Transposon tools hopping in vertebrates

Jun Ni, Karl J. Clark, Scott C. Fahrenkrug and Stephen C. Ekker

Abstract

In the past decade, tools derived from DNA transposons have made major contributions to vertebrate genetic studies from gene delivery to gene discovery. Multiple, highly complementary systems have been developed, and many more are in the pipeline. Judging which DNA transposon element will work the best in diverse uses from zebrafish genetic manipulation to human gene therapy is currently a complex task. We have summarized the major transposon vector systems active in vertebrates, comparing and contrasting known critical biochemical and *in vivo* properties, for future tool design and new genetic applications.

Keywords: *transposon; gene delivery; gene discovery; gene transfer vectors; vertebrates*

INTRODUCTION

DNA transposons are mobile genetic elements that move or transpose using a 'cut and paste' mechanism. Use of transposons in vertebrates was effectively launched with the high profile publication of the *Sleeping Beauty* (*SB*) transposon system [1]. In this study, a molecularly reconstructed system was developed, with the transposase catalyst provided *in trans* with the key DNA cargo tagged by flanking transposon end sequences. Based in part on the rich array of prior work with transposon elements from invertebrates and plants, this paradigm is still the primary transposon approach deployed today.

In the past decade, DNA transposons from many different gene families have now been shown to work well in diverse species. This review focuses on activities in vertebrate organisms or tissues, as such studies will benefit development and application of DNA transposon tools in gene discovery and gene delivery. With advantages of easy production, simpler handling and without pathogenicity, it is worthwhile to generate and optimize non-viral transposon vectors over the current preferentially used viral vectors [2]. As the biochemical properties of elements within a family can be very similar, we have organized current tools accordingly, with each element (Table 1) and transposon family (Table 2) summarized and compared.

Genetic applications can differentially leverage biochemical diversity in these transposon systems. 'Transposition efficiency' has been a limitation for DNA transposon elements compared to highly active viral methods. To generate transposon systems with improved activity is a constant effort in this field. 'Integration site preference' is an important consideration, as gene discovery applications can be enhanced by intragenic insertional preference, whereas such preference would be a perceived weakness for gene therapy uses. The DNA 'footprint' left behind after transposon integration and remobilization differentially impacts the use of these elements. For gene discovery work, a small DNA tag can be valuable to distinguish an engineered chromosome from an otherwise wild-type allele in that study. However, gene therapy applications would prefer as little disruption in the genome should an element hop more than once before landing in the final locus. High 'cargo capacity' is generally preferred, and transposon systems show robust differences in this key area. 'Host origin' is also an important constraint, as transposon vectors are normally used in heterologous organisms to avoid

Corresponding author. Stephen C. Ekker, Department of Biochemistry and Molecular Biology, Mayo Clinic, Rochester, MN 55905, USA. Tel: 507-284-5530; Fax: 507-284-9349; E-mail: ekker.stephen@mayo.edu

Jun Ni's research interests are on the biochemistry of transposons for vertebrate genetic applications.

Karl J. Clark has a long-standing interest in the development of vertebrate transposons including key work on *Sleeping Beauty* and *Passport*.

Scott C. Fahrenkrug's science focus includes advanced applications for the use of transposons for application to animal biotechnology.

Stephen C. Ekker is developing transposons for genetic applications in the zebrafish, *Danio rerio*.

Table I: Current DNA transposons for vertebrate gene transfer applications

Superfamily[a]	Species of origin[b]	Vertebrate transposition[c]	Insertion site preference[d]	Excision footprint[e]
Tc1/Mariner				
SB	salmonid[f]	Zebrafish, Xenopus, mouse germline, embryonic stem cell, liver and lung gene therapy and human cell culture	TA	Canonical: TACAGTA; TACTGTA Non-canonical, with deletions
FP	*R. pipiens*[f]	Human, hamster, *Xenopus*, fathead minnow and zebrafish cell lines	TA	Imprecise excision, generally+5 **TA**C$^{A}/_{T}$G**TA**
Hsmar1	*Homo sapiens*[f]	HeLa and zebrafish[g]	TA	Imprecise excision, generally+5 **TA**T$^{A}/_{T}$A**TA**
Minos	*D. hydei*	Human cell lines and mouse lymphatic tissues	TA	*Drosophila*: precise excison or canonical footprints (TcgagT or ActcgT); more heterogeneous in mammalians
Himar1	*H. irritans*	Human cell line	TA	Not determined
Passport	*P. platessa*	Human (HeLa, HTI080), monkey (vero), hamster (CHO), turkey (TT), chicken (DFI) and pig (PEGE) cell lines	TA	Not determined
hAT				
Tol2	*O. latipe*	Zebrafish, Xenopus, chicken, mouse and human cell lines	The 8 bp insertion site duplication	Imprecise deletion of 8 bp target site duplication
Tol1	*O. latipe*	Zebrafish, mouse cell line, human cell line and *C. elegans*	The 8 bp insertion site duplication	Imprecise
Ac/Ds	*Z. mays*	Zebrafish and human cell line	The 8 bp insertion site duplication	Imprecise
PIF/Harbinger				
Harbinger3-DR	*D. rerio*[f]	HeLa and zebrafish[g]	CWG	No footprint
piggyBac				
piggyBac	*Trichoplusia ni*	Mouse germ line, human cell lines and pig	TTAA	No footprint

[a]Elements are organized by genetic relationship within known transposon families.
[b]Organism of origin for the specified element. The salmonid origin for *SB* reflects its reconstructed origin using genomic DNA from multiple fish species.
[c]Lists some of the most notable experimental systems a particular element has been shown to exhibit transposition activity.
[d]Observations of the transposition insertion site sequence preference in native species and vertebrate organisms are noted.
[e]After remobilization from the donor site and repair of the double-strand break, traces of DNA sequences that are different from the wild-type loci could be left behind, termed 'excision footprint'. Excision footprints from each transposon are listed where known.
[f]Reconstructed/reanimated transposon.
[g]Excision data only were documented.

mobilizing related endogenous transposon sequences. With 10 transposons from four different superfamilies, genetic researchers have a rich array of diverse options at their disposal for vertebrate applications.

