DNA Transposition: Classes and Mechanisms

Category: Introductory

Lisa AM Braam, *University of Wisconsin-Madison, Madison, Wisconsin, USA*
William S Reznikoff, *University of Wisconsin-Madison, Madison, Wisconsin, USA*

Accepted for publication: August 1998

Keywords: * transposon * retrotransposon * retrovirus * transfer * integration

Transposons are mobile genetic elements that contribute to the dynamic nature of chromosomes. Mechanistically similar, transposable elements make up a diverse classification including bacterial insertion sequences and composite transposons and eukaryotic excision/insertion transposons and retroelements.

Introduction

Organisms have evolved complex 'machineries' of both protein and deoxyribonucleic acid (DNA) parts that cut and join DNA in an event called transposition. Transposition has a major role in the plasticity of the genome causing chromosome rearrangements, DNA insertion and deletion mutations and horizontal gene transfer. The process contributes to evolution. The first transposons to be discovered were the *Ac/Ds* controlling elements in studies of maize chromosome breakage through the work of Barbara McClintock. Since then, transposons have been found in virtually every species examined.

Transposons are mobile DNA units that can 'jump' from one place to another – for example, between chromosomes or on to a plasmid. For the most part, transposons coexist peacefully with host cells owing to regulation or inactivation; however, in some cases, transposition can be detrimental.

Transposable elements (Table 1) occur in both prokaryotic (nonnucleated) and eukaryotic (nucleated) cells. The spread of antibiotic resistance genes throughout bacterial populations is due, in part, to transposition. In eukaryotes, including single cell organisms, plants and even humans, the elements can be either of the strictly DNA form or can take on a more complex strategy of movement via a ribonucleic acid (RNA) intermediate. An elaborate example is a retrovirus that spends part of its life cycle outside of the host cell.

Table 1 Common transposable elements

Bacterial	Retroviral	Eukaryotic
IS elements	Avian sarcoma virus	Class I
IS*630*	Human immunodeficiency virus	LTR
IS*3* family		Ty1
IS*4/5* family		non-LTR
		LINE
Composite transposons		
TN*5*		Class II
		Tc/mariner superfamily
Complex transposons		P elements
Tn*7*		
Mu		
Tn*3*		
Conjugative 'transposons'		

Protein–DNA Complexes

A transposon 'machine' has the following basic parts: (1) transposon DNA; (2) a catalytic protein termed transposase; and (3) divalent metal ion necessary as a cofactor. Figure 1 illustrates a basic process called cut-and-paste transposition. Assembly occurs first by transposase binding to the transposon DNA ends. Then, transposase–transposase interactions bring the ends together to form a synapse. Usually, only after synapsis has occurred can the transposition reactions take place. The DNA is cut at the ends and is subsequently transferred to a new location called a target. Short gaps are left adjacent to each side of the transposon that, after repair by host enzymes, result in target duplication.

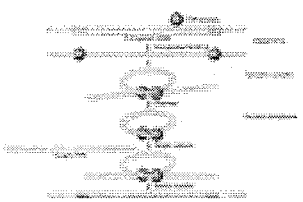

Figure 1
DNA transposition. The transposon ends are red, the target insertion site is yellow and the target DNA is green.

The transposon DNA element is autonomous in that it contains all of the necessary sequences for the production and assembly of the transposon 'machine'. The ends are inverted repeats that can be specifically recognized by the transposase. Internal to the ends resides a gene for the expression of the transposase itself. Other genes might be present that may or may not have a role in transposition – such as accessory factors or antibiotic resistance genes.

During transposition, transposase orchestrates many steps and, correspondingly, the protein contains multiple structural domains. A domain for site-specific DNA binding is usually located in the N-terminal part of the protein. In the body of the protein is a conserved primary sequence motif that is folded into a catalytic structural domain. This domain consists of a β sheet surrounded by α helices and brings three negatively charged amino acid side chains (two aspartates and one glutamate, the 'DDE' motif) into proximity, forming the active site. Finally, the transposase contains protein–protein interaction domains important during synapsis.

