OPEN ACCESS

Repository of the Max Delbrück Center for Molecular Medicine (MDC)
Berlin (Germany)
http://edoc.mdc-berlin.de/10224/

Transposon-mediated genome manipulation in vertebrates

Zoltán Ivics, Meng Amy Li, Lajos Mátés, Jef D. Boeke, Andreas Nagy, Allan Bradley, and Zsuzsanna Izsvák

Published in final edited form as:
Nature Methods. 2009 Jun ; 6(6): 415-422 | doi: 10.1038/nmeth.1332
Nature Publishing Group (U.S.A.) ►

OPEN ACCESS

Final Draft

Transposon-mediated genome manipulation in vertebrates

Zoltán Ivics[1], Meng Amy Li[2], Lajos Mátés[1], Jef D. Boeke[3], Andras Nagy[4,5], Allan Bradley[2], and Zsuzsanna Izsvák[1,6]

[1] Max Delbrück Center for Molecular Medicine, Berlin, Germany
[2] Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK
[3] High Throughput Biology Center, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA
[4] Mount Sinai Hospital, Samuel Lunenfeld Research Institute, Toronto, Canada
[5] University of Toronto, Department of Molecular Genetics, Toronto, Canada
[6] Institute of Biochemistry, Biological Research Center of the Hungarian Academy of Sciences, Szeged, Hungary

ABSTRACT | Transposable elements are DNA segments with the unique ability to move about in the genome. This inherent feature can be exploited to harness these elements as gene vectors for genome manipulation. Transposon-based genetic strategies have been established in vertebrate species over the last decade, and current progress in this field suggests that transposable elements will serve as indispensable tools. In particular, transposons can be applied as vectors for somatic and germline transgenesis, and as insertional mutagens in both loss-of-function and gain-of-function forward mutagenesis screens. In addition, transposons will gain importance in future cell-based clinical applications, including nonviral gene transfer into stem cells and the rapidly developing field of induced pluripotent stem cells. Here we provide an overview of transposon-based methods used in vertebrate model organisms with an emphasis on the mouse system and highlight the most important considerations concerning genetic applications of the transposon systems.

Transposable elements and transposition

Transposable elements are mobile genetic elements of which two classes are distinguished based on their respective transposition mechanisms. The mobility of class I elements or retrotransposons is achieved through an RNA intermediate of a 'copy-and-paste' mechanism, whereas class II or DNA transposons use a DNA-mediated, 'cut-and-paste' mode of transposition (Fig. 1a). The most abundant transposons in mammals are non–long terminal repeat retrotransposons represented by the long interspersed nuclear elements and the short interspersed nuclear elements. The major long interspersed nuclear elements in humans and rodents (*LINE-1* or *L1*) contain two open reading frames (Fig. 1b). These encode a nucleic acid binding protein and an enzyme with endonuclease and reverse transcriptase activity, respectively [1,2]. Endonuclease generates a single-stranded nick in the target DNA, and reverse transcriptase uses the nicked DNA to prime reverse transcription from the 3' end of the *L1* RNA [3,4].

Class II transposable elements that move in the host genome via a 'cut-and-paste' mechanism are simply organized; they encode a transposase protein in their simple genome flanked by inverted terminal repeats that carry transposase binding sites necessary for transposition (Fig. 1b). Transposition results in excision of the element from the DNA and subsequent integration into a new sequence environment. The transpositional process can easily be controlled by separating the transposase source from the transposable DNA, thereby creating a non-autonomous transposable element (Fig. 1c). In such a two-component system, the transposon can only move by *trans*-supplementing the transposase protein.

Transposons as DANN delivery tools

Transposons have been successfully used in invertebrate animal models, including *C. elegans* [5] and *Drosophila* [6] for transgenesis and insertional mutagenesis, but until the reactivation of the *Sleeping Beauty* (*SB*) transposon system in 1997 [7], there was no indication of DNA-based transposons in vertebrates sufficiently active for these purposes. Later on, other elements have been shown to catalyze efficient transposition in vertebrate model organisms; their characteristics are summarized in Table 1. The basic criteria for the applicability of a class II transposable element in any given model organism are (i) sufficient transpositional activity in the given species and (ii) a lack of endogenous copies in the targeted genome or other strategies that avoid mobilization of resident copies. Other practical considerations for the design of a particular gene transfer experiment include cargo capacity of the transposable element and integration site preference [8].

Capacity for cargo

Tolerance for cargo size varies greatly between transposable elements (Table 1). Members of the *Tc1*/*mariner* family, including *SB*, are inhibited by large size [9]. A particular modification of the *SB* transposon in this respect was the generation of a 'sandwich' transposon vector that has two complete *SB* elements flanking a transgene to be mobilized [10]. The sandwich *SB* vector enhanced transposition of large (>10 kb) transgene constructs and therefore probably is the method of choice for transgene constructs that would otherwise transpose poorly owing to their large size. The *piggyback* [11], *Tol1* [12] and *Tol2* [13,14] transposons appear to be more tolerant to larger cargo, allowing complex transgene designs to be incorporated within the transposon without sacrificing transposition efficiency.

Integration site preference

Where the transposon inserts can greatly influence the utility of transposon vectors for different applications. For example, human gene therapy protocols would require application of transposon vectors showing the least preference to target genes, for obvious safety considerations. On the contrary, mutagenesis screens can capitalize on elements that tend to land in genes. The insertion pattern of most transposons is nonrandom, with many 'hotspots' and 'cold regions' on a genome-wide

OPEN ACCESS

scale. Common hotspots represent the main limitation to full genome coverage with individual transposable element–based vectors. The preferences of particular elements to integrate into expressed genes versus noncoding DNA, and preferences for integration sites within genes are expected to be substantially different. Thus, in this respect, the utility of transposons for mutagenesis is greatly enhanced by the availability of multiple, alternative vector systems with distinct preferences for insertion (Table 1).

Transposons as vectors for transgenesis

Stable gene transfer into stem cells and generation of iPSCs

Transposon-based technologies can be used for gene transfer in cultured cells. For example, to integrate plasmid-based short hairpin RNA expression cassettes into chromosomes to obtain stable knockdown cell lines by RNA interference [15]. Furthermore, transposons are promising vectors for therapeutic gene delivery to facilitate clinical implementation of gene- and cell-based therapies [16]. A genetic screen in mammalian cells yielded a hyperactive SB transposase (SB100X) with ~100-fold enhancement in efficiency when compared to the first-generation transposase. SB100X supported 35–50% stable gene transfer in human $CD34^+$ hematopoietic stem or progenitor cells that were proficient in multilineage hematopoietic reconstitution after transplantation into immunodeficient mice [17]. Thus, the efficiency of stable gene transfer by this hyperactive SB system approaches that of viral methods and therefore may be developed into an efficient, simple and cheap method for genetic manipulation of stem cells and other primary cell types.