Tc1/Mariner FAMILY MEMBERS

The *Tc1/mariner* superfamily represents the most widespread DNA transposon family in nature, members having been described for all kingdoms of life [3]. Canonical representatives of this family were isolated from nematodes (*Tc1* [4]) and arthropods (*mariner* [5]). Since that time, numerous members of this family have been isolated, some having been harnessed for gene transfer in numerous species. However, only a selected few are able to efficiently transfer genes in vertebrate cells or embryos (see *Minos* [6] and *Himar* [7] subsequently), leading to the search for related family members in vertebrate genomes. Despite their being widespread, often at thousands of copies per genome, nearly all of the elements that have been characterized and recovered have been 'evolutionarily inactivated' [5]. Phylogenetically informed reconstruction of an inactive piscine element resulted in the isolation of the first active vertebrate *Tc1/mariner* element, *SB* [1]. From that time, multiple members of the *Tc1/mariner* family from vertebrates have been reconstructed using similar approaches, including *Frog Prince* [8] and *Hsmar1*

Table 2: Positive and negative features of DNA transposon superfamilies in major genetic applications

Superfamily	Gene delivery		Gene discovery	
	Advantages	Limitations	Advantages	Limitations
Tc1/Mariner	• Active in a variety of cells and tissues • Protocols available for *SB* gene therapy in animal models	• Limited cargo capacity • Requires empirical testing of diverse transposase concentrations for maximum gene delivery	• Little or no insertion site preference	• Transposition efficiency can be limiting in some systems
hAT	• Active in a variety of cells and tissues • High cargo capacity	• Limited protocols for gene therapy	• Modest insertion site preference	
PIF/Harbinger	• No footprint upon excision	• Modest transposition efficiency		• Insertion site may not be random
piggyBac	• Active in a variety of cells and tissues • High transposition efficiency • No footprint upon excision	• Prefers insertion inside transcribed loci	• Prefers insertion inside transcribed loci	

[9] transposons. In addition, intact endogenous elements have been discovered by comparative genome sequence analysis, one of which is widely represented in pleuronectid fish and was recently demonstrated as functionally competent for gene transfer in mammalian cells [10].

Tc1/mariner elements are ~1300–2400 bp in length and contain a single gene encoding a transposase enzyme that is flanked by terminal inverted repeats (IRs) [3]. Shared biochemical features of the *Tc1/mariner* family include a transposase that contains a 'DDE' motif [11]. This motif (two aspartic acid residues and a glutamic acid residue, the latter two separated by 35 residues) was first identified in retroviral integrases and bacterial IS (insertion sequence) transposases [12, 13]. This acidic amino acid triad has been inferred to be critical to the catalytic active site and involved in contacting divalent ions [14]. In the *Tc1/mariner* family, the second aspartic acid residue and the glutamic acid residue are separated by only 34 residues, and an aspartic acid residue is found at the position of the glutamic acid residue in the *mariner*-like elements.

These transposons are bracketed by inverted terminal repeats that bind the transposase. The length of these repeats vary; the simplest being <100 bp with a single transposase binding site per repeat [7, 15]. *Tc3* elements have IRs of >400 bp in length, each of which contains two binding sites, although the internal pair is not required for transposition [16]. A third group referred to as 'IR–DR' have a pair of binding sites containing short 15–20 bp direct repeats (DRs located at the ends of IRs that are 200–250 bp long, including *Minos* and *S* elements in flies [6, 17], *Quetzal* elements in mosquitos [18], *Txr* elements in frogs [19] and at least four *Tc1*-like transposon subfamilies in fish [10, 20]. *Tc1/mariner* elements invariably integrate into TA dinucleotides in a manner that has been generally described as random, although local sequence context and deformation apparently plays some role in target site selection [21, 22], and at least one family member displays some proclivity for landing in genes [10].

There is conflicting data regarding the sensitivity of transposition for various *Tc1/mariner* family members to cargo size. *In vitro* (cell free) and *in vivo* transposition of *Tc1* was decreased by 10- to 15-fold by simply increasing cargo size from ~2 kb to ~8 kb [16]. This limitation in cargo has also been observed for *Himar1* [23] and for the *SB* transposon.

Transfection of cultured mouse and human cells with *SB* transposons with different sized cargos has revealed a dramatic inverse relationship between transposon size and transposition [24, 25], although total plasmid size, not cargo size, has been hypothesized as responsible [26]. Interpretation of these experiments is in part confounded by size-dependent, differential transfection and resulting gene expression as described by Yin W *et al.*, 2005. In addition, transposition from genome-resident *SB* transposons [27, 28], or from DNA injected into the mouse pronucleus (S. Fahrenkrug, submitted for publication) does not seem to be as constrained by cargo size. Although the reason for these differences is unknown, it may be that nuclear localization of transposon–transposase complexes is size limiting, a barrier not experienced by transposons already resident in, or physically introduced to the nucleus.

