Sleeping Beauty Transposition

ZOLTÁN IVICS[1] and ZSUZSANNA IZSVÁK[2]

[1]Division of Medical Biotechnology, Paul Ehrlich Institute, Langen, Germany;
[2]Max Delbrück Center for Molecular Medicine, Berlin, Germany

ABSTRACT *Sleeping Beauty* (*SB*) is a synthetic transposon that was constructed based on sequences of transpositionally inactive elements isolated from fish genomes. *SB* is a Tc1/*mariner* superfamily transposon following a cut-and-paste transpositional reaction, during which the element-encoded transposase interacts with its binding sites in the terminal inverted repeats of the transposon, promotes the assembly of a synaptic complex, catalyzes excision of the element out of its donor site, and integrates the excised transposon into a new location in target DNA. *SB* transposition is dependent on cellular host factors. Transcriptional control of transposase expression is regulated by the HMG2L1 transcription factor. Synaptic complex assembly is promoted by the HMGB1 protein and regulated by chromatin structure. *SB* transposition is highly dependent on the nonhomologous end joining (NHEJ) pathway of double-strand DNA break repair that generates a transposon footprint at the excision site. Through its association with the Miz-1 transcription factor, the *SB* transposase downregulates cyclin D1 expression that results in a slowdown of the cell-cycle in the G1 phase, where NHEJ is preferentially active. Transposon integration occurs at TA dinucleotides in the target DNA, which are duplicated at the flanks of the integrated transposon. *SB* shows a random genome-wide insertion profile in mammalian cells when launched from episomal vectors and "local hopping" when launched from chromosomal donor sites. Some of the excised transposons undergo a self-destructive autointegration reaction, which can partially explain why longer elements transpose less efficiently. *SB* became an important molecular tool for transgenesis, insertional mutagenesis, and gene therapy.

Sleeping Beauty KISSED BACK TO LIFE – A SHORT HISTORY

Members of the Tc1/*mariner* superfamily are probably the most widespread DNA transposons in nature ([1]). However, these elements appear to be transpositionally inactive in vertebrates due to the accumulation of mutations. In an attempt to isolate potentially active copies, we surveyed a number of fish genomes for the presence of Tc1-like elements from 11 different species. In summary, all the Tc1-like elements that we ([2]) and others ([3], [4]) described from the different fish species were defective copies carrying inactivating mutations that accumulated over long evolutionary times. Nevertheless, careful sequence analysis allowed us to predict a consensus sequence that would likely represent an active archetypal sequence. We have engineered this sequence by eliminating the inactivating mutations from the transposase open reading frame. The resurrected synthetic transposon was named *Sleeping Beauty* (*SB*), in analogy of the Grimm brothers' famous fairy tale. *SB* can be identical or closely related to an ancient transposon that once successfully invaded several fish genomes, in part by horizontal transmission between species ([5]). The resurrection of *SB* was the first demonstration that ancient transposable elements can be brought back to life. Before this work was published in 1997, there was no indication that any DNA-based transposon was active in vertebrates. *SB* not only represents the first DNA-based transposon ever shown to be active in cells of vertebrates, but the first functional gene ever reconstructed from inactive, ancient genetic material, for which an active, naturally occurring copy either does not exist or has not yet been isolated.

STRUCTURAL AND FUNCTIONAL COMPONENTS OF THE *Sleeping Beauty* TRANSPOSON

The *SB* transposon has a simple structure. In its natural form, it consists of a single gene encoding the transposase polypeptide, the enzymatic factor of transposition, which is flanked by terminal inverted repeats (IRs) containing binding sites for the transposase [Fig. 1(A)] (5). The transposase gene can be physically separated from the IRs, and replaced by other DNA sequences [Fig. 1(B)]. This is because the transposase can mobilize transposons in *trans*, as long as they retain the IRs. *SB* transposes through a conservative, cut-and-paste mechanism, during which the transposable element is excised from its original location by the transposase, and is integrated into a new location.
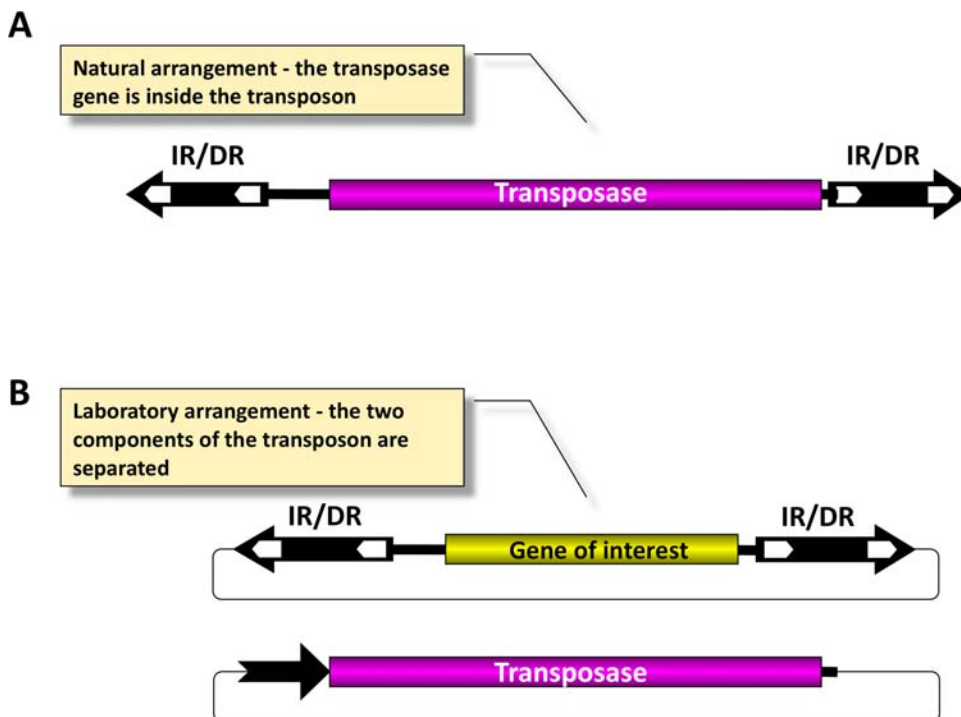
THE TRANSPOSASE

The DNA-binding domain

The overall domain structure of the transposase is conserved in the entire Tc1/*mariner* superfamily (1). Specific substrate recognition is mediated by an N-terminal, bipartite DNA-binding domain of the transposase (Fig. 2) (6, 7, 8). This DNA-binding domain has been proposed to consist of two helix-turn-helix (HTH) motifs, similar to the paired domain of some transcription factors in both amino acid sequence and structure (2, 8, 9). The modular paired domain has evolved versatility in binding to a range of different DNA sequences through various combinations of its subdomains (PAI+RED) (10). The origin of the paired domain is not clear, but phylogenetic analyses indicate that it might have been derived from an ancestral transposase (11).

The first of these HTH motifs has been crystallized in complex with double-stranded DNA corresponding to the termini of Tc3 transposons in *Caenorhabditis elegans* (12). The crystal structure indeed showed a HTH fold, and a dimer of transposase subunits bringing together the two DNA ends. The recently described NMR solution structure of the PAI subdomain of the *SB* transposase identified amino acid residues located in the second and third alpha helices forming the HTH motif to be involved in binding to DNA (Fig. 3) (13).

FIGURE 1 The *SB* transposon system. (A) Structure of the *SB* transposon. The central transposase gene (purple box) is flanked by terminal IRs (black arrows) that contain binding sites for the transposase (white arrows). (B) Gene transfer vector system based on *SB*. The transposase coding region can be replaced by a gene of interest (yellow box) within the transposable element. This transposon can be mobilized if a transposase source is provided in cells; for example, the transposase can be expressed from a separate plasmid vector containing a suitable promoter (black arrow). Reprinted from *Current Gene Therapy* (161) with permission from the publisher. doi:10.1128/microbiolspec.MDNA3-0042-2014.f1
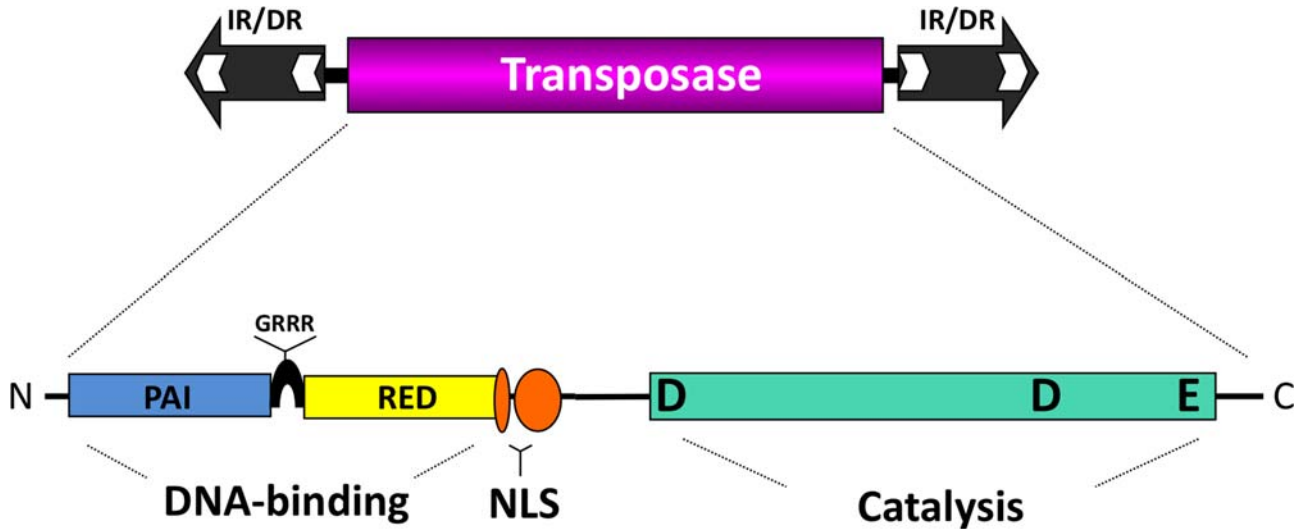
FIGURE 2 Structural and functional components of *SB*. On top, a schematic drawing of the transposon is shown. The terminal IR/DR (black arrows) contain two binding sites for the transposase (white arrows). The element contains a single gene encoding the transposase (purple box). The transposase has an N-terminal, bipartite, paired-like DNA-binding domain containing a GRRR AT-hook motif, an NLS, and a C-terminal catalytic domain. The DNA binding domain consists of a PAI and a RED subdomain containing helix-turn-helix DNA-binding motifs. The DDE amino acid triad is a characteristic signature of the catalytic domain that catalyzes the DNA cleavage and joining reactions. Reprinted from *Molecular Therapy* ([164](#)) with permission from the publisher. [doi:10.1128/microbiolspec.MDNA3-0042-2014.f2](#)

We found that a GRPR-like sequence (GRRR in *SB*) between the two HTH motifs is conserved in Tc1/*mariner* transposases ([Fig. 2](#)). The GRPR motif is similar to an AT-hook ([6](#)), and characteristic to homeodomain proteins ([14](#)). It mediates minor groove interactions with DNA in the case of the Hin invertase of *Salmonella* ([15](#)) and in the V(D)J recombination activating gene (RAG1) recombinase (see V(D)J chapter) ([16](#)).

Partially overlapping with the RED subdomain in the transposase is a bipartite nuclear localization signal (NLS) ([Fig. 2](#)), flanked by phosphorylation target sites of casein kinase II ([2](#)). Phosphorylation of these sites is a potential checkpoint in the regulation of transposition. The NLS indicates that these transposons, unlike murine retroviruses, can take advantage of the receptor-mediated transport machinery of host cells for nuclear uptake of their transposases.

The catalytic domain

The second major domain of the *SB* transposase has been referred to as the catalytic domain, because it is

FIGURE 3 Structures of the PAI subdomain of the *SB* transposase and the DNA-bound N-terminal DNA-binding subdomains of the Tc3 and Mos1 *mariner* transposases and the Pax5 transcription factor. Residues on the second and third alpha-helices of the *SB* PAI subdomain are directly involved in DNA-binding. Reprinted from *Protein Science* ([13](#)) with permission from the publisher. [doi:10.1128/microbiolspec.MDNA3-0042-2014.f3](#)
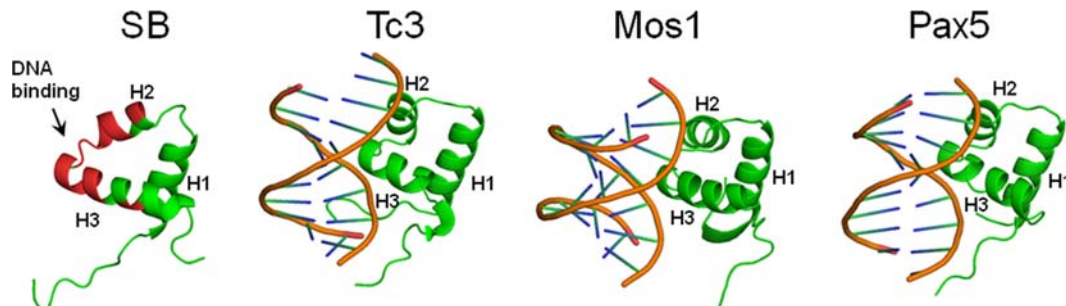
responsible for the DNA cleavage and joining reactions of transposition. The majority of known transposases, including *SB*, and retroviral integrases possess a well-conserved triad of amino acids, known as the aspartate–aspartate–glutamate, in short the DDE signature in their C-terminal catalytic domain ([Fig. 2](#)) ([17](#), [18](#), [19](#), [20](#)). These amino acids play an essential role in catalysis by coordinating, in general, two divalent cations necessary for activity. One metal ion acts as a Lewis acid, and stabilizes the transition state of the penta-coordinated phosphate, the other one acts as a general base and deprotonizes the incoming nucleophile during trans-esterification and strand transfer ([21](#)). The biologically relevant cation required for the catalytic steps of transposition is thought to be Mg^{++} ([22](#), [23](#)).

