Mobile DNA Elements

Daniel F. Simola

22 May, 2005

1 Element Classification

Transposable DNA elements are found in the genomes of all organisms, and have forged a mutual relationship with genomic DNA: elements need genes to survive generation to generation, but these elements also influence the evolution of genes and their function. Retrotransposable elements are thought to have arisen from retroviruses, and so there is a gradient of elements, from fully autonomous and active to deprecated and inactive: retrovirus → LTR retrotransposons → non-LTR autonomous → non-LTR nonautonomous. Many elements have terminal repeats (direct or inverted), which hybridize with each other, allowing a sequence loop to be created within the genomic DNA, which is then excised using a transposase enzyme or incised via a element repeat to genomic repeat hybridization.

There are two kinds of recombinase enzymes which are required for transposition in general

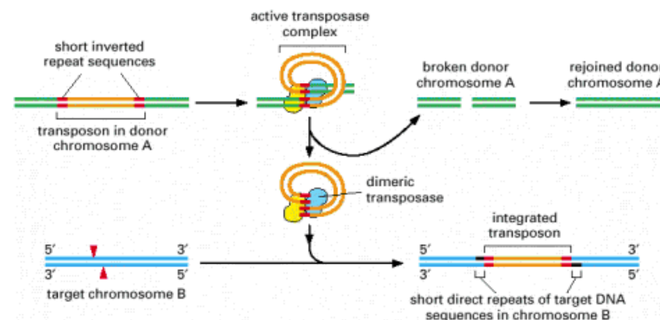
Conservative site-specific: No DNA degradation or re-synthesis occurs, although formation of a small heteroduplex is required - thus every nucleotide bond is preserved. Bacteriophage λ integrase, Cre (loxP), and Flp (FRT) recombinases

Transpositional site-specific: bacteriophage Mu, fly P-elements, Tc1/mariner (transposase) and LTR retrotransposon and retrovirus integrases. All member of this class share a hand-like protein structure. These guys have only modest target site selectivity, and no homology is required between the transposon termini and target insertion site.

1.1 Here is the breakdown of transposable element types

1.1.1 DNA-only transposons

Non-replicative transposons (Tc1-mariner, IS) migrate using a "cut and paste" mechanism with low insertion specificity, so they can reintegrate in any type of sequence (nearly random insertion), but are limited to local hopping from excision site. Active transposons encode a transposase enzyme between the inverted repeat termini, and the enzyme binds these termini and the target DNA to make the cut and paste. Insertions result in residual target site duplications (TSP), which can be detected as historical transposition events. The short inverted repeats at their termini, due to the loop structure and heteroduplex that is formed for excision. Since excision is a result of a double strand break, a rejoining event must occur, most of which are homologous, but nonhomologous events may occur.



Replicative DNA transposons operate in a similar fashion to their "cut-and-paste" friends, but go a step beyond and actually recruit a DNA polymerase and other enzymes to copy the original element and insert into another area of DNA. This method does not require a double strand break of the donor site.

1.1.2 Retrotransposons

These elements are copied into the DNA via a single strand DNA break, reverse transcription on the loose strand, and insertion of the cDNA element into a new location.

Replicative: The insertion mechanism transcribes an already integrated copy of DNA into RNA, which is reverse transcribed into cDNA and reintegrated into the genome. Thus this is replicative transposition.

- **LTR Retrotransposons:** These contain long terminal repeats and are very similar to retroviruses, both encoding functional gag (viral coat protein), pol (integrase), a reverse transcriptase, and RNAse H; retroviruses however contain a functional env (envelope) which allows a virus to transmigrate to other cells. Reverse transcription occurs in viral-like particles in the cytoplasm.

  These guys have specific target insertion sites, notably at or near promoters of active genes or within coding regions.