Where it has been examined, members of this superfamily have commonly been found to be sensitive to a phenomenon referred to as 'overproduction-' or 'overexpression inhibition', first described for the *mariner* element [29], wherein excessive transposase expression was found to inhibit transpositional transgenesis. The same phenomenon has been described for the *Himar1* [23], *SB* [10] and *Passport* [10] transposons, although Izsvak *et al*. [26] found the frequency of *SB* transposition to be directly proportional to the level of transposase expression. The mechanism of 'overproduction inhibition' is still poorly understood. One working model suggests that an optimal ratio of transposase molecule to its binding terminals in an effective complex falls into a small concentration window and excess transposase will quench the formation of the complex [30]. In numerous studies, the optimal concentration and molar ratio of transposon and transposase has been found to vary between cell types, suggesting either differing abundance of co-factors, or a difference in the expression or stability, of the transposase. Thus, application of these transposons as gene delivery vehicles requires a careful assessment of the conditions best suited for the particular application and cell type.

Sleeping Beauty

SB was the first widely deployed transposon for vertebrate applications and was molecularly reconstructed from ancient inactivated *mariner*-like elements [1]. *SB* has become the *de facto* reference system and has been used for diverse applications, including: cellular transgenesis (piscine, avian and mammalian), animal transgenesis (fish, frogs, mice, rats and pigs) and gene therapy, with over 175 publications describing its use at the time of this publication.

Some milestones of *SB* applications are highlighted here. *SB* was the first DNA transposon with high efficiency gene transfer into human cells [1], the first non-viral integrating gene therapy vector [31], used for the first non-viral somatic mutagenesis system [32], and deployed for the first transposon-based transgenic mammal [33]. *SB* was the first effective transposon-based system for trangenic fish [34] and insertional mutants [35] in zebrafish. Finally, *SB* was the basic system for the initial attempts at generating a chimeric transposase for targeting transposon integration [36].

Components of the *SB* system have undergone extensive engineering with the resulting elements showing increased activity from improvements to the transposon [37–39] and to the transposase ([25, 39–41]). Most assays for these hyperactive elements are focused on measurements using a HeLa tissue culture gene transfer method [1]. *SB* activity, like other *Tc1/mariner* family members, is subject to overexpression inhibition; as a result, the selection of transposase dose for maximal gene transfer is a critical consideration in using *SB* for best gene transfer results.

Methylation studies have demonstrated that chemical modification of *SB* transposons by cytosine-phosphodiester-guanidine (CpG) hypermethylation can significantly enhance their transposition rates [42]. The mechanism for this enhancement of transposition is currently unknown. Identifying the experimental scenarios where methylation can make a significant improvement in gene transfer by *SB* is an ongoing research area for this element and will potentially have implications for work in other transposons.

In a series of yet-unexplained observations in *Xenopus*, two research groups have noted a positive gene transfer rate for *SB* that did not appear to be due to standard transposition reactions, instead displaying non-canonical footprints [43, 44]. Given the abundance of related *Tc1/mariner* transposons in frog genomes, this may reflect interference by endogenous transposable elements, or this may be due to some other aspect of *SB* biology that is limited in frogs.

Frog Prince

The *Frog Prince* (*FP*) transposon was 'kissed' to life by Csaba Miskey in the laboratories of Zsuzsanna Izsvak and Zoltan Ivics [8] using the approach deployed for *SB*. For generating the *FP* transposon system, 10 inactive putative elements with complete open reading frames were identified from the frog *Rana pipiens* and used to produce a consensus transposase sequence. A single element with two amino acid substitutions different than the consensus (mFP) was used to produce a consensus transposase (*FP*). Interestingly, the mFP version of the transposase was inactive until correction to the consensus in *FP* [8]. *FP* is closely related to *Txr* elements found within *Xenopus* but differs significantly from other members of the *Tc1* family, notably *SB*. *FP* is capable of mobilizing *Txr* elements albeit at a reduced rate to *FP* transposons. In a limited analysis, the *FP* IRs allowed splicing to internal splice-acceptors suggesting the 214 bp IRs do not significantly interfere with splicing, making *FP* a potential transposon for gene discovery using gene-trapping technologies [8].

Like *SB*, *FP* appears to be active in a wide range of vertebrate cells, including HeLa, Cho-KI (hamster), A6 (*Xenopus laevis*), FHM (fat-head minnow) and PAC2 (zebrafish). The level of transposition in these cell lines was comparable or slightly improved relative to transposition of the original *SB* system.

Hsmar1

Miskey *et al*. [9] reactivated the human *mariner* 1 transposon (*Hsmar1*) from about 200 copies of full-length transposons that contained inactivating mutations. The consensus *Hsmar1* transposase sequence was not capable of promoting transposition in human cells. Therefore, 51 sequences from the human genome as well as mariner sequences from the *cecropia* transposase subfamily were aligned to determine the likely ancestral protein sequence for site-specific mutagenesis. An additional four amino acid substitutions were needed to make the Hsmar1 reconstructed ancestral protein (*HsMar1-Ra*). *HsMar1-Ra* was capable of excising transposons in HeLa cells and zebrafish embryos. Integration was shown by colony formation as well as examination of the integration sites in HeLa cells. *Hsmar1* integrates into TA dinucleotides and appears to distribute randomly throughout the genome. The levels of excision or transposition of *Hsmar1-Ra* was not directly compared to any other transposon during this study, although there was an ~20-fold increase in stable colony formation in HeLa cells.

Hsmar1-Ra is unlikely to be used for human applications like gene therapy because of its ability to mobilize both non-autonomous full *HsMar1* elements as well as the more prevalent (approximately 2500) *Hsmar1*-related mini-inverted repeat transposable elements (MITEs) found within the human genome [9]. How valuable *HsMar1* will be in non-human work is still unclear. It seems likely that *Hsmar1* will function in other systems including the zebrafish (where Miskey *et al*. demonstrated the ability of *Hsmar1-Ra* to cut or excise *Hsmar1* elements after injection into zebrafish embryos), but whether the transposition achieved is at a high enough rate for practical use remains to be tested. One advantage of *Hsmar1-Ra* is the short IRs of this class of *mariner*-type elements, as their relatively simple sequence requirements may facilitate the production of transposon vectors through the addition of IR sequences to a particular genetic cargo via a PCR reaction.

