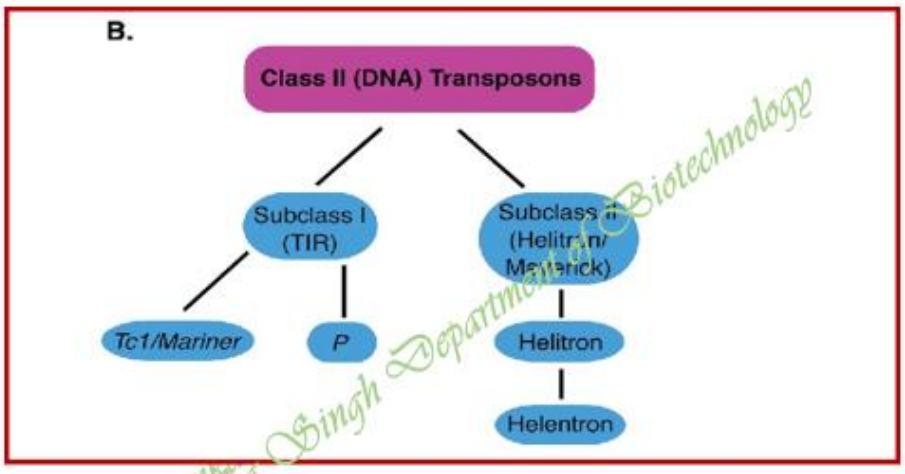
Chhatrapati Shahu Ji Maharaj University, Kanpur
MOLECULAR BIOLOGY
Dr ANNIKA SINGH
DEPARTMENT OF
BIOTECHNOLOGY

Drosophila melanogaster Transposons

- 20% of the *Drosophila melanogaster* genome consists of TEs, at least 30% of these elements are full length and believed to be active.
- *D. melanogaster* is a promising model organism for the study of eukaryotic TEs.



Drosophila melanogaster Transposons



No active DNA transposons have been identified in humans due to lack of functional transposases. 16% of the DNA transposons in *D. melanogaster* are full length and potentially active, including *1360*, *hobo*, *Bari1*, *pogo*, and *P* elements.

Chhatrapati Shahu Ji Maharaj University, Kanpur
MOLECULAR BIOLOGY
Dr ANNIKA SINGH
DEPARTMENT OF
BIOTECHNOLOGY

Drosophila melanogaster Transposons

- DNA transposons are often less than 5 kb in length and typically encode a single transposase gene
- DNA transposons are divided into 2 sub-classes based on their transposition mechanisms.
- **Sub-class I elements** utilize the canonical cut-and-paste mechanism of TIR transposon transposition, and are divided into several superfamilies: Tc1/*mariner*, PIF/*Harbinger*, hAT, Mutator, Merlin, Transib, P, *piggyBac*, and CACTA.
- **Sub-class II DNA transposons include *Helitron*** and *Maverick* elements
- *Helitrons*, mobilize by a different mechanism than TIR transposons, using rolling-circle replication with a single stranded DNA intermediate.
- *D. Melanogaster* has numerous active DNA transposons with full length TIRs and functional transposase genes.

TC1/*MARINER* SUPERFAMILY

- ***Bari* elements of the Tc1/*mariner* superfamily**.
- *D. melanogaster Bari1* elements are autonomous DNA transposons with short TIRs, usually less than 40 nucleotides in length
- non-autonomous *Bari1* elements with long TIRs have been identified in other *Drosophila* species.
- Transposition of *Bari1* elements and other TEs in the Tc1/*mariner* superfamily is initiated by interactions between one or more direct repeats in the TIRs and the element encoded transposase
- Dimerization of TIR bound transposases induces cleavage of the element from surrounding sequences
- Like *Tc1* elements, *Bari1* elements then target TA sites and integration results in the duplication of these nucleotides at both ends
- Target site duplications are characteristic of TIR transposon insertions and may be used to identify transposition events and distinguish between different families of TIR transposons.

Chhatrapati Shahu Ji Maharaj University, Kanpur
MOLECULAR BIOLOGY
Dr ANNIKA SINGH
DEPARTMENT OF
BIOTECHNOLOGY