Site-directed mutagenesis of the DDE residues in the related Tc3 transposase confirmed that these three amino acids are indeed essential for all catalytic activities ([24](#)). Similarly, a DAE variant of the *SB* transposase is catalytically dead ([25](#)). Although the crystal structure of the *SB* transposase catalytic domain is yet to be solved, the available Mos1 *mariner* structure from *Drosophila mauritiana* ([26](#)) is assumed to be the closest to model the structure of the *SB* transposase. Intriguingly, crystallographic analyses of the catalytic domains of transposases and some other proteins whose functions are not obviously related to transposition, such as RNaseH ([27](#)) or RuvC ([28](#)), have revealed a remarkably similar overall fold. Besides similarities, there are important differences between *mariner* and *SB* transposases. First, while *SB* has the characteristic DDE motif in its catalytic domain, this motif in *mariner*s is DDD. Interestingly, a change of the exceptional third D of *mariner*, turning the DDD into the canonical DDE, inactivates the transposase ([29](#)), and the reciprocal substitution in the *SB* transposase (DDD) is also inactive ([30](#)). A conserved glycine-rich subdomain can be found within the catalytic domains of *SB* and Tc1-like transposases ([5](#)). This glycine-rich subdomain is not present in *mariner*s or retroviral integrases, and its function is yet to be determined. The emerging picture reinforces the notion of a common structural motif that catalyzes polynucleotidyl transfer reactions in diverse biological contexts ([31](#), [32](#)), and that the different specificities in binding to DNA might have evolved by the apparent acquisition of different DNA-binding domains in the evolution of DDE/D recombinases ([18](#)).

Molecular evolution of the SB100X hyperactive *Sleeping Beauty* transposase

In evolutionary terms, the *SB* transposon was a successful element with the ability to colonize several fish genomes millions of years ago ([2](#)). However, even successful transposons have not been selected for the highest possible activity. On the contrary, there is strong selective pressure to avoid insertional mutagenesis of essential genes of their host. In an attempt to derive hyperactive transposase variants for advanced genetic engineering, amino acid substitutions spanning almost the entire *SB* transposase polypeptide have been screened for eliciting a change in catalytic activity. These amino acid replacements were conducted either by systematic alanine-scanning ([33](#)), by "importing" single amino acids or small (2 to 7 aa) blocks of amino acids from related transposases ([34](#), [35](#), [36](#)), and by rational replacement of selected amino acid residues based on charge ([35](#)). These approaches generated transposase variants with (i) no change in activity; (ii) reduced activity; or (iii) a relatively modest increase of transposition activity. The vast majority of the mutations have a neutral or a negative effect on transposition activity. The inactivating mutations generally map to the evolutionary conserved domains of the transposase. Surprisingly, some combinations of hyperactive mutations were found to result in a significant reduction of activity. Nevertheless, a strategy of identifying those hyperactive variants that acted in an additive or synergistic manner combined with a high throughput genetic screening of ~2,000 possible combinations yielded *SB* transposase variants with significantly enhanced activities ([30](#)). The most hyperactive version, *SB*100X, displays a ~100-fold hyperactivity when compared to the originally resurrected transposase ([Fig. 4](#)) ([30](#)). The hyperactivity of *SB*100X cannot be explained by altered transposase stability, nor by increased binding to the transposon IRs; instead, the particular combination of mutations in *SB*100X appears to affect the folding properties of the transposase ([30](#)). The use of the *SB*100X system yielded robust gene transfer efficiencies into human hematopoietic progenitors ([30](#), [37](#)), mesenchymal stem cells, muscle stem/progenitor cells (myoblasts), iPSCs ([38](#)), and T cells ([39](#)). These cells are relevant targets for stem cell biology and for regenerative medicine and gene- and cell-based therapies of complex genetic diseases. Moreover, the *SB*100X transposase enables highly efficient germline transgenesis in relevant mammalian models, including mice, rats, rabbits, and pigs ([40](#), [41](#), [42](#)).

The inverted repeats of the *Sleeping Beauty* transposon

Similarly to most transposon ends, the IRs of *SB* are composed of two functional parts. The 2 to 3 terminal base pairs of the ends are the recombinationally
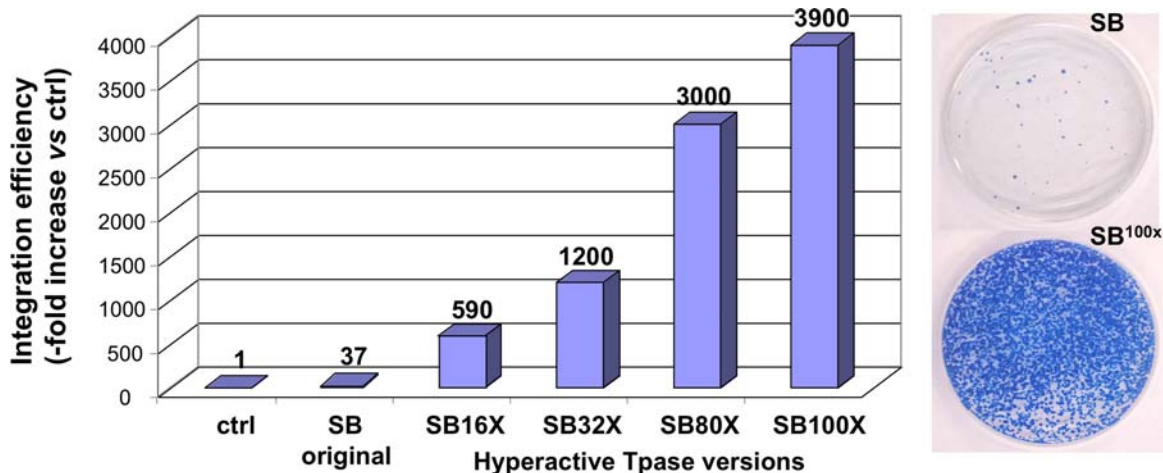
FIGURE 4 Comparison of different hyperactive versions of the *SB* transposase in transfected human HeLa cells. The chart shows the respective potential of transposase mutants to generate antibiotic-resistant cell colonies in human cell culture. The Petri dishes on the right show stained, antibiotic-resistant cell colonies obtained with the original *SB* transposase and with the *SB*100X hyperactive variant. doi:10.1128/microbiolspec.MDNA3 -0042-2014.f4

active sequences involved in the cleavage and the strand transfer reactions. The other functional part is situated within the IRs and it ensures the sequence-specific positioning of the transposase on the transposon ends.

The IRs of *SB* are 200 to 250 bp long and carry a pair of transposase-binding sites within the ends of each IR, characterized by short, 15 to 20 bp direct repeats (DRs) (Fig. 1). This special organization of IRs is termed IR/DR (1, 43), and can be found in numerous elements in the Tc1 transposon family, including the *Minos*, *S*, *Paris* and *Bari* elements in various *Drosophila* species (1, 44, 45, 46), *Quetzal* elements in mosquitos (47), at least three Tc1-like transposon subfamilies in fish (2), and *Txr*, *Eagle*, *Froggy*, and *Jumpy* transposons in *Xenopus* (48, 49). The spacing of about 200 bp between the outer and inner DRs is conserved in all elements within the IR/DR group, but the actual DNA sequences are not similar, suggesting convergent evolution of the IR/DR-type repeats. The IR/DR group differs significantly from Tc1 or the *mariner* elements, which are simpler and have repeats of less than 100 bp and a single transposase binding site per repeat (8, 50). The four DRs of *SB* are not identical; the outer ones are longer by two base pairs. The IRs are not identical either; the left IR contains a sequence motif called "half-DR" (HDR), which resembles the 3′-half of the transposase binding sites (6). A construct containing two left IRs transposes more efficiently than the wild-type transposon, but another version that has two right IRs has very poor mobility, indicating that the left and

right IRs are functionally distinct (6). The multiple binding sites of the IR/DR elements likely impose control over the timing and specificity of the transposition reaction (see below).

THE MOLECULAR MECHANISM OF *Sleeping Beauty* TRANSPOSITION AND ITS REGULATION BY HOST-ENCODED FACTORS

The typical "cut-and-paste" transposition process of *SB* can be divided into at least four major steps: (1) binding of the transposase to its sites within the transposon IRs; (2) formation of a synaptic complex in which the two ends of the elements are paired and held together by transposase subunits; (3) excision from the donor site; (4) reintegration at a target site (Fig. 5).

The wide phylogenetic distribution of the Tc1/*mariner* family suggested no or weak host factor requirement of the transposition reaction. Supporting this assumption, the activity of Tc1 was reconstructed *in vitro*, and the reaction required only the presence of the transposon and the transposase (51). In addition, the nematode Tc3 element was demonstrated to jump in zebrafish; however, this *trans*-species transposition reaction was not efficient (52). Despite earlier assumptions, *SB* transposition turned out to be highly dependent on cellular host factors (described in detail in sections below) (53, 54, 55, 56) and became an excellent model system to study transposon–host interactions in higher eukaryotes (57).
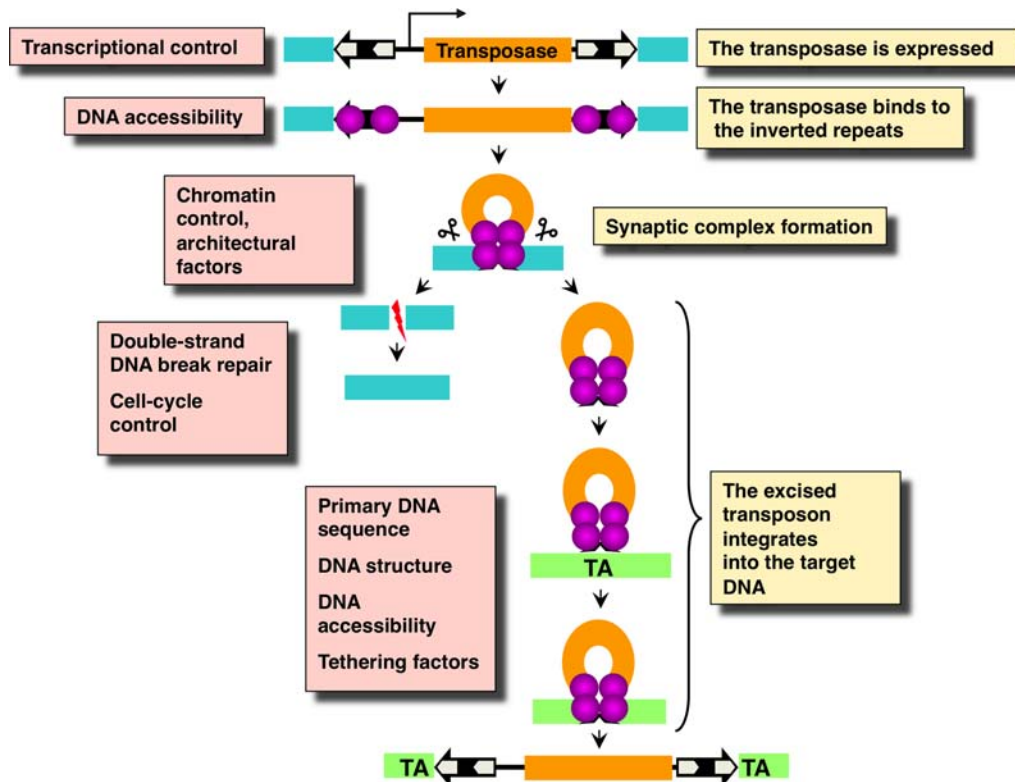
FIGURE 5 Mechanism and regulation of *SB* transposition. The transposable element consists of a gene encoding a transposase (orange box) bracketed by terminal IRs (solid black arrows) that contain binding sites of the transposase (white arrows) and flanking donor DNA (blue boxes). Transcriptional control elements in the 5′-UTR of the transposon drive transcription (arrow) of the transposase gene. The transposase (purple spheres) binds to its sites within the transposon IRs. Excision takes place in a synaptic complex, and separates the transposon from the donor DNA. The excised element integrates into a TA site in the target DNA (green box) that is duplicated and flanks the newly integrated transposon. On the right, the various steps of transposition are shown. On the left, mechanisms and host factors regulating each step of the transposition reaction are indicated. Reprinted from (57) with permission from the publisher. doi:10.1128/microbiolspec.MDNA3-0042 -2014.f5

SB can transpose in a wide range of vertebrate cells from fish to human, although the efficiency of transposition varies significantly (58), suggesting that differential interactions between the transposon and host-encoded factors may affect activity and eventually limit the host range. Indeed, transposition of *SB* seems to be restricted to vertebrates, with the exception of a chordate, *Ciona intestinalis* (59). Importantly, the identified host factors of *SB* are evolutionarily conserved in vertebrates and support *SB* transposition from fish to human. In summary, the regulation of *SB* transposition is mediated both by transposon- and host-encoded factors. Thus, the *SB* transposon has an intimate relationship with the host that likely modulates transposition at every step of the transposition reaction (Fig. 5).