Minos

Minos, named after the legendary Greek King, is a DNA transposon belonging to the *Tc1/mariner* superfamily [6]. *Minos* elements isolated from *Drosophila hydei* show high sequence homogeneity with characteristic features: about 1.8 kb in length; perfect inverted terminal repeats (IRs) of 254 bp; encode a putative single transposase gene with two exons and a 60 bp intron in between, which has been shown to be spliced out in the *D. melanogaster* germline. Southern blot analyses of various *D. hydei* strains strongly suggested that *Minos* transposons are active in the germline [45] and that *Minos*-like elements are widely spread among *Drosophila* species [46].

In *Drosophila*, the *P*-element is the most widely used transposon and has generated thousands of valuable insertion lines for various research purposes. However, *P*-element activity is very host-restricted and appears to be inactive in non-Drosophilids [47]. In contrast, *Minos* elements have demonstrated a much wider host spectrum, including *Drosophila* and non-Drosophilid insects (for review, see [48]).

In human HeLa cells, researchers were able to achieve high frequency of *Minos* transposition in stably transformed cells. They proposed that *Minos* can be used to potentially tag all genes of the human genome [49]. *Minos* elements are active in mouse lymphatic tissues and embryonic fibroblast cell lines [50]. *Minos* transposase was effective in

mediating marker gene transposition in the mouse germline [44].

Minos does not prefer any sequence beyond the di-nucleotide TA and tends to insert more frequently into introns than exons in both fly and mouse [49]. In the fruit fly, the excision of *Minos* could leave precise conversion at the insertion sites or 6 bp characteristic footprints [48]. However, in mammalian systems, the footprints appear more complex, potentially due to differences in the host chromatid-repair machinery in these organisms [49, 51].

Himar1

Himar1 is a natural element of the *Tc1/mariner* superfamily from the horn fly, *Haematobia irritans*. This element has been shown to be active in a number of animals, including mammals [52, 53]. Hyperactive variants have been developed as well [54]. How this element quantitatively compares to other members of this family in vertebrate applications is still largely unknown.

Passport

The *Passport* transposon was initially identified within the plaice (*Pleuronectes platessa*) genome by Leaver *et al*. in 2001 [55]. In that study, a complete transposase open reading frame was described from within the plaice genome with a high degree of similarity among isolated transposon sequences. This potentially active transposon was dubbed PPTN.

Recently, Clark *et al*. [56, 10] dissected the transposase and transposon IRs of the PPTN transposon and showed that it was a functional transposase active with native IR and transposase sequences in human cells. The transposon system was renamed *Passport* referring to its ability to give the right or privilege of passage, entry or acceptance—in this case into the genome. *Passport* is a *Tc1*-type element, but it differs significantly from both *SB* and *FP* in IR and transposase sequence. Like other *Tc1*-type elements, *Passport* has IRs of around 200 bp and exhibits transposition into a wide variety of cell types, including human, hamster, pig, monkey, turkey and chicken. *Passport* was directly compared to *SB* in human cells through a range of transposase activity. *Passport* was about 40% as active as an improved *SB* at their respective peaks. *Passport* was shown to exhibit overexpression inhibition dynamics that were shaped similar to *SB*; delivery of too little or too much transposase results in 40% or less peak activity.

As compared to other *Tc1* elements used in vertebrates, early data suggests that target-site selection is less random with a preference for integration into transcription units (~60%) and 6 of 27 integrations into human chromosome 12 (distributed) and 0 of 27 integrations into human chromosome 1. This is unexpected since the size of human chromosome 1 is nearly twice as long as chromosome 12. *Passport* could likely function as a new transposon system for the manipulation of vertebrate genomes.

hAT FAMILY MEMBERS

Transposition from *hAT* family members in maize was the first visible mobile genetic element activity noted in science [57]. Diverse active native elements have been identified in vertebrates, including two from the fish medaka (*Tol1* and *Tol2*; [58, 59]). Unlike studies with reconstructed *SB*, most *hAT* family work to date has been with autonomously active elements. Native *hAT* transposons are about 5 kbp in size, and common shared properties include high cargo capacity, no obvious primary sequence insertion site requirement and less influence of overexpression inhibition compared to *SB* and related *Tc1/mariner* elements. A common 8 bp duplication is noted at the insertion site for each of these elements, with no primary sequence constraint for integration.

Tol2

Most *hAT* transposon family work in vertebrates to date has been with derivatives of the native element from the medaka fish (*Oryzias latipe*), *Tol2* [58, 60]. The native *Tol2* transposon is ~5 kbp in size and encodes an *hAT* family transposase [61]. Minimized *Tol2* IR transposon ends sufficient for full transposition *in vivo* have been developed (called 'miniTol2'; [61–63]). Good cargo capacity has been described for *Tol2*, with little or no major decrease in activity with transposons >10 kbp compared to smaller, 2 kbp elements [62, 63]; larger elements can be robustly mobilized with *Tol2* in zebrafish (S. Ekker, submitted for publication). *Tol2* demonstrates perhaps the least functional constraint due to overexpression inhibition of the elements describe here, with a large range of the transposase resulting in near peak gene transfer activity due to transposition [62].

Tol2 is active in all vertebrates tested to date, including robust activity in zebrafish [64], *Xenopus* [65], mice and human tissue culture cells [66]. Favorable kinetics over *SB* makes this a current

favorite element for the generation of transgenic zebrafish [62], and notably a range of excellent pre-assembled vectors for work in this model organism (called the 'Tol2kit') is freely available [67, 68]. *Tol2* represents the current standard for *hAT* transposon family functionality and utility in vertebrates.