The recent discovery of induced pluripotent stem cells (iPSCs) holds enormous promise for future regenerative medicine. By expressing only four genes (encoding transcription factors Oct4, Sox2, Klf4 and c-Myc), somatic cells can be transformed to a pluripotent state with a developmental capacity similar to that of embryonic stem cells [18]. Initially, this could only be achieved by retro- or lentiviral transduction. However, owing to safety issues, permanent viral insertions limit the value of the resulting iPSCs for clinical applications. Two special features of the transposon systems make them a promising alternative. First, the efficiency of transgene insertion by the *piggyBac* transposon [19] and by the recently developed hyperactive variant of *SB* [17] is comparable to that of viral transduction. Second, because transposon excision is not always followed by pasting into a new genomic location, the 'cut' component of the transposition offers removal of the transgenes after completion of reprogramming. Transposition-mediated generation of mouse and human iPSCs and traceless removal of the reprogramming factors from the pluripotent cells have already been achieved by the *piggyBac* system [20,21]. One caveat that still remains is the possibility of the transposon to jump into a new location during the factor removal process. A way to solve this problem would be to develop a transposase that allows cutting but is deficient in pasting. Nevertheless, the transposon system–assisted reprogramming is looking forward to a bright future in regenerative medicine.

Transgenesis in oocytes and embryos

Classical methods to stably express foreign genes in vertebrates rely on microinjection of gene constructs into oocytes or fertilized eggs. Three main drawbacks of this method are the low rate of genomic integration (<10%), the integration of injected DNA as a concatemer that is prone to silencing [22], and that founders are predominantly mosaic for the transgene because integration generally occurs relatively late during embryonic development. All of these drawbacks can be circumvented by using transposition-mediated gene delivery as it increases the efficiency of chromosomal integration and facilitates single-copy insertion events. The injection of *in vitro*–synthesized mRNA as a transposase source can further enhance the efficiency of this technique because of the more rapid availability of the transposase, resulting in reduced transgene mosaicism in the embryo and therefore elevated germline transmission rates. This method has been used for germline transgenesis in *Ciona intestinalis* with *Minos* [23], to generate transgenic zebrafish with *Tol1* [24], transgenic zebrafish, medaka fish and *Xenopus* with *Tol2* [25,26] and *SB* [27-30], transgenic chicken with *Tol2* [31] and transgenic mice with *SB* [32-34] and *piggyback* [11]. The recently developed hyperactive SB100X transposase yields average transgenic frequencies of 45% in mouse embryos [17], indicating that this hyperactive transposon system can be developed into a simple and efficient tool for transgenesis in vertebrates.

Transposon-mediated forward genetic approaches

Gene targeting has been very successful in generating altered alleles of specific genes, particularly in embryonic stem cells, allowing individual gene function to be dissected at the cellular and whole-animal levels. However, this gene-by-gene approach does not facilitate gene discoveries related to a particular pathway of interest on a genome-wide scale. Genome-wide, forward insertional mutagenesis provides a powerful and high-throughput means to ascribe functions to genes associated with particular biological pathways. Insertional mutagenesis using engineered transposable elements can be one of the most productive and versatile approaches to disrupt and manipulate genes on a genome-wide scale. However, even if a transposable element inserts into a gene, it may not have a mutagenic effect. For example, intronic insertions are likely spliced out without having an effect on gene expression (Fig. 2a). Thus, various technologies have been established to enhance the mutagenicity as well as reporting capabilities of insertional vectors by 'trapping' transcription units (Fig. 2b–e).

Resessive genetic screens in embryonic stem cells

In cell culture systems, transposon delivery can be achieved by transfection of plasmid DNA containing the transposons or by mobilizing a chromosomally located transposon that has been placed in the genome by gene targeting or a prior transposition event. Transfection-based, 'plasmid-to-genome' delivery (Fig. 3a) yields relatively unbiased genome-wide integrations; however, careful titration of the amount of the donor and transposase plasmid is required to provide the appropriate

OPEN ACCESS

copy number per cell of the transposon. Intragenomic, 'genome-to-genome' mobilization (Fig. 3b) can be selected based on excision and reintegration [35], allowing efficient genome-wide mutagenesis and tight control over the copy number. Pluripotent mouse embryonic stem cells are attractive models for *in vitro* mutagenesis because they can differentiate into many cell types including the germline and because they are amenable to sophisticated genetic manipulation.

However, insertional mutagenesis in somatic cells is challenged by the diploid genome. Inactivation of both copies of a gene is nearly always required to evoke a phenotypic change, but the probability of generating bi-allelic mutations of a single locus by two independent 'hits' is extremely low. However, a system has been developed in embryonic stem cells that combines insertional mutagenesis with a *Blm*-deficient genetic background. *Blm*-deficient embryonic stem cells have a high rate of homologous recombination between homologous chromosomes, thereby promoting the conversion of single-allele mutations to bi-allelic mutations by loss of heterozygosity [36,37] (Supplementary Fig. 1 online). A typical recessive genetic screen using *Blm*-deficient embryonic stem cells (Supplementary Fig. 1) includes: (i) genome-wide mutagenesis and selection of mutants with the insertional mutagen; (ii) mutant pool propagation to provide sufficient generations for homozygote conversion; (iii) phenotype-driven screening of the biological pathway of interest to isolate candidate mutants; and (iv) candidate validation in terms of mutant locus identification, homozygosity status, phenotype rescue and functional relevance to the biology of interest. As discussed above, DNA transposons such as *piggyBac* have been shown to have a more random genome-wide distribution than retroviruses [38]. As a proof of principle, *piggyBac* mutagenesis in *Blm*-deficient embryonic stem cells was followed by screening for components involved in the DNA mismatch repair pathway. Four known components of the mismatch repair pathway genes were recovered [39], whereas in a previous retrovirus-based screen, only one known component and a new gene were isolated [40].