In general, transposition processes are of three types: cut-and-paste transposition, replicative transposition and retroviral-like DNA integration (Figure 2). All of these processes involve DNA cleavage and joining but are different in the extent of cleavage and in the type of host repair that takes place following transposition. Cut-and-paste transposition is characterized by the complete release of the transposon from the donor DNA sequence prior to target insertion. Thus, an excised transposon exists as an intermediate. In contrast during replicative transposition, the transposon is only partially released from the donor DNA. After transfer into a target, a complex remains that contains both the donor and target in addition to the transposon. Replication by the host cell produces a cointegrate molecule that has two copies of the transposon flanked by the donor and target DNA sequences. At this point, host recombination mechanisms or an element-encoded resolvase can act to separate the cointegrate into two DNA molecules, one containing the transposon with the target and one containing the transposon with the donor. Alternatively, this replicative transposition process can be diverted into a 'paste' type of insertion if host repair enzymes act on the target–transposon–donor intermediate to simply remove the donor DNA. The substrate for the third type of transposition process, retroviral-like DNA integration, is a linear DNA element that is processed by transposase to remove a short oligonucleotide before it is joined to a target site.

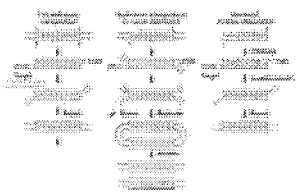

Figure 2
Comparison of transposition pathways. Grey segments represent transposable DNA elements. Small arrows indicate phosphodiester bond breakage, as discussed in the text (Basic reactions). Solid and dotted lines represent donor and target DNA, respectively.

Following transposition, the fate of the donor DNA varies. After replicative transposition, the donor remains as a copy. Cut-and-paste transposition presents a problem in that the plasmid or chromosome donor is left with DNA breaks. One solution to this problem is to coordinate transposition with host DNA replication so that one newly synthesized intact chromosome will persist. Apart from replication, multiple copies of transposon-carrying plasmids usually exist in any given bacterial cell and one can be lost safely. Furthermore, the cell may attempt to repair any DNA breaks by ligation.

Most transposons exhibit some specificity in target choice, although a few transpose randomly. In general, it is damaging to the host if the transposon inserts into genes. Therefore, some transposons have strategies for insertion into specific sites. Target recognition can be the result of specific protein–DNA interactions. Alternatively, the transposase may recognize structural aspects of the target DNA, for example, regions of bending.

The requirement for protein–protein interactions and protein–DNA complexes controls and coordinates the timing and precision of the transposition process. Synapsis, therefore, ensures that cleavage occurs on both ends at the same time. Furthermore, depending on the transposon, accessory proteins and accessory DNA

sequences might have a role in forming the synaptic complex and controlling transposition.

It is apparent that excessive transposition within a cell causes mutations and, thus, for transposon survival, attenuation of transposition is essential. Regulation of transposition is multifaceted. For example, the promotor for the transposase gene is often weak and may lie within or near the transposon ends where transposase binding can interfere with transcription. Some elements encode specific transcription repressors. Other transposases are regulated at the level of translation, for example, by the production of antisense RNA or by frameshifting strategies. In bacterial cells, the coupling of transcription and translation can be utilized; the DNA-binding domain of the transposase protein is produced first, at the N-terminus, and is presumed to bind to the inverted repeat ends even before translation is complete and before the C-terminus can interfere with DNA binding. This, combined with the fact that many transposases are inherently unstable, can cause a *cis* -acting phenotype in the bacterial cell such that transposase only acts on local DNA molecules. Post-translational regulation by transposon-encoded proteins is observed for some elements, involving competition for DNA binding sites or the formation of inactive complexes by protein–protein interactions. Also, a variety of host factors have been observed to participate at some level in transposition regulation.

Transposons make up a surprisingly large fraction of chromosomal DNA in some species. However, the elements are often cryptic due to inactive transposase coding sequences or defective transposon ends.

Basic Reactions

The basic catalytic reactions of transposition are DNA cleavage and DNA strand transfer. During cleavage, a water molecule within the active site serves as a nucleophile and attacks a phosphodiester bond in the DNA backbone at each transposon end. This reaction exposes a 3'OH group on the transferred strand (Figures 2 and 3). As part of cut-and-paste transposition, the nontransferred strand is also hydrolysed. Some of the simple insertion sequences utilize a hairpin intermediate in order to accomplish blunt cleavage at each end (Figure 3). Such a scheme allows a single active site to recognize and cleave both strands even though the strands initially have opposite polarities. Other transposon systems probably use a more complicated arrangement of transposase molecules to release the nontransferred strand. Subsequently, strand transfer occurs. The exposed 3'OH of the transferred strand acts as a nucleophile to attack a target phosphodiester bond (Figure 3). Strand transfer closely resembles cleavage except that the nucleophile is the DNA 3'OH, not water. Cleavage and strand transfer occur within the active site (separately) on the same or a nearby transposase molecule inside the synaptic complex. No high-energy cofactors, such as adenosine triphosphate (ATP), are required; in the strand transfer step, one phosphodiester bond is created for each one eliminated at the site of attack.