Tol1

Tol1 was identified as an active DNA transposon from the medaka genome [58, 69]. The autonomous *Tol1* element is ~4.4 kbp, encoding a transposase gene with three exons. The transcribed mRNA is ~2.9 kbp coding for a putative functional transposase of 851 amino acids with sequence similarity to *hAT* superfamily [70].

The activity of *Tol1* transposase to deliver cargos flanking by *Tol1* IRs has been analyzed in detail. Minimized *Tol1* IRs with 157 bp from the left and 106 bp from the right termini maintain full transposition efficiency [71]. Cargo capacity was also tested, with DNA fragments up to 22.1 kbp in size were transposed without internal deletion or aberration [71], consistent with previous observation that transposons from *hAT* family have demonstrated large cargo delivery capacity in general compared to those from other families [62]. *Tol1*, like *Tol2*, has a wide range of activity in a variety of host species, including good activity in human (HeLa) and mouse cells (NIH/3T3) [71] and *Caenorhabditis elegans* [59]. Excision activity of *Tol1* has been reported in the frog *X. laevis* [72]. Most recently, *Tol1*-derived vectors have been shown to actively transpose in zebrafish germline and pass to the next generation [73].

So far, *Tol1* biochemistry parallels work with the widely used *Tol2* transposon in zebrafish. *Tol1* and *Tol2* do not cross react to the cis-sequences of each other, indicating *Tol1* as a good candidate for developing genetic tools in vertebrates and further mutagenesis of insertional mutants generated by *Tol2* [73].

Activator/Dissociation

Over 50 years ago, the first transposable system *Ac/Ds* (*Activator/Dissociation*) was discovered by Barbara McClintock from *Zea mays* (maize) [57]. Being one of the founding members of the *hAT* superfamily, the *Ac* element is 4.5 kb in length and displays all the common molecular features of this family, including short terminal IRs (11 bp), encoding a putative DDE-domain containing transposase and causing 8 bp host duplication upon insertion [74]. *Ds* elements are the natural non-autonomous versions of *Ac*, containing the IRs and cis-required sequences but lacking the ability to produce a functional transposase [75]. Their transpositions depend on the presence of the *Ac* element or *Ac* transposase.

The transposition of *Ac/Ds* system in heterologous plants has been demonstrated in both dicot and monocot (for review see [76]). Later, *Saccharomyces cerevisiae* became an even more distant species that can host *Ac/Ds* 'cut and paste' activity [77]. Most recently, the possibility of employing a plant transposon in a vertebrate was exploited. A chimeric *Ac* transposase (NLS-Tpase), composed of an animal-origin nuclear localization signal (NLS) fused to the N-terminal of the truncated maize *Ac* lacking the fist 102 amino acids (Tpase 103–807), demonstrated *Ds* element transposition in zebrafish in both somatic and germline cells [78]. This modified *Ac* (NLS K5E-Tpase) effectively inserts *Ds* element DNA in a human embryonic kidney cell line as well [78].

The *Ac* transposase behaves similarly in zebrafish and plant species in various transposition assays [76, 78]. However, it is unknown whether the regulation of *Ac* observed in the plant kingdom still holds in vertebrates, such as different dosage effects in different species [79, 80], methylation of transposase binding sites [81, 82] and certain host factors effects [83]. More careful analysis of *Ac* activity in more diverse vertebrates should provide more information on the mechanisms of transposition of the *hAT* family.

PIF/Harbinger FAMILY

The *PIF/Harbinger* transposon family includes elements from plants and animals, and the *Harbinger3.DR* is a molecularly reconstructed element of this family from the zebrafish (*Danio rerio*) genome [84]. Besides containing a gene encoding transposase, *Harbinger.DR* has a second gene encoding a Myb-like protein involved in either DNA binding or protein–protein interactions. This transposon has short 12 bp IRs. Insertion site preference indicates a strong requirement for CAG or CTG trinucleotides, with an even larger, 15 bp total consensus sequence noted [84]. This element does not appear to leave a footprint upon excision [84].

piggyBac FAMILY

piggyBac [85] is the founding member of its class of transposon and was first described as a DNA

hitchhiker in an invertebrate culture system. This element has been effectively harnessed, with minimized sequences identified [86]. *piggyBac* is active in many vertebrates tested to date [87–90]. Key biochemical properties include a sequence requirement at integration ('TTAA'; [85]), and an apparent preference for integration into transcriptionally active genomic DNA [90]. Remobilization of this element often results in a chromosome with no residual genetic change yielding no observable footprint [85]. A more modest overexpression inhibition profile compared to *SB* has been described [89].

CONCLUSION

DNA transposons are extremely effective tools for genetic modifications of vertebrates. With molecular engineering skills activating evolutionarily quiescent genes, or minimizing autonomously active elements, we now have an array of vectors from diverse organisms showing robust activity in a variety of model systems. The complementary biochemical properties encoded by these vector systems open the door to many new vertebrate gene discovery and gene transfer applications. Table 2 has summarized the positive and negative features of each transposon superfamily for prospective applications.

Key Points

- Ten different DNA transposons are now in use for vertebrate gene transfer applications.
- Some transposons systems are harnessed native elements, some are reanimated from evolutionarily defunct elements.
- These vectors are from diverse gene families with related, unique and complementary biochemical activities.
- Transposon systems are being continuously engineered for better and new properties.
- Transposons are powerful vectors for gene transfer and gene discovery applications in vertebrates.

FUNDING

Mayo Foundation and National Institutes of Health grant (DA14546 to S.C.E.).