Recessive, loss-of-function screens in the germline in vivo

Screens using DNA transposons

In a two-component experimental system, in which transposition is controlled by *trans*-supplementation of the transposase (Fig. 1c), two transgenic stocks are generated: a 'jumpstarter' strain that expresses the transposase and a 'mutator' strain that typically carries nonautonomous transposons equipped with gene trap cassettes [41] (Fig. 4). These two stocks are crossed to bring the two components of the transposon system together, and transposition of the gene trap transposons is expected to occur in the sperm cells of F1 double-transgenic males (referred to as 'seed' mice; Fig. 4). Such males are repeatedly crossed to wild-type females to segregate the different insertion events in their sperm cells in separate F2 animals (Fig. 4). For the detection of gene trap insertions *in vivo*, fluorescent reporters such as GFP have been widely used. Transposon insertion sites can easily be established by PCR protocols from genomic DNA isolated from GFP-positive pups.

SB has been successfully used in mice that expressed the transposase either ubiquitously [42-45] or in the male germline [46]. Recently, *SB*-based insertional mutagenesis was also established in the rat by using essentially the same experimental approach [47,48]. In the mouse system, up to 90% of the F2 progeny can carry transposon insertions [43], and a single sperm of a seed male can contain, on average, two insertion events [42]. The germline of such a founder was estimated to contain approximately 10,000 different mutations [44]. Notably, transposition of gene trap transposons identified mouse genes with ubiquitous and tissue-specific expression patterns, and mutant or lethal phenotypes were easily obtained by generating homozygous mice [44,45,49,50]. Other studies [51] showed that local saturation mutagenesis of a genomic region is a realistic goal using the *SB* transposon system with a chromosomally resident transposon donor site. Insertional mutagenesis with *SB* in the germline of mice and rats has been approached with mutator lines containing transposon donor loci containing many (up to several hundred) copies of the transposon vector in the form of concatemeric arrays [42,44,47,48]. However, recombination between newly transposed transposon copies and the donor concatemer could lead to unwanted genomic rearrangements [50]. The SB100X hyperactive transposase is ~120-fold more active in mobilizing single-copy, chromosomally resident transposons than the wild-type transposase [17] and may thus potentially eliminate the need for concatemeric donor sites in genetic screens.

The *Minos* transposon has also been shown to mobilize in mice by transposase expression in oocytes using *Zp3* [52] and in lymphocytes using *Cd2* promoters [53]. *PiggyBac* has been used in mice [11,54], and the activity of the *Tol2* element has already been demonstrated in mouse embryonic stem cells [55] and *in vivo* in the mouse liver [13]. In zebrafish, *SB* and *Tol2* have been shown to be useful for insertional mutagenesis in coinjection experiments [27,56,57]. As discussed above, the availability of a battery of vector systems based on diverse transposable elements will undoubtedly increase genome coverage in mutagenesis screens.

Screens using the L1 retrotransposon

Mutagenesis screens using the *L1* retrotransposon in mice are similar to the scheme shown in Figure 4, except that only a single transgenic stock carrying a transcriptionally (and hence transpositionally) active *L1* element needs to be established. One system, described in detail, is based on an *ORFeus* transgene driven by a constitutive promoter and marked by a retrotransposition indicator cassette, in which a GFP marker is disrupted by an intron (Supplementary Fig. 2 online). Germline insertion frequency was estimated to be about 30%, and the genomic distribution of *de novo* retrotransposon insertions revealed ~28% of the events occurring in RefSeq genes and a uniform ditribution of intragenic insertions along the targeted genes [60].

From the perspective of their use as mutagenesis tools, *L1* retrotransposons have several potential advantages. (i) Because of their 'copy-and-paste' mechanism of retrotransposition, the donor copy of the element is stable. (ii) Since donor elements can be driven by cellular promoters that are not transcribed, it is possible to design them so that they transpose only once (Supplementary Fig. 2). (iii) Retrotransposition can be controlled by the

OPEN ACCESS

extent of RNA expression, for example, by using Cre-*loxP* technology [61]. One feature of *L1* that can pose problems to its implementation as a mutagenesis tool is that ~90% of the progeny transposition events are associated with rearrangements (typically, 5' truncations), and the complexities of these structures can create problems in determining sites of new insertion.

Dominant, gain-of-function screens in the soma

Targeted over- and/or misexpression screens in somatic tissues of mice using *SB* has been shown to be especially useful for the generation of experimental cancers in animal models. Though this approach is similar to the application of retroviruses, transposable elements allow the recovery of tumors in tissues previously not amenable to such genomic approaches, including the liver and the brain. The 'oncogene trap' *SB* transposon (Fig. 2e) can induce loss-of-function mutations in tumor suppressor genes as well as gain-of-function overexpression of proto-oncogenes near the genomic insertion sites. Mutator lines with transposon donor loci containing many copies (25–358 copies) of the transposon vector in the form of concatemeric arrays [62,63] are crossed with stocks that express the SB transposase to generate double-transgenic mice (Supplementary Fig. 3 online). Experimental tumors develop in somatic tissues of the double-transgenic mice as a result of dominant mutations. In the published studies, somatic mobilization of the oncogene trap transposons accelerated tumor formation (mostly sarcomas) in a p19Arf-deficient cancer-predisposed genetic background [62] as well as the formation of leukemia and medulloblastoma in wild-type mice [63]. The next step in the procedure is to isolate the transposon insertions from tumor samples (Supplementary Fig. 3) by using high-throughput PCR methods [64] and determine which one(s) are causative with respect to tumor formation by using common insertion site analysis. This analysis identifies repeated occurrence of insertions in particular genes in independent tumor samples. Candidate oncogenes are validated by transgenic models (Supplementary Fig. 3), for which transposons can be applied as powerful gene vectors [65]. To devise customized screens for cancer development, a current approach is to establish mouse lines conditionally expressing the transposase by using Cre recombinase–inducible transposase alleles. This approach has been elegantly applied to conditionally express the SB transposase in the liver and in the epithelial cells of the gastrointestinal tract in experimental mice by tissue-specific expression of Cre [66,67]. The screens yielded genetic loci associated with hepatocellular carcinoma and colorectal cancer, and several of these are potential new targets for therapeutic intervention.