Transcriptional regulation of the *Sleeping Beauty* transposon

Some of the 5′-untranslated regions (UTRs) upstream of the initiation codon of the transposase gene contain promoter motifs (60), suggesting that they might have functions associated with control of transposition activity. However, previous studies did not reveal an internal promoter in the Tc1 element; instead they showed that the elements are transcribed by read-through transcription from *C. elegans* genes (61). The left IR of *SB* is separated from the transposase coding sequence by a 160-bp stretch of DNA [Fig. 6(A)] with no apparent function in the transposition reaction (35). As measured by transient luciferase reporter assays, transcription driven by the 5′-UTR of *SB* is ~18-fold higher
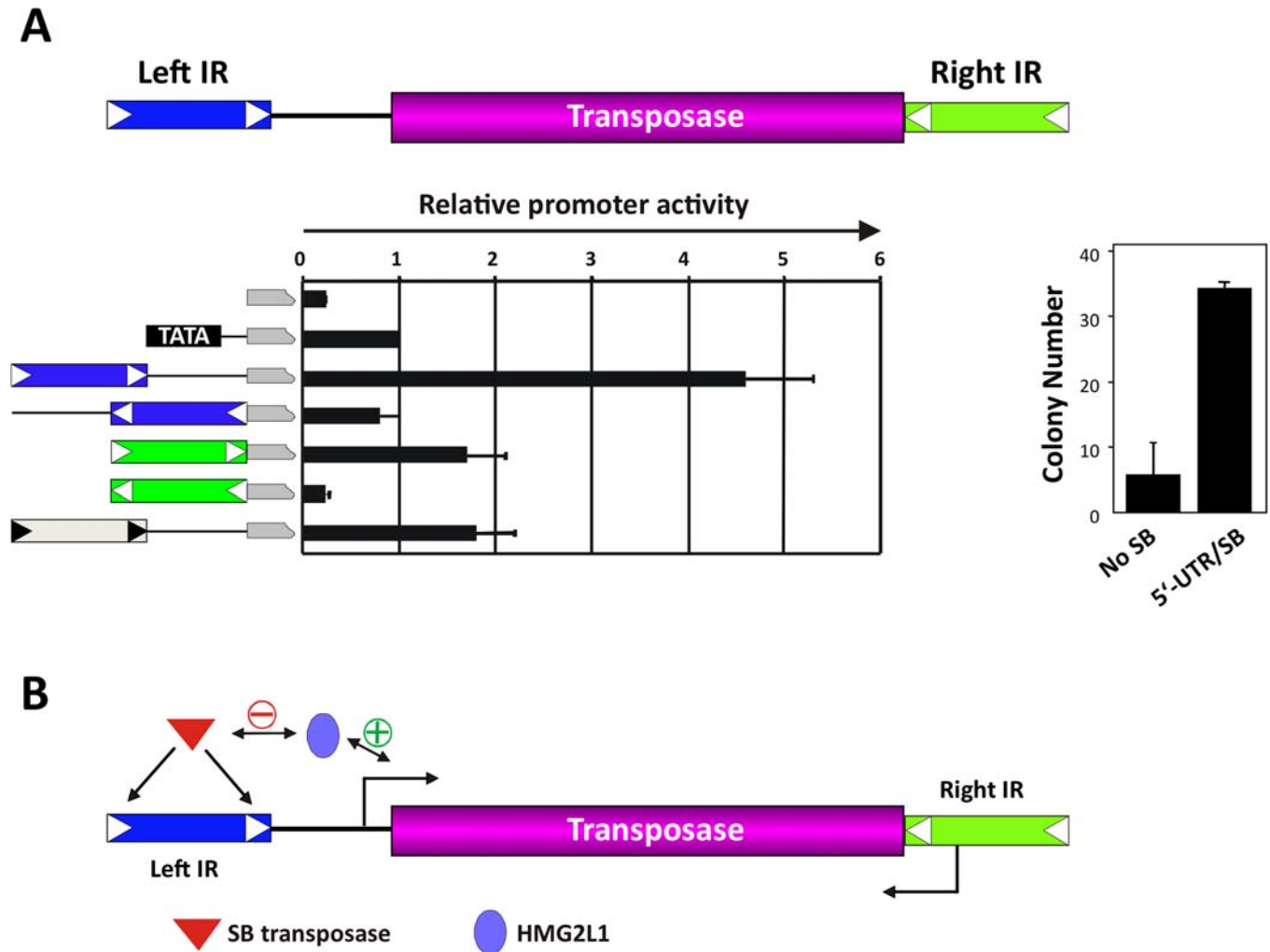
FIGURE 6 The UTRs of the *SB* transposon exhibit moderate, directional promoter activities. (A) Transcriptional activities residing within the *SB* transposon. On top, a schematic drawing of the transposon is shown. The terminal IRs contain two binding sites for the transposase (white arrows). The element contains a single gene encoding the transposase (purple box). Relative promoter activities as determined by transient luciferase assays in HeLa cells. Activity of a minimal promoter (TATA-box) control was arbitrarily set to value 1. Transposon sequences flanking the transposase gene were placed in front of a luciferase reporter gene in two possible orientations (in the case of the 5′-UTR, the luciferase gene precisely replaces the transposase coding region). Blue box: left IR of *SB*; green box: right IR of *SB*; beige box: left IR of *Frog Prince*; black lines connecting the IRs and the luciferase gene represent transposon sequences directly upstream of the transposase coding regions. The 5′-UTR of *SB* can drive transposase expression at a level sufficient for the detection of chromosomal transposition events in cultured cells. A neo-tagged *SB* transposon plasmid was cotransfected together with an *SB* expression construct, in which the transposase is expressed from the 5′-UTR of the transposon or with an empty cloning vector. The difference in numbers of G418-resistant cell colonies is evidence for transposition. (B) A model for transcriptional regulation of the *SB* transposase gene. In the wild-type, natural transposon, the central transposase gene (purple box) is flanked by UTRs that include the left and right inverted repeats (IRs, blue and green boxes, respectively) that contain binding sites for the transposase (white arrows). Arrows indicate the direction of transcription that is initiated within the UTRs. HMG2L1 upregulates, whereas *SB* transposase downregulates transcription from the 5′-UTR. Reprinted from *Molecular Therapy* ([54](#)) with permission from the publisher. [doi:10.1128/microbiolspec.MDNA3-0042-2014.f6](#)

than transcription of a promoter-less sequence, ~4.6-fold higher than transcription driven by a TATA-box minimal promoter, and ~2.5-fold higher than transcription driven by the 5′-UTR of the closely related *Frog Prince* (*FP*) transposon [Fig. 6(A)] (54). The 5′-UTR drives expression of the *SB* transposase at a level sufficient to detect *SB* transposition in a colony-forming transposition assay in HeLa cells [Fig. 6(A)]. The right IR can also drive expression towards the inside of the element, but at a lower efficiency than the 5′-UTR [Fig. 6 (A)]. Convergent transcription of *SB* transposons raises the possibility of the formation of transposon-specific double-stranded RNA molecules that may serve as triggers of transposon regulation by RNA interference (54, 62). Indeed, transgene expression from genomically integrated *SB* copies was enhanced in the presence of the p19 protein (derived from the tomato bushy stunt viruses) that suppresses RNA interference (63).

A cellular host factor, HMG2L1 (alias HMGXB4), was identified as a physical interacting partner of the *SB* transposase and shown to upregulate transcription from the 5′-UTR of *SB* by 10- to 15-fold (54) [Fig. 6(B)]. HMG2L1 is an HMG-box DNA-binding domain-containing protein that shares structural similarity with lymphocyte enhancer-binding factor 1 (LEF-1), sex-determining region Y (SRY), and SRY-related HMG-box protein 4 (SOX4) transcription factors (64). HMG-box transcription factors specifically bind their target DNA through their HMG-box domains, and regulate transcription of target genes (64). Indeed, in addition to its interaction with the transposase, HMG2L1 also interacts with the *SB* transposon DNA *in vivo*, as shown by chromatin immunoprecipitation (ChIP) experiments (54). Interestingly, co-expression of the *SB* transposase with HMG2L1 has a repressing effect on transcription by the 5′-UTR [Fig. 6 (B)]. Thus, transposase expression in the context of the naturally occurring *SB* transposon is subject to negative feedback regulation, with the transposase acting as a transcriptional repressor. This model postulates a sensitive balance in the regulation of transposase expression that is calibrated by transposase concentrations in the cell: low concentrations allow more transposase to be made, whereas high concentrations lead to shutting off transposase expression.

Specific DNA-binding by the *Sleeping Beauty* transposase

Similar to the DNA-binding domain of the transposase, the binding sites also have a bipartite structure in which the 3′-part of the binding site is recognized by the PAI subdomain, whereas the 5′-sequences interact with the RED subdomain of the transposase (6). Specificity of DNA-binding is predominantly determined by base-specific interactions mediated by the PAI subdomain (6).

All four binding sites within the IR/DR structure are required for *SB* transposition (58). The paired-like DNA-binding domain forms tetramers in complex with transposase binding sites (6). The inner DRs are more strongly bound by the transposase than the outer DRs (53, 65). Recent NMR data confirmed differential binding of the DNA-binding domain of the *SB* transposase to the inner and outer DRs (13). The PAI subdomain also binds to the HDR motif within the left IR of *SB* and mediates protein–protein interactions with other transposase subunits (6). Thus, the PAI subdomain is proposed to have at least three distinct functions: interaction with both the DRs and the HDR motif, and transposase oligomerization. In cooperation with the main DNA-binding domain, the GRRR motif was shown to function as an AT-hook, contributing to specific substrate recognition (6). Although part of the NLS is included in the RED subdomain, it does not appear to contribute to DNA recognition. Domain swapping experiments have shown that primary DNA-binding is not sufficient to determine specificity of the transposition reaction (6). Zebrafish *Tdr1* elements are closely related to *SB*, but are not mobilized by the *SB* transposase. Comparison of the transposase binding site sequences of *SB* and *Tdr1* elements revealed main differences in the 5′-half of the DRs. This sequence is contacted by the RED subdomain, indicating that the function of the RED is to enforce specificity at a later step in transposition. Substrate recognition of *SB* transposase is therefore sufficiently specific to prevent activation of transposons of closely related subfamilies.

Ordered assembly of synaptic complexes and the role of HMGB1 in *Sleeping Beauty* transposition

A uniform requirement among transposition reactions is the formation of a nucleoprotein complex, before transposon excision can take place. This very early step, synaptic complex assembly, is the process by which the two ends of the elements are paired and held together by transposase subunits (Fig. 5). The necessary factors that are required for synaptic complex assembly of *SB* include the complete IRs with four transposase binding sites, the HDR motif, and tetramerization-competent transposase. These tetrameric complexes form only if all the four binding sites are present and they are in the

proper context. The HDR motif is important but not essential in transposition and therefore can be viewed as a transpositional enhancer that, together with the PAI subdomain of the transposase, stabilizes complexes formed by a transposase tetramer bound at the IR/DR. In contrast to the *Mu* transposase, where the two specificities of binding to the enhancer and to the recombination sites are encoded in two distinct domains (66), the paired-like region of *SB* transposase combines these two functions in a single protein domain. As described above, the *SB* transposase preferentially binds the inner DRs within the transposon IRs. This suggests that the unequal strengths of transposase binding and the positions of the DRs within the IRs are required for ordered assembly of transposase-DNA complexes at the ends of the transposon that has a fundamental effect on the outcome of the transposition reaction.

The high mobility group protein HMGB1 is required for efficient *SB* transposition in mammalian cells (53). HMG family members have been shown to also be required for V(D)J recombination (67, 68). HMGB1 is an abundant, nonhistone, nuclear protein associated with eukaryotic chromatin and has the ability to bend DNA (69). *SB* transposition is significantly reduced in HMGB1-deficient mouse cells (53). This effect was complemented by expressing HMGB1 and HMGB2, but not with the more distantly related HMGA1 protein. Overexpression of HMGB1 in wild-type cells enhanced transposition, indicating that HMGB1 is a limiting factor of transposition (53). HMGs have low affinity for standard B-form DNA, and interactor proteins need to guide them to certain sites (69). *SB* transposase was found to interact with HMGB1 *in vivo* and to form a ternary complex with the transposase and transposon DNA (Fig. 7) (53), suggesting that the transposase may actively recruit HMGB1 to transposon DNA via protein–protein interactions. Considering the significant drop of transposition activity in HMGB1-deficient cells, the role of HMGB1 in transposition is a critical one.

HMGB1 was proposed to promote communication between DNA motifs within the transposon that are otherwise distant to each other, including the DRs, the transpositional enhancer, and the two IRs (53). However, as mentioned above, physical proximity of the DRs is not sufficient for *SB* transposition; a highly specific configuration of functional DNA elements within the IRs has a critical importance. It was also found that HMGB1 enhances transposase binding to both DRs, but its effect is significantly more pronounced at the inner sites (53). It appears, therefore, that the order of events that take place during the very early steps of transposition is binding
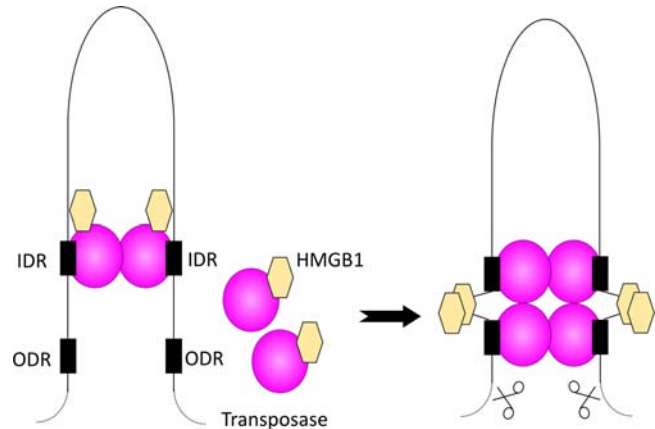

FIGURE 7 A model for the role of HMGB1 in *SB* synaptic complex formation. *SB* transposase (pink spheres) recruits HMGB1 (dotted hexagons) to the transposon IRs. First, HMGB1 stimulates specific binding of the transposase to the inner binding sites (IDRs). Once in contact with DNA, HMGB1 bends the spacer regions between the DRs, thereby assuring correct positioning of the outer sites (ODRs) for binding by the transposase. Cleavage (scissors) proceeds only if complex formation is complete. The complex includes the four binding sites (black boxes) and a tetramer of the transposase. Reprinted from *Nucleic Acids Res* (53) with permission from the publisher. doi:10.1128/microbiolspec.MDNA3-0042-2014.f7

of transposase molecules first to the inner sites and then to the outer sites. The pronounced effect of HMGB1 on binding of the transposase to the inner sites suggests that HMGB1 enforces ordered assembly of a catalytically active synaptic complex (Fig. 7). Indeed, interference with this sequence of events by replacing the outer transposase binding sites with the inner sites abolishes *SB* transposition (65). This ordered assembly process probably controls that cleavage at the outer sites occurs only if all the previous requirements have been fulfilled.