References

1. Ivics Z, Hackett PB, Plasterk RH, *et al*. Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. *Cell* 1997;**91**: 501–10.
2. Ohlfest JR, Freese AB, Largaespada DA. Nonviral vectors for cancer gene therapy: prospects for integrating vectors and combination therapies. *Curr GeneTher* 2005;**5**:629–41.
3. Plasterk RH, Izsvak Z, Ivics Z. Resident aliens: the Tc1/mariner superfamily of transposable elements. *Trends Genet* 1999;**15**:326–32.
4. Liao LW, Rosenzweig B, Hirsh D. Analysis of a transposable element in Caenorhabditis elegans. *Proc Natl Acad Sci USA* 1983;**80**:3585–9.
5. Jacobson JW, Medhora MM, Hartl DL. Molecular structure of a somatically unstable transposable element in Drosophila. *Proc Natl Acad Sci USA* 1986;**83**:8684–8.
6. Franz G, Savakis C. Minos, a new transposable element from *Drosophila hydei*, is a member of the Tc1-like family of transposons. *Nucleic Acids Res* 1991;**19**:6646.
7. Lampe DJ, Churchill ME, Robertson HM. A purified mariner transposase is sufficient to mediate transposition in vitro. *EMBO J* 1996;**15**:5470–9.
8. Miskey C, Izsvak Z, Plasterk RH, *et al*. The Frog Prince: a reconstructed transposon from *Rana pipiens* with high transpositional activity in vertebrate cells. *Nucleic Acids Res* 2003;**31**:6873–81.
9. Miskey C, Papp B, Mates L, *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* 2007;**27**:4589–600.
10. Clark KJ, Carlson DF, Leaver MJ, *et al*. Passport, a native Tc1 transposon from flatfish, is functionally active in vertebrate cells. *Nucleic Acids Res* (in press).
11. Doak TG, Doerder FP, Jahn CL, *et al*. A proposed superfamily of transposase genes: transposon-like elements in ciliated protozoa and a common "D35E" motif. *Proc Natl Acad Sci USA* 1994;**91**:942–6.
12. Fayet O, Ramond P, Polard P, *et al*. Functional similarities between retroviruses and the IS3 family of bacterial insertion sequences? *Mol Microbiol* 1990;**4**:1771–7.
13. Khan E, Mack JP, Katz RA, *et al*. Retroviral integrase domains: DNA binding and the recognition of LTR sequences. *Nucleic Acids Res* 1991;**19**:851–60.
14. Craig NL, Craigie R, Gellert M, Lambowitz A. (eds). *Mobile DNA II*. American Society for Microbiology, Published by ASM Press, 2002.
15. Vos JC, Plasterk RH. Tc1 transposase of Caenorhabditis elegans is an endonuclease with a bipartite DNA binding domain. *EMBO J* 1994;**13**:6125–32.
16. Fischer SE, van Luenen HG, Plasterk RH. Cis requirements for transposition of Tc1-like transposons in *C. elegans*. *Mol Gen Genet* 1999;**262**:268–74.
17. Merriman PJ, Grimes CD, Ambroziak J, *et al*. S elements: a family of Tc1-like transposons in the genome of Drosophila melanogaster. *Genetics* 1995;**141**:1425–38.
18. Ke Z, Grossman GL, Cornel AJ, *et al*. Quetzal: a transposon of the Tc1 family in the mosquito Anopheles albimanus. *Genetica* 1996;**98**:141–7.
19. Lam WL, Seo P, Robison K, *et al*. Discovery of amphibian Tc1-like transposon families. *J Mol Biol* 1996;**257**:359–66.
20. Izsvak Z, Ivics Z, Garcia-Estefania D, *et al*. DANA elements: a family of composite, tRNA-derived short interspersed DNA elements associated with mutational activities in zebrafish (Danio rerio). *Proc Natl Acad Sci USA* 1996;**93**: 1077–81.

