Transposable Elements

Chapter · January 2010

CITATIONS
95

READS
17,169

4 authors, including:

Uttar Kumar Tomar
Arid Forest Research Institute, India
28 PUBLICATIONS **421** CITATIONS

SEE PROFILE

CHAPTER 20

Transposable Elements

G. Tripathi[1*], B.M. Sharma[1], J. Shasmal[1] and U.K. Tomar[2]
[1]Depart. of Zoology, J.N.V. University, Jodhpur– 42 001
E-mail: drgst@rediffmail.com
[2]Arid Forest Research Institute, Jodhpur–342 005
E-mail: uktomar@afri.res.in

Deoxyribonuleic acid (DNA) was earlier believed to be a static entity. But now it is known that certain DNA segments can move from one location to another on the genome creating a genetic instability. They are known as transposable elements, transposons, translocatable elements, mobile sequence, movable sequence, insertion elements, jumping elements, parasitic elements or selfish DNA. In 1940s Barbara McClintock was first to identify transposable genetic elements, when DNA was not even considered to be the genetic material, while investigating the inheritance of pigmentation in the common corn plant, *Zea mays*. She reported that the variegation in pigments of kernel is due to action certain distinct genetic units which can move from one site to another on different chromosomes of maize. The placement of these genetic units into genes turns the expression of genes 'on' or 'off', regulating thereby their expression. She called them as 'controlling elements'. She was awarded the Noble prize in 1983 for her work. About 30 years after the discovery, a new class of mutation was observed in the genome of the common intestinal bacterium, *Escherichia coli*. It was found that the mutation was caused by insertion of certain DNA segments into the bacterial genome. These nucleotide sequences of DNA were called insertion sequences or IS elements. Initially they were recognized as small fragments of DNA inserted into genes (Jordan *et al.*, 1968; Shapiro, 1969), but later as performed movable elements (Malamy *et al.*, 1972).

Some genes are responsible for causing resistance to antibiotics in bacteria. These genes are able to move from one DNA molecule to another. In 1974, Hedges and Jacob showed that a DNA segment carrying antibiotic-resistance gene can be transposed from one DNA to another,

*Corresponding author

and called as transposable segment transposable element (Tn element) or transposons (Hedges and Jacob, 1974). After 1976 the research began to look into the mechanism of transposition of these mobile elements. It was soon found that during transposition the mobile elements were replicated, as a result of which one copy remained in the old position and the replicated one moved to the host genome.

Transposons encode the enzymes 'transposonases' or 'transposases', required for transposition of mobile genetic elements. Any alteration in the transposonase gene may abolish the transposition process. Another group of Tn element (e.g., Tn 3) express or encodes a repressor protein, which can bind to the transposonase and its own promoter region of the DNA. This repressor protein is considered as a site-specific recombination enzyme, acting as a part-time regulatory molecule.

There are two general groups of transposons. One that mobilizes via DNA and another has ability to make DNA copies of their RNA transcripts and these copies may integrate at new sites in the genome. The latter is related to retroviruses. Here we have mainly described the transposons that mobilize via DNA. The frequency of transcription varies from transposon to transposon. The overall rate of transcription is approximately 10^{-4} to 10^{-3} per element per generation.

INSERTION SEQUENCES

Insertion sequences (IS) are simple transposition modules. They are small elements (less than 2000 base pairs) encoding only transposition determinants. Like transposons, IS elements are capable of inserting themselves randomly into a DNA molecule and also have inverted terminal repeats of about 30 nucleotide base pairs (bp). It has been reported that the presence of an IS in a plasmid also acts as a transposon. Besides these, an IS has other functional similarities with the transposon. For instance, the incorporation of an IS into host DNA is associated with the replication of few bp in the recipient DNA. Integration of IS element with DNA molecule causes mutation which can be restored to normal after removing the insertion sequence. It also abolishes the function of the gene of an operon where it has been integrated because the IS element comprises nonsense and terminator sequences which cause rho- factor-dependent termination of transcription (Nevers and Saedlor, 1977). Insertion of an IS into a gene diminishes the activities of other genes downstream (5'-3') in the direction of transcription due to disturbances created in the continuity of nucleotide bp in the DNA.

Insertion sequences are categorized into non-homologous classes as IS1, IS2, IS3, IS4 and IS5. They appear to be normal components of chromosomes and plasmids of the gram-negative bacteria. For example, the genome of *Escherichia coli* possesses about eight copies of IS1, five copies of IS2 and three copies of IS3. The elements IS2 and IS3 have been reported to be present on various drug-resistance episomes. The smallest transposable elements (750-1500 bp) are known as 'transposition modules' or 'IS-like modules'. They encode only genetic determinants relevant to promoting and regulating their own transposition. They are IS1-IS5, IS10, IS102, etc. Insertions occur at a level comparable to spontaneous mutation rate, usually 10^{-5} to 10^{-7} per generation. Reversion (by excision of IS element) is infrequent with a range of rates of 10^{-6} to 10^{-10} per generation, which is approximately 10^{5} times less frequent than insertion (Lewin, 1994).

The composite transposons are elements, which encode not only transposition functions but other functions unrelated to transposition, e.g., antibiotic resistance. The drug-resistance

causing transposable elements (Tn elements) have 4,500 nucleotide base pairs (Tn1, Tn2, Tn3) to 20,000 base pairs (Tn4). These Tn elements (Tn1 to Tn4) carry ampicillin-resistance and contain the inverted repeat sequences of 140 bp, while Tn5 (kanamycin-resistance) and Tn10 (tetracycline-resistance) have longer inverted repeat sequences (eg. Tn5, 2450 bp; and Tn10, 1400 bp). However, Tn9 (chloramphenicol-resistance) comprises a direct repeat sequences of about 750 bp. The repeat sequences of Tn9 and Tn10 are structurally more or less similar to the IS3 and IS1 respectively. Certain composite transposons are flanked by two copies of IS either in direct or in inverted orientation (Fig. 1A). There are other complex transposons not flanked by IS elements (Fig. 1B). A transposons having long, inverted repeat sequences can be visualized by electron microscope. When DNA strands of a transposon are separated and allowed to self-anneal, the complementary bases at both the ends join with one another, forming double-stranded circle 'lollipop' (Fig. 2). These lollipops are recognizable in electron micrograph.

BASIC EVENTS DURING TRANSPOSITION

The ends of transposons contain two copies of IS elements having terminal inverted-repeats. It has been documented that the incorporation of any gene between two copies of IS induces the transfer of gene into other DNA molecule through non-homologous recombination process. The transposition involves three steps: (i) Duplication of mobile elements; (ii) duplication at insertion site; and (iii) DNA rearrangements.

Duplication of Mobile Elements

When a mobile genetic element translocates from one plasmid (donor site) to another plasmid (recipient site), the formation of an intermediate (or byproduct?) takes place. The intermediate is composed of one copy of both the plasmids (donor and recipient) linked together by two copies of the mobile element. It infers that the mobile elements are replicated during transposition. Therefore, each element comprises exactly the same set of non-permuted DNA sequences as present in the donor DNA or on the same DNA molecule. The nucleotides present at both ends of a transposon join with the broken segments of DNA where it is going to be inserted. For example, there are three groups of transposons in gram-negative bacteria: IS modules, Tn3 group, Mu and its derivatives. It has been shown that transposition is a

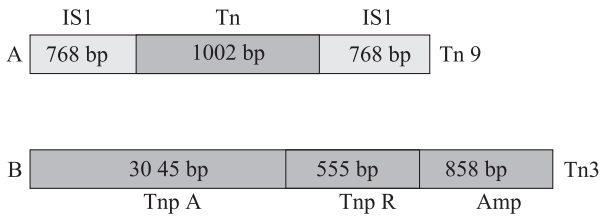

Fig. 1. Two types of composite transposons (Tn elements), one flanked by IS sequences (A) and the other in devoid of IS (B). Tn9 has a short inverted repeat of 23 bp but it is bound by direct repeat of IS1. Chloramphenicol (Cam)-resistance region is present on Tn element. Tn3 has three parts, two (TnpA and TnpR) of which an transposable and the third is ampicillin (Amp)-resistance region which shows gene amplification during resistance production.