Projections

One obvious immediate application of transposon-based technologies is germline transgenesis in laboratory animals. Transgenic animal facilities worldwide could immediately adapt their standard operating procedure to the use of transposon-based plasmid vectors for highly efficient and reliable production of laboratory stocks. Furthermore, the current transposon technologies are immediately testable in large animal species of agricultural and biotechnological importance, including cattle, sheep and pig. The recently developed SB100X hyperactive transposon system yields unprecedented stable gene transfer efficiencies after nonviral gene delivery into therapeutically relevant primary cell types, including stem cells, and thus may facilitate the clinical implementation of *ex vivo* and *in vivo* gene therapies. Additionally, it is now becoming amenable to create libraries of gene knockouts and to thereby establish new models of human disease for therapeutic and pharmaceutical intervention in species in which embryonic stem cell and homologous recombination–based knockout technology has not been established. For example, quantitative trait loci implicated in cardiovascular diseases could be dissected in the future using transposon-mediated insertional mutagenesis in the rat system, the preferred model for cardiovascular biology [68]. Finally, recent advances in iPSC reprogramming should facilitate the identification of genetic determinants involved in physiological or pathological pathways in cells derived from patients with specific genetic diseases [69]. Thus, transposon-based technologies have enormous potential to develop powerful genomic tools with the vision of creating a bridge between physiology and genetics.

Corresponding Author

Zoltán Ivics, zivics@mdc-berlin.de

References

1. Feng, Q., Moran, J.V., Kazazian, H.H., Jr. & Boeke, J.D. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell 87, 905–916 (1996).

2. Mathias, S.L., Scott, A.F., Kazazian, H.H. Jr., Boeke, J.D. & Gabriel, A. Reverse transcriptase encoded by a human transposable element. Science 254, 1808–1810 (1991).

3. Luan, D.D., Korman, M.H., Jakubczak, J.L. & Eickbush, T.H. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. Cell 72, 595–605 (1993).

4. Cost, G.J., Feng, Q., Jacquier, A. & Boeke, J.D. Human L1 element target-primed reverse transcription *in vitro*. EMBO J. 21, 5899–5910 (2002).

5. Zwaal, R.R., Broeks, A., van Meurs, J., Groenen, J.T. & Plasterk, R.H. Target-selected gene inactivation in *Caenorhabditis elegans* by using a frozen transposon insertion mutant bank. Proc. Natl. Acad. Sci. USA 90, 7431–7435 (1993).

6. Thibault, S.T. *et al*. A complementary transposon tool kit for *Drosophila melanogaster* using *P* and *piggyBac*. Nat. Genet. 36, 283–287 (2004).

   Demonstration that full genome coverage in *Drosophila* cannot be done by using a single transposon owing to distinct preferences for genomic insertion sites.

7. Ivics, Z., Hackett, P.B., Plasterk, R.H. & Izsvak, Z. Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. Cell 91, 501–510 (1997).

   First report on the resurrection of an extinct transposon and the first class II transposon ever reported to jump in a vertebrate cell.

8. Mates, L., Izsvak, Z. & Ivics, Z. Technology transfer from worms and flies to vertebrates: transposition-based genome

manipulations and their future perspectives. Genome Biol. 8 (Suppl. 1), S1 (2007).

9. Izsvak, Z., Ivics, Z. & Plasterk, R.H. Sleeping Beauty, a wide host-range transposon vector for genetic transformation in vertebrates. J. Mol. Biol. 302, 93–102 (2000).

10. Zayed, H., Izsvak, Z., Walisko, O. & Ivics, Z. Development of hyperactive *Sleeping Beauty* transposon vectors by mutational analysis. Mol. Ther. 9, 292–304 (2004).

11. Ding, S. *et al*. Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. Cell 122, 473–483 (2005).

Report describes efficient transposition of the insect *piggyBac* transposon in mice.

12. Koga, A. *et al*. The Tol1 element of medaka fish is transposed with only terminal regions and can deliver large DNA fragments into the chromosomes. J. Hum. Genet. 52, 1026–1030 (2007).

13. Balciunas, D. *et al*. Harnessing a high cargo-capacity transposon for genetic applications in vertebrates. PLoS Genet. 2, e169 (2006).

14. Urasaki, A., Morvan, G. & Kawakami, K. Functional dissection of the Tol2 transposable element identified the minimal cis-sequence and a highly repetitive sequence in the subterminal region essential for transposition. Genetics 174, 639–649 (2006).

15. Kaufman, C.D., Izsvak, Z., Katzer, A. & Ivics, Z. *Frog Prince* transposon-based RNAi vectors mediate efficient gene knockdown in human cells. J. RNAi Gene Silencing 1, 97–104 (2005).

16. Ivics, Z. & Izsvak, Z. Transposons for gene therapy! Curr. Gene Ther. 6, 593–607 (2006).

17. Mátés, L. *et al*. Molecular evolution of a hyperactive *Sleeping Beauty* transposase enables robust stable gene transfer in vertebrates. Nat. Genet. advance online publication, doi: doi:10.1038/ng.343 (3 May 2009).

Report of a hyperactive *Sleeping Beauty* transposase that mediates stable gene transfer at an efficiency similar to that of viral vectors in primary human stem cells and in mice.

18. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126, 663–676 (2006).

19. Cadinanos, J. & Bradley, A. Generation of an inducible and optimized piggyBac transposon system. Nucleic Acids Res. 35, e87 (2007).

20. Woltjen, K. *et al*. *piggyBac* transposition reprograms fibroblasts to induced pluripotent stem cells. Nature 458, 766–770 (2009).

This and the next paper show how to use the *piggyBac* transposon as a vector for iPSC reprogramming.

21. Yusa, K., Rad, R., Takeda, J. & Bradley, A. Generation of transgene-free induced pluripotent mouse stem cells by the piggyBac transposon. Nat. Methods 6, 363–369 (2009).

22. Henikoff, S. Conspiracy of silence among repeated transgenes. Bioessays 20, 532–535 (1998).

23. Sasakura, Y., Awazu, S., Chiba, S. & Satoh, N. Germ-line transgenesis of the Tc1/*mariner* superfamily transposon *Minos* in *Ciona intestinalis*. Proc. Natl. Acad. Sci. USA 100, 7726–7730 (2003).

24. Koga, A., Cheah, F.S., Hamaguchi, S., Yeo, G.H. & Chong, S.S. Germline transgenesis of zebrafish using the medaka Tol1 transposon system. Dev. Dyn. 237, 2466–2474 (2008).

25. Kawakami, K., Shima, A. & Kawakami, N. Identification of a functional transposase of the Tol2 element, an Ac-like element from the Japanese medaka fish, and its transposition in the zebrafish germ lineage. Proc. Natl. Acad. Sci. USA 97, 11403–11408 (2000).