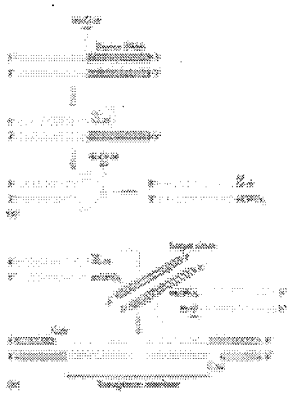

Figure 3
Basic reactions. (a) During cleavage at each end, an activated water molecule serves as a nucleophile to attack a phosphodiester bond. The opposite strand may or may not be released, depending on the transposon. In some systems, cleavage occurs via a hairpin intermediate as depicted. When the hairpin is resolved by hydrolysis, the transposon is left with a blunt end. (b) Strand transfer occurs via staggered insertions, one in each target strand. The result will be single-stranded gaps adjacent to each end that, when repaired, result in target site duplications.

Strand transfer occurs via insertions that are staggered relative to each other, one in each target DNA strand (Figure 3). As a result, short gaps (2–14 bp in length) will be left adjacent to each newly inserted end, the 'signature' that a transposition event has just occurred. Host repair, by DNA polymerization and ligation, fills in the gaps and produces direct repeats (duplications) flanking the transposon.

Cleavage and strand transfer take place inside the transposase catalytic domain where the negatively charged active site residues of the DDE motif coordinate divalent metal ions. The metal ions, probably Mg^{2+} physiologically, serve to activate the nucleophile and/or to facilitate lability of the phosphodiester bond at the site of nucleophilic attack. Also, positively charged residues near the active site probably play important roles for substrate DNA positioning.

Transposition can be distinguished from other types of DNA cutting-and-joining mechanisms by the absence of any protein–DNA covalent bonds during the process. Nontransposition mechanism examples include: site-specific recombination (such as the integration of phage λ in *Escherichia coli*); homologous recombination (important during eukaryotic meiosis and for DNA repair); topoisomerase-mediated DNA unwinding (alters DNA supercoiling); and a type of rolling circle DNA replication process in bacteria.

The mechanisms of transposition are highly conserved in nature and transposases are part of a superfamily of nucleic acid phosphoryl transfer enzymes. This superfamily includes ribonuclease H and a DNA processing enzyme, Ruv C, that resolves DNA junctions following homologous recombination in *E. coli* . Mechanistically similar to transposition is the process of V(D)J recombination that rearranges and joins various gene segments for the production of antibodies and T-cell receptors in human immune cells.

Bacterial Transposons: Overview

Transposons are widespread in bacteria and a large variety have been isolated. Mobile DNA elements have been associated with bacterial virulence and pathogenicity, and thus provide some advantages for their hosts in spite of the potential for chromosomal damage. In addition, transposons can jump on to conjugal plasmids and, in doing so, disseminate genes throughout populations and between species.

Some of the best-studied examples of transposition are found in Gram-negative bacteria, especially *E. coli* . The elements tend to fall into three classes: insertion sequences (IS elements), composite transposons and complex transposons. Exceptionally, 'conjugative transposons' form a distinct class of nontransposon mobile elements in that they move by a mechanism unlike transposition.

IS elements

Bacterial IS elements are compact transposons that encode only those factors essential for transposition. Interestingly, certain IS types can occur in different species, suggesting horizontal transfer of these pervasive mobile elements. One subgroup, the IS*630* family, is related to a eukaryotic transposable element family, *Tc/mariner* , that, surprisingly, has also crossed species barriers. IS sequences range in size from approximately 800 bp to 2500 bp. They contain short inverted repeat ends (9–40 bp), the ends sometimes differing slightly from each other. Many IS elements have host protein-binding sites, typically near or overlapping the ends, that either facilitate or decrease transposition. Most IS elements transpose by a cut-and-paste mechanism but some, for example IS *6* , form cointegrate intermediates as part of replicative transposition.

To date, many hundreds of IS elements have been isolated. The largest subgroups include the IS*3* , IS*4* and IS*5* families.

3 family

This family bears striking similarities to retroviruses. The transposase genes contain a 'DD(35)E' signature within the catalytic motif. Thus, the second aspartate residue and the glutamate residue are separated by 35 amino acids. Furthermore, at the transposon 3' ends, essential CA dinucleotides are found adjacent to the cleavage site. The transposases are produced as fusion proteins owing to a programmed translational frameshifting process that also attenuates transposition.

