REVIEW ARTICLE

Foldback Intercoil DNA and the Mechanism of DNA Transposition

Byung-Dong Kim[1,2]*

[1]Department of Plant Science, Seoul National University, Seoul 151-921, Korea,
[2]The Korean Academy of Science and Technology, Seongnam 463-808, Korea

Foldback intercoil (FBI) DNA is formed by the folding back at one point of a non-helical parallel track of double-stranded DNA at as sharp as 180° and the intertwining of two double helixes within each other's major groove to form an intercoil with a diameter of 2.2 nm. FBI DNA has been suggested to mediate intra-molecular homologous recombination of a deletion and inversion. Inter-molecular homologous recombination, known as site-specific insertion, on the other hand, is mediated by the direct perpendicular approach of the FBI DNA tip, as the attP site, onto the target DNA, as the attB site. Transposition of DNA transposons involves the pairing of terminal inverted repeats and 5–7-bp tandem target duplication. FBI DNA configuration effectively explains simple as well as replicative transposition, along with the involvement of an enhancer element. The majority of diverse retrotransposable elements that employ a target site duplication mechanism is also suggested to follow the FBI DNA-mediated perpendicular insertion of the paired intercoil ends by non-homologous end-joining, together with gap filling. A genome-wide perspective of transposable elements in light of FBI DNA is discussed.

Introduction

The transposable element (TE) was first described by McClintock [1] as a controlling element that jumps from one position to another in the maize chromosome in the mid-1940s, long before the discovery by Watson and Crick of the double helix structure of DNA as a genetic element [2]. It was a time when, after the rediscovery of Mendel's Law of Heredity in 1900, the cytogenetic study of chromosomes was at the forefront of genetics and when genes were thought to be beads on a string located on the chromosome. It was in the late 1970s and early 1980s when insertion sequence (IS) elements and transposons (Tn) were discovered and found to be similar to McClintock's controlling elements, and her AC/DS elements were confirmed at the DNA sequence level as a TE [3]. Salient features of DNA transposable elements (dTEs), including terminal inverted repeats (TIRs), target site duplication (TSD), the transposase gene, and simple and replicative transposition mechanisms have been well established by extensive molecular biology and biochemistry

studies [4]. Then, retrotransposable elements (rTEs), including long terminal repeat (LTR)-retrotransposons, non-LTR-retrotransposons, and other retro-elements, were being added to the repertoire since the 1980s to make the picture diverse and complicated [5]. Many new families of TEs are being added, especially from eukaryotes, even by computational screening in the post-genome era, which has necessitated a new classification based on their structures and mechanisms of transposition [6].

The presence of foldback intercoil (FBI) DNA was first reported by Kim in 1985 [7] and shown by a space filling model to mediate intra-molecular homologous recombination of inversions and deletions. It was further shown in 1987 that FBI DNA can mediate inter-molecular DNA rearrangements, such as site-specific insertions, at the foldback tip and DNA transpositional integration at the intercoil end of the detached dTE [8]. It will be examined in this review how FBI DNA mediation of transposition can be extended to different classes and families of rTEs. Readers are suggested to refer to many good reviews available on the

classification, structures, genetic contents, and function of the gene products and on the mechanisms of transposition with excellent diagrammatic illustrations [5, 9-12]. This review will minimize duplicate descriptions and focus on the mechanistic features that are relevant to the application of FBI DNA to the mechanisms of DNA transposition.

Structures of TEs and Mechanisms of Transposition

To build the basis for comparison and understand transposition mechanisms, it is necessary to consider TE structures based on the presence or absence of the following factors: 1) RNA intermediates before transposition, 2) TIRs or LTRs, 3) transposase or reverse transcriptase (RT) and other proteins related to autonomous or non-autonomous transposition, 4) TSD, and 5) other features, like DNA replication modes, hairpin structure, and enhancers.

Transposition of dTEs involves direct movement of the dTE intermediate either by a simple cut-and-paste mechanism, as in Tn5 [13], Tn7 [14], and Tn10 [15], or by replicative transposition, in which case one copy stays in the original site and a daughter dTE copy appears in a distant location [16-18]. Step-wise description of the transposition events includes the requirement and pairing of both dTE ends [13, 19], binding of the transposase proteins to the terminal sequences, double-strand breaks (DSBs) of the dTE terminals [14] and target ends, ligation of the 3'-OH donor end and the 5'-P target end, and gap repair of the single-strand gaps on the target that were generated by DSB [20]. In the case of the Mu bacteriophage, the gap can be repaired while the entire transposon gets replicated in a cointegrate mechanism [21, 22].