26. Hamlet, M.R. *et al*. Tol2 transposon-mediated transgenesis in *Xenopus tropicalis*. Genesis 44, 438–445 (2006).

27. Davidson, A.E. *et al*. Efficient gene delivery and gene expression in zebrafish using the *Sleeping Beauty* transposon. Dev. Biol. 263, 191–202 (2003).

28. Hermanson, S., Davidson, A.E., Sivasubbu, S., Balciunas, D. & Ekker, S.C. Sleeping Beauty transposon for efficient gene delivery. Methods Cell Biol. 77, 349–362 (2004).

29. Sinzelle, L. *et al*. Generation of trangenic *Xenopus laevis* using the Sleeping Beauty transposon system. Transgenic Res. 15, 751–760 (2006).

30. Grabher, C. *et al*. Transposon-mediated enhancer trapping in medaka. Gene 322, 57–66 (2003).

31. Sato, Y. *et al*. Stable integration and conditional expression of electroporated transgenes in chicken embryos. Dev. Biol. 305, 616–624 (2007).

32. Dupuy, A.J. *et al*. Mammalian germ-line transgenesis by transposition. Proc. Natl. Acad. Sci. USA 99, 4495–4499 (2002).

33. Wilber, A. *et al*. RNA as a source of transposase for sleeping beauty-mediated gene insertion and expression in somatic cells and tissues. Mol. Ther. 13, 625–630 (2006).

34. Carlson, C.M., Frandsen, J.L., Kirchhof, N., McIvor, R.S. & Largaespada, D.A. Somatic integration of an oncogene-harboring Sleeping Beauty transposon models liver tumor development in the mouse. Proc. Natl. Acad. Sci. USA 102, 17059–17064 (2005).

35. Luo, G., Ivics, Z., Izsvak, Z. & Bradley, A. Chromosomal transposition of a Tc1/mariner-like element in mouse embryonic stem cells. Proc. Natl. Acad. Sci. USA 95, 10769–10773 (1998).

36. Luo, G. *et al*. Cancer predisposition caused by elevated mitotic recombination in Bloom mice. Nat. Genet. 26, 424–429 (2000).

37. Yusa, K. *et al*. Genome-wide phenotype analysis in ES cells by regulated disruption of Bloom's syndrome gene. Nature 429, 896–899 (2004).

38. Wang, W. *et al*. Chromosomal transposition of PiggyBac in mouse embryonic stem cells. Proc. Natl. Acad. Sci. USA 105, 9290–9295 (2008).

39. Wang, W., Bradley, A. & Huang, Y. A piggyBac transposon-based genome-wide library of insertionally mutated Blm-deficient murine ES cells. Genome Res (2009).

40. Hansen, G.M. *et al*. Large-scale gene trapping in C57BL/6N mouse embryonic stem cells. Genome Res. 18, 1670–1679 (2008).

41. Takeda, J., Izsvak, Z. & Ivics, Z. Insertional mutagenesis of the mouse germline with Sleeping Beauty transposition. Methods Mol. Biol. 435, 109–125 (2008).

42. Dupuy, A.J., Fritz, S. & Largaespada, D.A. Transposition and gene disruption in the male germline of the mouse. Genesis 30, 82–88 (2001).

43. Horie, K. *et al*. Efficient chromosomal transposition of a Tc1/mariner- like transposon Sleeping Beauty in mice. Proc. Natl. Acad. Sci. USA 98, 9191–9196 (2001).

44. Horie, K. *et al*. Characterization of Sleeping Beauty transposition and its application to genetic screening in mice. Mol. Cell. Biol. 23, 9189–9207 (2003).

OPEN ACCESS

45. Carlson, C.M. *et al*. Transposon mutagenesis of the mouse germline. Genetics 165, 243–256 (2003).

46. Fischer, S.E., Wienholds, E. & Plasterk, R.H. Regulated transposition of a fish transposon in the mouse germ line. Proc. Natl. Acad. Sci. USA 98, 6759–6764 (2001).

47. Kitada, K. *et al*. Transposon-tagged mutagenesis in the rat. Nat. Methods 4, 131–133 (2007).

48. Lu, B. *et al*. Generation of rat mutants using a coat color-tagged Sleeping Beauty transposon system. Mamm. Genome 18, 338–346 (2007).

49. Yae, K. *et al*. Sleeping beauty transposon-based phenotypic analysis of mice: lack of Arpc3 results in defective trophoblast outgrowth. Mol. Cell. Biol. 26, 6185–6196 (2006).

50. Geurts, A.M. *et al*. Gene mutations and genomic rearrangements in the mouse as a result of transposon mobilization from chromosomal concatemers. PLoS Genet. 2, e156 (2006).

51. Keng, V.W. *et al*. Region-specific saturation germline mutagenesis in mice using the Sleeping Beauty transposon system. Nat. Methods 2, 763–769 (2005).

This paper shows that local hopping of the *Sleeping Beauty* transposon can be applied for chromosomal region-specific saturation mutagenesis in mice.

52. Drabek, D. *et al*. Transposition of the *Drosophila hydei* Minos transposon in the mouse germ line. Genomics 81, 108–111 (2003).

53. Zagoraiou, L. *et al*. In vivo transposition of Minos, a *Drosophila* mobile element, in mammalian tissues. Proc. Natl. Acad. Sci. USA 98, 11474–11478 (2001).

54. Wu, S., Ying, G., Wu, Q. & Capecchi, M.R. Toward simpler and faster genome-wide mutagenesis in mice. Nat. Genet. 39, 922–930 (2007).

55. Kawakami, K. & Noda, T. Transposition of the Tol2 element, an Ac-like element from the Japanese medaka fish Oryzias latipes, in mouse embryonic stem cells. Genetics 166, 895–899 (2004).

56. Kawakami, K. *et al*. A transposon-mediated gene trap approach identifies developmentally regulated genes in zebrafish. Dev. Cell 7, 133–144 (2004).

57. Parinov, S., Kondrichin, I., Korzh, V. & Emelyanov, A. Tol2 transposon-mediated enhancer trap to identify developmentally regulated zebrafish genes in vivo. Dev. Dyn. 231, 449–459 (2004).

58. Sivasubbu, S. *et al*. Gene-breaking transposon mutagenesis reveals an essential role for histone H2afza in zebrafish larval development. Mech. Dev. 123, 513–529 (2006).