21. Liu G, Geurts AM, Yae K, *et al*. Target-site preferences of Sleeping Beauty transposons. *J Mol Biol* 2005;**346**:161–73.
22. Geurts AM, Hackett CS, Bell JB, *et al*. Structure-based prediction of insertion-site preferences of transposons into chromosomes. *Nucleic Acids Res* 2006;**34**:2803–11.
23. Lampe DJ, Grant TE, Robertson HM. Factors affecting transposition of the Himar1 mariner transposon in vitro. *Genetics* 1998;**149**:179–87.
24. Karsi A, Moav B, Hackett P, *et al*. Effects of insert size on transposition efficiency of the sleeping beauty transposon in mouse cells. *Mar Biotechnol (NY)* 2001;**3**:241– 5.
25. Geurts AM, Yang Y, Clark KJ, *et al*. Gene transfer into genomes of human cells by the sleeping beauty transposon system. *Mol Ther* 2003;**8**:108–17.
26. Izsvak Z, Ivics Z, Plasterk RH. Sleeping Beauty, a wide host-range transposon vector for genetic transformation in vertebrates. *J Mol Biol* 2000;**302**:93–102.
27. Luo G, Ivics Z, Izsvak Z, *et al*. Chromosomal transposition of a Tc1/mariner-like element in mouse embryonic stem cells. *Proc Natl Acad Sci USA* 1998;**95**:10769–73.
28. Horie K, Yusa K, Yae K, *et al*. Characterization of Sleeping Beauty transposition and its application to genetic screening in mice. *Mol Cell Biol* 2003;**23**:9189–207.
29. Lohe AR, Hartl DL. Autoregulation of mariner transposase activity by overproduction and dominant-negative complementation. *Mol Biol Evol* 1996;**13**:549–55.
30. Hackett PB. Integrating DNA vectors for gene therapy. *Mol Ther* 2007;**15**:10–2.
31. Yant SR, Meuse L, Chiu W, *et al*. Somatic integration and long-term transgene expression in normal and haemophilic mice using a DNA transposon system. *Nat Genet* 2000;**25**: 35–41.
32. Dupuy AJ, Akagi K, Largaespada DA, *et al*. Mammalian mutagenesis using a highly mobile somatic Sleeping Beauty transposon system. *Nature* 2005;**436**:221–6.
33. Dupuy AJ, Clark K, Carlson CM, *et al*. Mammalian germline transgenesis by transposition. *Proc Natl Acad Sci USA* 2002;**99**:4495–9.
34. Davidson AE, Balciunas D, Mohn D, *et al*. Efficient gene delivery and gene expression in zebrafish using the Sleeping Beauty transposon. *Dev Biol* 2003;**263**:191–202.
35. Sivasubbu S, Balciunas D, Davidson AE, *et al*. Gene-breaking transposon mutagenesis reveals an essential role for histone H2afza in zebrafish larval development. *Mech Dev* 2006;**123**:513–29.
36. Yant SR, Huang Y, Akache B, *et al*. Site-directed transposon integration in human cells. *Nucleic Acids Res* 2007;**35**:e50.
37. Cui Z, Geurts AM, Liu G, *et al*. Structure-function analysis of the inverted terminal repeats of the sleeping beauty transposon. *J Mol Biol* 2002;**318**:1221–35.
38. Izsvak Z, Khare D, Behlke J, *et al*. Involvement of a bifunctional, paired-like DNA-binding domain and a transpositional enhancer in Sleeping Beauty transposition. *J Biol Chem* 2002;**277**:34581–8.
39. Yant SR, Park J, Huang Y, *et al*. 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* 2004;**24**:9239–47.
40. Baus J, Liu L, Heggestad AD, *et al*. Hyperactive transposase mutants of the Sleeping Beauty transposon. *Mol Ther* 2005; **12**:1148–56.
41. Zayed H, Izsvak Z, Walisko O, *et al*. Development of hyperactive sleeping beauty transposon vectors by mutational analysis. *Mol Ther* 2004;**9**:292–304.
42. Yusa K, Takeda J, Horie K. Enhancement of Sleeping Beauty transposition by CpG methylation: possible role of heterochromatin formation. *Mol Cell Biol* 2004;**24**: 4004–18.
43. Doherty JR, Johnson Hamlet MR, Kuliyev E, *et al*. A flk-1 promoter/enhancer reporter transgenic Xenopus laevis generated using the Sleeping Beauty transposon system: an *in vivo* model for vascular studies. *Dev Dyn* 2007;**236**: 2808–17.
44. Sinzelle L, Vallin J, Coen L, *et al*. Generation of trangenic *Xenopus laevis* using the Sleeping Beauty transposon system. *Transgenic Res* 2006;**15**:751–60.
45. Franz G, Loukeris TG, Dialektaki G, *et al*. Mobile Minos elements from Drosophila hydei encode a two-exon transposase with similarity to the paired DNA-binding domain. *Proc Natl Acad Sci USA* 1994;**91**:4746–50.
46. Arca B, Savakis C. Distribution of the transposable element Minos in the genus Drosophila. *Genetica* 2000; **108**:263–7.
47. O'Brochta D A, Handler AM. Mobility of P elements in drosophilids and nondrosophilids. *Proc Natl Acad Sci USA* 1988;**85**:6052–6.
48. Pavlopoulos A, Oehler S, Kapetanaki MG, *et al*. The DNA transposon Minos as a tool for transgenesis and functional genomic analysis in vertebrates and invertebrates. *Genome Biol* 2007;**8**:S2.
49. Klinakis AG, Zagoraiou L, Vassilatis DK, *et al*. Genome-wide insertional mutagenesis in human cells by the Drosophila mobile element Minos. *EMBO Rep* 2000;**1**: 416–21.
50. Zagoraiou L, Drabek D, Alexaki S, *et al*. *In vivo* transposition of Minos, a Drosophila mobile element, in mammalian tissues. *Proc Natl Acad Sci USA* 2001;**98**:11474–8.
51. Drabek D, Zagoraiou L, deWit T, *et al*. Transposition of the Drosophila hydei Minos transposon in the mouse germ line. *Genomics* 2003;**81**:108–11.
52. Zhang L, Sankar U, Lampe DJ, *et al*. The Himar1 mariner transposase cloned in a recombinant adenovirus vector is functional in mammalian cells. *Nucleic Acids Res* 1998;**26**: 3687–93.
53. Felsheim RF, Herron MJ, Nelson CM, *et al*. Transformation of Anaplasma phagocytophilum. *BMC Biotechnol* 2006;**6**:42.
54. Keravala A, Liu D, Lechman ER, *et al*. Hyperactive Himar1 transposase mediates transposition in cell culture and enhances gene expression in vivo. *Hum Gene Ther* 2006;**17**: 1006–18.
55. Leaver MJ. A family of Tc1-like transposons from the genomes of fishes and frogs: evidence for horizontal transmission. *Gene* 2001;**271**:203–14.
56. Clark KJ, Carlson DF, Foster LK, *et al*. Enzymatic engineering of the porcine genome with transposons and recombinases. *BMC Biotechnol* 2007;**7**:42.
57. Mc CB. The origin and behavior of mutable loci in maize. *Proc Natl Acad Sci USA* 1950;**36**:344–55.