The IS*3* family probably transposes by an interesting variation on traditional transposition pathways. In IS*911* , transposon circles appear as intermediates. The circles are formed by a two-step process in which a 3'OH is exposed on one transposon end only, even though synapsis is required. The 3'OH undergoes strand transfer into a target near the other end, resulting in a circular intermediate that can be processed by host enzymes into a double-stranded circle. The circle has juxtaposed ends separated by a short sequence. The arrangement creates a strong promoter for the transposase gene that is then transcribed at high frequency leading to elevated transposition. As soon as this happens, the process downregulates itself because transposase binds to the ends and initiates a more familiar form of transposition cleavage/strand transfer, disrupting the promotor.

4 and IS*5* families

The IS*4* and IS*5* families are closely related. The best-characterized members are IS*10* and IS*50* of the IS*4* family.

The cut-and-paste transposition of IS*10* is highly characterized. Probably monomers of transposase, 412 amino acids each, bind to each transposon end (22 bp in length). Thus, two transposase molecules likely form the active synaptic complex. Cleavage produces blunt ends – that is, both strands are cut at the same position in the DNA sequence. The 3' strand is cut first, then a hairpin intermediate is formed, completely releasing the transposon end, and finally the hairpin is resolved by hydrolysis. Subsequently, the excised transposon captures and then inserts into a 9-bp target site. Therefore, after transposition, the 9-bp sequence is duplicated adjacent to the inserted transposon. During transposition, at each end, one transposase molecule is responsible for all of the transposition reactions that occur at that end.

Transposases of the IS*4* and IS*5* families do not have the DD(35)E catalytic motif of the IS*3* family and of retroviral integrases. For example, IS*50* transposase has a DD(138)E motif. Nevertheless, the catalytic domain of the IS*50* transposase has the same overall structure as retroviral integrases, as has been determined by crystallographic structural studies.

Composite (compound) transposons

Composite transposons are composed of two insertion sequences arranged in inverted orientation relative to each other. The IS elements flank other genes that are not necessary for transposition. For example, Tn*5* transposon consists of two IS*50* elements, IS*50* R and IS*50* L, differing only by a single point mutation. Between the IS sequences is an operon encoding resistance to kanamycin, bleomycin and streptomycin. Transposition of Tn*5* occurs via pairs of end sequences, called 'outside ends', one from each IS element. Because it inserts with relatively low target specificity, Tn*5* has been widely used as a tool in genetic studies.

Tn*5* has a unique regulatory scheme. IS*5* *0* R encodes an inhibitor protein in the same reading frame as the transposase but initiating from a downstream start site. This protein (missing an N-terminal DNA-binding region) inhibits transposition through nonproductive multimerization with transposase.

Complex transposition systems

Complex transposons of bacteria encode accessory proteins, in addition to transposase, that function during synaptic complex assembly or target capture or that assist in the resolution of intermediates. The best-characterized complex transposons are Tn*7* , Mu and Tn*3* .

7

Tn*7* transposon moves by cut-and-paste transposition; however, the architecture of synapsis is complex. Multiple proteins are encoded by the transposon, including two transposases that act in concert. Transposase A is responsible for 5' strand release only, whereas transposase B cuts the 3' strand and carries out strand transfer. Furthermore, proper synaptic complex assembly occurs only with assistance from other transposon proteins and requires target binding before cleavage can start.

Unlike the cut-and-paste elements IS*10* and IS*50* , Tn*7* does not produce blunt ends upon cleavage but leaves a 5' three base overhang. Tn*7* chooses targets carefully and inserts into a single chromosomal site with the help of accessory proteins. In addition, Tn *7* exhibits a phenomenon called transposition immunity that prevents insertion of the ends into the transposon itself. Immunity is mediated by an element-encoded nonspecific DNA binding protein that has ATPase activity.

Mu is a bacteriophage that lysogenizes into a host chromosome as part of its life cycle. Unlike bacteriophage λ, which only integrates into specific chromosomal sites using a site-specific recombination mechanism, Mu inserts via a transposition mechanism and can choose numerous target sites. As with Tn *7* , multiple proteins are produced by Mu coding sequences; however, only one transposase is necessary for the catalytic activities. Unusually, the transposon ends contain tandem repeats of transposase-binding sites and the active synaptic complex consists of a transposase tetramer. Prior to cleavage, a distant DNA site called the enhancer must also be brought into the complex. The synaptic complexes are highly interwoven, such that monomers bound to one DNA end reach across to catalyse transposition reactions at the opposite end. Mu moves by replicative transposition and, in a manner similar to Tn*7*, shows target immunity.