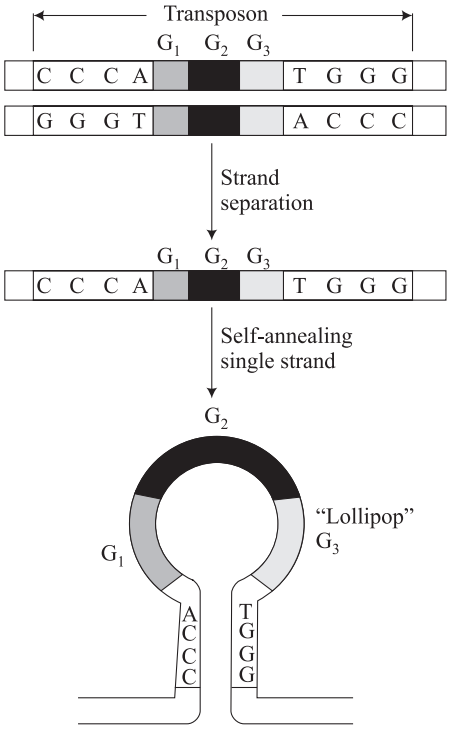

Fig. 2. Self-annealing of inverted repeat sequences produces a circle, the 'lollipop' which is visible under an electron microscope. Here, lollipop contains three, genes (G1, G2, G3) in the single stranded loop.

replicative process in the case of Tn3 family in which donor and target (recipient) replications fuse to produce a 'cointegrate' as an intermediate, which comprises two copies of Tn3, one at each junction between donor and target nucleotide sequences. The site-specific recombination between the two copies of a transposon can generate simple insertions from the cointegrate. Though the heterogenous collection of IS also promotes the fusion of replicons to form the cointegrate, the frequency is very low (1% of the frequency of simple insertions). Once the cointegrate is formed, it remains stable, i.e., no site-specific recombination occurs. If cointegrate is not broken down by recombination, how are simple insertions of an IS generated? Simple insertions are thought to be produced directly during incorporation of a single insertion sequence copy at a largest site of the DNA molecule. It is supported by the observation in the case of Tn5 elements that the cointegrates are produced by simple insertions from the dimers of donor replications rather than by replicons fusion between a monomer of donor and the target DNA.

It is thus evident that a fused structure or cointegrate is formed during transposition process. At the same time it remains unclear whether the cointegrate is a necessary intermediate or it is only a byproduct. In fact it is questionable because an enzyme should be present to resolve the fused structure if the cointegrate is a necessary intermediate. The function of this enzyme could be substituted by homologous recombination between two copies of the genetic element only in those cells which are proficient in recombination process.

Undoubtedly, the phenomenon of transposition also occurs in recombination-deficient cells mediated by an enzyme 'resolvase'. However, this enzyme is absent in several transposable elements except a single class of transposons. Another important fact reveals that a kind of molecular rearrangement is done to know the mechanism of transposition by DNA sequencing so that transposed sequences are ultimately generated. The possibility of error is not ignored, rather it has been suggested that the apparent duplication of the mobile genetic element during transfer may be due to the loss of donor chromosome after excision of a transposon-, 'suicidal transposition'.

Duplication at Insertion Site

In addition to replication of transposon, the replication of a short (3-12bp) DNA sequence occurs at the insertion site in the target (recipient) DNA. The nucleotide sequence of target DNA is found to have some duplicated nucleotides after transposition of the element, and the duplicated segments flank the transposon, i.e., the transposon integrated between duplicated DNA segments. The number of duplicated nucleotides is characteristic of particular element i.e., different nucleotides are duplicated for each insertion because the target DNA sequences are different at different insertion sites. The target DNA repeats in the case of IS1, IS2, IS3, IS4, and IS5 are 9,5,3 or 4, 11 or 12, and 4 bp respectively. Similarly, all TnA elements (Tn1, Tn2, Tn3, Tn501, Tn551, Tn801, Tn951, Tn1701, Tn2602 etc.) show target DNA repeat of 5 bp. The presence of short target DNA repeats indicates that the target positions on both the strands of DNA molecule. An experimentally constracted element is not flanked by direct repeats, indicating that the direct repeats do not play any structural role during transposition. Thus insertion of a transposon into a host DNA involves breakage of the target DNA not at random but in a staggered fashion, and a transposon gets inserted into and ligated with the cleaved or protruding chains of the target DNA molecule. It results into two gaps at both ends of target DNA, which are repaired by a small DNA duplication. The creation of flanking duplications is

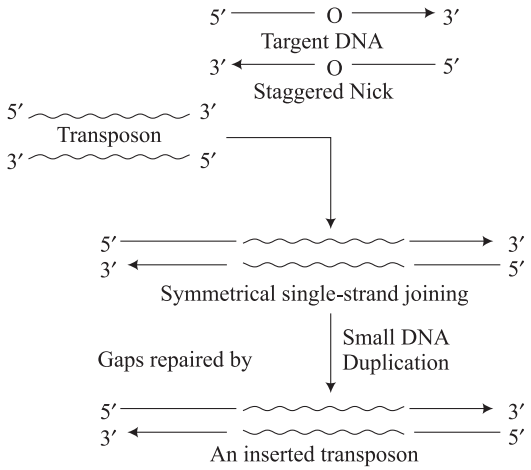

Fig. 3. Insertion of a transposon showing the precise and symmetrical joining of the transposon's ends to a target DNA molecule which is broken by staggered cuts at specific sites (0).

a characteristic of all prokaryotic transposons. The ligation of a transposon to the target DNA is always precise and symmetrical. The initial joining of the two is likely to be symmetrically-related single-stranded ligation at each end of the inserted element (Fig.3).

The duplication of a sequence of nine nucleotide pairs in host DNA by insertion of IS1 has been observed. Incorporation of such transposable element causes nine-nucleotide duplications. It is found that the staggered cleavage of recipient DNA occurs at positions 5, 9, or 11 nucleotide apart on opposite DNA stretches. After insertion of IS1 into the recipient DNA, the filling-in of gaps starts on both the strands of DNA. The process requires the synthesis of short single-strand fragments of complementary nucleotide sequences and in turn shows the nucleotide-sequence duplications (Fig. 4).