In summary, the IR/DR-type organization of IRs introduces a higher-level regulation into the transposition process. These elements might have evolved novel "built in regulatory checkpoints" to enforce synapsis prior to catalysis, thereby ensuring a higher level of accuracy and fidelity during the transposition process compared to transposons with simply structured IRs (53, 56, 70). The repeated transposase binding sites, their dissimilar affinity for the transposase, and the effect of HMGB1 to differentially enhance transposase binding to the inner sites are all important for a geometrically and timely orchestrated formation of synaptic complexes, which is a strict requirement for the subsequent catalytic steps of transposition. Such strictly regulated assembly of catalytically primed complexes could suppress unpaired reaction products or promiscuous synapses of distant ends of the transposon (70).

Regulation of transposon excision by DNA CpG methylation

CpG methylation of chromosomal DNA leads to formation of heterochromatin and is known to decrease or inhibit transpositional activity of diverse transposons ([71](#)). Surprisingly, CpG methylation of the *SB* transposon was found to enhance transpositional activity in mouse embryonic stem (ES) cells ([72](#)). It was subsequently found that the enhancing effect of CpG methylation is not restricted to *SB* but, rather, seems to be an intrinsic feature associated with the characteristic IR/DR structure of the *SB*, *Frog Prince*, and *Minos* elements ([73](#)). At which step(s) of cut-and-paste transposition does the effect of CpG methylation manifest? Because (i) CpG methylation has no effect on *SB* transposase binding, (ii) CpG methylation induces the formation of a condensed chromatin structure (heterochromatin), and (iii) DNA compaction by protamine enhances transposition, a model was proposed in which CpG methylation and subsequent chromatin condensation aids synaptic complex formation. Because heterochromatin formation results in tight packaging of DNA and histones, DNA sites that are usually far away from each other – for example, the two transposase binding sites in the IRs – might be brought closer together ([Fig. 8](#)). The physical proximity of the inner and outer binding sites might assist the formation of transposase dimers as soon as they bind, thereby facilitating the formation of a catalytically active synaptic complex. Thus, similarly to the effect of HMGB1, conformational changes of the excising transposon may greatly influence

FIGURE 8 A model for the enhancing effect of a compact chromatin structure on *SB* transposition. Euchromatin contains DNA wrapped around nucleosomes in a "beads-along-a-string"-like conformation (upper panel). Transposase subunits bound within the transposon IRs are separated by 166 bp DNA. Heterochromatin (lower panel), characterized by DNA CpG methylation and specific histone tail modifications, e.g., trimethylated lysine 9 of histone H3, features a higher histone : DNA ratio. Positioning of a nucleosome between the transposase binding sites (small orange arrows) shortens the distance between these sites, thereby facilitating the formation of a transposase dimer per IR and subsequent assembly of the synaptic complex. Reprinted from *Mobile DNA* ([73](#)) with permission from the publisher. doi:10.1128/microbiolspec.MDNA3-0042-2014.f8
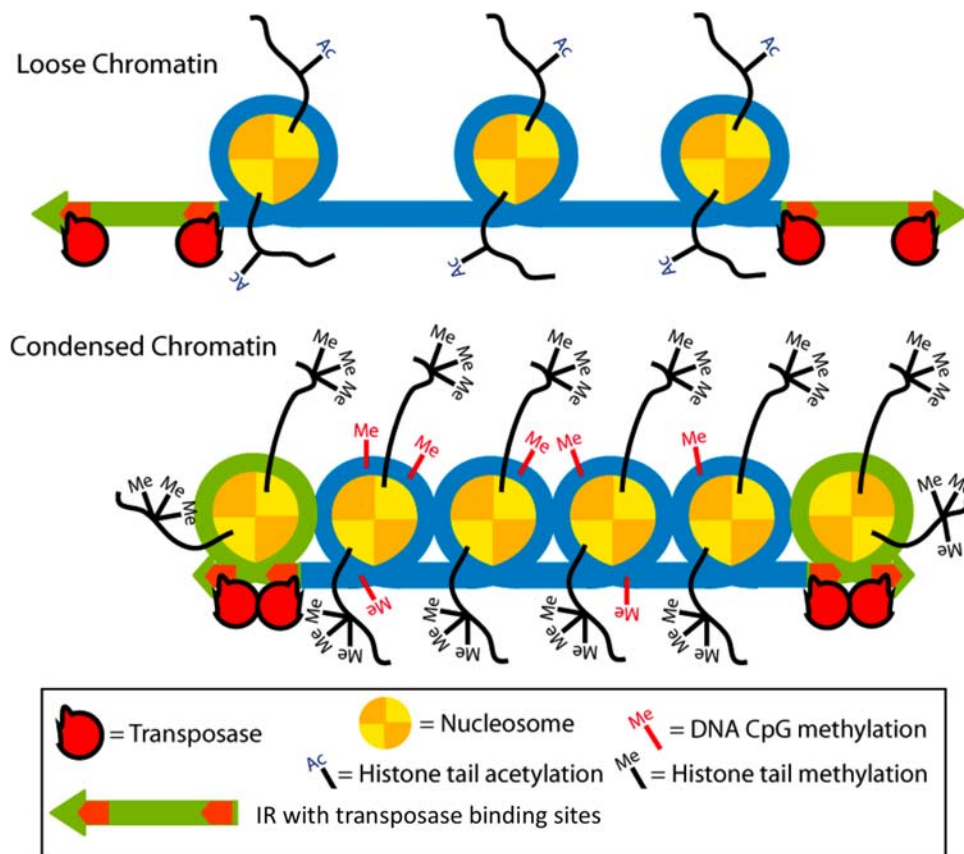
the efficiency of transposition. Assuming that the transposase source is provided by a transcriptionally active element located in euchromatin, host-cell-induced CpG methylation/heterochromatin-based silencing of transposons can be offset by a higher transposition efficiency out of condensed chromatin, thereby constituting a potential mechanism of *SB* and other similar-structured transposons to escape CpG methylation-mediated silencing imposed by the host.

Excision of the *Sleeping Beauty* transposon and double-strand break repair of excision sites

The key process of all transposon excision is the exposure of the 3′-OH groups of the transposon ends, which will later be used at the strand transfer reaction for integration (74). Every DNA strand cleavage in all transposition reactions is a transposase-catalyzed, Mg^{++}-dependent hydrolysis of the phosphodiester bonds of the DNA backbone, executed by a nucleophilic molecule. All the DDE recombinases catalyze similar chemical reactions (75), which begin with a single-strand nick that generates a free 3′-OH group. In the case of the first strand cleavage the nucleophile is H_2O (74). During cut-and-paste transposition, nicking of the element is followed by the cleavage of the complementary DNA strand. To catalyze second strand cleavage, DDE recombinases developed versatile strategies (76). Most DDE transposases use a single active site to cleave both strands of DNA at the transposon end via a DNA–hairpin intermediate (77, 78, 79, 80, 81). For example, in V(D)J recombination the single-strand nick is converted into a double-stranded break (DSB) by a transesterification reaction in which the free 3′-OH attacks the opposite strand, thereby creating a hairpin intermediate at the donor site (82, 83). Tn5 and Tn10 transposons also transpose *via* a hairpin intermediate, with the difference that the hairpin is on the transposon and not on the flanking DNA (84, 85). However, *mariner* (86) and *SB* (55) transposition does not proceed through a hairpin intermediate, and the exact mechanism of second-strand cleavage remains unknown. Thus, in the absence of a hairpin intermediate, *mariner* and *SB* transposases likely cleave the two strands of the DNA at each transposon end by sequential hydrolysis reactions.

Strand cleavage can occur at different positions relative to the transposon ends. The position of 5′-cleavage of the second strand required for the liberation of the element occurs directly opposite the 3′-cleavage site in V(D)J recombination (82) and for the bacterial Tn5

(87) and Tn10 elements (85) (thereby generating blunt ended products). In the case of the Tc1/*mariner* elements the non-transferred strand is cleaved a few nucleotides within the transposon [two nucleotides for the Tc1 and Tc3 elements (24, 50) and three nucleotides inside the element in the case of *mariner* and *SB* (Fig. 9) (50, 88)]. Thus, *SB* transposon excision leaves behind three-nucleotide-long 3′-overhangs (Fig. 9), which are processed by the DNA repair mechanisms of the cells leaving a transposon "footprint" at the transposon donor site (see below).

The DSBs generated by transposon excision are repaired either by the nonhomologous end joining pathway (NHEJ) or by homologous recombination (HR) (89, 90). The main factors that mediate NHEJ are a complex of DNA ligase IV and Xrcc4 and the DNA-dependent protein kinase (DNA-PK), a serine/threonine protein kinase (91, 92). DNA-PK consists of a catalytic subunit (DNA-PKcs) and a DNA-binding subunit termed Ku. The Ku heterodimer (composed of Ku70 and Ku80) binds to DNA ends and facilitates DSB repair by recruiting DNA-PKcs and additional factors such as the Xrcc4/DNA ligase IV complex to the site of damage.

The prominent pathway of repairing transposon excision sites in somatic mammalian cells is NHEJ, which generates transposon "footprints" that are identical to the first or last 2 to 4 nucleotides of the transposon in Tc1/*mariner* transposition (88, 93). NHEJ of the three-nucleotide-long 3′-overhangs left behind by *SB* excision generates a 3-bp footprint (Fig. 9) (55, 93). Factors of the NHEJ pathway of DSB repair, including Ku70 and DNA-PKcs, are required for *SB* transposition by repairing the transposon excision sites (55). NHEJ components have also been shown to be required for efficient retroelement integration and V(D)J recombination (92, 94). Ku70 physically interacts with the *SB* transposase (55), suggesting that it might be involved in shepherding excision site repair to NHEJ.

NHEJ and HR have overlapping roles in maintaining chromosomal integrity in vertebrate cells (95), and they can serve as alternative pathways for repair of the same DSB (96). Although the NHEJ pathway of DSB repair plays a dominant role in repair of transposon excision sites in somatic cells, the dependence of *SB* transposition on NHEJ factors is not absolute. In contrast to V(D)J recombination (92), HR can also be involved in excision site repair during *SB* transposition (55). Similarly, both NHEJ and HR play significant roles in the repair of DSBs generated by Tc1 excision in *C. elegans* (97), in P-element transposition in *Drosophila* (98, 99), and in Ty1 retrotransposon integration in yeast (100,
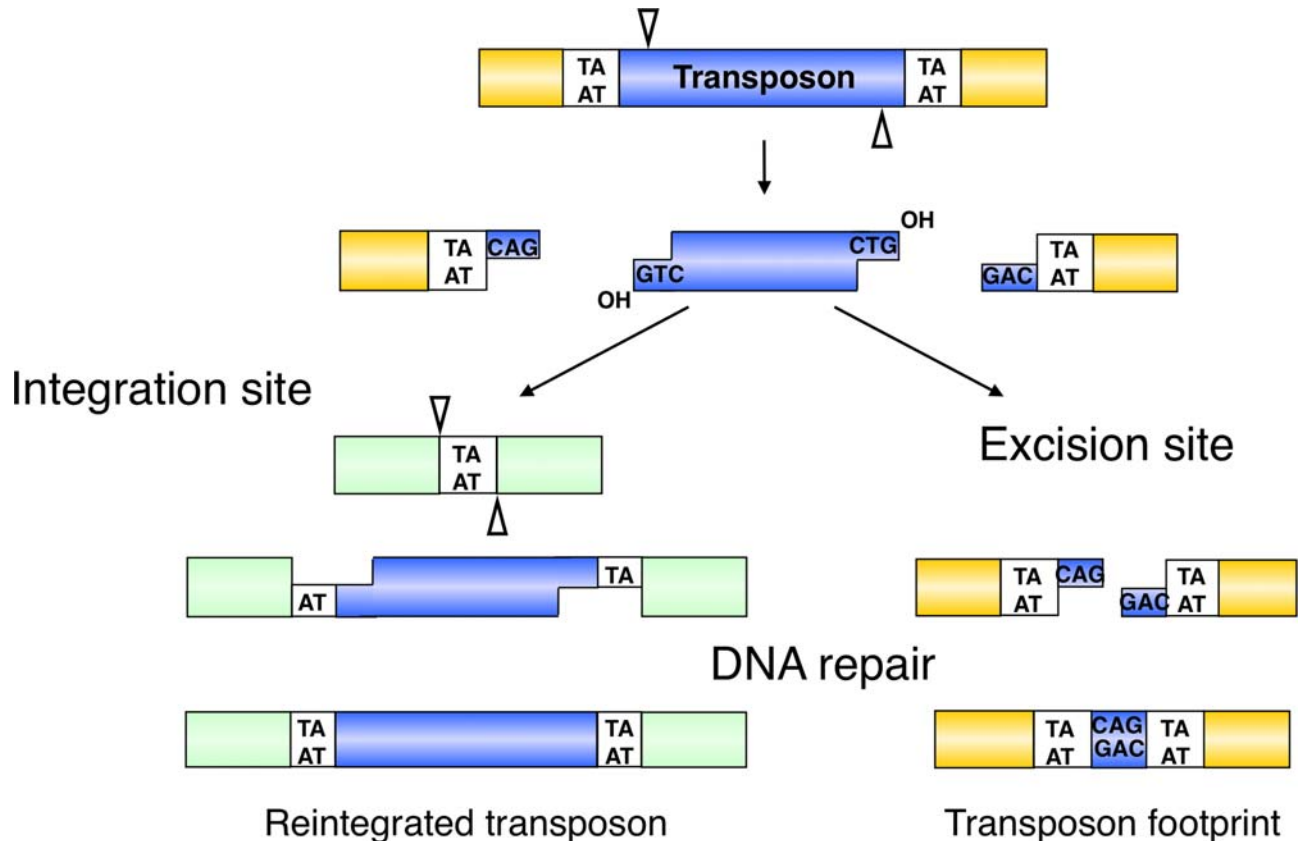
FIGURE 9 Molecular events during cut-and-paste transposition. The transposase initiates the excision of the transposon with staggered cuts and reintegrates it at a TA target dinucleotide. The single-stranded gaps at the integration site as well as the double-strand DNA breaks in the donor DNA are repaired by the host DNA repair machinery. After repair, the target TA is duplicated at the integration site, and a small footprint is left behind at the site of excision. The footprint is generated by the NHEJ pathway of DSB repair, and the central A:A mismatch is likely repaired by the mismatch repair system of the cell. Reprinted from *CMLS* (165) with permission from the publisher. doi:10.1128/microbiolspec.MDNA3-0042-2014.f9

101). These observations support the view that DSBs generated by radiation, V(D)J recombination, retroviral integration, and DNA transposition require overlapping but different factors for repair. The interplay between the repair factors and the recombination machineries probably determines how mechanistically similar processes can produce different products.