3

Tn*3* encodes one transposase and moves by replicative transposition; however, whereas Mu relies on host proteins to resolve cointegrate intermediates, Tn*3* encodes a special protein for this function, the resolvase, that resolves the cointegrate into two plasmids. The resolvase works by a site-specific recombination mechanism.

Conjugative transposons

Certain conjugative elements also encode functions for the specific chromosomal integration of the element DNA sequence upon conjugation into a recipient bacterial cell. Thus, these elements were originally termed 'transposons' because they moved from one DNA sequence to another. However, the elements do not encode transposases and do not move by a transposition mechanism; instead, they integrate by site-specific recombination.

Retroviral Integration

Retroviruses probably evolved from transposons having acquired the ability to move from one chromosome in one host cell to another chromosome in another host cell. Thus, in addition to the basic functions necessary for transposition, retroviruses also encode a variety of other factors important for virus formation and infectivity. Retroviruses are found in vertebrate animals. In some mammals such as chickens, retroviruses can cause tumours. In humans, the human immunodeficiency virus (HIV)-1 retrovirus infects cells of the immune system causing an immunodeficiency syndrome.

An essential step in the life cycle of a retrovirus is replication. In order to replicate, a retrovirus must be able to insert a copy of itself into a host chromosome. This process is called integration. The fundamental aspects of integration occur by a transposition mechanism.

A retrovirus contains genes *gag* , *pol* and *env* that are flanked by long terminal repeats (LTRs). The *gag* gene encodes structural virion capsid proteins; the *pol* gene encodes several activities as a polypeptide precursor, including reverse transcriptase and integrase; and the *env* gene encodes envelope proteins essential for infectious virus formation.

After infection of a host cell, the retrovirus unpacks the single-stranded viral RNA genome. Within a large nucleoprotein complex, viral reverse transcriptase produces a double-stranded DNA copy of the genome. The DNA is bundled into a preintegration complex (PIC) along with viral and host proteins. Then, PIC translocates to the nucleus where integration can take place.

The DNA copy in the PIC is held in a configuration similar to a transposon synaptic complex with the LTRs acting as the ends, although the exact architecture remains to be elucidated. As with the IS*3* family of insertion sequences, a conserved 3' CA end sequence is critical for transposition. The active integration complex within the cell is elaborate, extending over many base pairs at each LTR end. Some target preferences for insertion are observed, including exposed nucleosomal DNA and transcriptionally active regions. Also, a cellular 'immunity' factor has been identified that protects against self-integration.

Integration occurs by DNA cutting and joining in a mechanism identical to transposition (Figure 2), although, traditionally, the transposase is called an integrase and strand transfer is referred to as integration. Cleavage releases a dinucleotide from each 3' strand to expose a 3'OH. The 3'OH then mediates strand transfer into a target site. Just as with other transposons, the insertions occur in a staggered fashion. In the case of HIV-1, this stagger is 5 bp in length. Repair of the resulting gaps leaves 5-bp duplications adjacent to each side of the integrated element, which is now called a provirus.

The structure of the HIV-1 integrase catalytic domain has been solved by X-ray crystallography and is similar to other transposases that have been studied by this technique: avian sarcoma virus (ASV) retroviral integrase and two bacterial transposases, Mu and IS*50* . As predicted, the HIV-1 integrase catalytic domain contains the DD(35)E signature shared with IS*3* transposases. Thus, this particular catalytic domain has been conserved by nature for various transposition purposes. The domain is also predicted for eukaryotic excision/insertion transposases and retrotransposon integrases.

Eukaryotic Transposons

Eukaryotic mobile elements are classified based on whether or not an RNA intermediate is involved in

transposition. Class I elements have RNA intermediates, whereas class II elements do not. All eukaryotic species contain mobile elements and, in some cases, the elements serve useful functions for the host.

Class I

The class I elements are retroelements that require reverse transcription of RNA into DNA prior to transposition. Retroelements are of two types, LTR and non-LTR. In some species, including humans, retroelements make up a significant fraction of the genome especially LINE and SINE retroelements.

