

***Passport*, a native Tc1/mariner transposon from flatfish is functionally active in vertebrate cells**

Karl J. Clark^{1,2,5}, Daniel F. Carlson^{1,2}, Michael J. Leaver³, Linda K. Foster¹, and Scott C. Fahrenkrug^{1,2,4}

1. Department of Animal Science at the University of Minnesota, St. Paul, MN 55108, USA.
2. The Arnold and Mabel Beckman Center for Transposon Research at the University of Minnesota, Minneapolis, MN, USA
3. Institute of Aquaculture, University of Stirling, Stirling FK9 4LA, UK
4. The University of Minnesota Animal Biotechnology Center
5. Present address: Karl J. Clark, Department of Biochemistry and Molecular Biology, Mayo Clinic, Rochester, MN 55902, USA

For correspondence:

Scott Fahrenkrug
University of Minnesota
495 AniSci/VetMed
1988 Fitch Ave.
St. Paul, MN 55108
Email: fahre001@umn.edu

Abstract

The Tc1/mariner family of DNA transposons is widespread across fungal, plant and animal kingdoms, and thought to contribute to the evolution of their host genomes. To date, an active Tc1 transposon has not been identified within the native genome of a vertebrate. We demonstrate that *Passport*, a native transposon isolated from a fish (*Pleuronectes platessa*), is active in a variety of vertebrate cells. In transposition assays, we found that the *Passport* transposon system improved stable cellular transgenesis by 40-fold, has an apparent preference for insertion into genes, and is subject to overproduction inhibition like other Tc1 elements. *Passport* represents the first vertebrate Tc1 element described as both natively intact and functionally active, and given its restricted phylogenetic distribution, may be contemporaneously active. The *Passport* transposon system thus complements the available genetic tools for the manipulation of vertebrate genomes, and may provide a unique system for studying the infiltration of vertebrate genomes by Tc1 elements

Introduction

Mobilization of transposons is hypothesized to contribute to the evolution of genomes by several mechanisms, including; imperfect repair after excision, insertional mutagenesis, changes in the regulation of adjacent gene expression, and gene duplication or exon shuffling (1,2). *Tc1/mariner* elements are widely distributed, being found in species from animal, plant and fungal kingdoms (3,4). Transposons in this family contain a transposase gene flanked by inverted terminal repeats (ITRs) and are mobilized by a cut-and-paste mechanism (3). The *Tc1/mariner* transposases belong to a larger family of enzymes,

including bacterial transposases, retroviral integrases, and V(D)J recombinase, all of which are characterized by a DDE or in the case of mariner and mariner-like elements a DDD motif that is involved in polynucleotidyl transfer reactions (3).

Tc1/mariner elements can be active in the soma and the germline. Therefore, regulation of transposition is required for host viability, and by extension, transposon persistence (5). Evolutionary periods of transpositional activity are thus interspersed with periods of stochastic loss (6) and ‘vertical inactivation’ of transposons, wherein only defective versions are found within the genome, containing frame-shifts, deletions and missense mutations. Nonetheless, representatives of this family of transposons have been demonstrated to be active in nematodes (7,8) and arthropods (6,9,10). In contrast, the biology of *Tc1/mariner* elements in vertebrate cells/genomes is understudied. Despite being present at thousands of copies per genome, neither active *Tc1/mariner* transposition, nor functionally intact elements have been identified in vertebrate genomes. Instead, active vertebrate *Tc1/mariner* elements have been synthetically created by phylogeny informed reanimation of inactive transposons. The *Sleeping Beauty* (SB) transposon derived from salmonid fish represents the inaugural representative of *Tc1/mariner* transposon reanimation in vertebrates (11), and has been subsequently engineered to hyperactivity for genetic applications including transpositional transgenesis (TnT) and gene therapy. Additional transposons from amphibians (*Frog Prince*) and humans (*HsMar1*) have been similarly reanimated (12,13).

Leaver previously described a *Tc1*-like transposon (PPTN) from flatfish that is present at 200–300 copies in the plaice (*Pleuronectes platessa*) genome (14,15). This element belongs to the inverted repeat/direct repeat (IR/DR) group of *Tc1*-like elements as described by Izsvák et al. (16), and has 74–81% identity to related but distinct elements represented in the genomes of Atlantic salmon (*Salmo salar*) and frogs (*Rana temporaria*). These elements appeared to be absent in the genomes of other fish species. The fact that these related mobile elements are found in the genomes of phylogenetically distant animals yet absent in more closely related species, led Leaver to hypothesize that this transposon family was distributed by horizontal transfer (14,17). Additionally, the identification of structurally intact native transposons suggested the tantalizing possibility that these *Tc1*-like transposons might be active.