58. Koga A, Inagaki H, Bessho Y, *et al*. Insertion of a novel transposable element in the tyrosinase gene is responsible for an albino mutation in the medaka fish, Oryzias latipes. *Mol Gen Genet* 1995;**249**:400–5.
59. Kodama K, Takagi S, Koga A. The Tol1 element of the medaka fish, a member of the hAT transposable element family, jumps in Caenorhabditis elegans. *Heredity* 2008;**101**: 222–27.
60. Koga A, Sakaizumi M, Hori H. Transposable elements in medaka fish. *Zoolog Sci* 2002;**19**:1–6.
61. Kawakami K, Koga A, Hori H, *et al*. Excision of the tol2 transposable element of the medaka fish, Oryzias latipes, in zebrafish, Danio rerio. *Gene* 1998;**225**:17–22.
62. Balciunas D, Wangensteen KJ, Wilber A, *et al*. Harnessing a high cargo-capacity transposon for genetic applications in vertebrates. *PLoS Genet* 2006;**2**:e169.
63. 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* 2006;**174**:639–49.
64. Kawakami K, Takeda H, Kawakami N, *et al*. A transposon-mediated gene trap approach identifies developmentally regulated genes in zebrafish. *Dev Cell* 2004;**7**:133–44.
65. Hamlet MR, Yergeau DA, Kuliyev E, *et al*. Tol2 transposon-mediated transgenesis in Xenopus tropicalis. *Genesis* 2006;**44**:438–45.
66. Kawakami K. Tol2: a versatile gene transfer vector in vertebrates. *Genome Biol* 2007;**8**:S7.
67. Kwan KM, Fujimoto E, Grabher C, *et al*. The Tol2kit: a multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. *Dev Dyn* 2007;**236**:3088–99.
68. Villefranc JA, Amigo J, Lawson ND. Gateway compatible vectors for analysis of gene function in the zebrafish. *Dev Dyn* 2007;**236**:3077–87.
69. Tsutsumi M, Imai S, Kyono-Hamaguchi Y, *et al*. Color reversion of the albino medaka fish associated with spontaneous somatic excision of the Tol-1 transposable element from the tyrosinase gene. *Pigment Cell Res* 2006;**19**: 243–7.
70. Koga A, Shimada A, Kuroki T, *et al*. The Tol1 transposable element of the medaka fish moves in human and mouse cells. *J Hum Genet* 2007;**52**:628–35.
71. Koga A, Higashide I, Hori H, *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* 2007;**52**:1026–30.
72. Hikosaka A, Koga A. PCR detection of excision suggests mobility of the medaka fish Tol1 transposable element in the frog *Xenopus laevis*. *Genet Res* 2007;**89**:201–6.
73. Koga A, Cheah FS, Hamaguchi S, *et al*. Germline transgenesis of zebrafish using the medaka Tol1 transposon system. *Dev Dyn* 2008;**237**:2466–74.
74. Müller-Neumann M, Yoder JI, Starlinger P. The DNA sequnce of the transposable element Ac of Zea mays L. *Mol Gen Genet* 1984;**198**:19–24.
75. Doring HP, Tillmann E, Starlinger P. DNA sequence of the maize transposable element Dissociation. *Nature* 1984; **307**:127–30.
76. Haring MA, Rommens CM, Nijkamp HJ, *et al*. The use of transgenic plants to understand transposition mechanisms and to develop transposon tagging strategies. *Plant Mol Biol* 1991;**16**:449–61.
77. Weil CF, Kunze R. Transposition of maize Ac/Ds transposable elements in the yeast *Saccharomyces cerevisiae*. *Nat Genet* 2000;**26**:187–90.
78. Emelyanov A, Gao Y, Naqvi NI, *et al*. Trans-kingdom transposition of the maize dissociation element. *Genetics* 2006;**174**:1095–104.
79. Hehl R, Baker B. Properties of the maize transposable element Activator in transgenic tobacco plants: a versatile inter-species genetic tool. *Plant Cell* 1990;**2**:709–21.
80. Jones JD, Carland FM, Maliga P, *et al*. Visual detection of transposition of the maize element activator (Ac) in tobacco seedlings. *Science* 1989;**244**:204–7.
81. Chomet PS, Wessler S, Dellaporta SL. Inactivation of the maize transposable element Activator (Ac) is associated with its DNA modification. *EMBO J* 1987;**6**:295–302.
82. Fedoroff NV. About maize transposable elements and development. *Cell* 1989;**56**:181–91.
83. Kunze R, Starlinger P. The putative transposase of transposable element Ac from Zea mays L. interacts with subterminal sequences of Ac. *EMBO J* 1989;**8**:3177–85.
84. Sinzelle L, Kapitonov VV, Grzela DP, *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* 2008;**105**:4715–20.
85. Fraser MJ, Ciszczon T, Elick T, *et al*. Precise excision of TTAA-specific lepidopteran transposons piggyBac (IFP2) and tagalong (TFP3) from the baculovirus genome in cell lines from two species of Lepidoptera. *Insect Mol Biol* 1996;**5**: 141–51.
86. Elick TA, Lobo N, Fraser MJ, Jr. Analysis of the cis-acting DNA elements required for piggyBac transposable element excision. *Mol Gen Genet* 1997;**255**:605–10.
87. Ding S, Wu X, Li G, *et al*. Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. *Cell* 2005;**122**:473–83.
88. Lobo NF, Fraser TS, Adams JA, *et al*. Interplasmid transposition demonstrates piggyBac mobility in vertebrate species. *Genetica* 2006;**128**:347–57.
89. Wang W, Lin C, Lu D, *et al*. Chromosomal transposition of PiggyBac in mouse embryonic stem cells. *Proc Natl Acad Sci USA* 2008;**105**:9290–5.
90. Wilson MH, Coates CJ, George AL, Jr. PiggyBac transposon-mediated gene transfer in human cells. *Mol Ther* 2007;**15**:139–45.