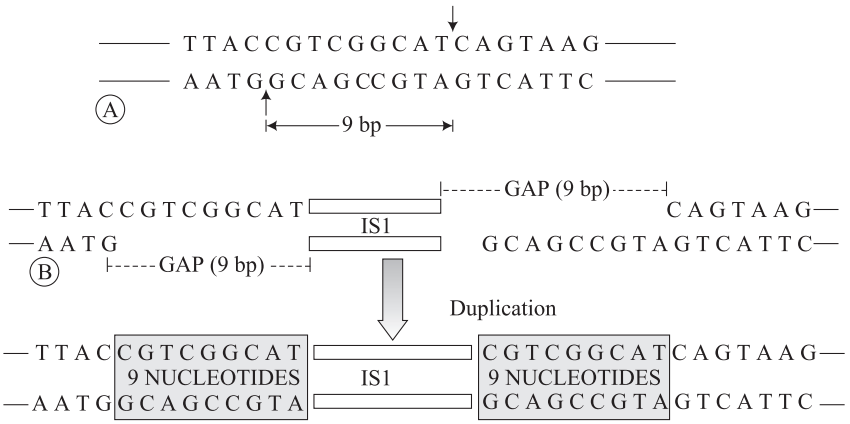

Fig. 4. Significance of nucleotide-sequence duplication in the recipient DNA molecule. Two strands of the DNA are cut at staggered sites (↑) which are 9 nucleotides apart from each other (A). Subsequently, an IS element (IS1) incorporates at the cut-sites (B) leaving 9 bases unpaired (gaps). Further, 9-nucleotide duplications of target DNA occur at both sides of the IS1 to fill-up the gaps. Thus, IS1 insertion generates a nine-nucleotide duplication.

Non-replicative Transposition

Transposition occurs by both replicative and nonreplicative ways. In case of non replicative transposition the transposon moves as a physical entity directly from one site to another and is conserved. Here the transposing element is lost from the donor site. After nonreplicative transposition the donor is either repaired or destroyed. Tn10 and Tn5 use the nonreplicative pattern of transposition. This type of transposition requires only transposase enzyme and not resolvase. Conservative transposition is a type of nonreplicative event in this case the mobile genetic element is excised from the donor site and inserted into a target site by a series of events in which every nucleotide bond is conserved. It resembles the mechanism of lambda integration. Transposes of such elements are related to λ integrate family. This way of transposition can mediate transfer of large elements so also of donor DNA from one bacterium to another. Actually, they are more appropriately regarded as episomes rather than transposons. Earlier they are classified as transposons.

Some transposons use only one type of pathway (replicative or nonreplicative) for transposition. However, other may utilized both pathways. The elements IS1 and IS903 use

both pathways. Similarly, the ability of phase Mu to turn to either type of pathway from common intermediate has been established (Lewin, 1994). Replicative transposition is the only mode of mobility of the TnA family. TnA family includes several related transposons and consists of large transposons (45 kw) of which Tn3 and Tn1000 are well known.

DNA Rearrangements

Insertion of a mobile genetic element into the target DNA initiates replication in the transposed and the target nucleotide sequences leading to the formation of double-stranded DNA segments separately. Each of the double strands comprises one old and one new (i.e., replicated) DNA stretch. The reciprocal recombination then occurs between the two copies of the transposon and inserts the element at a new genetic site, consequently regenerating the donor molecule (Fig. 5). The exact mechanism of such a recombination process is still not

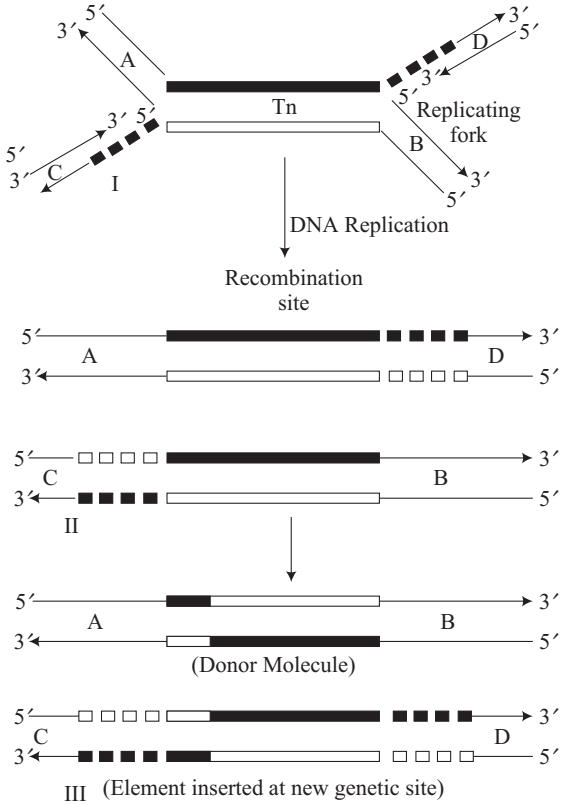

Fig. 5. Insertion of an element at a new genetic site by reciprocal recombination process. (1) : Replication of donor and recipient nucleotides after insertion. The donor segment originally comes from the A,B Genetic sites of a DNA molecule and is incorporated into recipient DNA between C,D sites; (II) : Two copies of a transposable element, each having one replicated strand, and (III) : Reciprocal recombination between these two copies of transposon incorporates the Tn element at a new genetic site and reproduces the donor DNA molecule (Cohen and Shapiro, 1980).

known. In case of Tn3, it is mediated through the nucleotide sequences present within the element. However, it is well known that such reciprocal recombination (non-homologous) does not require proteins which are needed in general homologous recombination process.

In addition to simple transposition process (Fig. 6A), the transposons produce a number of different types of DNA rearrangements involving the linkage of one or more transposon ends to a new target site. All types of DNA rearrangements can be explained on the basis of six types of events (1-6 in Fig. 6) mediated by transposons. Of these 1 and 3 are intermolecular events, while 2 and 4 are intramolecular in nature. Further in type (5), the incorporation of genetic element occurs without changing the intervening genes (e.g., ABC). It is not known whether this is due to inter- or intra-molecular transposition. The proper combination can lead to the formation of structures (products) similar to those generated by transposition alone. These structures are interconvertible. For instance, the homologous recombination can reduce the structure in type (5) to the structure in type (4), and can also interconvert the structures (5) and (2). Likewise, two circular DNA molecules each having a single copy of transposon from a cointegrate-type of structure by homologous recombination. It can be reverted through reverse reaction to generate two circular DNA molecules, each containing a copy of the element. Thus the homologous recombination process may be the most efficient way to produce some of the structures shown in Fig. 6 but under certain specific conditions.

From the above it is evident that IS are discrete DNA segments that can move from one genomic site to another and may cause genetic rearrangements. An interesting question arises whether these rearrangements are associated with duplication of IS that leads to rearrangements. In this connection, a strong evidence is provided by IS903, and it has been shown that any IS-mediated transcriptional recombination can occur by an efficient replicative mechanism (Weinert *et al.*, 1983). Based on several observations, Shapiro (1969) pointed out that the transposable elements bring together unrelated chromosomal segments to generate a variety of structural rearrangements including fusion, depletion, inversion and transposition. Hence the transposon play an important role in structural reshuffling of cellular DNA.

Tn rearrangements shown in Fig. 6 (e.g., 2, 3 and 5) generate products in which chromosomal (non-transposon) markers are bound by duplicate copies of an element. Such stretches are capable of their own transposition on the DNA molecules as discrete genetic units. These products come under the category of composite transposable elements which contain some accessory genes. In several cases the composite elements undergo the process of deletion or gene mutation to eliminate deleterious or unnecessary regions, thus refining the expression of accessory genes. Not only this rather composite elements can generate any of the structures shown in Fig. 6, but also the inner ends of IS modules can interact with a new target site to produce additional varieties of DNA rearrangements (Fig. 6:6). Several types of Tn elements have been identified which encode the accessory genetic determinants. For example, many transposons encoding genes for antibiotic resistance, heat-stable enterotoxin and lactose utilization have been shown in prokaryotes particularly in bacteria. The frequencies of many transposon-promoted genetic rearrangements range from 10^{-10} to 10^{-7} per generation.