We have developed the native PPTN transposon as a binary non-autonomous system, henceforth referred to as ‘*Passport*’, and have demonstrated it as functionally competent for transposition. We have tested the *Passport* transposon system in multiple vertebrate cell lines and provide molecular evidence of transposition. Our data indicates that *Passport* is active, and when mobilized prefers to integrate into the transcriptional units of genes. An expanded analysis of the phylogeny of *Passport* reveals that in addition to the presence of this transposon in other flatfish, including flounder (*Platichthys flesus*) and turbot (*Scophthalmus maximus*) (18), *Passport* shares a high degree of similarity with other recently reported transposons, including; *Eagle*, *Glan* and *Barb* (19,20). Combined, these observations suggest that this structurally and functionally intact transposon family may remain active in vertebrate lineages, and given a preference for integration into genes, may be contemporaneously active in vertebrate genome evolution.

MATERIALS AND METHODS

Vector construction

Sequence information, maps and material requests for these constructs can be found on our web site [<http://primer.ansci.umn.edu/fahrenkruglab>]. pPTnP-GeN. pPTnP-GeN was produced by cloning a 3.4 kb XmaI to NheI fragment of pKT2P-GeN (21), which contained the human PGK promoter and

miniintron, EGFP, the encephalomyocarditis virus internal ribosome site, neomycin phosphotransferase, and the rabbit beta-globin poly(A) signal, into pPTn2-SE. pPTn2-SE. Using T3-rev [TCTCCCTTAGTGAGGGT TAATT] and T7-rev [TCTCCCTATAGTGAGTCGTAT TA] primers a 102-bp PCR product of pKT2-SE that provides T7 and T3 polymerase binding sites orientated towards the ITRs of the PTn transposon and separated by a short multiple cloning site was cloned into the MscI site of prePTn1(-1). prePTn2(-1) was made by cloning a 0.65 kb BamHI to KpnI fragment of pCR4-PPTN2A into pK-A3 opened from KpnI to BamHI. pCR4-PPTN2A was created by topo cloning a 0.65 kb PCR product amplified from prePPTN2(-2) using oligos PPTN-F1 (BamHI) [AAGGATCCGATTACAGTGCCCTGCATAAGTAT] and PPTN-R2 (KpnI) [AAGGTACCGATTACAGTGC CTTGCATAAGTATT] into pCR4-Topo (Invitrogen). prePPTN2(-2) was created by amplifying the majority of pBluKS-PPTN5 (14) with oligos PPTN-OL2 [CCATCTT TGTTAGGGTTCACAGTA] and PPTN-OR1 [CCA GGTTCTACCAAGTATTGACACA]. The PCR fragment was then self-ligated to produce an empty transposon with a single MscI site in its interior. pKUb-PTs. pKUb-PTs was made by replacing the SB11 gene in pKUb-SB11 with PTs by cloning a 1.0 kb BamHI to NheI fragment from pCR4-PTs into pKUb-SB11 from NheI to BamHI. pCR4-PTs1 was made by cloning a PCR fragment of pBluKS-PPTN4 (14) amplified with primers CDS-PTs-F1 [AAAGCTAGCATGAA GACCAAGGAGCTCACC] and CDS-PTs-R1 [AAGGA TCCTCAATACTTGGTAGAAC] into pCR4-Topo (Invitrogen). pKC-PTs. The PTs coding region was placed behind the mCAGs promoter by cloning a 1.0 kb NheI to EcoRI fragment of pKUb-PTs containing the transposase into pK-mCAG opened from EcoRI to NheI. pK-mCAG was made by cloning the mCAG promoter from pSBTmCAG (22) as a 0.96-kb SmaI to EcoRI (filled) fragment into pK-SV40(A)x2 opened with AflII (filled).

pKUb-SB11. The construction of pKUb-SB11 has previously been described (21).

pKC-SB11. pKC-SB11 was made by cloning a 1.05-kb NheI to EcoRI fragment from pKUb-SB11 into pK-mCAG (21) opened from EcoRI to NheI.

pCMV-Bgal. pCMV-Bgal is available from Clontech (Mountainview, CA, USA) as pCMVb.

pPTnP-PTK. A 2.7 kb PvuII to PvuII fragment of pKPPTK_TS (21) was cloned into the EcoRV site of pPTn-RV to make pPTnP-PTK. pPTn-RV was made by cloning KJC-Adapter 4 [TCTCCCTTAGTGAGGGTTAATT GATATCTAACGACTCACTATAGGGAGA] into the MscI site of prePPTn2(-1) creating T7 and T3 polymerase binding sites orientated out towards the ITR of the PTn transposon and separated by an EcoRV site.