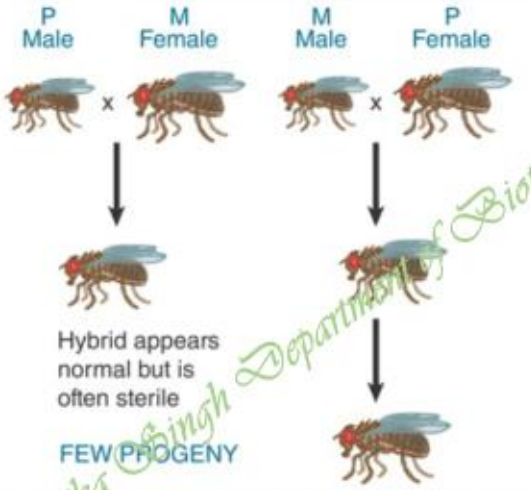
P elements

- *P* elements are the best-studied DNA transposons in the *D. melanogaster* genome.
- Full-length autonomous *P* elements are 2.9 kb in length with 31 bp TIRs and 4 exons that encode a transposase when spliced.
- Like other TIR transposons, *P* elements utilize a cut-and-paste mechanism of transposition and create target site duplications upon insertion
- *P* elements are unique, however, in their abilities to amplify themselves in *Drosophila* germline cells due to preferential insertion at regions of the genome that bind the origin recognition complex and function as replication origins.
- By transposing during S phase from replicated genomic regions to un-replicated regions, *P* elements are copied, amplifying their presence in the genome with the assistance of the host DNA repair machinery.
- These elements belong to the same class as *pogo* and *hobo* elements, and play a significant role in hybrid dysgenesis syndrome, a phenomenon observed in the progeny of hybrid crosses of certain *Drosophila* strains.
-

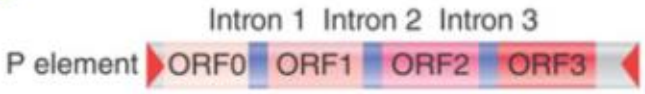
Hybrid Dysgenesis

- *Drosophila* strains are defined as either P type or M type, depending on whether hybrid dysgenesis results from crosses with the paternal or maternal parent.
- The phenomena observed in this syndrome include high rates of mutation, recombination, and sterility in the F_1 hybrids of only P type male crosses with M type females
- When flies from two of these strains are crossed, the progeny display "dysgenic traits"—a series of defects including mutations, chromosomal aberrations, distorted segregation at meiosis, and reduced fertility.
- The appearance of these correlated defects is called **hybrid dysgenesis**.
- Two systems responsible for hybrid dysgenesis have been identified in *D. melanogaster*.
- In the first, flies are divided into the **types I (inducer) and R (reactive)**.
- **Reduced fertility** is seen in crosses of I males with R females, but not in the reverse direction.
- In the second system, flies are divided into the two types, P (paternal contributing) and M (maternal contributing).
- A cross between a P male and an M female causes dysgenesis, but the reverse cross does not.

Chhatrapati Shahu Ji Maharaj University, Kanpur
MOLECULAR BIOLOGY
Dr ANNIKA SINGH
DEPARTMENT OF
BIOTECHNOLOGY



Hybrid dysgenesis is asymmetrical; it is induced by P male × M female crosses, but not by M male × P female crosses.



Helitrons

- Helitrons belong to a unique subclass of DNA transposons with a distinct mechanism of transposition.
- Unlike other DNA transposons, Helitrons lack TIRs and encode a DNA helicase and replicator initiator (Rep) protein with nuclease and ligase functions, resembling the machinery of rolling-circle replicons
- A subclass of Helitrons, called Helentrons, encode an additional apurinic-apyrimidinic endonuclease and may also mobilize non-autonomous Helentron-associated interspersed elements (HINEs).
- Helitrons are abundant in plant genomes and have been identified in many other eukaryotic genomes, including D. melanogaster in which 1% of the genome consists of non-autonomous Helitrons.
- Drosophila interspersed nuclear element- 1 (DINE-1), the most abundant TE in the *D. melanogaster* genome, is a non-autonomous Helentron, distinct from HINEs due to the presence of unique structural features such as inverted repeats.
- Helitrons utilize a rolling-circle replication mechanism of transposition, which has recently been validated by experiments conducted with the Helraiser

Chhatrapati Shahu Ji Maharaj University, Kanpur
MOLECULAR BIOLOGY
Dr ANNIKA SINGH
DEPARTMENT OF
BIOTECHNOLOGY

Retrotransposons Copia

- Retrotransposons, are the most abundant class of TEs in the D. melanogaster genome.
- Retrotransposons include LTR retrotransposons, non-LTR retrotransposons (LINEs and LINE-like elements), short interspersed nuclear elements (SINEs), and other similar TEs.
- LTR and non-LTR retrotransposons use similar mechanisms of transposition and regulation.
- Many of these retrotransposons are classified as endogenous retroviruses, or errantiviruses in Drosophila, as they either arose from retroviruses that lost infectivity or LTR retrotransposons that acquired env genes from exogenous sources
- The retrotransposon pol gene encodes a polyprotein, typically consisting of a protease, an integrase, and a reverse transcriptase (RT) with an RNase H domain and DNA polymerase activity