Another example of DNA rearrangement has been reported from immunoglobulin (Ig) genes in higher organisms. Immunoglobulins (Igs) have two short or light (L) chains and two long or heavy (H) chains. Each chain possesses a N-terminal which constitutes the variable (V) region, and a C-terminal which has the constant (C) region. V and C regions are coded by two separate genes (V and C genes). Two types of L chain-kappa (*k*) and lambda (λ) are found

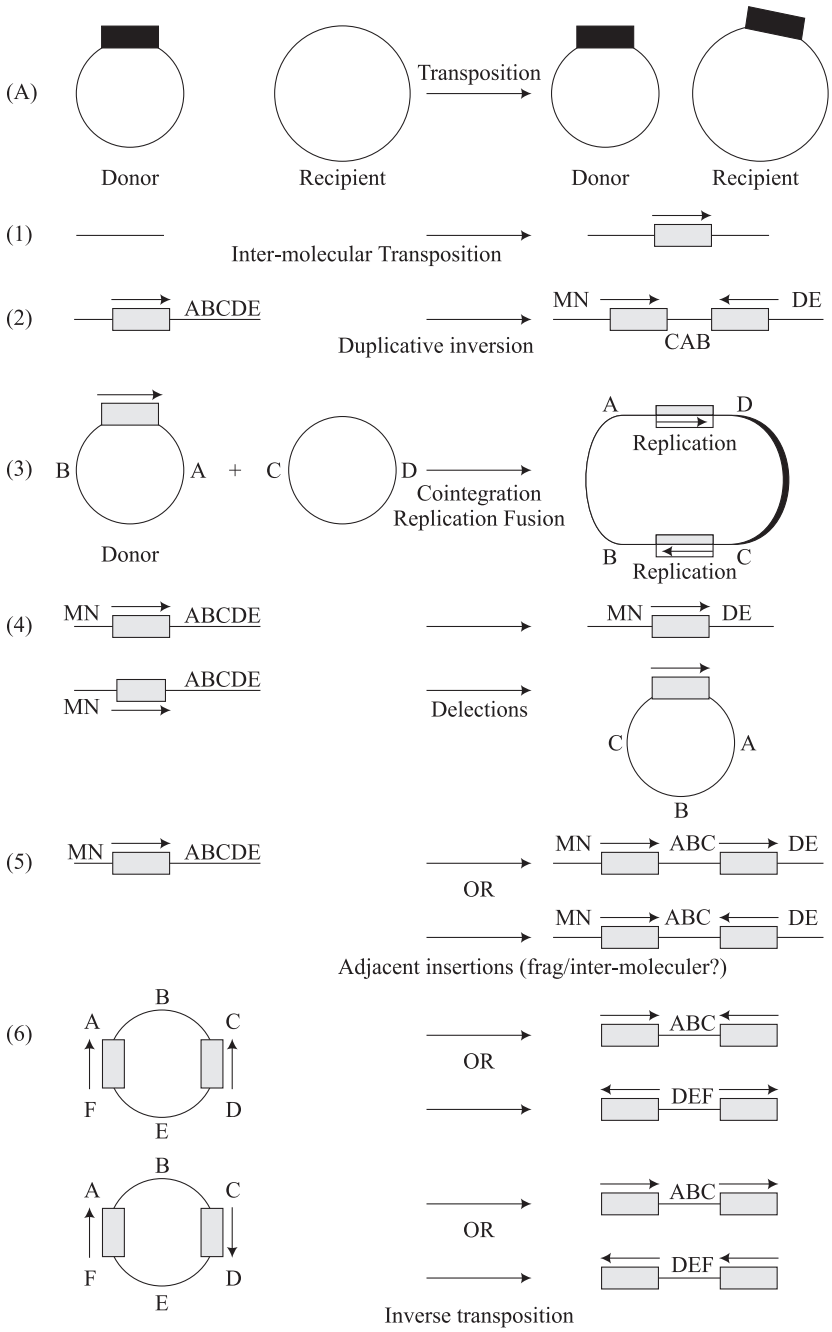

Fig. 6. Transposon-promoted genetic (DNA) rearrangements [(A) : Simple transposition of a genetic element between two different DNA molecules (donor and recipient). The numbers 1-6 indicate the varieties of transposon-induced events or rearrangements in which 1-5 are caused by a single transposon while 6 is due to a composite transposon.

which differ from each other in the amino acid (a. a) sequence of their C regions. Both the L chains of an Ig are always of one type; either *k* or λ. Likewise, H chains are of five types – alpha (α), delta (δ), epsilon (ε), gamma (γ) and mu (μ) based on their C region a.a. sequence. On the basis of their backbone, Igs are classified into five groups, IgA, IgD, IgE, IgG and IgM. The major immunoglobulins of the body are IgG, IgM and IgA. There are separate V and C genes for L and H chains. During differentiation of B-cells, transposition occurs in nucleotide sequences to bring a specific V gene into the proximity of a specific C gene (Fig. 7). This DNA rearrangement leads to transcription of a single mRNA from both V and C genes, which is translated into a L or a H chain from their respective mRNA molecules. In the body several thousand types of Igs are produced, which are necessary to combat millions of antigens (Ags). This is an important feature of the immune system to generate from a relatively small amount of genetic material a thousand types of antibody (Ab) molecules with different Ag binding specificities. Three families of Ig genes are recognized: (a) one consists of 20 to 30 genes coding for different V regions of both L and H chains; (b) nearby in the same chromosome is a second family of genes consisting of 2 to 4 genes coding for two C regions of L chain *k* and λ; and (c) the third family codes for the C-regions of various classes and subclasses of H chains (α, δ, ε, γ, μ).

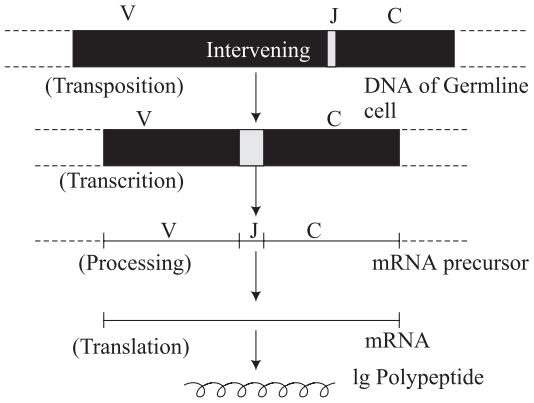

Fig. 7. Translocation of V and C genes through excision of intervening sequences and joining of the two by J segment. [Functional integration of V and C regions occurs by excision of the J, and forming an Ig polypeptide from V and C genes].

The genes within each family are linked but the families are not. On the basis of the above description it can be asked as to how an animal can produce about a million types of Igs? The mechanism for the expression of specific Ig genes must include selection of only one out of the two genes coding for *k* and λ; and the L chains should comprise a specific V gene alongwith one of the five classes of C genes for H chains (α, δ, ε, γ, μ). Thus antibody diversity arises by selection of one out of many V genes to combine with one out of some J genes to form the V region, which in turn, combines with one out of three to four C genes. In case of H chain one out of five classes, and many subclasses of C genes are selected. Therefore, 'generation of diversity of Ig' (God Ig) is by simple permutation and combination of many V, J and C genes. Hence the transpositions and rearrangements of certain specific DNA sequences in the immunoglobulin genes lead to the production of millions of antibodies which have much evolutionary significance.