59. Balciunas, D. *et al*. Enhancer trapping in zebrafish using the Sleeping Beauty transposon. BMC Genomics 5, 62 (2004).

60. An, W. *et al*. Active retrotransposition by a synthetic L1 element in mice. Proc. Natl. Acad. Sci. USA 103, 18662–18667 (2006).

61. An, W. *et al*. Conditional activation of a single-copy L1 transgene in mice by Cre. Genesis 46, 373–383 (2008).

62. Collier, L.S., Carlson, C.M., Ravimohan, S., Dupuy, A.J. & Largaespada, D.A. Cancer gene discovery in solid tumours using transposon-based somatic mutagenesis in the mouse. Nature 436, 272–276 (2005).

This and the following paper demonstrated that the *Sleeping Beauty* transposon can be used as an efficient somatic mutagen to drive tumorigenesis in mice.

63. Dupuy, A.J., Akagi, K., Largaespada, D.A., Copeland, N.G. & Jenkins, N.A. Mammalian mutagenesis using a highly mobile somatic Sleeping Beauty transposon system. Nature 436, 221–226 (2005).

64. Largaespada, D.A. & Collier, L.S. Transposon-mediated mutagenesis in somatic cells: identification of transposon-genomic DNA junctions. Methods Mol. Biol. 435, 95–108 (2008).

65. Su, Q. *et al*. A DNA transposon-based approach to validate oncogenic mutations in the mouse. Proc. Natl. Acad. Sci. USA 105, 19904–19909 (2008).

66. Keng, V.W. *et al*. A conditional transposon-based insertional mutagenesis screen for genes associated with mouse hepatocellular carcinoma. Nat. Biotechnol. 27, 264–274 (2009).

67. Starr, T.K. *et al*. A transposon-based genetic screen in mice identifies genes altered in colorectal cancer. Science 323, 1747–1750 (2009).

68. Aitman, T.J. *et al*. Progress and prospects in rat genetics: a community view. Nat. Genet. 40, 516–522 (2008).

69. Park, I.H. *et al*. Disease-specific induced pluripotent stem cells. Cell 134, 877–886 (2008).

70. Yant, S.R. *et al*. High-resolution genome-wide mapping of transposon integration in mammals. Mol. Cell. Biol. 25, 2085–2094 (2005).

71. Miskey, C., Izsvak, Z., Plasterk, R.H. & Ivics, Z. The Frog Prince: a reconstructed transposon from Rana pipiens with high transpositional activity in vertebrate cells. Nucleic Acids Res. 31, 6873–6881 (2003).

72. Miskey, C. *et al*. 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 (2007).

73. Clark, K.J., Carlson, D.F., Leaver, M.J., Foster, L.K. & Fahrenkrug, S.C. Passport, a native Tc1 transposon from flatfish, is functionally active in vertebrate cells. Nucleic Acids Res. 37, 1239–1247 (2009).

74. Wilson, M.H., Coates, C.J. & George, A.L., Jr. PiggyBac transposon-mediated gene transfer in human cells. Mol. Ther. 15, 139–145 (2007).

75. Emelyanov, A., Gao, Y., Naqvi, N.I. & Parinov, S. Trans-kingdom transposition of the maize dissociation element. Genetics 174, 1095–1104 (2006).

76. Sinzelle, L. *et al*. Transposition of a reconstructed Harbinger element in human cells and functional homology with two transposon-derived cellular genes. Proc. Natl. Acad. Sci. USA 105, 4715–4720 (2008).

77. Han, J.S. & Boeke, J.D. A highly active synthetic mammalian retrotransposon. Nature 429, 314–318 (2004).

The base composition of an *L1* element was altered leading to highly efficient retrotransposition in mammalian cells.

78. Cost, G.J. & Boeke, J.D. Targeting of human retrotransposon integration is directed by the specificity of the L1 endonuclease for regions of unusual DNA structure. Biochemistry 37, 18081–18093 (1998).

OPEN ACCESS

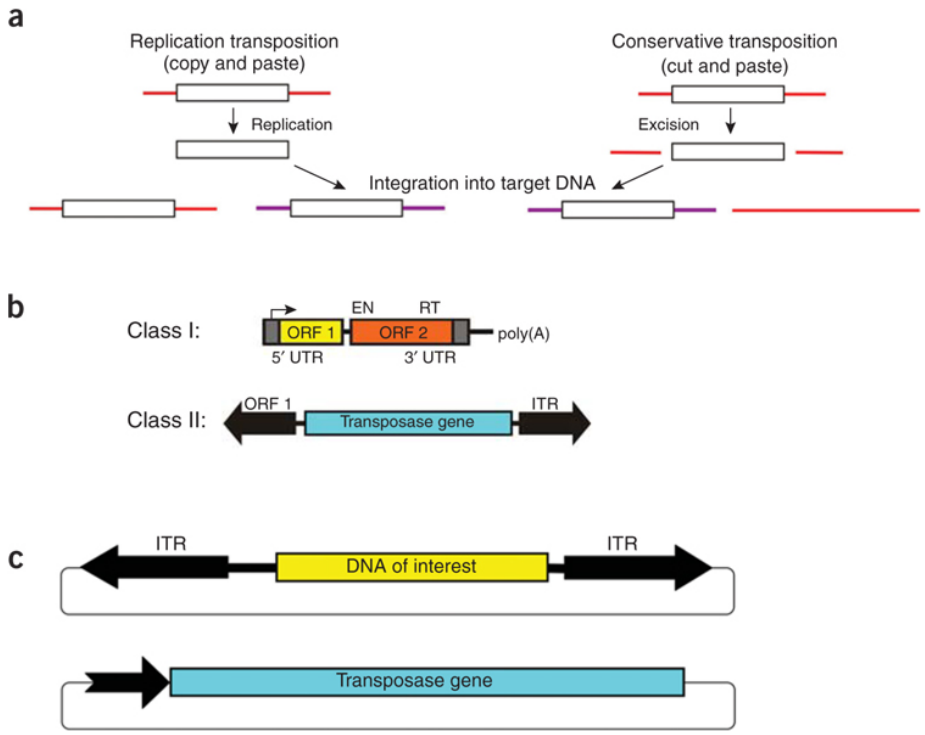

Figure 1. Mechanism of transposition and general organization of class I and class II transposable elements (**a**) Replicative transposition involves amplification of the element by copying through transcription followed by reverse transcription. The newly synthesized copy is inserted elsewhere in the genome, but the donor element does not move. During conservative transposition, the element is excised from the donor DNA (red), and integrates into a new target DNA (purple). (**b**) Class I non–long terminal repeat retrotransposons consist of a 5' untranslated region (UTR) that has promoter activity (arrow) that drives transcription of the element-encoded genes. Open reading frame (ORF) 1 encodes a nucleic acid binding protein. ORF 2 encodes an endonuclease (EN) and a reverse transcriptase (RT). The element has a poly(A) tail. Class II DNA transposons contain a central transposase gene flanked by inverted terminal repeats (ITRs). (**c**) In a gene transfer vector system based on a class II DNA transposon, the transposase coding region can be replaced by a DNA of interest. 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.