- **non-LTR Retrotransposons:** Two types which lack LTRs instead bearing 3' polyA sequence. Sequence is often truncated 5', and have variable length TSDs.
  - **Autonomous:** (LINE, L1) These are long (4-6kb) and have 2 ORFs, one encoding a nucleic acid binding protein, the other an endonuclease and reverse transcriptase; thus 'autonomous'. Insertion site specificity is *cis* and due to cut size of endonuclease. Their process of reverse transcription is weird and unclear, and occurs "in place" in genomic DNA, using a method "target primed reverse transcription".
  - **Non-autonomous:** (Alu) These guys are a type of SINE (short interspersed nucleotide repeat) and lack ORFs, thus requiring a full complement of host proteins, but are very similar to autonomous non-LTRs, except their ORFs have become inactivated through mutation and truncation events. SINEs are mobilized in *trans* by LINEs. Alus make up 10% of the human genome.

- **Distinct retroelements:** One type is similar to LTR retrotransposons, but lack the protease and integrase, and either LTR or polyA regions. The other type has an intron in the reverse transcriptase ORF. Processed pseudogenes are another type of element, which are thought to have been created by active L1 elements which copy live mRNA, but lack a promoter and may become truncated upon genomic insertion.

2 Enzyme types

transposases excise a target DNA sequence from larger sequence

recombinases swap target and acceptor DNA via common sequence motif (eg Cre operates on loxP sites)

integrases splice a target DNA into acceptor DNA

nucleases nick DNA - either single or double strand nicks

3 Influence of retrotransposons in genome evolution

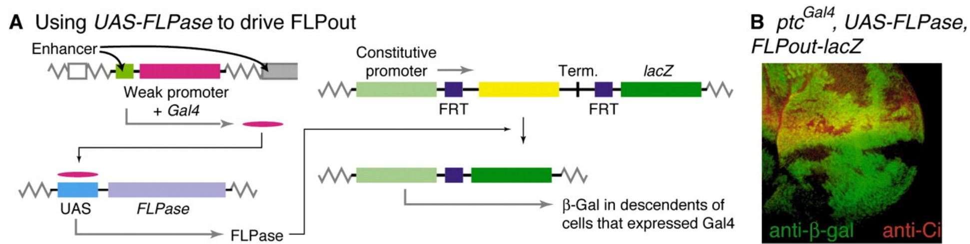
L1 elements have had strong destructive influence in the structure of mammalian genomes through insertion and rearrangement due to homologous recombination. Humans have about 100 active L1s, versus the 3000 in mice, and so murine L1s are responsible for a lot of mutation. L1s mobilize the non-autonomous retroelements as well. L1s can produce large genomic deletions due to homologous recombination. Alu pairs produce 400kb blocks of segmental duplication/deletion/inversion (often a segmental duplication is flanked by Alus), and are implicated in genomic diseases.

L1 elements are also constructive by repairing double strand breaks (endonuclease-independent insertion), effecting exon/promoter shuffling (L1 polyA cleavage signal is weak, so sequence downstream of L1 polyA also excised), generating active chimeric retrogenes (mostly small nuclear RNAs), constructing new exons by sacrificing their sequence, affecting gene expression (5' UTR contains antisense promoter), and introducing foreign but now critical enzymes (telomerase, immunoglobulin recombinase activating proteins).

4 Control of mobile elements

Genomes are able to prevent the mobility of elements using co-suppression via siRNA and methylation.

Humans take control of mobile elements all the time, such as P-elements in the fly and cre-lox for homologous knock outs in mouse. Below is an example of the modern use of transposable elements. Here a fly P-element is excised in a controlled fasion using the Gal4-USA inducible expression system. The P-element is flanked by FRT sequences, and FLPase is activated following expression of GAL4, whose coding sequence is inserted via an enhancer trap.



5 Reference

Most of this information was adapted from the review paper by Haig Kazazian and from *Molecular Biology of the Cell*. For more information about the specific classes of repeats and their phenotypic effects, check out `http://www.neuro.wustl.edu/neuromuscular/mother/dnarep.htm`