MECHANISM OF TRANSPOSITION

The process of excision and reintegration of the transposable element at another site is termed as transposition. The ends of a transposon is essential to conduct the transposition process and its deletion causes transposition defect. Almost all nucleotide bps present in the terminal inverted repeats at the ends of the elements Tn3 and IS10 are important to contribute the transpositional information and the latter element (IS10) also contains recognition sites for DNA or protein interactions. So there must be some specific enzymatic mechanism to identify or recognize the inverted-repeat ends and for the precise breakage of DNA segments. The gene responsible for the synthesis of these specific enzyme is located in the transposon itself. The mutation in this gene abolishes the transposition i.e., it changes the ability of the element to act as a transposable element. Mutations at another region of Tn3 lead to increased frequency of Tn3 transfer or movement into different plasmids. The mutated region may have a gene which alters the transposition capacity of Tn3 element! This gene of Tn3 element is known as 'transposonase gene', and the normal transposonase gene encodes an enzyme required for the transposition i.e., transposonase (a 100000 dalton peptide). Tn3 also contains a repressor gene encoding a repressor molecule (a peptide of 21000 dalton) which represses transposition of Tn3. The repressor molecule regulates transcription of Tn3. The repressor molecule regulates transcription of the transposonase gene as well as the synthesis of the repressor molecule (Gill *et al.*, 1979; Chou *et al.*, 1979). If repressor gene is mutated, there will be a constitutive transcription of transposonase gene to synthesize the transposonase enzyme. Thus the repressor molecule regulates the transcription at the promotor (P) sites of transposonase as well as repressor genes, and both the genes are transcribed in opposite direction (Gill *et al.*, 1979; Chou *et al.*, 1979). The Tn3 comprises three coding regions : (i) gene (bla) for the enzyme betalactamase which is responsible for ampicillin-resistance, (ii) gene coding transposonase, and (iii) gene encoding for a repressor protein. The repressor protein controls transcription of both the transposonase and repressor genes (Fig. 8).

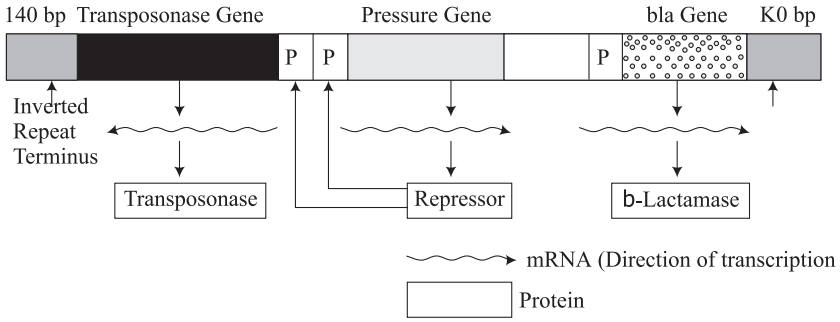

Fig. 8. Functional components of Tn3 transposon which is composed of three genes (transposonase, repressor and *bla* genes) [Transcription of transposonase and repressor genes is bidirectional to each other. The repressor molecule regulates transcription of both the transposonase and repressor genes through their promotor regions or promoters (P). β-lactamase is encoded from the *bla* gene and confers resistance to ampicillin as well as related antibiotics.

In some cases it has been suggested that the insertion site and orientation of transposon are decided by the ends of Tn element. For instance, if the ends of IS5 element is coupled to a promotor in an 'out-to-in' fashion, it directs the replication of recipient DNA. This is referred to as 'inceptor' activity. The inceptors are suggested to be the sites where transcription terminates to produce a primer for further DNA synthesis. In certain cases the ends of transposons possess active promoters, rho-termination sequences, Pribnow box like sequences or the sequences which can produce new promoters by DNA rearrangement. These indicate that the nucleotide sequences of the ends are A-T rich. Although the role of RNA polymerase in transposition is not known, it is speculated on the basis of above findings that the interaction of RNA polymerase and transposon ends could be important for transposition.

The insertion of transposon is not random. It can insert itself at multiple sites on a recipient DNA molecule. Certain parts of DNA are prone to multiple insertions. These specific regions of DNA are called as 'hot-spots'. Insertion, deletion and even inversion in Tn10 element occur regularly at hot-spots, and one such hot-spot has been shown per 1000 bp of the target DNA molecule. But it has also been documented that the Tn3 is inserted preferentially in the regions of recipient DNA, which are similar to the nucleotide sequences of the inverted-repeat ends of the transposon. Such insertions have been observed even in those bacterial cells which are unable to synthesize a protein (*recA* protein) needed for homologous recombination. Thus it appears that recognition of homologous DNA sequences may play some role in determining the frequency as well as site-specificity of transposon-associated recombinations (Cohen and Shapiro, 1980).

ROLE OF Tn ELEMENTS IN REGULATION OF GENE EXPRESSION

Insertion of transposable elements in DNA molecule controls the expression of adjacent genes. In some cases the genes are 'turned-off' while in others they are 'turned-on' by these elements. Hence the transposable elements act as switches to regulate the expression of genes.

Gene Turn-off

Many IS elements (IS-like modules) contain transcriptional terminators which turn-off the gene lying distal to the insertion sites. This causes polar mutation or polar effect (polarity) on the expression of genes present in an operon. Several such controls have been identified in classical operons (*gal* and *lac*) (Besemer and Herpers, 1977). Polarity of some IS elements induces the generation or synthesis of *rho*-termination factor for transcription, and the polar effect can be fully or partially relieved in *rho*-strains (Besermer and Herpers, 1977). Similarly, the polar mutations caused by insertions of the elements in lambda phage can also be suppressed by antitermination function (Kleckner *et al.*, 1978). It has also been suggested that the polar effects may be due to the stop codons present within the IS-element. So the multiplicity of stop signals and the absence of a nearby translation-reinitiation signals may lead to strong polarity.

Gene Turn-on

In some cases, insertion of the IS elements in DNA molecule activates the transcription of adjacent genes i.e., it turns-on the adjacent genes. Such regulation of genes by IS2 and IS3 interaction has been examined. Strong as well as weak turn-ons have been shown by IS2

insertions in both orientations. Tn5 insertion presents a weak turn-on of the gene. It has been reported (Saedler *et al.*, 1974) that IS2 activates the adjacent genes by the read-through transcription from a promotor region of the genes. However, it has been finally suggested that the transcriptional activation of adjacent gene is due to the formation of a new signal by juxtaposition of the element and target DNA sequences.

Several other evidences reveal that promoters present within IS elements are involved in the transcriptional activation of adjacent genes. For example, inversion of IS2 elements acts as a switch to control the expression of an adjacent gene (Fig. 9), and inversion of the loop is the result of a crossover between the two inverted repeat sequences (II) which can be reverted (I) by second crossover. The loop comprises a transcription terminator (T) and a promotor (P). Under normal condition (orientation in I), transcription of the promotor PG initiates the transcription of gene G. This transcription is terminated at T which ultimate inactivates the function of G gene. However, under inverted condition (orientation II) P starts transcription and in turn G is transcribed without initiation at the promotor PG. As shown in Fig. 10 the IS2 contains a strong transcriptional T and a P. So it may act as a switch for turning the genes on or off depending on its orientation.