OPEN ACCESS

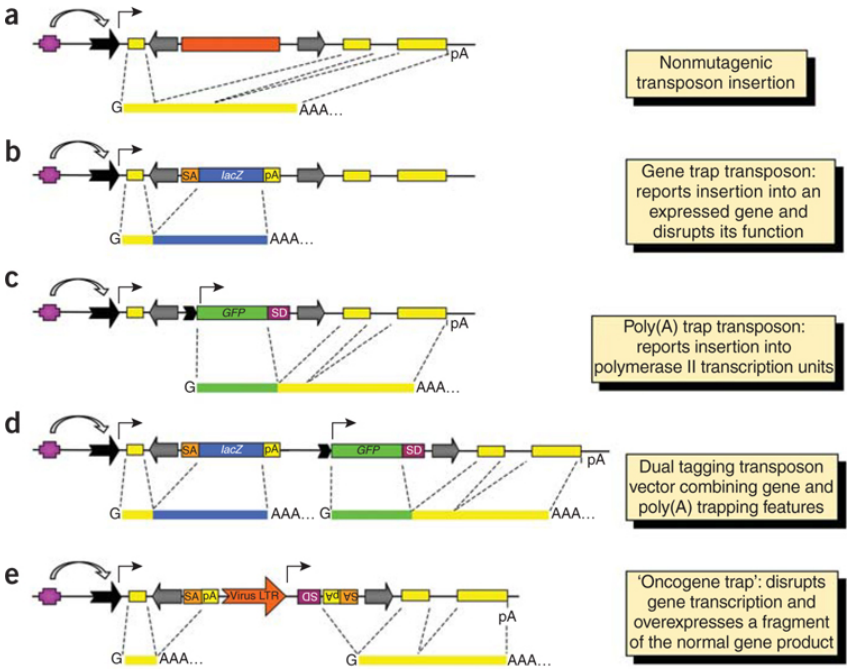

Figure 2. Summary of the basic gene trapping strategies. (**a**) A hypothetical transcription unit is depicted with an upstream regulatory element (purple), a promoter (black arrow), three exons (yellow) and a polyadenylation signal (pA). G denotes the 5' guanine cap, and AAA... indicates the poly(A) tail. An intronic transposon insertion is typically not mutagenic, because the transposon is spliced out from the primary RNA transcript together with the targeted intron sequences. Transposon inverted repeats are indicated by gray arrows. (**b**) Gene trapping cassettes contain a splice acceptor (SA) followed by a reporter gene and a pA. Thus, the expression of the reporter follows the expression pattern of the trapped gene. (**c**) Polyadenylation traps contain a promoter followed by a reporter gene and a splice donor (SD) site, but they lack a pA signal. Therefore, reporter gene expression depends on splicing to downstream exons of a polymerase II transcription unit containing a pA. (**d**) The 'dual tagging' vectors are based on both gene and poly(A) trapping of a targeted transcription unit. (**e**) The oncogene trap contains SA signals followed by pA signals in both orientations to disrupt transcription, as well as a strong, viral enhancer and promoter that drives transcription toward the outside of an inserted transposon and thereby overexpresses a gene product.

OPEN ACCESS

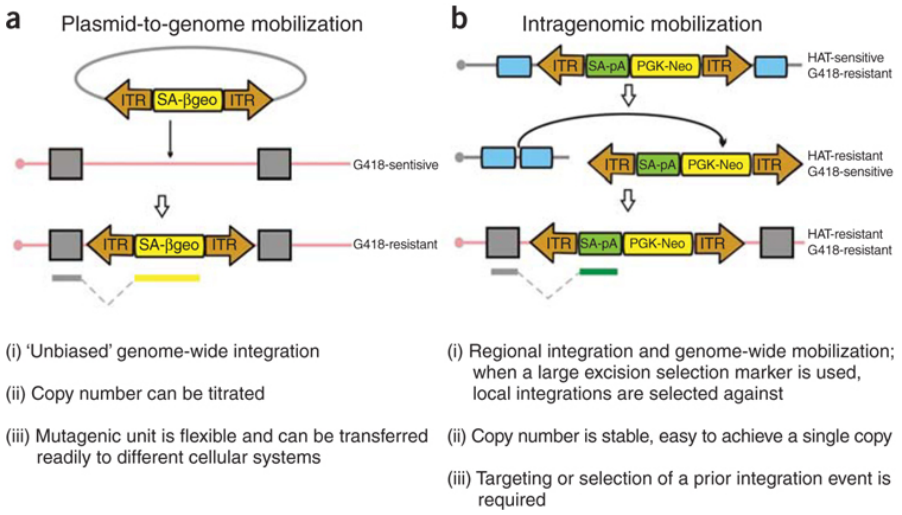

Figure 3. Transposon delivery methods in embryonic stem cells. (**a**) Gene trap–based loss-of-function mutagenesis is shown here as an example. In plasmid-to-genome mobilization, cells with mutagenic transposon insertions can be selected in the antibiotic G418 for the expression of the gene trap cassette (yellow) containing a splice acceptor (SA) and a β-galactosidase–neomycin fusion gene (β geo) cloned between the inverted terminal repeats (ITR) of the transposon. (**b**) In intragenomic mobilization, upon transposase expression, the transposon is excised from the donor site and reintegrates at a different genomic location. Enrichment of such cells can be achieved by selecting for transposition excision and reintegration. The gene-breaking cassette (green) contains a splice acceptor (SA) followed by a poly(A) signal (pA). The selection marker (yellow) is a phosphoglycerate kinase (PGK) promoter–driven neomycin resistance (Neo) gene. Using expression of the *Hprt1* locus as an excision selection marker, cells with the transposon excised from the donor site (blue) will be *Hprt1*-proficient and therefore resistant to hypoxanthine, aminopterine and thymidine (HAT).