Transposition of rTEs involves an RNA intermediate that is reverse-transcribed by RT into a cDNA. This double-stranded linear DNA functions as the direct precursor for integration by an integrase protein into a new location in the host chromosome in a 'copy and paste' manner, as in the 'cut-and-paste' transposition reaction in dTEs [20, 23].

TIRs are a hallmark of dTEs, as LTRs have been that of rTEs until they were joined by non-LTR retrotransposons, which have a 5'-untranslated region (5' UTR) and 3' UTR, as typified by LINE-1 (L1), or which has neither of the terminal elements, as typified by the Alu element. It becomes clear that the presence of terminal repeats in the TEs is no longer a common requirement for the transposition reaction.

The description of the protein factors above will suffice for this review, except for the long interspersed nuclear element (LINE) open reading frames (ORFs). The retrotransposition and integration of LINE have been viewed as a coupled process, called target- primed reverse-transcription [5, 24].

However, this model has recently been enhanced by the finding that L1 reverse- transcription does not require base pairing between the primer and template [25]. Both the RT and endonuclease domains in L1 are encoded by the same ORF. Therefore, the cDNA that was reverse-transcribed from the mRNA expressed from the genomic copy of LINE is inserted into the host genome [9].

TSDs are generated in almost all classes and families of TEs, except for the Helitron dTE superfamily [6, 9, 10]. The size of the TSD varies from 2 to 11, but the sequences are not conserved. The size is unique to superfamilies and can be used as a diagnostic feature. The ubiquitous presence of TSDs, which are generated by gap fillings (GFs) after TE integration on the target [26], reflects the unity in transposition reactions, regardless of dTE or rTE, and the presence or absence of terminal repeats [22]. Yet, the meaning of the small oligomer size of the duplicated target is seldom discussed in terms of DNA structure. This aspect will be dealt with below with the FBI DNA mechanism.

Helitron, a eukaryotic dTE, is considered, together with Polintons, a third class of TEs for many unique aspects [9]. Helitron does not have TIRs but has 5'-TC and CCTR-3' termini instead. They contain 16–20-bp hairpins 10–12 nucleotides inside of the 3'-end and transpose precisely between the 5'-A and T-3' with no TSD generation. It is known to adopt replicative rolling circle transposition [27]. Helitrons indiscriminately capture and mobilize gene sequences and may impact hybrid vigor or heterosis in maize [28]. The hairpin is presumed to serve as the terminator of rolling circle replication [9]. Since the hairpin sequence corresponds to the FBI DNA motif, it will be interesting to consider whether it can not only serve as a terminator or rolling circle replication but also serve as a replication origin for its own replication and for frequent gene capture activities, which the Helitron is capable of.

Polinton was discovered and characterized based on bioinformatics studies [27]. Polinton is 15–20 kb long, with a 6-bp TSD and 100–1,000-bp TIR at both ends. Polintons code up to 10 proteins, including a family of B DNA polymerases (POLBs), a retroviral-like integrase, an A transposase, and an adenoviral-like cysteine protease. Polintons, however, with their structural characteristics and presence of TSD, should follow non-homologous end-joining (NHEJ) with the GF transposition mechanism, as ordinary dTEs. What makes it unique is the presence of a self-encoded POLB and a short 1–3-bp terminal tandem repeat that supports a protein-primed self-synthesis mechanism for Polinton propagation. This mechanism calls for an extrachromosomal single-stranded Polinton intermediate that forms a racket-like structure and follows POLB-mediated replication to become double-stranded before being integrated into the

host genome [9].