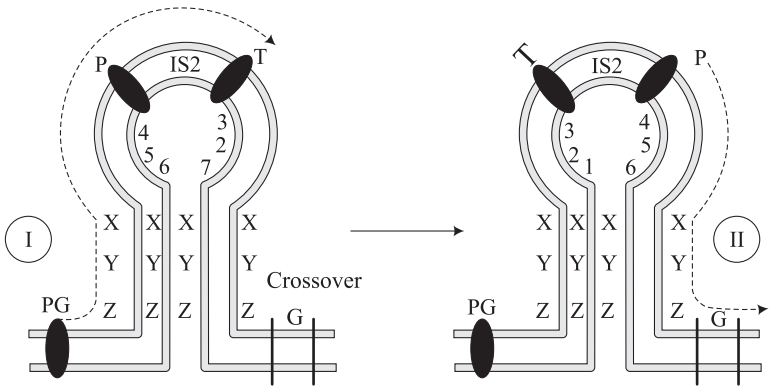

Fig. 9. Inversion of IS2 acting as a switch to control the transcription of an adjacent gene [Xx, Yy and Zz are the complementary base pairs of inverted repeat sequences. A crossover between these inverted repeat sequences forms an inversion of the loop which can be restored by another crossover. In orientation (I), the transcription of the promotor PG is terminated at T and in-turn inactivates the function of G gene. In case opposite orientation (II), P initiates transcription and G is transcribed without the transcriptional initiation at PG, thus turn-on the gene.

TRANSPOSITION OF ANTIBIOTIC RESISTANT GENES

It is well established that bacteria possess several antibiotic resistant genes in their genomes. They contain plasmids which are responsible to transmit different antibiotic resistances to a new bacterial cell. However, it was not clear until the discovery of transposition phenomenon as to how a number of antibiotic resistant genes were located on a single plasmid-DNA molecule? The answer to the above question is that the drug resistance-determinant segments of plasmids evolve as collections of transposons (Fig. 10), each carrying a gene conferring resistance to one or more antibiotics (Cohen and Shapiro, 1980). The plasmid, carrying several resistance genes, seems to be generated by the joining of a resistance-determinant segment and

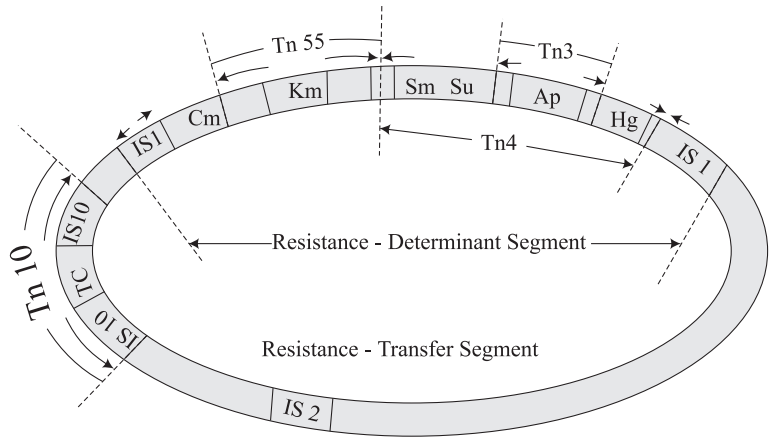

Fig. 10. A plasmid carrying several resistance genes represents the collection of transposons [Genes imparting resistance to chloramphenicol (Cm), kanamycin (Km), streptomycin (Sm), sulfonamide (Su), ampicillin (Ap, and mercury (Hg) are clustered on the resistance-determinant segment which is composed of many transposons (Tn55, Tn4, Tn3). Tn3 is present within the Tn4 element. While the transposon (Tn10) encoding resistance to tetracycline (Tc) is located on the resistance-transfer segment (Cohen and Shapiro, 1980).

resistance-transfer segment, the junction containing IS1 elements. Such junctions are the sites where the two segments (resistance-determinant and resistance-transfer segments) sometimes dissociate reversibly at a specific site. Each transposon present on the plasmid can be transferred at another site independently. IS elements are present at both the dissociable sites as well as at the sites where the interaction of plasmid and chromosomal DNA occurs to induce the transfer of genome between different bacterial cells. It has been shown that the antibiotic-resistant transposons can move even among the different bacterial species. For instance, certain DNA sequences similar to the Tn3 elements have been observed to be responsible for penicillin resistance in other bacterial species. The elements such as Tn1, Tn3, Tn501, and Tn1771 exhibit the phenomenon of transposition which is related to antibiotic resistance. Plasmid carrying one copy of an element does not allow the insertion of another copy of the same element in the plasmid. This inhibition of another or secondary insertion occurs to a lesser degree in cases of intramolecular transposition. The inhibition of secondary insertion is shown to be *cis*-acting. In Tn3 element, Tnp A end is essential for causing immunity and the immunity can occur even in the absence of Tnp R segment.

A GENERAL SURVEY OF TRANSPOSONS IN ORGANISMS

Transposons are present both in prokaryotic as well as eukaryotic cells. They are normal constituents of many bacterial genomes, plasmids and bacteriophages. Similarly, transposons have been identified in yeasts, maize, nematodes, fruitflies and sea urchins. IS2 to IS5 are present in *E. coli* K12 while IS1 in other strains of *E. coli* and many enterobacteria. ISR1 is located in the genome of *Rhizobium lupini* and IS8 in RP4 plasmid. Tn1 and Tn2 are present in plasmids of *Pseudomonas, Salmonella, N. gonorrhoea* and *H. influenzae*. Tn3, Tn4 and Tn2350 are located

within the bacterial plasmids R1 drd 19, whereas Tn5 is present in *Klebsiella* plasmid JR67 and Tn6 in JR72. Tn10 is the resistance-determinant of R100, Tn9 and Tn981 (Tn204) are the products of deletion derived from Tn2671. Sources for TnD (TnD:TnA and TnD:Tn cam) are *H. influenzae*.

DNA of a temperate bacterial virus (bacteriophage) is inserted at multiple sites of the bacterial genome leading to several types of mutations in the bacterium. The phage Mu is refered to as 'mutator'. The transposition of Mu DNA causes several rearrangements of the host DNA molecule. Several such studies have shown that Mu is a transposable element. Another example of Tn element is the retrovirus (RNA tumour virus) which induces cancer in higher organisms after infection either by inserting itself adjacent to the chromosomal oncogene leading to its expression or by picking up an oncogene from cell and inserting it into a more active site of the same or of different cell. First, the genomic RNA of the retrovirus produces a double-stranded (ds) DNA molecule by reverse trascriptase, which gets integrated to host DNA and modulates the expression of oncogene which is located adjacent to the insertion site. The oncogene inactivates certain other genes after transposition to another site and causes DNA rearrangement. Retrovirus also contains short terminal inverted repeats of nucleotide sequences. Due to these properties sometimes retroviruses are also considered as transposons. This is an example virus Tn elements integrated into the eukaryotic DNA.

A transposable element (Tcl) has been studied in the genomes of two closely related strains of *Caenorhabditis elegans*. It is found that there are approximately 30 copies of Tcl in the Bristol strain whereas 200 to 300 copies are present in the Bergerac strain. Most of them are structurally highly conserved. This is an example of Tn element having strain-specific differences. The distribution of transposons appears to be ubiquitous, hence it may not be possible to describe them in all organisms. Here we will discuss about transposons in some eukaryotic organisms which are considered standard for the related genetic studies.

(i) Yeast

There is a family (Ty elements) of dispersed repetitive DNA sequences in *Saccharomyces cerevisiae*. Tyl and Ty917 are the two major classes of Ty elements and most of them fall into one of the two classes. Tyl is composed of 35 copies of certain DNA sequences having a total length of 5600bp. Many copies have related nucleotide sequences but not the identical sequences. Each copy has 250bp known as S. There are 100 copies of S in the genome, out of which 70 copies are associated with Tyl sequences and 30 copies are situated somewhere else. Direct repetitions of S sequences resemble the structure of some bacterial Tn elements. The function and mechanism of transposition of Tyl mobile element are not clearly known. It may be said that Ty element transcribes into two poly $(A)^+$ RNA species, but whether the Ty RNA is translated into protein or not, is not known.