Cell culture and transposition assays

HT1080, HeLa, CHO-K1, NIH-3T3 and Vero cells are available from ATCC. TT and DF1 cells were kind gift from the laboratory Dr Douglas Foster, University of Minnesota (23,24). The isolation of PEGE cells has been described previously (21). CHO-K1 cells were grown in DMEM-F12 while all other cell lines were cultured with DMEM. Both mediums were enriched with 10% FBS, 1% Penn/Strep, and 1% L-Glutamine. PEGE cells were also enriched with insulin at 10 mg/ml. Transposition assays were carried out after seeding cells in six-well plates to achieve 60–80% confluence prior to transfection with DNA complexed with TransIT-LT1 transfection reagent (Mirus Bio Corporation, WI, USA).

Transfections were carried out according to manufacturer's instructions with a ratio of 3:1 lipid:DNA. Two days after transfection, cells were isolated from their wells with trypsin and collected by centrifugation. Two replicates of 30 000 cells were plated on 100mm dishes and selected in the appropriate selectable media. HT1080 cells were selected in 600 mg/ml of G418. For puromycin selection, HT1080, HeLa, Cho-K1, NIH-3T3, Vero, TT1, DF1 and PEGE cells were selected under 0.65, 0.4, 8.0, 1.5, 1.8, 0.35, 0.8 and 0.3 mg/ml puromycin, respectively. After colony formation, typically 9–12 days under selection, colonies were stained with methylene blue and counted.

Southern hybridization

Genomic DNA from independent clones derived after transfection with *Passport* transposons (pPTnP-PTK) and *Passport* transposase (pKC-PTs) was isolated using standard methods. Approximately 10 mg of DNA was digested with AseI and run on a 0.7% agarose gel. The DNA was transferred to a positively charged nylon membrane using 10X SSC and standard methods. The membrane was hybridized with a random primed fragment of pKP-PTK-TS isolated after digestion with XmaI. This probe contains the bulk of the puromycin-thymidine kinase gene, about 1.5 kb.

Cloning junction fragments

Blocked linker-mediated PCR was performed as described (21) except that DNA was obtained from colonies of cells that had been dried and stained with methylene blue. Briefly, genomic DNA was digested with a cocktail of restriction enzymes, including XbaI, NheI, AvrII and SpeI. The DNA was ligated to a blocked linker made by annealing the oligos primerette-long [CCTCCACTAC GACTCACTGAAGGGCAAGCAGTCCTAACAAACC ATG] and blink-XbaI [50P-CTAGCATGGTTGTTAGG ACTGCTTGC-30P]. Nested PCR was performed on the ligated DNA to specifically amplify junctions between the *Passport* transposon and genomic DNA. The transposon-specific primers for the primary PCR included PTn- IRDR(L)-O1 [GTGTTGGTCCATTACATAAACTCAC GATGAA] or PTn-IRDR(R)-O1 [GGGTGAATACTT ATGCACCCAACAGATG], transposon-specific primers for the secondary PCR reactions included PTn-IRDR(L)- O2 [GCATGACAAAATGTAGAAAAGTCCAAAGG] or PTn-IRDR(R)-O2 [CAGTACATAATGGGAAAAA GTCCAAGGG].

Phylogenetic analysis

The 1626 bp DNA sequence of PPTN (*Passport*) was used to query the entire ENSEMBL (www.ensembl.org) genome database using BLASTN. Consensus DNA sequences were derived, as described by Leaver (14), from a minimum of seven of the most similar sequences from each genome. Deduced consensus transposase amino acid sequences were aligned using ClustalW and phylogenetic trees generated as described (14). The Atlantic salmon (*Salmo salar*) and rainbow trout (*Oncorhynchus mykiss*) EST and tentative consensus cDNA databases (<http://compbio.dfci.harvard.edu/tgi/>) were also interrogated with PPTN using BLASTN and sequences assembled into consensus polypeptides as described for genome sequences.