Sleeping Beauty transposase modulates cell-cycle progression through interaction with Miz-1

The Myc-interacting zinc finger protein 1 (Miz-1) transcription factor (102) was identified as an interactor of the *SB* transposase in a yeast two-hybrid screen (103). Miz-1 is a transcriptional regulator of the cyclin D1 gene (102), and it downregulates the cyclin D1 promoter

resulting in slower cell growth. Decreased cellular levels of cyclin D1 prevent cells from entering the S phase, resulting in cell-cycle arrest in the G1 phase (104). Strikingly, through its physical association with Miz-1, the *SB* transposase seems to downregulate cyclin D1 expression in human cells, resulting in a cell-cycle slowdown (Fig. 10) (103).

The likely biological significance of our finding is that, by inducing a temporary G1 delay, the *SB* transposase potentiates the involvement of NHEJ to repair transposition-inflicted DNA damage (55). Indeed, a delay in the G1/S transition and S phase progression by cell-cycle checkpoints is thought to facilitate DNA repair to avoid replication and subsequent propagation of potentially hazardous mutations. In eukaryotic cells, the NHEJ and HR pathways are complementary but act at different stages of the cell-cycle: NHEJ is preferentially active in
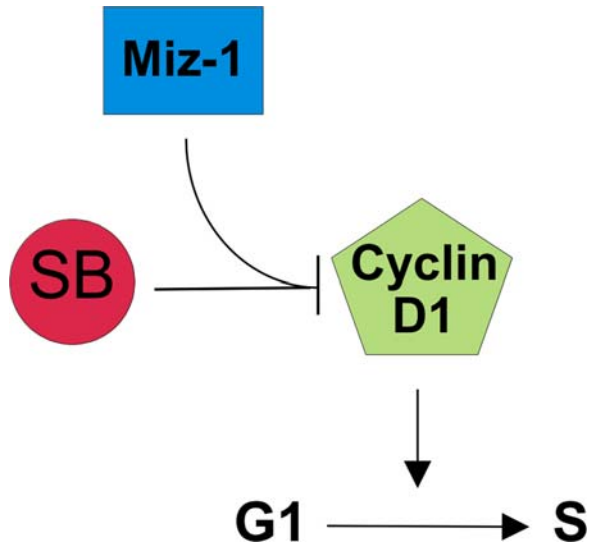
FIGURE 10 The *SB* transposase modulates cell-cycle progression through interaction with Miz-1. The *SB* transposase, through its interaction with Miz-1, downregulates cyclin D1 expression, which results in an inhibition of the G1/S transition of the cell-cycle. Reprinted from *PNAS* ([103](#)) with permission from the publisher. doi:10.1128/microbiolspec.MDNA3-0042-2014.f10

the G1 and early S phases ([105](#)), whereas HR is active in the late S and G2 phases ([95](#)). Accordingly, there is increasing evidence for a correlation between the particular pathway used for the repair of transposon-induced DNA damage and the cell-cycle stage where recombination occurs. This is nicely illustrated by gene rearrangements through V(D)J recombination, which is tightly linked to the G1 phase of the cell-cycle and to NHEJ ([92](#), [106](#)). As described above, DSBs generated by *SB* transposition are preferentially repaired by the NHEJ pathway ([55](#), [107](#)), and the *SB* transposase physically interacts with the Ku DNA-binding subunit of DNA-PK, a key component of the NHEJ machinery ([55](#)). The data suggest a model in which *SB* transposase induces a cyclin D1-dependent G1 slowdown in proliferating cells through interaction with Miz-1, thereby ensuring that transposon-induced DNA damage is repaired by NHEJ. In nature, preferential use of NHEJ for the repair of transposon-induced DSBs might help avoid homologous recombination events between dispersed copies of transposable elements in the genome, thereby assisting the maintenance of genomic stability.

Other parasitic genetic elements, including HIV-1 ([108](#)), Herpes simplex virus ([109](#)), cytomegalovirus ([110](#)), Epstein-Barr virus ([111](#)), Kaposi's sarcoma-associated herpesvirus ([112](#)), and mouse hepatitis virus ([113](#)), have also developed versatile strategies to perturb the cellular machinery to maximize their chance for survival and

propagation. Thus, overriding the normal cell-cycle program seems to be a shared strategy of parasitic genetic elements.

Transposon integration: target site selection properties of *Sleeping Beauty*

The second step of the transposition reaction is the transfer of the exposed 3′-OH transposon tip to the target DNA molecule by transesterification ([Fig. 9](#)). Similarly to the initial DNA cut, the strand transfer is executed by a nucleophilic attack. In this case, the 3′-OH groups of the already liberated transposon ends serve as a nucleophile that couples the element to the target, without previous target DNA cleavage. As a result, the transposon ends are covalently attached to staggered positions: one of the transposon ends joining to one of the target strands, the other end joining to a displaced position of the other target strand. Due to this staggered fashion of the strand joining reaction, and because the inserted element has 3′-overhangs, the integration is flanked by single-stranded gaps ([Fig. 9](#)). DNA repair at these gaps restores the terminal nucleotides of the inserted transposon and generates a characteristic duplication of the target site flanking the element that is called target site duplication (TSD) ([Fig. 9](#)). *SB* transposition almost exclusively occurs at TA dinucleotides ([5](#), [114](#), [115](#), [116](#), [117](#), [118](#)), and *SB* integrants therefore are flanked by TA TSDs ([Fig. 9](#)), which are molecular hallmarks of *SB* transposition. At very low frequencies, non-TA target sites were also found at insertions generated by *SB*100X ([115](#), [119](#)). About 75% of *SB* transposon excision events are coupled to chromosomal integration ([93](#)) and no extrachromosomal, excised molecules are readily detectable. The ~25% loss of excised transposons might be due to challenge of productive transposition by suicidal autointegration, i.e., when the transposon integrates into itself (see below).

The genome-wide insertion pattern of most transposons is non-random, showing characteristic preferences for insertion sites at the primary DNA sequence level and 'hotspots' and 'cold regions' on a genome-wide scale. Sequences responsible for target site selection of the bacterial Tn*10* transposon and retroviruses have been mapped to the core catalytic domain of the transposase or integrase, respectively ([120](#), [121](#)). However, despite the implication that the conserved catalytic DDE domain is responsible for locating the target site, no common pattern of integration emerges on the primary DNA sequence level. *SB* displays considerable specificity in target site selection at the primary DNA sequence level: in addition to the highly preferred TA,

a palindromic AT-repeat consensus sequence in an AT-rich sequence context [Fig. 11(A)] with bendability and hydrogen bonding potential was found to constitute preferred target sites (122). It was shown that a characteristic deformation of the DNA sequence may be a recognition signal for target selection (118). This

FIGURE 11 Genomic insertion preferences of *SB*. (A) Consensus sequence of *SB* insertion sites. Seqlogo analysis and nucleotide probability plot of *SB* insertion sites in HeLa cells. Twenty base pairs upstream and downstream of the TA target sites were analyzed. The y-axis represents the strength of the information, with 2 bits being the maximum for a DNA sequence. (B) Relative frequencies of insertions into genes by retroviruses and transposons. The top portions of the graphs indicate an over-representation of genic insertions as compared to random. Part (A) reprinted from *Molecular Therapy* (126) with permission from the publisher. Part (B) reprinted from *BioEssays* (170) with permission from the publisher. doi:10.1128/microbiolspec.MDNA3-0042-2014.f11

deformation, and the likelihood a particular TA will be targeted by *SB*, can be computationally predicted ([118]), which may allow a theoretical assessment of the likelihood of transposon insertions in particular genomic regions ([123]). These results indicate that a combination of particular physical properties generates a spatial optimum of the DNA for transposase interaction. This pattern of structural preference is conserved in the Tc1/ *mariner* family and in other relatively randomly integrating transposons in the DDE recombinase family ([122]). However, these factors cannot be the only determinants of target site selection because the Tc1 and Tc3 elements have different insertion profiles in *C. elegans* ([124]). Therefore, it appears that there exist at least two levels of selection that together determine how favorable a particular DNA sequence is for transposon insertion. Physical properties of the DNA primarily specify a set of sequences in a genome that are in a spatial optimum to receive a transposon insertion, whereas the ability of the transposase to efficiently interact with such sequences specifies a subset within these sites where insertions occur.

In contrast to the considerable specificity at the primary DNA sequence level, *SB* integration can be considered fairly random on the genomic level ([114], [115], [116], [117], [119], [122], [125]). Roughly one-third of *SB* insertions in mouse and human cells occur in transcribed regions [Fig. 11(B)], and because genes cover about one-third of the genome, such frequency suggests neither preference for nor disfavoring of insertion into genes. *SB* shows no pronounced preference for inserting into transcription units or transcriptional regulatory regions of genes, and the vast majority of those insertions that occur in genes are located in introns ([114], [115], [116], [117], [119], [122], [125]). The transcriptional status of targeted genes apparently does not influence the integration profile of *SB* ([125]). This is in marked contrast to target site distributions of several other transposons including *Tol2* [Fig. 11(B)] ([117], [126]), *TcBuster* ([115]), *SPIN* ([115]), and *piggyBac* [Fig. 11(B)] ([115], [117], [126], [127], [128]). These all show significant difference from random insertion with respect to favored integration into genes and near chromatin marks characteristic of active transcription units (e.g., H3K27 acetylation and H3K4 monomethylation) and disfavored integration near marks characteristic of inactive chromatin (e.g., H3K27 trimethylation).

The random genomic distribution of *de novo SB* insertions can be observed when the transposon DNA is introduced into the nucleus by extrachromosomal gene delivery, including plasmid vectors ([114], [115], [116], [117], [119], [122], [125]), integration-deficient lentiviral vectors (IDLVs) ([114]), adenovirus vectors ([129]), herpesvirus

vectors ([130]), and adeno-associated vectors ([131]). In these cases, transposition takes place from the extrachromosomal vector into the genome. However, target site selection properties of *SB* when launched from a chromosomal site are markedly different and are governed by "local hopping". Local hopping describes a phenomenon of chromosomal transposition in which transposons have a preference for landing into *cis*-linked sites in the vicinity of the donor locus. Local hopping seems to be a shared feature of "cut-and-paste" transposons. However, the actual extent of hopping to linked chromosomal sites and the interval of local transposition varies. For example, the P-element transposon of *Drosophila* prefers to insert within ~100 kb of the donor site at a rate ~50-fold higher than in regions outside that interval ([132]). Similarly, 30 to 80% of *SB* transposition events were found to re-insert locally on either side of the transposon donor locus ([93], [133], [134], [135], [136], [137], [138], [139]). In contrast to the P-element, *SB* seems to have a much larger local transposition interval (in the megabase range), but the targeted window for local reinsertion appears to be dependent on the donor locus and can range from <4 Mb ([140]) to 5 to 15 Mb ([134]). Local hopping of *SB* was also observed within 3 Mb of the donor locus in *Xenopus* ([141]).

The local hopping feature not only differs between different transposons, but a given transposon may show great variations in local hops in different hosts and in different donor loci even in the same host. For example, local hopping of the *Ac* element in tomato seems overall to be less prevalent than in maize ([142], [143]), and there are species-specific differences in its tendency for local hopping out of different transposon donor loci ([144]). This variation in local hopping of the same element could possibly be explained by varying affinity of the transposase for unknown, chromatin-associated factors in different hosts ([145]).

Self-disruptive autointegration of *Sleeping Beauty*

In the process of productive transposition, the excised DNA molecule integrates into a new genomic location. However, in principle, the excised transposon molecule could reinsert, in a self-disruptive process, into its own genome. This suicidal transposition event is called autointegration, which has been observed with the bacterial systems Tn*10* ([146]) and *Mu* ([147]), with Ty1 retrotransposons in yeast ([148]), and with retroviruses ([149]). Apparently, the *SB* transposon is also subject to autointegration ([56]), thereby compromising integration of a fraction of excised transposon molecules.

The conserved size of Tc1-like elements is between 1.6 and 1.8 kb. The efficacy of transposition usually correlates negatively with increasing size of the transposon ([58](), [150](), [151](), [152]()), and such an effect was also observed with the bacterial transposons IS*1* ([153]()) and Tn*10* ([154]()) and with *piggyBac* ([56]()). Since larger transposons contain more potential target sites, they could be particularly attractive targets for autointegration. Indeed, increasing size was found to sensitize both *SB* and *piggyBac* transposition for autointegration ([56]()). However, the competition between autointegration and productive transposition is unlikely to be the only factor responsible for sensitivity to size as transposon excision,

a step prior to integration, is already affected by the size of the *SB* transposon ([56]()).