Non-LTR retrotransposons

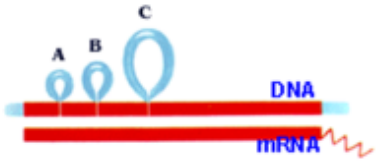
- Non-LTR retrotransposons, or LINE-like elements, have been classified into over 100 families, separated into 28 clades and 6 groups: R2, L1, RTE, I, Jockey and Rand I.
- Non-LTR retrotransposons are structurally similar to LTR retrotransposons, but often lack some of the ORFs and protein domains encoded by LTR retrotransposons and do not contain LTRs at their 3' and 5' ends.
- The absence of LTRs suggests that these elements interact with their encoded proteins differently than LTR retrotransposons and may utilize different mechanisms of transposition.
- While non-LTR and LTR retrotransposons encode similar proteins and often generate target site duplications, the reverse transcription and integration events of non-LTR retrotransposon mobilization are unique, at least for the R2 group of elements that often lack promoters and only encode a single ORF with RT and endonuclease activities

Chhatrapati Shahu Ji Maharaj University, Kanpur
MOLECULAR BIOLOGY
Dr ANNIKA SINGH
DEPARTMENT OF
BIOTECHNOLOGY

SPLIT OR INTERRUPTED GENE

A split or interrupted gene is defined as a gene consisting of introns and exons. Removal (splicing) of the intron(s) from a primary transcript (pre-mRNA) is essential for creating a mRNA.

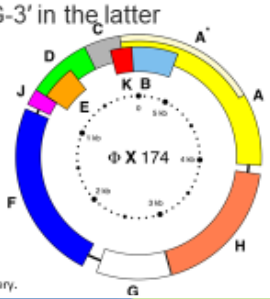
- Richard J. Roberts and Phillip A. Sharp (1993 Nobel Prize) discovered the existence of split genes in the adenovirus, when they examined a hybridized nucleic acid molecule made between a adenoviral mRNA and its template DNA in the electron microscope.
- They observed that the mRNA was much shorter in length and thus was not encoded as an equal colinear segment of the DNA molecule
- Instead, large loops of unhybridized DNA (A, B and C in figure) were seen.
- Their interpretation was that the mature messenger RNA was derived from *four discountinuous segments* on the viral DNA.
- The segments retained in the mRNA they called *exons* and the intervening sequences (A, B and C), which are excised during mRNA processing and maturation, are called *introns*.



Overlapping Genes

Overlapping Genes

- A DNA fragment (segment) can code for more than one gene product by using different reading frames or different initiation codons.
- This phenomenon of overlapping genes is widespread in DNA and RNA viruses, as well as in organelles, and bacteria, also known in nuclear eukaryotic genomes (Makalowska *et al.*, 2005).
- Overlapping genes can also arise by the use of the complementary strand of a gene; for example, genes specifying $tRNA^{ILE}$ and $tRNA^{GLN}$ in the human mitochondrial genome are located on different strands and there is a three-nucleotide overlap between these that reads 5′-CTA-3′ in the former and 5′-TAG-3′ in the latter (Makalowska *et al.*, 2005).



*This picture is a work by Emmanuel Douzery.

Chhatrapati Shahu Ji Maharaj University, Kanpur
MOLECULAR BIOLOGY
Dr ANNIKA SINGH
DEPARTMENT OF
BIOTECHNOLOGY

Cryptic genes

- Cryptic genes are phenotypically silent DNA sequences, not normally expressed during the life cycle of an individual.
- They may, however, be activated in a few individuals of a large population by mutation, recombination, insertion elements, or other genetic mechanisms.
- They differ from other silent DNA sequences, collectively referred to as **selfish DNA**, which spread in the genome by making additional copies of themselves and survive at the expense of the host, in most cases without conferring any phenotypic advantage to the organism.
- Cryptic genes also differ from **pseudogenes** which are homologous copies of an active gene but have a large number of accumulated mutations so that their reversion into an active form is not usually possible.
- Example: A cryptic gene for alcohol dehydrogenase (Adh IV) has been identified in yeast
- Genes for β-glucoside utilization are cryptic in E. coli and Salmonella

*MITALI MUKERJI and S. MAHADEVAN | Genet., Vol. 76, Nmnber 2, August 1997, pp. 147-[illegible]

Dr. Annika Singh Department of Biotechnology

References:

1. **Molecular Biology Of The Gene seventh edition James D. Watson**
2. **Lewin's Genes X - Benjamin Lewin, Jocelyn Krebs, Stephen**