It has been proposed that eukaryotic dTEs follow three types of transposition mechanisms according to different replication modes: cut-and-paste for most dTEs, rolling circle replication for Helitrons, and synthetic replication for Polintons [9]. In fact, there seems to be many different cases of transposon-related replication that need to be clarified in the future. As mentioned earlier, continuation of DNA replication of an inserted TE molecule, like in Mu co-integration, may also fill the gap to generate TSDs. It should be emphasized that DNA transposition has been grouped in two: simple, conservative, or 'cut-and-paste' mechanism and replicative transposition [16, 17, 29]. In Tn5, at the outer ends, ori-C-like 9-bp sequences are found, which may serve to direct host replication functions to the ends of the element during transposition [17]. Transposition of Tn5 may involve replication proceeding from both ends in a symmetrical manner. This type of mechanism may differ from the polarized replication of other elements [17]. A synthetic mechanism is proposed for Polinton propagation to explain protein-primed replication of a racket-like structure of a single-stranded Polinton [9]. If it gets to be understood that the heteroduplex structure that is derived from FBI DNA may serve as the template of DNA replication, many of the diverse cases of DNA replication may be clarified (see below).

FBI DNA

FBI DNA is formed by the folding back at one point of a non-helical parallel track of double-stranded DNA at as sharp as 180° and the intertwining of the bent helixes within each other's major groove to form an intercoil with a diameter of 2.2 nm (Fig. 1A). FBI DNA could be considered a double-stranded version of a hairpin or stem-and-loop structure of RNA and single-stranded DNA. But, such sharp

bending and formation of a four-stranded intercoil are not considered possible with a double helical B-DNA with a diameter of 2.0 nm [8]. In an extended definition, FBI DNA may contain a loop and a much longer stem under biological conditions. Intercoil DNA is designated to distinguish itself from supercoil DNA, which is thicker than 2.2 nm. Intercoil DNA is also termed to avoid the use of four-stranded DNA, which has the connotation of four-stranded base pairing among homologous repeats. It needs to be emphasized that intercoil DNA is formed by two intertwining duplexes of any sequence, but when homology is met at a certain region, homologous recombination may proceed via heteroduplex formation with the help of the proper enzymes involved.

FBI DNA was first formulated based upon a rare but unique electron microscopy (EM) configuration of plant mitochondrial DNA (mtDNA) molecules and space-filling models [7, 8]. Once the FBI DNA structure was realized with the space filling model, it was possible to perform homologous recombination with direct contact of the repeat sequences. The intra-molecular homologous recombination of a deletion and inversion occurs via 1) synapsis of the repeats in the intercoil, 2) heteroduplex formation by 90° base flipping, and 3) nick-and-close resolution of the sugar phosphate backbone crossing over [7]. Inter-molecular homologous recombination, known as site-specific insertion, on the other hand, is mediated by direct contact between the attP site at the tip of the FBI DNA and the attB site on the target DNA [8]. Perpendicular approach of the foldback tip onto the target site effectively resolves the year-long topological puzzle encountered in four-stranded base pair winding of two circular DNA molecules and explains that 6- or 7-bp attP and attB sites correspond to the diameter of the approaching FBI DNA tip and the target width. DNA transposition is explainable by the perpendicular approach of the detached transposon ends in intercoil form to the host target, as described in detail below.
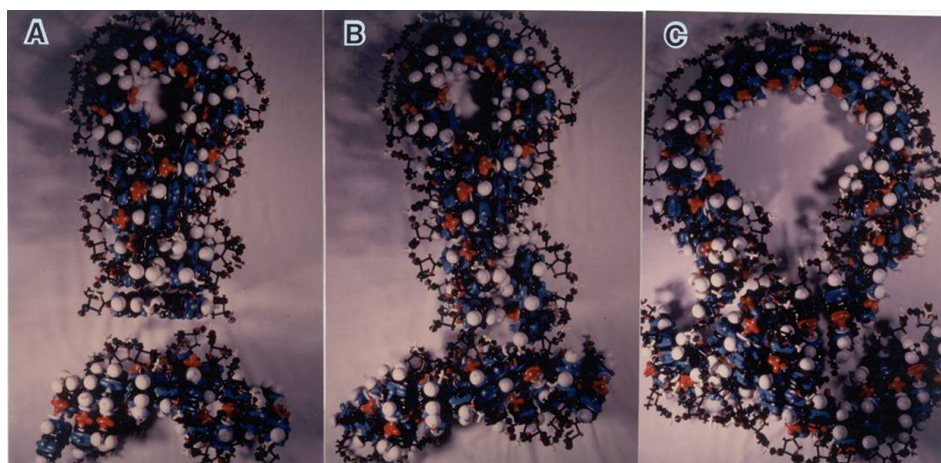

Fig. 1. (A–C) Space filling model of foldback intercoil DNA for DNA transposition (Adopted from Kim. *KAST Rev Mod Sci Technol* 2007;3:65-79 [8], with permission of Korean Academy of Science and Technology).