A host-encoded protein, barrier-to-autointegration factor (BAF or BANF1) has been identified by its ability to protect retroviruses from autointegration ([155](), [156](), [157](), [158](), [159]()). Intriguingly, BANF1 also inhibited transposon autointegration of *SB* and was detected in higher-order protein complexes containing the transposase in human cells ([56]()). Thus, the *SB* transposon seems to be able to recruit phylogenetically conserved cellular factors such as BANF1 that protects against self-disruption in a new environment (a human cell is a naïve host for *SB*). In fact, BANF1 might be an ideal cellular factor for

FIGURE 12 Broad applicability of *SB* transposon-based gene vectors in vertebrate genetics. Reprinted from *Mobile DNA* ([168]()) with permission from the publisher. [doi:10.1128/microbiolspec.MDNA3-0042-2014.f12]()
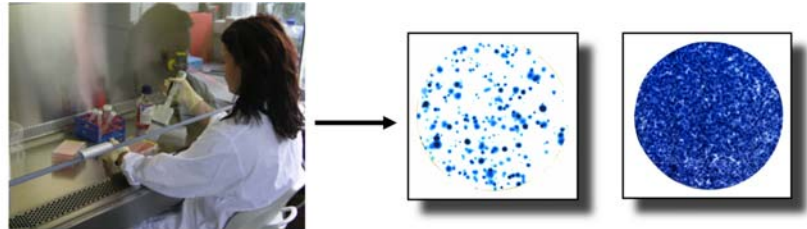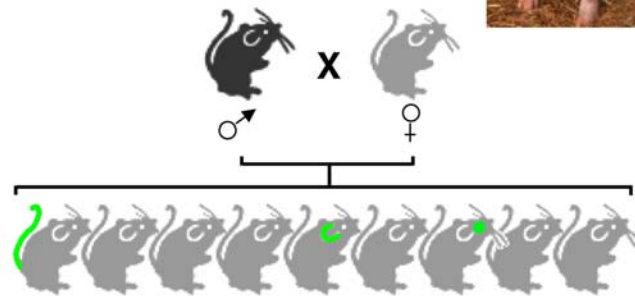
integrating elements in higher eukaryotes. As in retroviral integration ([155](), [156](), [157](), [158]()), BANF1 may compact the transposon genome to be a less accessible target for autointegration, thereby promoting the chromosomal integration step.

Sleeping Beauty AS A GENETIC TOOL

Transposons can be viewed as natural DNA transfer vehicles that, similar to integrating viruses, are capable of efficient genomic insertion. Transposition can be controlled by conditionally providing the transposase component of the transposition reaction. Thus, a DNA of interest (be it a fluorescent marker, an shRNA expression cassette, a mutagenic gene trap, or a therapeutic gene construct) cloned between the IR sequences of a transposon-based vector can be utilized for stable genomic insertion in a regulated and highly efficient manner. This methodological paradigm opened up a number of avenues for genome manipulations in vertebrates, including transgenesis for the generation of transgenic cells in tissue culture, the production of germline-transgenic animals for basic and applied research, forward genetic screens for functional gene annotation in model species, and therapy of genetic disorders in humans ([Fig. 12]()). *SB* was the first transposon ever shown to be capable of gene transfer in vertebrate cells, and recent results confirm that *SB* supports a full spectrum of genetic engineering including transgenesis, insertional mutagenesis, and therapeutic somatic gene transfer, both *ex vivo* and *in vivo*. The first clinical application of the *SB* system will help to validate both the safety and efficacy of this approach. Applications of the *SB* system fall outside the scope of this chapter, and readers are referred to recent review articles ([160](), [161](), [162](), [163](), [164](), [165](), [166](), [167](), [168](), [169](), [170](), [171](), [172]()).

ACKNOWLEDGMENTS

REFERENCES

1. Plasterk RH, Izsvak Z, Ivics Z. 1999. Resident aliens: the Tc1/mariner superfamily of transposable elements. *Trends Genet* **15:**326–332.

2. Ivics Z, Izsvák Z, Minter A, Hackett PB. 1996. Identification of functional domains and evolution of Tc1-like transposable elements. *Proc Natl Acad Sci U S A* **93:**5008–5013.

3. Radice AD, Bugaj B, Fitch DH, Emmons SW. 1994. Widespread occurrence of the Tc1 transposon family: Tc1-like transposons from teleost fish. *Mol Gen Genet* **244:**606–612.

4. Goodier JL, Davidson WS. 1994. Tc1 transposon-like sequences are widely distributed in salmonids. *J Mol Biol* **241:**26–34.

5. Ivics Z, Hackett PB, Plasterk RH, Izsvak Z. 1997. Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. *Cell* **91:**501–510.

6. Izsvák Z, Khare D, Behlke J, Heinemann U, Plasterk RH, Ivics Z. 2002. Involvement of a bifunctional, paired-like DNA-binding domain and a transpositional enhancer in *Sleeping Beauty* transposition. *J Biol Chem* **277:**34581–34588.

7. Pietrokovski S, Henikoff S. 1997. A helix-turn-helix DNA-binding motif predicted for transposases of DNA transposons. *Mol Gen Genet* **254:**689–695.

8. Vos JC, Plasterk RH. 1994. Tc1 transposase of *Caenorhabditis elegans* is an endonuclease with a bipartite DNA binding domain. *EMBO J* **13:**6125–6132.

9. Franz G, Loukeris TG, Dialektaki G, Thompson CR, Savakis C. 1994. Mobile Minos elements from Drosophila hydei encode a two-exon transposase with similarity to the paired DNA-binding domain. *Proc Natl Acad Sci U S A* **91:**4746–4750.

10. Czerny T, Schaffner G, Busslinger M. 1993. DNA sequence recognition by Pax proteins: bipartite structure of the paired domain and its binding site. *Genes Dev* **7:**2048–2061.

11. Breitling R, Gerber JK. 2000. Origin of the paired domain. *Dev Genes Evol* **210:**644–650.

12. van Pouderoyen G, Ketting RF, Perrakis A, Plasterk RH, Sixma TK. 1997. Crystal structure of the specific DNA-binding domain of Tc3 transposase of *C. elegans* in complex with transposon DNA. *EMBO J* **16:**6044–6054.

13. Carpentier CE, Schreifels JM, Aronovich EL, Carlson DF, Hackett PB, Nesmelova IV. 2014. NMR structural analysis of Sleeping Beauty transposase binding to DNA. *Protein Sci* **23:**23–33.

14. Gehring WJ, Qian YQ, Billeter M, Furukubo-Tokunaga K, Schier AF, Resendez-Perez D, Affolter M, Otting G, Wuthrich K. 1994. Homeodomain-DNA recognition. *Cell* **78:**211–223.

15. Feng JA, Johnson RC, Dickerson RE. 1994. Hin recombinase bound to DNA: the origin of specificity in major and minor groove interactions. *Science* **263:**348–355.

16. Spanopoulou E, Zaitseva F, Wang FH, Santagata S, Baltimore D, Panayotou G. 1996. The homeodomain region of Rag-1 reveals the parallel mechanisms of bacterial and V(D)J recombination. *Cell* **87:**263–276.

17. Kulkosky J, Jones KS, Katz RA, Mack JP, Skalka AM. 1992. Residues critical for retroviral integrative recombination in a region that is highly conserved among retroviral/retrotransposon integrases and bacterial insertion sequence transposases. *Mol Cell Biol* **12:**2331–2338.

18. Doak TG, Doerder FP, Jahn CL, Herrick G. 1994. A proposed superfamily of transposase genes: transposon-like elements in ciliated protozoa and a common "D35E" motif. *Proc Natl Acad Sci U S A* **91:**942–946.

19. Kim DR, Dai Y, Mundy CL, Yang W, Oettinger MA. 1999. Mutations of acidic residues in RAG1 define the active site of the V(D)J recombinase. *Genes Dev* **13:**3070–3080.

20. Landree MA, Wibbenmeyer JA, Roth DB. 1999. Mutational analysis of RAG1 and RAG2 identifies three catalytic amino acids in RAG1 critical for both cleavage steps of V(D)J recombination. *Genes Dev* **13:**3059–3069.

21. Haren L, Ton-Hoang B, Chandler M. 1999. Integrating DNA: transposases and retroviral integrases. *Annu Rev Microbiol* **53:**245–281.

22. Bujacz G, Alexandratos J, Wlodawer A, Merkel G, Andrake M, Katz RA, Skalka AM. 1997. Binding of different divalent cations to the active site of avian sarcoma virus integrase and their effects on enzymatic activity. *J Biol Chem* **272:**18161–18168.

23. Goldgur Y, Dyda F, Hickman AB, Jenkins TM, Craigie R, Davies DR. 1998. Three new structures of the core domain of HIV-1 integrase: an active site that binds magnesium. *Proc Natl Acad Sci U S A* **95:**9150–9154.

24. **van Luenen HG, Colloms SD, Plasterk RH.** 1994. The mechanism of transposition of Tc3 in C. elegans. *Cell* **79:**293–301.

25. **Yant SR, Meuse L, Chiu W, Ivics Z, Izsvak Z, Kay MA.** 2000. Somatic integration and long-term transgene expression in normal and haemophilic mice using a DNA transposon system. *Nat Genet* **25:**35–41.

26. **Richardson JM, Colloms SD, Finnegan DJ, Walkinshaw MD.** 2009. Molecular architecture of the Mos1 paired-end complex: the structural basis of DNA transposition in a eukaryote. *Cell* **138:**1096–1108.

27. **Katayanagi K, Miyagawa M, Matsushima M, Ishikawa M, Kanaya S, Ikehara M, Matsuzaki T, Morikawa K.** 1990. Three-dimensional structure of ribonuclease H from E. coli. *Nature* **347:**306–309.

28. **Ariyoshi M, Vassylyev DG, Iwasaki H, Nakamura H, Shinagawa H, Morikawa K.** 1994. Atomic structure of the RuvC resolvase: a holliday junction-specific endonuclease from E. coli. *Cell* **78:**1063–1072.

29. **Lohe AR, De Aguiar D, Hartl DL.** 1997. Mutations in the mariner transposase: the D,D(35)E consensus sequence is nonfunctional. *Proc Natl Acad Sci U S A* **94:**1293–1297.

30. **Mates L, Chuah MK, Belay E, Jerchow B, Manoj N, Acosta-Sanchez A, Grzela DP, Schmitt A, Becker K, Matrai J, Ma L, Samara-Kuko E, Gysemans C, Pryputniewicz D, Miskey C, Fletcher B, Vandendriessche T, Ivics Z, Izsvak Z.** 2009. Molecular evolution of a novel hyperactive Sleeping Beauty transposase enables robust stable gene transfer in vertebrates. *Nat Genet* **41:**753–761.

31. **Bushman FD, Engelman A, Palmer I, Wingfield P, and Craigie R.** 1993. Domains of the integrase protein of human immunodeficiency virus type 1 responsible for polynucleotidyl transfer and zinc binding. *Proc Natl Acad Sci U S A* **90:**3428–3432.

32. **Dyda F, Hickman AB, Jenkins TM, Engelman A, Craigie R, Davies DR.** 1994. Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. *Science* **266:**1981–1986.

33. **Yant SR, Park J, Huang Y, Mikkelsen JG, Kay MA.** 2004. Mutational analysis of the N-terminal DNA-binding domain of sleeping beauty transposase: critical residues for DNA binding and hyperactivity in mammalian cells. *Mol Cell Biol* **24:**9239–9247.

34. **Geurts AM, Yang Y, Clark KJ, Liu G, Cui Z, Dupuy AJ, Bell JB, Largaespada DA, Hackett PB.** 2003. Gene transfer into genomes of human cells by the sleeping beauty transposon system. *Mol Ther* **8:**108–117.

35. **Zayed H, Izsvak Z, Walisko O, Ivics Z.** 2004. Development of hyperactive sleeping beauty transposon vectors by mutational analysis. *Mol Ther* **9:**292–304.

36. **Baus J, Liu L, Heggestad AD, Sanz S, Fletcher BS.** 2005. Hyperactive transposase mutants of the Sleeping Beauty transposon. *Mol Ther* **12:**1148–1156.

37. **Xue X, Huang X, Nodland SE, Mates L, Ma L, Izsvak Z, Ivics Z, LeBien TW, McIvor RS, Wagner JE, Zhou X.** 2009. Stable gene transfer and expression in cord blood-derived CD34+ hematopoietic stem and progenitor cells by a hyperactive Sleeping Beauty transposon system. *Blood* **114:**1319–1330.

38. **Belay E, Matrai J, Acosta-Sanchez A, Ma L, Quattrocelli M, Mates L, Sancho-Bru P, Geraerts M, Yan B, Vermeesch J, Rincon MY, Samara-Kuko E, Ivics Z, Verfaillie C, Sampaolesi M, Izsvak Z, Vandendriessche T, Chuah MK.** 2010. Novel hyperactive transposons for genetic modification of induced pluripotent and adult stem cells: a nonviral paradigm for coaxed differentiation. *Stem Cells* **28:**1760–1771.

39. **Jin Z, Maiti S, Huls H, Singh H, Olivares S, Mates L, Izsvak Z, Ivics Z, Lee DA, Champlin RE, Cooper LJ.** 2011. The hyperactive Sleeping Beauty transposase SB100X improves the genetic modification of T cells to express a chimeric antigen receptor. *Gene Ther* **18:**849–856.

40. **Ivics Z, Garrels W, Mates L, Yau TY, Bashir S, Zidek V, Landa V, Geurts A, Pravenec M, Rulicke T, Kues WA, Izsvak Z.** 2014. Germline transgenesis in pigs by cytoplasmic microinjection of Sleeping Beauty transposons. *Nat Protoc* **9:**810–827.

41. **Ivics Z, Hiripi L, Hoffmann OI, Mates L, Yau TY, Bashir S, Zidek V, Landa V, Geurts A, Pravenec M, Rulicke T, Bosze Z, Izsvak Z.** 2014. Germline transgenesis in rabbits by pronuclear microinjection of Sleeping Beauty transposons. *Nat Protoc* **9:**794–809.

42. **Ivics Z, Mates L, Yau TY, Landa V, Zidek V, Bashir S, Hoffmann OI, Hiripi L, Garrels W, Kues WA, Bosze Z, Geurts A, Pravenec M, Rulicke T, Izsvak Z.** 2014. Germline transgenesis in rodents by pronuclear microinjection of Sleeping Beauty transposons. *Nat Protoc* **9:**773–93.

43. **Izsvák Z, Ivics Z, Hackett PB.** 1995. Characterization of a Tc1-like transposable element in zebrafish (*Danio rerio*). *Mol Gen Genet* **247:**312–322.

44. **Franz G, Savakis C.** 1991. Minos, a new transposable element from Drosophila hydei, is a member of the Tc1-like family of transposons. *Nucleic Acids Res* **19:**6646.