These retroelements are highly similar to retroviruses both in arrangement and in mechanism of transposition. The fundamental difference is that LTR retrotransposons do not encode the *env* gene and, therefore, do not synthesize infectious virions. For example, the Ty1 elements of yeast *Saccharomyces cerevisiae* rely on host functions to transcribe an RNA copy of the DNA element and to translate various element-encoded proteins, including reverse transcriptase and integrase. Virus-like particles are produced that, presumably, facilitate reverse transcription (of the RNA copy into DNA) and PIC formation, leading to integrase-mediated transposition back into the host genome.

Also referred to as retroposons, non-LTR elements lack the terminal repeats of retrotransposons; nor do they encode a transposition integrase. For example, LINE elements (long interspersed elements) instead encode an endonuclease that can introduce a staggered cleavage at a target site to expose 3'OH groups. The key to retroposon integration is reverse transcription of the RNA element, beginning at the exposed DNA 3'OH of the target, so that as the DNA copy is produced by reverse transcriptase, it is simultaneously attached to the target. In *Drosophila melanogaster* , the ends of chromosomes (the telomeres) are maintained by this type of retroposition process. Retroposition is fundamentally different from transposition in that end cleavage and strand transfer do not occur; however, the process leaves short duplications adjacent to the inserted element, as observed with transposons.

Class II

The class II elements need no RNA intermediates and move by cut-and-paste transposition. Class II transposons are also referred to as excision/insertion elements.

Tc/mariner superfamily

These elements resemble bacterial insertion sequences both in size and organization. In fact, the *Tc/mariner* family is related to the bacterial IS*630* family.

Tc/mariner elements have been found in many species, including fungi, insects, plants, fish and mammals, even humans, and are ancient in origin. Surprisingly, *Tc/mariner* family members have apparently crossed species barriers in a spectacular fashion by horizontal transfer, although the precise mechanisms of transmission are not known; for example, evolutionarily related transposons have been found in both insect and mammalian species, clearly distant relatives of evolution!

Members of this family encode a single functional transposase gene. The Tc1 transposase from the worm species *Caenorhabditis elegans* catalyses cleavage of the the element leaving a 2-bp 3' overhang at each end and catalyses insertion into TA target sequences exclusively.

Some species contain inactive copies of *Tc/mariner* transposons. The inactivation is thought to result, in part, from selective pressure that has caused mutations in the transposase genes. Amazingly, it is possible

to reconstruct an active transposase based on an analysis of many defective versions from different species, as has been done using eight species of fish. Because they are relatively species-insensitive, *Tc/mariner* transposases are attractive as potential eukaryotic genetic tools.

The P elements were originally identified in the fly species *D. melanogaster* in studies of hybrid dysgenesis mutations that occur in germline cells. The elements vary in size, the largest approximately 3000 bp in length, and are species-specific, such that different *Drosophila* species contain different P elements, although some species contain no P elements at all.

P element movement is tightly controlled and occurs only in the germline, owing to regulation at the level of transposase gene pre-mRNA splicing.

Both the transposase and the transposon end sequences are unusual in P elements. Transposase binds near, but not on, the terminal inverted repeat ends at an internal adjacent site. The cleavage reactions produce recessed 5' ends. Although P elements move by a traditional cut-and-paste mechanism, the transposase requires guanosine 5'- triphosphate (GTP) as a cofactor (although not GTP hydrolysis) for unknown reasons.

Further Reading

Berg DE and Howe MM (eds) (1989) *Mobile DNA.* Washington, DC: American Society for Microbiology.

Chaconas G, Lavoie BD and Watson MA (1996) DNA transposition: jumping gene machine, some assembly required. *Current Biology* **6**: 817–820.

Craig NL (1995) Unity in transposition reactions. *Science* **270**: 253–254.

Dawson A and Finnegan DJ (1998) Sleeping beauty awakes. *Nature Biotechnology* **16**: 20–21.

Finnegan DJ (1997) Transposable elements: how non-LTR retrotransposons do it. *Current Biology* **7**: R245–R248.

Grindley NDF and Leschziner AE (1995) DNA transposition: from a black box to a color monitor. *Cell* **83**: 1063–1066.

Lewis SM and Gillian EW (1997) The origins of V(D)J recombination. *Cell* **88**: 159–162.

Polard P and Chandler M (1995) Bacterial transposases and retroviral integrases. *Molecular Microbiology* **15**: 13–23.

Sherratt DJ (ed.) (1995) *Mobile Genetic Elements.* Oxford: Oxford University Press.

Skalka AM (1993) Retroviral DNA integration: lessons for transposon shuffling. *Gene* **135**: 175–182.