RESULTS

Native *Passport* is competent for transposition in cells of diverse vertebrate origin

The SB, Frog Prince and HsMar1 transposon systems are active in a wide array of vertebrate cells (11–13), although to differing degrees. In order to assess the ability and ubiquity of *Passport* function, we undertook an analysis of TnT in human (HeLa, HT1080), monkey (Vero), pig (PEGE), hamster (CHO), mouse (3T3), chicken (DF1) and turkey (TT) cells using a *Passport* transposon containing a puromycin thymidine kinase fusion protein (25) driven by the mouse PGK promoter (pPTnP-PTK). Cells were transfected with the pPTnPPTK transposon and a *Passport* transposase expression construct (pKC-PTs) at a Tn:Ts molar ratio of 1:0.5, or with the molar equivalent of pCMV-bgal. Following transfection, replicates of \sim 30 000 cells were plated and selected in puromycin, fixed, stained and enumerated. In all cases, *Passport*-dependent TnT resulted in the generation of a number of puromycin-resistant colonies far exceeding that observed for controls lacking transposase, in the case of HeLa cells reflecting at least a 40-fold enhancement (Figure 1). As with other transposon systems (12,26), TnT varied between cell types (as did background- resistant colony formation), although comparing relative transpositional activity across cell lines may be confounded by the fact that transfections were conducted under identical conditions that may be suboptimal for some cell lines. Nonetheless, native *Passport* is functional in cells from a broad sampling of vertebrate species.

Passport is sensitive to overproduction inhibition

Overproduction inhibition, where excess transposase reduces the rate of transposition, is a hallmark of *Tcl/mariner* elements (5). We thus undertook an analysis of this effect for *Passport* and compared its sensitivity to that of the well-characterized SB transposon system (27,28). A series of transfections was performed with varying ratios of transposase to transposon vector to measure the effect