DNA Transposition Explained by FBI DNA

Transposition of a dTE molecule in FBI DNA configuration is tested by a space filling model, as presented in Fig. 1 and as a detailed diagram in Fig. 2 [30]. The TE molecule, detached from the host by a DSB in a blunt end cut, approaches, as shown in the diagram, at a perpendicular angle to the target site. Both 3'-OH ends of the TE ligate with the 5'-P ends of the staggered-cut target, leaving the overhang target ends single-stranded, which would be repaired by gap filling. As shown in the diagram (Fig. 2), when the TE ends approach the target on the minor groove side, the target size would be 5 bp on the minor groove side and 7 bp on the major groove side. This explains the meaning of the average size of the 5–7-bp target duplication, as it matches the diameter of the incoming intercoil DNA, 2.2 nm. In reality, the target sizes vary between 2 to even over 10 bp in a TE superfamily-specific manner. This TSD size variation could be explained by several factors—namely, local squeeze and stretch of the DNA double helix; unequal cutting positions of the TE ends, as exemplified in Tn7 by the 3-nt overhang by staggered DSB [14]; the characteristics of the endonucleases involved; and even temporary loosening of the intercoil end during a very elaborate and concerted coordination of the transpososome structure [8].

In earlier days, when IS and Tn, and even the LTR retrotransposons, were the major players of transposition, it was the presence of terminal repeats that attracted attention for explaining the transpositional mechanism, as in homo-logous recombination. Therefore, it was somewhat puzzling to explain the cases of non-LTR retrotransposons with asymmetric termini sequences. But, it is now TSD that attracts attention as a common factor for the transposition mechanism of all TEs, except for Helitrons. It is the structural nature of the intercoil, which is formed by intertwining duplexes, regardless of the homology between the two termini. When both ends of a TE synapse in the form of an intercoil, are cut by a DSB off the host duplex strand, and approach the target in perpendicular direction, their 3'-OH ends ligate to the staggered 5'-P ends of the broken target. The resulting overhang single-strand DNA gaps on both flanking ends of the target are repaired to generate the accompanying TSD. The significance of the terminal repeats is then the homologous recombination that is required for further transactions of recombination, replication, and transcription.

FBI DNA and Enhancer Concept for DNA Transposition

The enhancer is a cis element of DNA that affects, as a third factor, DNA transactions, such as transcription, homologous recombination, and transposition [31-34]. It has remained somewhat of an enigma, because the presence of an enhancer element and its binding proteins at a distance in a sequence and orientation non-specific manner enhances the efficiency of the reaction by 1,000 times.

Rather unique configurations of rigid rod-like DNA with a big protein complex at one end [35, 36 Fig. 4-2] may shed
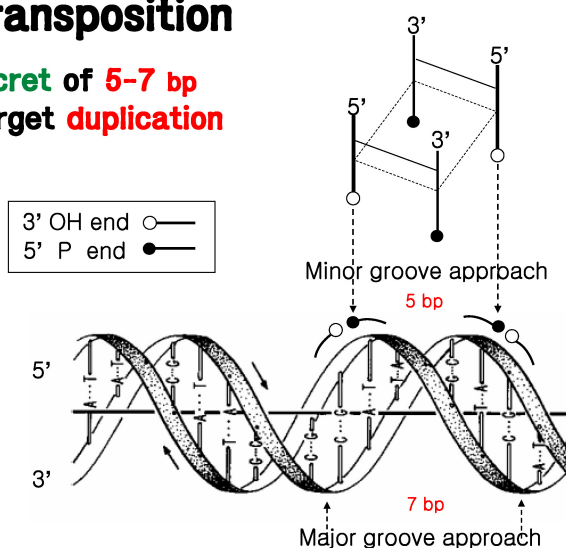