45. **Merriman PJ, Grimes CD, Ambroziak J, Hackett DA, Skinner P, Simmons MJ.** 1995. S elements: a family of Tc1-like transposons in the genome of *Drosophila melanogaster*. *Genetics* **141:**1425–1438.

46. **Moschetti R, Caggese C, Barsanti P, Caizzi R.** 1998. Intra- and interspecies variation among *Bari*-1 elements of the *melanogaster* species group. *Genetics* **150:**239–250.

47. **Ke Z, Grossman GL, Cornel AJ, Collins FH.** 1996. *Quetzal*: A transposon of the Tc1 family in the mosquito *Anopheles albimanus*. *Genetica* **98:**141–147.

48. **Lam WL, Seo P, Robison K, Virk S, Gilbert W.** 1996. Discovery of amphibian Tc1-like transposon families. *J Mol Biol* **257:**359–366.

49. **Sinzelle L, Pollet N, Bigot Y, Mazabraud A.** 2005. Characterization of multiple lineages of Tc1-like elements within the genome of the amphibian Xenopus tropicalis. *Gene* **349:**187–196.

50. **Lampe DJ, Churchill ME, Robertson HM.** 1996. A purified *mariner* transposase is sufficient to mediate transposition *in vitro*. *EMBO J* **15:**5470–5479.

51. **Vos JC, De Baere I, Plasterk RH.** 1996. Transposase is the only nematode protein required for *in vitro* transposition of Tc1. *Genes Dev* **10:**755–761.

52. **Raz E, van Luenen HG, Schaerringer B, Plasterk RH, Driever W.** 1998. Transposition of the nematode Caenorhabditis elegans Tc3 element in the zebrafish Danio rerio. *Curr Biol* **8:**82–88.

53. **Zayed H, Izsvak Z, Khare D, Heinemann U, Ivics Z.** 2003. The DNA-bending protein HMGB1 is a cellular cofactor of *Sleeping Beauty* transposition. *Nucleic Acids Res* **31:**2313–2322.

54. **Walisko O, Schorn A, Rolfs F, Devaraj A, Miskey C, Izsvak Z, Ivics Z.** 2008. Transcriptional activities of the Sleeping Beauty transposon and shielding its genetic cargo with insulators. *Mol Ther* **16:**359–369.

55. **Izsvák Z, Stuwe EE, Fiedler D, Katzer A, Jeggo PA, Ivics Z.** 2004. Healing the wounds inflicted by sleeping beauty transposition by double-strand break repair in mammalian somatic cells. *Mol Cell* **13:**279–290.

56. **Wang Y, Wang J, Devaraj A, Singh M, Jimenez Orgaz A, Chen JX, Selbach M, Ivics Z, Izsvak Z.** 2014. Suicidal autointegration of sleeping beauty and piggyBac transposons in eukaryotic cells. *PLoS Genet* **10:**e1004103.

57. **Walisko O, Jursch T, Izsvak Z, Ivics Z.** 2008. Transposon-host cell interactions in the regulation of Sleeping Beauty transposition, p 109–132. *In* Volff J-N, Lankenau D-H (ed), *Transposons and the Dynamic Genome*, Springer Berlin Heidelberg.

58. **Izsvák Z, Ivics Z, Plasterk RH.** 2000. *Sleeping Beauty*, a wide host-range transposon vector for genetic transformation in vertebrates. *J Mol Biol* **302:**93–102.

59. **Hozumi A, Mita K, Miskey C, Mates L, Izsvak Z, Ivics Z, Satake H, Sasakura Y.** 2013. Germline transgenesis of the chordate Ciona intestinalis with hyperactive variants of sleeping beauty transposable element. *Dev Dyn* **242:**30–43.

60. **Leaver MJ.** 2001. A family of Tc1-like transposons from the genomes of fishes and frogs: evidence for horizontal transmission. *Gene* **271:**203–214.

61. **Sijen T, Plasterk RH.** 2003. Transposon silencing in the Caenorhabditis elegans germ line by natural RNAi. *Nature* **426:**310–314.

62. **Moldt B, Yant SR, Andersen PR, Kay MA, Mikkelsen JG.** 2007. Cis-acting gene regulatory activities in the terminal regions of sleeping beauty DNA transposon-based vectors. *Hum Gene Ther* **18:**1193–1204.

63. **Rauschhuber C, Ehrhardt A.** 2012. RNA interference is responsible for reduction of transgene expression after Sleeping Beauty transposase mediated somatic integration. *PLoS One* **7:**e35389.

64. **Bewley CA, Gronenborn AM, Clore GM.** 1998. Minor groove-binding architectural proteins: structure, function, and DNA recognition. *Annu Rev Biophys Biomol Struct* **27:**105–131.

65. **Cui Z, Geurts AM, Liu G, Kaufman CD, Hackett PB.** 2002. Structure-function analysis of the inverted terminal repeats of the *Sleeping Beauty* transposon. *J Mol Biol* **318:**1221–1235.

66. **Leung PC, Teplow DB, Harshey RM.** 1989. Interaction of distinct domains in Mu transposase with Mu DNA ends and an internal transpositional enhancer. *Nature* **338:**656–658.

67. **van Gent DC, Hiom K, Paull TT, Gellert M.** 1997. Stimulation of V(D)J cleavage by high mobility group proteins. *EMBO J* **16:**2665–2670.

68. **Agrawal A, Schatz DG.** 1997. RAG1 and RAG2 form a stable post-cleavage synaptic complex with DNA containing signal ends in V(D)J recombination. *Cell* **89:**43–53.

69. **Bustin M.** 1999. Regulation of DNA-dependent activities by the functional motifs of the high-mobility-group chromosomal proteins. *Mol Cell Biol* **19:**5237–5246.

70. **Claeys Bouuaert C, Liu D, Chalmers R.** 2011. A simple topological filter in a eukaryotic transposon as a mechanism to suppress genome instability. *Mol Cell Biol* **31:**317–327.

71. **Yoder JA, Walsh CP, Bestor TH.** 1997. Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet* **13:**335–340.

72. **Yusa K, Takeda J, Horie K.** 2004. Enhancement of Sleeping Beauty transposition by CpG methylation: possible role of heterochromatin formation. *Mol Cell Biol* **24:**4004–4018.

73. **Jursch T, Miskey C, Izsvak Z, Ivics Z.** 2013. Regulation of DNA transposition by CpG methylation and chromatin structure in human cells. *Mobile DNA* **4:**15.

74. **Mizuuchi K.** 1992. Polynucleotidyl transfer reactions in transpositional DNA recombination. *J Biol Chem* **267:**21273–21276.

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

76. **Turlan C, Chandler M.** 2000. Playing second fiddle: second-strand processing and liberation of transposable elements from donor DNA. *Trends Microbiol* **8:**268–274.

77. **Bischerour J, Chalmers R.** 2009. Base flipping in tn10 transposition: an active flip and capture mechanism. *PLoS One* **4:**e6201.

78. **Bischerour J, Lu C, Roth DB, Chalmers R.** 2009. Base flipping in V(D)J recombination: insights into the mechanism of hairpin formation, the 12/23 rule, and the coordination of double-strand breaks. *Mol Cell Biol* **29:**5889–5899.

79. **Mitra R, Fain-Thornton J, Craig NL.** 2008. piggyBac can bypass DNA synthesis during cut and paste transposition. *EMBO J* **27:**1097–1109.

80. **Zhou L, Mitra R, Atkinson PW, Hickman AB, Dyda F, Craig NL.** 2004. Transposition of hAT elements links transposable elements and V(D)J recombination. *Nature* **432:**995–1001.

81. **Hencken CG, Li X, Craig NL.** 2012. Functional characterization of an active Rag-like transposase. *Nat Struct Mol Biol* **19:**834–836.

82. **Gellert M.** 2002. V(D)J recombination: RAG proteins, repair factors, and regulation. *Annu Rev Biochem* **71:**101–132.

83. **van Gent DC, Mizuuchi K, Gellert M.** 1996. Similarities between initiation of V(D)J recombination and retroviral integration. *Science* **271:**1592–1594.

84. **Bhasin A, Goryshin IY, Reznikoff WS.** 1999. Hairpin formation in Tn5 transposition. *J Biol Chem* **274:**37021–37029.

85. **Kennedy AK, Guhathakurta A, Kleckner N, Haniford DB.** 1998. Tn10 transposition via a DNA hairpin intermediate. *Cell* **95:**125–134.

86. **Richardson JM, Dawson A, O'Hagan N, Taylor P, Finnegan DJ, Walkinshaw MD.** 2006. Mechanism of Mos1 transposition: insights from structural analysis. *EMBO J* **25:**1324–1334.

87. **Goryshin IY, Reznikoff WS.** 1998. Tn5 in vitro transposition. *J Biol Chem* **273:**7367–7374.

88. **Miskey C, Papp B, Mates L, Sinzelle L, Keller H, Izsvak Z, Ivics Z.** 2007. The Ancient Mariner Sails Again: Transposition of the Human Hsmar1 Element by a Reconstructed Transposase and Activities of the SETMAR Protein on Transposon Ends. *Mol Cell Biol* **27:**4589–4600.

89. **Engels WR, Johnson-Schlitz DM, Eggleston WB, Sved J.** 1990. High-frequency P element loss in *Drosophila* is homolog dependent. *Cell* **62:**515–525.

90. **Lohe AR, Timmons C, Beerman I, Lozovskaya ER, Hartl DL.** 2000. Self-inflicted wounds, template-directed gap repair and a recombination hotspot. Effects of the mariner transposase. *Genetics* **154:**647–656.

91. **Durocher D, Jackson SP.** 2001. DNA-PK, ATM and ATR as sensors of DNA damage: variations on a theme? *Curr Opin Cell Biol* **13:**225–231.

92. **Jackson SP, Jeggo PA.** 1995. DNA double-strand break repair and V(D)J recombination: involvement of DNA-PK. *Trends Biochem Sci* **20:**412–415.

93. **Luo G, Ivics Z, Izsvak Z, Bradley A.** 1998. Chromosomal transposition of a Tc1/mariner-like element in mouse embryonic stem cells. *Proc Natl Acad Sci U S A* **95:**10769–10773.

94. **Daniel R, Katz RA, Skalka AM.** 1999. A role for DNA-PK in retroviral DNA integration. *Science* **284:**644–647.

95. **Takata M, Sasaki MS, Sonoda E, Morrison C, Hashimoto M, Utsumi H, Yamaguchi-Iwai Y, Shinohara A, Takeda S.** 1998. Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. *EMBO J* **17:**5497–5508.

96. **Richardson C, Jasin M.** 2000. Coupled homologous and nonhomologous repair of a double-strand break preserves genomic integrity in mammalian cells. *Mol Cell Biol* **20:**9068–9075.

97. **Plasterk RH.** 1991. The origin of footprints of the Tc1 transposon of *Caenorhabditis elegans*. *EMBO J* **10:**1919–1925.

98. **Gloor GB, Moretti J, Mouyal J, Keeler KJ.** 2000. Distinct P-element excision products in somatic and germline cells of Drosophila melanogaster. *Genetics* **155:**1821–1830.

99. **Nassif N, Penney J, Pal S, Engels WR, Gloor GB.** 1994. Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. *Mol Cell Biol* **14:**1613–1625.

100. **Downs JA, Jackson SP.** 1999. Involvement of DNA end-binding protein Ku in Ty element retrotransposition. *Mol Cell Biol* **19:**6260–6268.

101. **Sharon G, Burkett TJ, Garfinkel DJ.** 1994. Efficient homologous recombination of Ty1 element cDNA when integration is blocked. *Mol Cell Biol* **14:**6540–6551.

102. **Peukert K, Staller P, Schneider A, Carmichael G, Hanel F, Eilers M.** 1997. An alternative pathway for gene regulation by Myc. *EMBO J* **16:**5672–5686.

103. **Walisko O, Izsvak Z, Szabo K, Kaufman CD, Herold S, Ivics Z.** 2006. Sleeping Beauty transposase modulates cell-cycle progression through interaction with Miz-1. *Proc Natl Acad Sci U S A* **103:**4062–4067.

104. **Baldin V, Lukas J, Marcote MJ, Pagano M, Draetta G.** 1993. Cyclin D1 is a nuclear protein required for cell cycle progression in G1. *Genes Dev* **7:**812–821.

105. **Lee SE, Mitchell RA, Cheng A, Hendrickson EA.** 1997. Evidence for DNA-PK-dependent and -independent DNA double-strand break repair pathways in mammalian cells as a function of the cell cycle. *Mol Cell Biol* **17:**1425–1433.

106. **Lee J, Desiderio S.** 1999. Cyclin A/CDK2 regulates V(D)J recombination by coordinating RAG-2 accumulation and DNA repair. *Immunity* **11:**771–781.

107. **Yant SR, Kay MA.** 2003. Nonhomologous-end-joining factors regulate DNA repair fidelity during *Sleeping Beauty* element transposition in mammalian cells. *Mol Cell Biol* **23:**8505–8518.

108. **Emerman M.** 1996. HIV-1, Vpr and the cell cycle. *Curr Biol* **6:**1096–1103.

109. **Lomonte P, Everett RD.** 1999. Herpes Simplex Virus type 1 immediate-early protein Vmw110 inhibits progression of cells through mitosis and from G1 into S phase of the cell cycle. *J Virol* **73:**9456–9467.

110. **Lu M, Shenk T.** 1999. Human cytomegalovirus UL69 protein induces cells to accumulate in G1 phase of the cell cycle. *J Virol* **73:**676–683.

111. **Cayrol C, Flemington EK.** 1996. The Epstein-Barr virus bZIP transcription factor Zta causes G0/G1 cell cycle arrest through induction of cyclin-dependent kinase inhibitors. *EMBO J* **15:**2748–2759.

112. **Izumiya Y, Lin S-F, Ellison TJ, Levy AM, Mayeur GL, Izumiya C, Kung H-J.** 2003. Cell cycle regulation by Kaposi's sarcoma-associated herpesvirus K-bZIP: direct interaction with Cyclin-CDK2 and induction of G1 growth arrest. *J Virol* **77:**9652–9661.