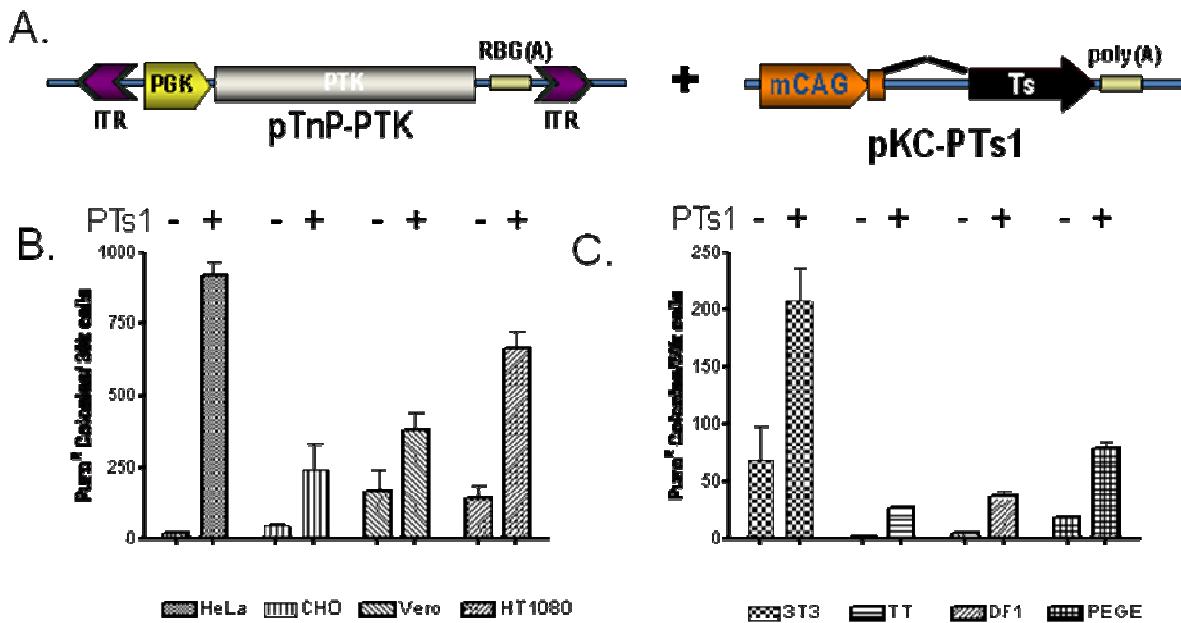


Figure 1. *Passport* functions in cells from a wide variety of vertebrate sources. (A) A *Passport* transposon that expresses Puromycin phosphotransferase was co-transfected with a source of *Passport* transposase, pKC-PTs1 (+PTs) or pCMV-Bgal (-PTs). Cells were selected in puromycin and stable colonies were counted. (B) HeLa, CHO, Vero and HT1080 cells displayed an increase in stable colony formation with the addition of *Passport* transposase. (C) 3T3, TT, DF1 and PEGE cells produced fewer colonies under these transfection conditions; however, the addition of *Passport* transposase significantly improved colony

formation. The addition of transposase rather than beta-galactosidase significantly increased colony formation in all cell types ($P<0.05$, except CHO where $P=0.07$).

of increasing transposase concentration on the rate of transposition. In addition, two promoters [human UbC (29) and mCAG (22,30)] were used to drive expression of the *Passport* transposase across a broad range of transposase levels (Figure 2). In HT1080 cells, gene expression from the mCAGs promoter is 5- to 10-fold higher than from the UbC promoter (data not shown). A constant amount of transposon (pPTnP-GeN, 75 fmol) was co-transfected with transposase vector containing either the UbC or mCAGs promoter (pKC-PTs or pKUb-PTs at a Tn:Ts molar ratio of 1:0.2, 1:0.5, 1:1, 1:2 or 1:5 corresponding to 15, 37.5, 75, 150 and 375 fmol of transposase plasmid). The total amount of transfected DNA was kept at 2 mg by supplementing with pCMVbgal DNA. To compare to the SB transposon system, identical reactions were performed with an SB transposon (pKT2P-GeN) and SB11 transposase expressed from the UbC and mCAGs promoters (pKUb-SB11 and pKCSB11). Following transfection, two replicates of 30 000 cells were plated and selected in G418 for 10–14 days, fixed, stained and the resulting colonies enumerated. Our previous studies indicated that a molar ratio of 1:1 SB transposon to SB transposase expressed from the human UbC promoter resulted in near-optimal transposition rates for the SB transposon system. Therefore to correct for any variation in transfection or selection, a 1:1 ratio of pKT2P-GeN:pKUb-SB11 was included as an internal standard for each day of transfection. The relative sensitivity of the two transposon systems to overproduction inhibition is presented in Figure 2B and C, where colony formation is expressed relative to the contemporary pKT2P-GeN:pKUb-SB11 internal standard. As shown in Figure 2B, the hyperactive SB system resulted in more than twice as many colonies as the native *Passport* system (Figure 2C) at their respective optimal Tn:Ts ratios. As expected, the SB transposon system is sensitive to overproduction inhibition. The peak transpositional activity for *Passport* was observed using a 1:5 ratio of pPTnP-GeN:pKUb-PTs1 or a 1:0.2 ratio of pPTnPGeN: pKC-PTs1, beyond which increasing transposase expression resulted in reduced transposition, indicating that *Passport* is indeed susceptible to overproduction inhibition.

Molecular characterization of *Passport* transposition

To validate transposition we examined the number of integration events per cellular clone by Southern analysis. Transposition is supported by hybridizing fragments of varying lengths, corresponding to genomic restriction sites at varying distances from the transposon insertion sites (Figure 3A). Non-transpositional DNA integration results in the formation of multi-copy concatemers (31) that are expected to result in a predictable restriction enzyme fragment derived from sites within the transposon vector (Figure 3B). The Southern analysis of DNA isolated from 15 HT1080 clones revealed that *Passport* indeed had transposed into the human genome, with one to four integrations per cellular clone (Figure 3C). Clones 4, 5 and 9 also contain a hybridizing band near the predicted size of a concatemer, although low signal intensity suggests low copy inserts not inconsistent with transposition. To further verify TnT by *Passport*, and to characterize the insertion target sites and preferences within HT1080 cells, junction fragments between the transposon and host genome were cloned and sequenced. *Passport*, like other Tc1 transposons, is expected to integrate into a TA dinucleotide and cause target-site duplication of the TA sequence at the ITR boundary. Table 1 lists 27 independent insertion events identified in HT1080 cells, demonstrating integration of the transposon into a TA within the human genome, and validating genuine transposition. The genomic location of each transposon insertion was determined by comparison of the cloned junction sequence to the human genome using Blastn (32). Insertions were dispersed across the human genome (Table 1). However, insertions did not appear to be

completely random as chromosome 1, which is twice as long as chromosome 12, has no integrations, whereas chromosome 12 has six integration events. In addition, the cloned junctions were found in transcription units in 63% of the cases, which is inconsistent with a completely random integration profile.