Fig. 2. Diagrammatic illustration of a perpendicular insertion of a transposon in the foldback intercoil DNA configuration (DNA double helix modified from Klug. *J Mol Biol* 2004;335:3-26 [30], with permission of Elsevier Ltd.).
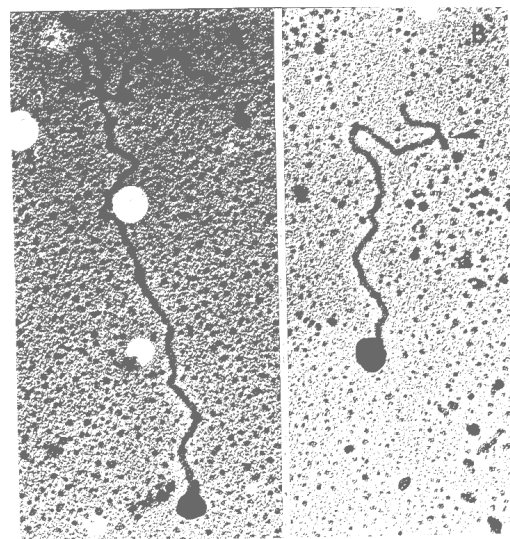

Fig. 3. Transmission electron microscopy pictures of rigid rod-like DNA that includes unique foldback intercoil DNA configurations (Adopted from Kim. *KAST Rev Mod Sci Technol* 2007;3:65-79 [8], with permission of Korean Academy of Science and Technology).
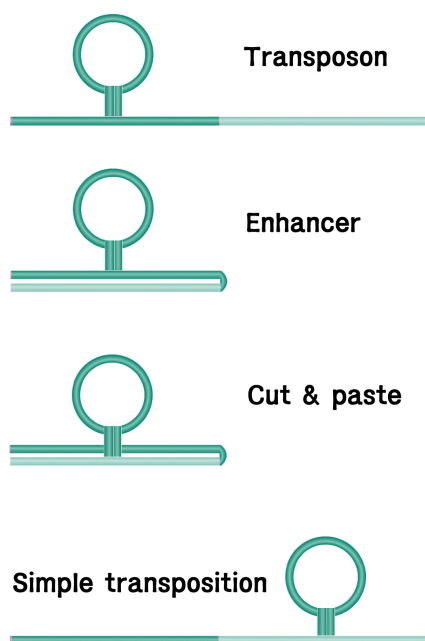
Fig. 4. Diagrammatic illustration of a simple conservative or cut-and-paste transposition in terms of foldback intercoil DNA and an enhancer concept (Adopted from Kim. *Foldback Intercoil DNA* 2008 [36], with permission of Free Academy Press).

Fig. 5. Diagrammatic illustration of a replicative transposition in terms of foldback intercoil DNA and an enhancer concept (Adopted from Kim. *Foldback Intercoil DNA* 2008 [36], with permission of Free Academy Press).
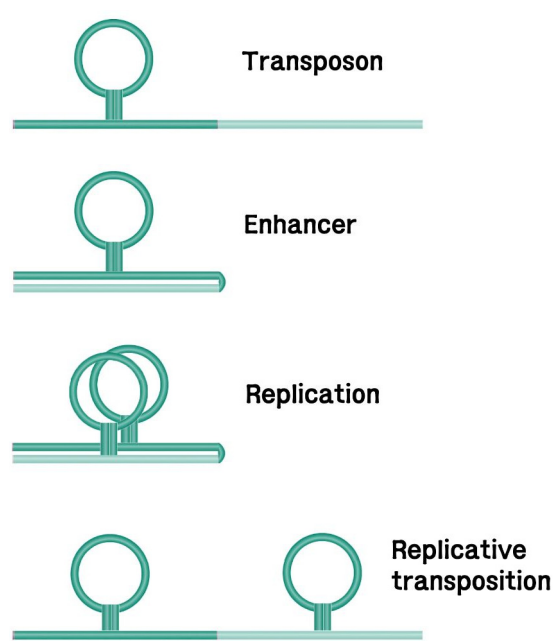
light on the enhancer concept and give us an idea of what actually happens under cellular conditions (Fig. 3). Folding back of a DNA duplex at a distant point would bring an enhancer site into juxtaposition with the base of a TE, which has also branched out as FBI DNA (Figs. 4 and 5). Proteins binding to the enhancer site and to both ends of the TE and any other additional factors together would form a tight DNA-protein complex, called a transpososome. This transpososome would provide a tight and stable surgical platform for a series of highly concerted precision cut-and- paste reactions: namely, DSB of the TE termini, sealing of the host duplexes by NHEJ, staggered cuts on the target site, NHEJ of the TE 3'-OH ends and target 5'-P ends, and finally, gap-filling of the 5–7-nucleotide single-strand portions to flank the transposed TE.