113. **Chen CJ, Makino S.** 2004. Murine coronavirus replication induces cell cycle arrest in G0/G1 phase. *J Virol* **78:**5658–5669.

114. **Moldt B, Miskey C, Staunstrup NH, Gogol-Doring A, Bak RO, Sharma N, Mates L, Izsvak Z, Chen W, Ivics Z, Mikkelsen JG.** 2011. Comparative genomic integration profiling of Sleeping Beauty transposons mobilized with high efficacy from integrase-defective lentiviral vectors in primary human cells. *Mol Ther* **19:**1499–1510.

115. **Li X, Ewis H, Hice RH, Malani N, Parker N, Zhou L, Feschotte C, Bushman FD, Atkinson PW, Craig NL.** 2013. A resurrected mammalian hAT transposable element and a closely related insect element are highly active in human cell culture. *Proc Natl Acad Sci U S A* **110:**E478–487.

116. **Voigt K, Gogol-Doring A, Miskey C, Chen W, Cathomen T, Izsvak Z, Ivics Z.** 2012. Retargeting sleeping beauty transposon insertions by engineered zinc finger DNA-binding domains. *Mol Ther* **20:**1852–1862.

117. **Ammar I, Gogol-Doring A, Miskey C, Chen W, Cathomen T, Izsvak Z, Ivics Z.** 2012. Retargeting transposon insertions by the adeno-associated virus Rep protein. *Nucleic Acids Res* **40:**6693–6712.

118. **Liu G, Geurts AM, Yae K, Srinivasan AR, Fahrenkrug SC, Largaespada DA, Takeda J, Horie K, Olson WK, Hackett PB.** 2005. Target-site preferences of Sleeping Beauty transposons. *J Mol Biol* **346:**161–173.

119. **de Jong J, Akhtar W, Badhai J, Rust AG, Rad R, Hilkens J, Berns A, van Lohuizen M, Wessels LF, de Ridder J.** 2014. Chromatin landscapes of retroviral and transposon integration profiles. *PLoS Genet* **10:**e1004250.

120. **Junop MS, Haniford DB.** 1997. Factors responsible for target site selection in Tn10 transposition: a role for the DDE motif in target DNA capture. *EMBO J* **16:**2646–2655.

121. **Katzman M, Sudol M.** 1995. Mapping domains of retroviral integrase responsible for viral DNA specificity and target site selection by analysis of chimeras between human immunodeficiency virus type 1 and visna virus integrases. *J Virol* **69:**5687–5696.

122. **Vigdal TJ, Kaufman CD, Izsvak Z, Voytas DF, Ivics Z.** 2002. Common physical properties of DNA affecting target site selection of sleeping beauty and other Tc1/mariner transposable elements. *J Mol Biol* **323:**441–452.

123. **Geurts AM, Hackett CS, Bell JB, Bergemann TL, Collier LS, Carlson CM, Largaespada DA, Hackett PB.** 2006. Structure-based prediction of insertion-site preferences of transposons into chromosomes. *Nucleic Acids Res* **34:**2803–2811.

124. **van Luenen HG, Plasterk RH.** 1994. Target site choice of the related transposable elements Tc1 and Tc3 of Caenorhabditis elegans. *Nucleic Acids Res* **22:**262–269.

125. **Yant SR, Wu X, Huang Y, Garrison B, Burgess SM, Kay MA.** 2005. High-resolution genome-wide mapping of transposon integration in mammals. *Mol Cell Biol* **25:**2085–2094.

126. **Grabundzija I, Irgang M, Mates L, Belay E, Matrai J, Gogol-Doring A, Kawakami K, Chen W, Ruiz P, Chuah MK, VandenDriessche T, Izsvak Z, Ivics Z.** 2010. Comparative analysis of transposable element vector systems in human cells. *Mol Ther* **18:**1200–1209.

127. **Huang X, Guo H, Tammana S, Jung YC, Mellgren E, Bassi P, Cao Q, Tu ZJ, Kim YC, Ekker SC, Wu X, Wang SM, Zhou X.** 2010. Gene transfer efficiency and genome-wide integration profiling of Sleeping Beauty, Tol2, and piggyBac transposons in human primary T cells. *Mol Ther* **18:**1803–1813.

128. **Li MA, Pettitt SJ, Eckert S, Ning Z, Rice S, Cadinanos J, Yusa K, Conte N, Bradley A.** 2013. The piggyBac transposon displays local and distant reintegration preferences and can cause mutations at noncanonical integration sites. *Mol Cell Biol* **33:**1317–1330.

129. **Zhang W, Muck-Hausl M, Wang J, Sun C, Gebbing M, Miskey C, Ivics Z, Izsvak Z, Ehrhardt A.** 2013. Integration profile and safety of an adenovirus hybrid-vector utilizing hyperactive sleeping beauty transposase for somatic integration. *PLoS One* **8:**e75344.

130. **de Silva S, Mastrangelo MA, Lotta LT, Jr., Burris CA, Izsvak Z, Ivics Z, Bowers WJ.** 2010. Herpes simplex virus/Sleeping Beauty vector-based embryonic gene transfer using the HSB5 mutant: loss of apparent transposition hyperactivity in vivo. *Hum Gene Ther* **21:**1603–1613.

131. **Zhang W, Solanki M, Muther N, Ebel M, Wang J, Sun C, Izsvak Z, Ehrhardt A.** 2013. Hybrid adeno-associated viral vectors utilizing transposase-mediated somatic integration for stable transgene expression in human cells. *PLoS One* **8:**e76771.

132. **Tower J, Karpen GH, Craig N, Spradling AC.** 1993. Preferential transposition of Drosophila P elements to nearby chromosomal sites. *Genetics* **133:**347–359.

133. **Fischer SE, Wienholds E, Plasterk RH.** 2001. Regulated transposition of a fish transposon in the mouse germ line. *Proc Natl Acad Sci U S A* **98:**6759–6764.

134. **Carlson CM, Dupuy AJ, Fritz S, Roberg-Perez KJ, Fletcher CF, Largaespada DA.** 2003. Transposon mutagenesis of the mouse germline. *Genetics* **165:**243–256.

135. **Horie K, Yusa K, Yae K, Odajima J, Fischer SE, Keng VW, Hayakawa T, Mizuno S, Kondoh G, Ijiri T, Matsuda Y, Plasterk RH, Takeda J.** 2003. Characterization of Sleeping Beauty transposition and its application to genetic screening in mice. *Mol Cell Biol* **23:**9189–9207.

136. **Liang Q, Kong J, Stalker J, Bradley A.** 2009. Chromosomal mobilization and reintegration of Sleeping Beauty and PiggyBac transposons. *Genesis* **47:**404–408.

137. **Kokubu C, Horie K, Abe K, Ikeda R, Mizuno S, Uno Y, Ogiwara S, Ohtsuka M, Isotani A, Okabe M, Imai K, Takeda J.** 2009. A transposon-based chromosomal engineering method to survey a large cis-regulatory landscape in mice. *Nat Genet* **41:**946–952.

138. **Dupuy AJ, Fritz S, Largaespada DA.** 2001. Transposition and gene disruption in the male germline of the mouse. *Genesis* **30:**82–88.

139. **Ruf S, Symmons O, Uslu VV, Dolle D, Hot C, Ettwiller L, Spitz F.** 2011. Large-scale analysis of the regulatory architecture of the mouse genome with a transposon-associated sensor. *Nat Genet* **43:**379–386.

140. **Keng VW, Yae K, Hayakawa T, Mizuno S, Uno Y, Yusa K, Kokubu C, Kinoshita T, Akagi K, Jenkins NA, Copeland NG, Horie K, Takeda J.** 2005. Region-specific saturation germline mutagenesis in mice using the Sleeping Beauty transposon system. *Nat Methods* **2:**763–769.

141. **Yergeau DA, Kelley CM, Kuliyev E, Zhu H, Johnson Hamlet MR, Sater AK, Wells DE, Mead PE.** 2011. Remobilization of Sleeping Beauty transposons in the germline of Xenopus tropicalis. *Mobile DNA* **2:**15.

142. **Belzile F, Yoder JI.** 1992. Pattern of somatic transposition in a high copy Ac tomato line. *Plant J* **2:**173–179.

143. **Osborne BI, Corr CA, Prince JP, Hehl R, Tanksley SD, McCormick S, Baker B.** 1991. Ac transposition from a T-DNA can generate linked and unlinked clusters of insertions in the tomato genome. *Genetics* **129:**833–844.

144. Knapp S, Larondelle Y, Rossberg M, Furtek D, Theres K. 1994. Transgenic tomato lines containing Ds elements at defined genomic positions as tools for targeted transposon tagging. *Mol Gen Genet* **243:**666–673.

145. Kunze R, Weil CF. 2002. The hAT and CACTA superfamilies of plant transposons, p 565–610. *In* Craig NL, Craigie R, Gellert M, Lambowitz AM (ed), *Mobile DNA II*, ASM Press, Washington, DC.

146. Benjamin HW, Kleckner N. 1989. Intramolecular transposition by Tn10. *Cell* **59:**373–383.

147. Maxwell A, Craigie R, Mizuuchi K. 1987. B protein of bacteriophage mu is an ATPase that preferentially stimulates intermolecular DNA strand transfer. *Proc Natl Acad Sci U S A* **84:**699–703.

148. Garfinkel DJ, Stefanisko KM, Nyswaner KM, Moore SP, Oh J, Hughes SH. 2006. Retrotransposon suicide: formation of Ty1 circles and autointegration via a central DNA flap. *J Virol* **80:**11920–11934.

149. Shoemaker C, Hoffman J, Goff SP, Baltimore D. 1981. Intramolecular integration within Moloney murine leukemia virus DNA. *J Virol* **40:**164–172.

150. Fischer SE, van Luenen HG, Plasterk RH. 1999. Cis requirements for transposition of Tc1-like transposons in C. elegans. *Mol Gen Genet* **262:**268–274.

151. Lampe DJ, Grant TE, Robertson HM. 1998. Factors affecting transposition of the *Himar*1 *mariner* transposon *in vitro*. *Genetics* **149:**179–187.

152. Karsi A, Moav B, Hackett P, Liu Z. 2001. Effects of insert size on transposition efficiency of the *Sleeping Beauty* transposon in mouse cells. *Mar Biotechnol* **3:**241–245.

153. Chandler M, Clerget M, Galas DJ. 1982. The transposition frequency of IS1-flanked transposons is a function of their size. *J Mol Biol* **154:**229–243.

154. Morisato D, Way JC, Kim HJ, Kleckner N. 1983. Tn10 transposase acts preferentially on nearby transposon ends in vivo. *Cell* **32:**799–807.

155. Lee MS, Craigie R. 1994. Protection of retroviral DNA from autointegration: involvement of a cellular factor. *Proc Natl Acad Sci U S A* **91:**9823–9827.

156. Lee MS, Craigie R. 1998. A previously unidentified host protein protects retroviral DNA from autointegration. *Proc Natl Acad Sci U S A* **95:**1528–1533.

157. Suzuki Y, Craigie R. 2002. Regulatory mechanisms by which barrier-to-autointegration factor blocks autointegration and stimulates intermolecular integration of Moloney murine leukemia virus preintegration complexes. *J Virol* **76:**12376–12380.

158. Mansharamani M, Graham DR, Monie D, Lee KK, Hildreth JE, Siliciano RF, Wilson KL. 2003. Barrier-to-autointegration factor BAF binds p55 Gag and matrix and is a host component of human immunodeficiency virus type 1 virions. *J Virol* **77:**13084–13092.

159. Lin CW, Engelman A. 2003. The barrier-to-autointegration factor is a component of functional human immunodeficiency virus type 1 preintegration complexes. *J Virol* **77:**5030–5036.

160. Mates L, Izsvak Z, Ivics Z. 2007. Technology transfer from worms and flies to vertebrates: transposition-based genome manipulations and their future perspectives. *Genome Biol* **8 Suppl 1:**S1.

161. Ivics Z, Izsvak Z. 2006. Transposons for gene therapy! *Curr Gene Ther* **6:**593–607.

162. Ivics Z, Izsvák Z. 2004. Transposable elements for transgenesis and insertional mutagenesis in vertebrates: a contemporary review of experimental strategies. *Methods Mol Biol* **260:**255–276.

163. Ivics Z, Li MA, Mates L, Boeke JD, Nagy A, Bradley A, Izsvak Z. 2009. Transposon-mediated genome manipulation in vertebrates. *Nat Methods* **6:**415–422.

164. Izsvák Z, Ivics Z. 2004. Sleeping beauty transposition: biology and applications for molecular therapy. *Mol Ther* **9:**147–156.

165. Miskey C, Izsvak Z, Kawakami K, Ivics Z. 2005. DNA transposons in vertebrate functional genomics. *Cell Mol Life Sci* **62:**629–641.

166. Takeda J, Izsvak Z, Ivics Z. 2008. Insertional mutagenesis of the mouse germline with Sleeping Beauty transposition. *Methods Mol Biol* **435:**109–125.

167. VandenDriessche T, Ivics Z, Izsvak Z, Chuah MK. 2009. Emerging potential of transposons for gene therapy and generation of induced pluripotent stem cells. *Blood* **114:**1461–1468.

168. Ivics Z, Izsvak Z. 2010. The expanding universe of transposon technologies for gene and cell engineering. *Mobile DNA* **1:**25.

169. Ivics Z, Izsvak Z. 2011. Nonviral gene delivery with the sleeping beauty transposon system. *Hum Gene Ther* **22:**1043–1051.

170. Izsvák Z, Hackett PB, Cooper LJ, Ivics Z. 2010. Translating Sleeping Beauty transposition into cellular therapies: victories and challenges. *BioEssays* **32:**756–767.

171. Swierczek M, Izsvak Z, Ivics Z. 2012. The Sleeping Beauty transposon system for clinical applications. *Expert Opin Biol Ther* **12:**139–153.

172. Ammar I, Izsvak Z, Ivics Z. 2012. The Sleeping Beauty transposon toolbox. *Methods Mol Biol* **859:**229–240.