OPEN ACCESS

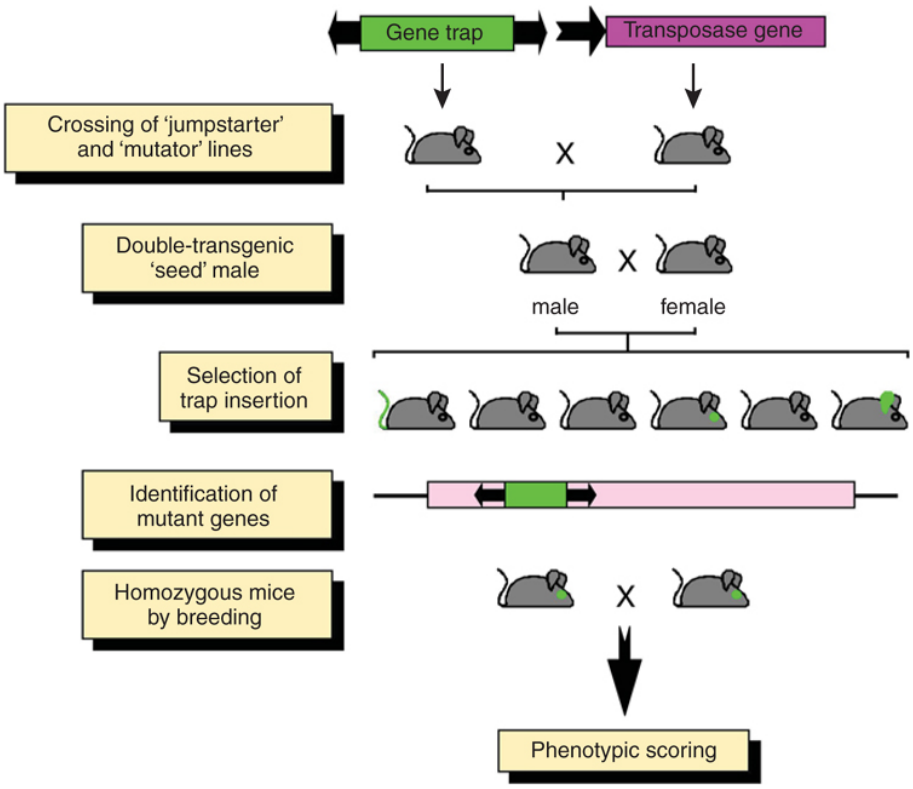

Figure 4. *In vivo* germline mutagenesis of the mouse with transposable elements. Breeding of 'jumpstarter' and 'mutator' stocks induces transposition in the germline of double-transgenic 'seed' males. The transposition events that take place in germ cells are segregated in the offspring. Animals with transposition events need to be bred to homozygosity to visualize the phenotypic effects of recessive mutations. Mutant genes can easily be cloned by different PCR methods making use of the inserted transposon as a unique sequence tag.

OPEN ACCESS

Table 1. The most important characteristics of transposons that are active in vertebrates.

Transposon name and source	Transposon family	Tolerated cargo size	Target site	Chromatic integration pattern	Systems tested
SB, reconstructed from fish [7]	*Tc1/mariner*	Increased cargo size exponentially decreases the efficiency of transposition in cultured cells [9]	TA	No preference for genes; gene hits dominantly in introns [70]	Cultured vertebrate cell lines [9], mouse [17,34,42-46,49,50,62,63], rat [47,48], zebrafish [27,59], medaka fish [30] and Xenopus [29]
Frog Prince, reconstructed from Rana pipiens (Northern leopard frog) [71]	*Tc1/mariner*	Possibly similar to other *Tc1/mariner* transposons	TA	Highly efficient gene trapping in tissue culture cells [71]; gene hits dominantly in introns	Cultured vertebrate cell lines and zebrafish embryos [71]
Minos, Drosophila hydei	*Tc1/mariner*	Possibly similar to other *Tc1/mariner* transposons	TA	No preference for genes; gene hits dominantly in introns	Cultured human cells, mouse tissues [52,53] and Ciona intestinalis [23]
Hsmar1, reconstructed from human [72]	*Tc1/mariner*	Possibly similar to other *Tc1/mariner* transposons	TA	Similar to SB [72]	Cultured human cells and zebrafish embryos [72]
Passport, Pleuronectes platessa (plaice) [73]	*Tc1/mariner*	Possibly similar to other Tc1/mariner transposons	TA	May have a preference for transcription units [73]	Mammalian and avian cell culture [73]
piggyBac, Trichoplusia ni (cabbage looper moth)	*piggyBac*	Efficiency drops for >9.1 kilobase cargo in pronucleus-injected mice [11]	TTAA	Preference for transcription units [74]	Mammalian cell culture, including mouse embryonic stem cells and mouse tissues [11,38,39,54]
Tol1 *Oryzias latipes* (medaka fish)	*hAT*	>20 kilobase cargo known to move but at reduced efficiency [12]	8-base-pair heterogenic sequences	Unknown but, similar to other hAT elements, may prefer the 5' regions of genes	Mammalian cell culture and zebrafish embryos [24]
Tol2, *Oryzias latipes* (medaka fish)	*hAT*	>10 kilobase transposons jump efficiently in human cells and zebrafish embryos [13,14]	8-base-pair heterogenic sequences	Unknown but, similar to other hAT elements, may prefer the 5' regions of genes	Cultured vertebrate cell lines, including mouse embryonic stem cells [55], zebrafish [25,56,57], Xenopus [26] and chicken embryos [31]
Ac/Ds *Zea mays* (maize)	*hAT*	At least 6.5 kilobases in zebrafish embryos [75]	8-base-pair heterogenic sequences	May have a preference for transcription units [75]	Mammalian cell culture and zebrafish embryos [75]
Harbinger3_DR, reconstructed from *Danio rerio* (zebrafish) [76]	*PIF/Harbinger*	Not tested experimentally	Preferentially inserts into a 15-base-pair consensus target sequence [76]	Unknown	Cultured human and zebrafish cells [76]
ORFeus, synthetic mouse *L1* retrotransposon [77]	*L1*	5' truncations are frequent	Preference for (A+T)-rich sequences [78]	~30% of insertions in genes [60]	Mouse and human cells [60]