(ii) Fruitfly

There are many types of transposable elements in fruitfly *Drosophila*, in which *Copia*, 297, 412, P-elements are found. *Copia* and 412 are dispersed in the genome and each is present in approximately 35 copies. These are bound by short terminal direct repeat sequences. The

nucleotide stretch of *Copia* is 5000bp long having direct repeats of about 300bp. The length of 412 sequence is approximately 7300bp with direct repeats of 500bp. The sequence of 297 is not well known. The *White* mutation in *Drosophila* can be caused by insertion of *Copia* element. Another important example of transposons in *Drosophila* is confined to P-elements which are present in some lines and absent in others. In *Drosophila*, transposition of mobile elements generates a number of genetic disturbances known as 'hybrid dysgenesis' (Kidwell *et al.*, 1977) which comprises increased mutational rate, hybrid sterility and induction of chromosomal rearrangements. Hybrid dysgenesis is found only by a cross made between wild males and laboratory female flies, so the germ line of hybrids is affected. Hybrids produced in the first generation are morphologically normal but show reduced fertility and the progeny of these hybrids have other dysgenic traits or characters. P-elements of wild males cause hybrid dysgenesis in eggs of laboratory females lacking P-elements (M-female) and thus constitute a P-M system. Hence P-elements are found in P-stocks (30-50 elements/haploid genome) and absent in M-stocks. It has been suggested that hybrid dysgenesis involves interaction between paternal genome and maternal cytoplasm. P-element is believed to encode repressors of the transposition, and the P-element transposition is suppressed by the cytoplasm of females.

When certain genes cloned into P-element are injected in embryos of early stage the injected genes get inserted at multiple sites of the chromosomes of embryos and their expressions are in a regulated fashion. Through the expressions of these genes are different from that in mammalian cells, in which the clones are transplanted, it is not clear whether this is due to the differences in animals or transposons-guided regulation.

(iii) Maize

The transposition was first observed in the corn, *Zea mays*. At present much is known about the genetics and physiology of 'controlling elements' in maize. It was first discovered by McClintock that there are two genes in *Zea mays*, known as dissociator (Ds) and activator (Ac), which produce a drastic effect on colour-expression of the kernels. Ds affects the expression of genes adjacent to its location, while Ac acts on genes which are far apart from its position. Ac has 4300bp sequence and more than 30 copies of Ac derivatives are present in the genome which are defective in transposition. Simple internal deletion of Ac produces Ds hence, these are Ac derivatives. Complex structures having more than one Ds element can also be generated from these Ac.

An active maize Adhl-F gene, a Ds induced mutation of this gene and two independent Ac-induced revertants alleles have been described (Sutton *et al.*, 1984). Ds mutant has a 405bp insertion flanked by a direct repeat (8bp) which is a duplication of 8bp. The insertion sequence is A-T rich. It is suggested that the 405bp sequence is a Ds element. Low level of Adhl mRNA has been reported in Ds mutants. About 30 sequences are shown in the genome of maize, which are related to Ds element. It is significant to note that several elements have been found in various other plants having similar termini but different internal DNA sequences.

RETROPOSONS

Retroviruses have ability to insert DNA copies (proviruses) of an RNA viral genome into the chromosome of a host cell. Some eukaryotic transposons are related to retroviral proviruses in

their organization. They transpose through RNA intermediates. Such mobile genetic elements are referred to as retro–transposons or retroposons. They range from retroviruses that freely infect host cell to the sequences that have transposed via RNA. The sequences do not themselves have ability to transpose that is why they are transposed via RNA. Life cycle of retroviruses shows transposon like events. An interesting part of life cycle in the occurrence of transducing viruses. The transducing viruses are variants that have acquired cellular sequences substituted for part of the viral sequence. The part of viral sequence is replaced by the *v-onc* gene, as a result *Gag-v-onc* whether speciation is due to transposable elements. It needs an protein is synthesized instead of usual Gag, Pol and Env proteins. The transformed virus is replication defective and cannot sustain and infective cycle by itself Ty elements of yeast are similar to retrovirses.

Retroposons can be divided into viral and non-viral superfamilies. Viral superfamilies retroposones have capacity to transpose because they code for reverse transcriptase and/or integrase activities. This retroposon differs from retrovirus in not passing through an independent infectious form. The non-viral superfamilies retroposons originated in RNA sequences. They do not code for protein that have transposition functions and the transposition occurs by an enzyme system coded elsewhere. A significant part of the moderately repetitive DNA of mammlian genomes consists of retroposons (Lewin, 1994).

ROLE OF TRANSPOSONS IN SPECIATION

The possible role of transposable elements has been ascribed to hybrid incompatibility and speciation (Ish-Horowicz, 1982). Many mutations caused by hybrid imcompatibility (hybrid dysgenesis) are genetically unstable. Two such separate dysgenic systems have been recognized. They are P-M and I-R systems which are induced by two different mobile elements. Much is known about the P-M system in *Drosophila* which is caused by insertion of P-elements. It is found that several mutations can be induced in the *White* locus by hybrid dysgenesis, in which five out of seven mutants are due to insertions of homologous sequences of differing length. It has been shown that dysgenesis leads to a high frequency transposition of P-elements into M chromosomes. Some of the P-elements behave like P chromosomes. For instance, they can induce sterility after crossing with M females. Thus there is a strong correlation between the number of acquired P-elements and the ability of M chromosome to produce dysgenic sterility. In most of the cases phenotypes accompanied with hybrid dysgenesis act specially to mobilize P-elements (Ish-Horowicz, 1982).

Few reports indicate that in the population of *D. melanogaster*, dysgenesis may be a step towards speciation (Schaefer *et al.*, 1979). It is a speculated that the different elements generated in isolated populations can lead to hybrid sterility and reproductive isolation which, in turn, lead to speciation (Ish-Horowicz, 1982). But it is not exactly clear extensive species-specific studies of mobile elements to explore a clear-cut concept about the role of transposable elements in speciation.

EVOLUTIONARY SIGNIFICANCE OF TRANSPOSONS

Evolution of transposon has occurred either into confer some selective advantage to organisms under varying conditions or they are just selfish sequences of DNA (selfish DNA) and survive

because of a replicative advantage over the host genome. The first view indicates transposons as natural tools for genetic manipulations leading to variability or diversity among organisms for natural selection. In higher organisms, transposons play an important role for modifying gene activity during development, which shows a clear selective advantage. The second view holds that any advantage in hosts may be entirely incidental. Both the opinions regarding advantage conferred on the host can be explained. In a population of organisms residing in an antibiotic environment, transposon can confer adaptive advantage by causing antibiotic resistance. Transposition can also turn-on some silent metabolic pathways which may give an adaptive advantage to the population. This may be considered as incidental advantage.

Evans (1984) has thrown light on some evolutionary discussions about selfish or parasitic DNA. According to him an element, which imposes any cost on its own host, will be eliminated by natural selection (Evans, 1984). It creates the selection pressure which can be counteracted either by a mechanism for horizontal transfer (i.e., virus-like infection of cells) or by *de novo* production of new elements at a high rate. On the other hand, it has been suggested that during zygote formation, the parasitic elements capable of duplicative transposon are able to 'colonize' new chromosomes spreading them to new genomes by a process other than the horizontal transfer. Several studies incorporating negative selection into the model of selfish DNA suggests that the P-elements in *Drosophila* represents demonstrable negative effect on the fitness of the yet they have spread rapidly within the natural population. Therefore, the parasitic DNA is not only possible, rather highly probable, in eukaryotic organisms.