In replicative transposition (Fig. 5), there are two possibilities [21]. First, the gap-filling reaction during TSD may continue on to the TE molecule by displacement or polarized replication, even though the direction of replication may be reversible [20]. Second, as shown in the diagram, TIRs in the FBI DNA configuration may transform into heteroduplexes and serve as primers for bidirectional replication of the internal TE, thus keeping one copy at the original site and depositing the other copy to the target site in juxtaposition. Inversion of the replicated TE may result from homologous recombination, as predicted [7, 16].

Conclusions and Perspectives

TSD serves as a common landmark of unity on transposition mechanisms for both dTEs and rTEs. The only exception so far is the Helitron superfamily, which lacks TSD and employs a rolling circle replication model. As long as a TSD is generated during TE integration into the host chromosome, it indicates that the immediate intermediate of integration is in a double-stranded DNA form, regardless of RNA or single-stranded DNA being the initial intermediate.

From the EM pictures of FBI and rigid rod configurations of native DNA (Fig. 3), an ideal configuration of a transpososome may be envisioned. It becomes apparent that vertical insertion of the DSB ends of the TE in an intercoil alignment is probably the best-fitting mechanism of DNA transposition. Small sizes of the target, around 6–7 bp, encountered in TSD as well as in attP and attB in site-specific insertion all point to the fact that they correspond to the diameter of FBI DNA 2.2 nm.

The transpososome, made of FBI DNA configurations of the TE and the enhancer element, may effectively abolish lingering mechanistic puzzles about keeping four broken strands of the TE and the host from flying apart and answer how a TE molecule jumps onto a new target site at a distance. Replicative transposition is explained likewise, even utilizing the terminal homology as a possible origin of replication. It would be, however, difficult, if not impossible, to

reproduce such a transpososome complex *in vitro* by FBI DNA configuration, which is critically needed for EM observations and crystallography analysis.

There are four main modes of transactions by which the FBI DNA motif may be utilized in DNA functions in the cell: alpha deletion, omega inversion, needle point site-specific insertion, and NHEJ-GF integration of TE [8]. With such an abundant presence in strategically important positions in the genome and very dynamic functional roles, it is hoped that FBI DNA receives the proper recognition and interpretation for a better understanding of the life process.