SUMMARY

Certain DNA sequences are capable of inserting themselves at specific sites on the chromosome(s) of the same or different cell(s). They are widely known as 'transposable elements' or 'transposons' because they transpose the blocks of genetic material back and forth. Transposition of these genetic elements causes multiple rearrangements in DNA leading to unstable mutations. Thus transposons control the expression of specific genes by turning them on or off. Transposons ignore the rules of general recombination by joining unrelated DNA segments. In addition to transposons that mobilize via DNA, there are also retroviruses and retroposons. Reverse trancription is the unifying mechanism for retroposons. Retroviruses are regarded from the perspective of the free viral (RNA) form, while retroposons are regarded from the stance of the genome (duplex DNA) form. Retroviruses have genomes of single-stranded RNA that are replicated through a double stranded DNA intermediate. An individual retrovirus contain two copies of its genome. The mobile genetic elements occur both in pro- and eukaryotes. They create genetic diversity (variation) and ultimately lead to evolution of newer traits. In addition to their medical importance in producing resistance to antibiotics, transposons are speculated to play an important role in speciation.

ACKNOWLEDGEMENTS

Authors are thankful to the Council of Scientific and Industrial Research, New Delhi for granting permission to reproduce major portion of an article published in Journal of Scientific and Industrial Research.

REFERENCES

Besermer, J. and Herpers, M. (1977) Suppression of polarity of insertion mutations within the gal operon of *Escherichia coli.* Mol. Gen. Genet., 151 : 295-304.

Chou, J., Lemaux, P.G. Casadaban, M.J. and Cohen, S.N. (1979) Transposition protein of Tn3: identification and characterization of an essential repressor-controlled gene product. Nature, 282 : 801-806.

Cohen, S.N. and Shapiro, J.A. (1980) Transposable genetic elements. Scient. Am., 242 : 36-45.

Evans, R.V. (1984) Limits on the selfishness of DNA. Nature, 310 : 456.

Gill, R.E., Heffron, F. and Falkow, S. (1979) Identification of the protein encoded by the transposable element Tn3 which is required for its transposition. Nature, 282 : 797-801.

Hedges, R.W. and Jacob, A.E. (1974) Transposition of ampicillin resistance from RP4 to other replicons. Mol. Gen. Genet., 132 : 31-40.

Isch-Horowicz, D. (1982) Transposable elements, hybrid incompatibility and speciation. Nature, 299 : 676-677.

Jordan, E., Saedler, H. and Stralinger, P. (1968) 0^o and strong-polar mutations in the gal operon are insertions. Mol. Gen. Genet., 102 : 353-363.

Kidwell, M.G., Kidwell, J.F. and Sved, J.A. (1977) Hybrid dysgenesis in Drosophila melanogaster : A syndrome of aberrant traits including mutation, sterility and male recombination. Genet., 86 : 813-833.

Kleckner, N. Barker, D.F., Ross, D.G., Botstein, D., Swan, J.A. and Zubeau, M. (1978) Properties of the translocatable tetracycline-resistance element Tn10 in *Escherichia coli* and bacteriophage lambda. Genet., 90 : 427-450.

Lewin, B. (1994) Genes V. Oxford University Press, Oxford, U.K. pp 1272.

Malamy, M.H., Fiandt, M. and Szybalski, P. (1972) Electron microscopy of polar insertions in the lac operon of *Escherichia coli*. Mol. Gen. Genet., 102 : 353-363.

Nervers, P. and Saedlor, H. (1977) Transposable genetic elements as agents of gene instability and chromosomal rearrangements. Nature, 268 : 109-115.

Saedler, H., Reif, H.J., Hu, S. and Davidson, N. (1974) IS2, A genetic element for turn-off and turn-on of gene activity in *Escherichia coli.* Mol. Gen. Genet., 132 : 265-289.

Schaefer, B.E., Kidwell, M.G. and Fausto-Sterling, A. (1979) Hybrid dysgenesis in Drosophila melanogaster : Morphological and cytological studies of ovarian dysgenesis. Genet., 92 : 1141-1152.

Shapiro, J.A. (1969) Mutations caused by the insertion of genetic material into the galactose operon of *Escherichia coli.* J. Mol. Biol., 40 : 93-105.

Sutton, W.D., Gerlach, W.L., Schwartz, D. and Peacock, W.J. (1984) Molecular analysis of Ds controlling element mutations at the *adh*1 locus of Maize. Science, 223 : 1265-1268.

Weinert, T.A., Schaus, N.A. and Grindley, N.D.F. (1983) Insertion sequence duplication in transpositional recombination. Science, 222 : 755-765.

Suggested Readings

1. Genes V (1994) Lewin, B. Oxford University Press, Oxford, New York.
2. Tripathi, G. and Shukla, S.P. (1986) Transposons. J. Scient. Ind. Res. (India), 45 : 21-23.
3. Text Book of Biochemistry (2002) Ed. T.M. Devlin, 5^{th} Edition, John-Wiley & Sons, Inc., Publication, New York.

REVIEW QUESTIONS

Long Questions

1. Define transposon and discuss insertion sequences in detail.
2. Explain basic events during transposition of DNA segment.
3. Distinguish between replicative and nonreplicative transposition.
4. Describe DNA rearrangement after insertion of a mobile genetic element
5. Elaborate mechanism of transposition in DNA.
6. Write an essay on general survey of transposon in organisms.
7. Discuss transposition in antibiotic resistance genes.
8. Write short note on the following :
   (a) Role of transposons in speciation.
   (b) Evolutionary significance of transposons.
   (c) Retroposons

Multiple Choice Questions

1. Transposon is
   (a) A static entity of genome
   (b) A loci at which no gene resides
   (c) An independent element like plasmid DNA
   (d) A discrete mobile sequence of the genome
2. There are two classes of transposons viz.,
   (a) DNA and mRNA
   (b) rRNA and DNA
   (c) Mobile DNA sequences and retroposons
   (d) Retroviruses and static genome
3. Transposons is a genetic restructuring which does not rely on
   (a) Donor sequences
   (b) Any relationship between donor and recipient sequences
   (c) Any relationship among additional sequences
   (d) Recipient sequences
4. Insertion sequences are
   (a) Simple transposition molecules
   (b) Direct terminal repeats
   (c) Inverted terminal repeats
   (d) Abnormal constituents of bacterial chromosomes
5. The rate of transposition per element per generation is
   (a) 10^{-7} to 10^{-5} (b) 10^{-10} to 10^{-8}
   (c) 10^{-4} to 10^{-3} (d) 10^{-8} to 10^{-6}

6. The enzyme which help in transposition is
   (a) Protein kinase (b) Transposase
   (c) Catalase (d) Superoxide dismutase
7. Transposition occurs by
   (a) Only replicative pathway
   (b) Only nonreplicative pathway
   (c) Both replicative and nonreplicative pathway
   (d) Only conservative pathway
8. TnA transposition requires
   (a) Transposase (b) Lamda integrates
   (c) Resolvase (d) Transposase and resolvase
9. Retroposon are
   (a) Only retroviruses
   (b) DNA sequences of eukaryotes
   (c) Only retrotransposons
   (d) Retroviruses infecting host cell to sequences that transpose via RNA
10. There are two superfamilies of retroposons viz
    (a) viral and nonviral (b) viral and bacterial
    (c) bacterial and fungal (d) nonviral and bacterial

Answers

1. (d)	2. (c)	3. (b)	4. (a)	5. (c)	6. (b)
7. (c)	8. (d)	9. (d)	10. (a)		