Acknowledgments

References

1. McClintock B. The significance of responses of the genome to challenge. In: *Nobel Lectures Physiology or Medicine 1981-1990* (Frangsmyr T, Lindsten J, eds.). Singapore: World Scientific Pub. Co., 1993. pp. 180-199.
2. Watson JD, Crick FH. Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature* 1953;171:737-738.
3. Fedoroff N, Wessler S, Shure M. Isolation of the transposable maize controlling elements Ac and Ds. *Cell* 1983;35:235-242.
4. Craig NL, Craigie R, Gellert M, Lambowitz AM. *Mobile DNA II*. Washington, DC: ASM Press, 2002.
5. Kazazian HH Jr. Mobile elements: drivers of genome evolution. *Science* 2004;303:1626-1632.
6. Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, *et al*. A unified classification system for eukaryotic transposable elements. *Nat Rev Genet* 2007;8:973-982.
7. Kim BD. Four-stranded DNA: an intermediate of homologous recombination and transposition. *Korean J Breed* 1985;17:453-466.
8. Kim BD. The foldback intercoil DNA: its implication on DNA structures and functions. *KAST Rev Mod Sci Technol* 2007;3:65-79.
9. Jurka J, Kapitonov VV, Kohany O, Jurka MV. Repetitive sequences in complex genomes: structure and evolution. *Annu Rev Genomics Hum Genet* 2007;8:241-259.
10. Mills RE, Bennett EA, Iskow RC, Devine SE. Which transposable elements are active in the human genome? *Trends Genet* 2007;23:183-191.
11. Kim YJ, Lee J, Han K. Transposable elements: no more 'Junk DNA'. *Genomics Inform* 2012;10:226-233.
12. Lee SI, Kim NS. Transposable elements and genome size variations in plants. *Genomics Inform* 2014;12:87-97.
13. Steiniger-White M, Rayment I, Reznikoff WS. Structure/function insights into Tn5 transposition. *Curr Opin Struct Biol* 2004;14:50-57.
14. Bainton R, Gamas P, Craig NL. Tn7 transposition *in vitro* proceeds through an excised transposon intermediate generated by staggered breaks in DNA. *Cell* 1991;65:805-816.
15. Haniford DB, Chelouche AR, Kleckner N. A specific class of IS10 transposase mutants are blocked for target site interactions and promote formation of an excised transposon fragment. *Cell* 1989;59:385-394.
16. Shapiro JA. Molecular model for the transposition and replication of bacteriophage Mu and other transposable elements. *Proc Natl Acad Sci U S A* 1979;76:1933-1937.
17. Johnson RC, Reznikoff WS. DNA sequences at the ends of transposon Tn5 required for transposition. *Nature* 1983;304:280-282.
18. Parks AR, Li Z, Shi Q, Owens RM, Jin MM, Peters JE. Transposition into replicating DNA occurs through interaction with the processivity factor. *Cell* 2009;138:685-695.
19. Richardson JM, Colloms SD, Finnegan DJ, Walkinshaw MD. Molecular architecture of the Mos1 paired-end complex: the structural basis of DNA transposition in a eukaryote. *Cell* 2009;138:1096-1108.
20. Mizuuchi K. Transpositional recombination: mechanistic insights from studies of mu and other elements. *Annu Rev Biochem* 1992;61:1011-1051.
21. Craigie R, Mizuuchi K. Mechanism of transposition of bacteriophage Mu: structure of a transposition intermediate. *Cell* 1985;41:867-876.
22. Craig NL. Unity in transposition reactions. *Science* 1995;270:253-254.
23. Miyao A, Tanaka K, Murata K, Sawaki H, Takeda S, Abe K, *et al*. Target site specificity of the Tos17 retrotransposon shows a preference for insertion within genes and against insertion in retrotransposon-rich regions of the genome. *Plant Cell* 2003;15:1771-1780.
24. Luan DD, Korman MH, Jakubczak JL, Eickbush TH. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell* 1993;72:595-605.
25. Kulpa DA, Moran JV. Cis-preferential LINE-1 reverse transcriptase activity in ribonucleoprotein particles. *Nat Struct Mol Biol* 2006;13:655-660.
26. Grindley ND, Reed RR. Transpositional recombination in prokaryotes. *Annu Rev Biochem* 1985;54:863-896.
27. Kapitonov VV, Jurka J. Self-synthesizing DNA transposons in eukaryotes. *Proc Natl Acad Sci U S A* 2006;103:4540-4545.
28. Lal S, Oetjens M, Hannah LC. Helitrons: enigmatic abductors and mobilizers of host genome sequences. *Plant Sci* 2009:176;181-186.
29. Derbyshire KM, Grindley ND. Replicative and conservative transposition in bacteria. *Cell* 1986;47:325-327.
30. Klug A. The discovery of the DNA double helix. *J Mol Biol* 2004;335:3-26.
31. Craig NL. The mechanism of conservative site-specific recombination. *Annu Rev Genet* 1988;22:77-105.
32. Kahmann R, Rudt F, Koch C, Mertens G. G inversion in bacteriophage Mu DNA is stimulated by a site within the invertase gene and a host factor. *Cell* 1985;41:771-780.
33. Johnson RC, Simon MI. Hin-mediated site-specific recombi-

nation requires two 26 bp recombination sites and a 60 bp re-combinational enhancer. *Cell* 1985;41:781-791.

34. Huber HE, Iida S, Arber W, Bickle TA. Site-specific DNA inversion is enhanced by a DNA sequence element in cis. *Proc Natl Acad Sci U S A* 1985;82:3776-3780.

35. Kim BD, Lee KJ, DeBusk AG. Linear and 'lasso-like' structures of mitochondrial DNA from *Pennisetum typhoides*. *FEBS Lett* 1982;147:231-234.

36. Kim BD. *Foldback Intercoil DNA*. Seoul: Free Academy Press, 2008.