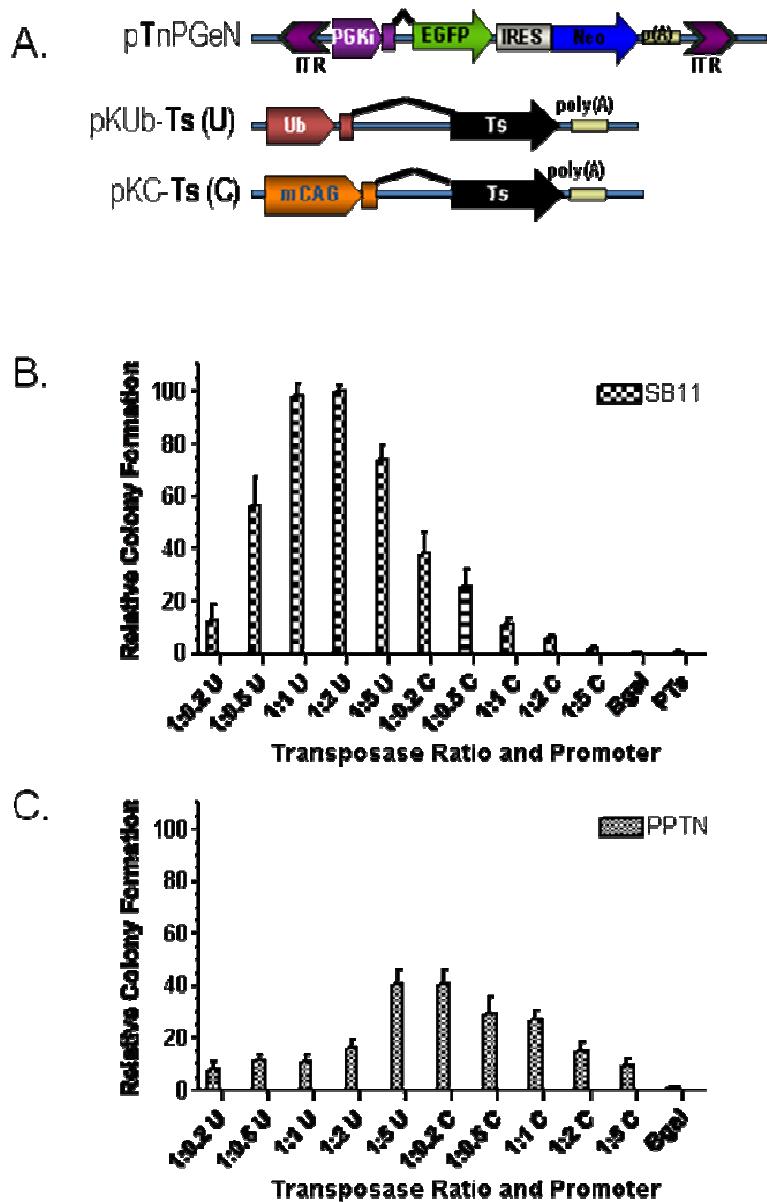


Figure 2. Examination of overproduction inhibition. (A) To examine the effect of transposase dose on transposition rates, a constant amount of pTnP-GeN (75 fmol) was co-transfected with five different molar ratios of transposase expression vector driven by either the human UbC promoter (pKUb-Ts) or the mCAGs promoter (pKC-Ts), where T and Ts generically refer to either SB or Passport components. In all cases the total amount of DNA transfected was adjusted to 2 mg by the addition of the appropriate amount of pCMV-Bgal. After transfection and selection in G418, colonies were counted and the data compared to an internal reference transfection of SB at a ratio of 1:1U. The raw data for the internal reference transfection came from a total of 30 replicates and ranged from 68 to 324, with a median of 150 and a mean of 170 (data not shown). The relative transposition

efficiencies confirm overproduction inhibition of (B) the SB transposon system and (C) demonstrate overproduction inhibition of the Passport transposon system. Error bars represent the SE.

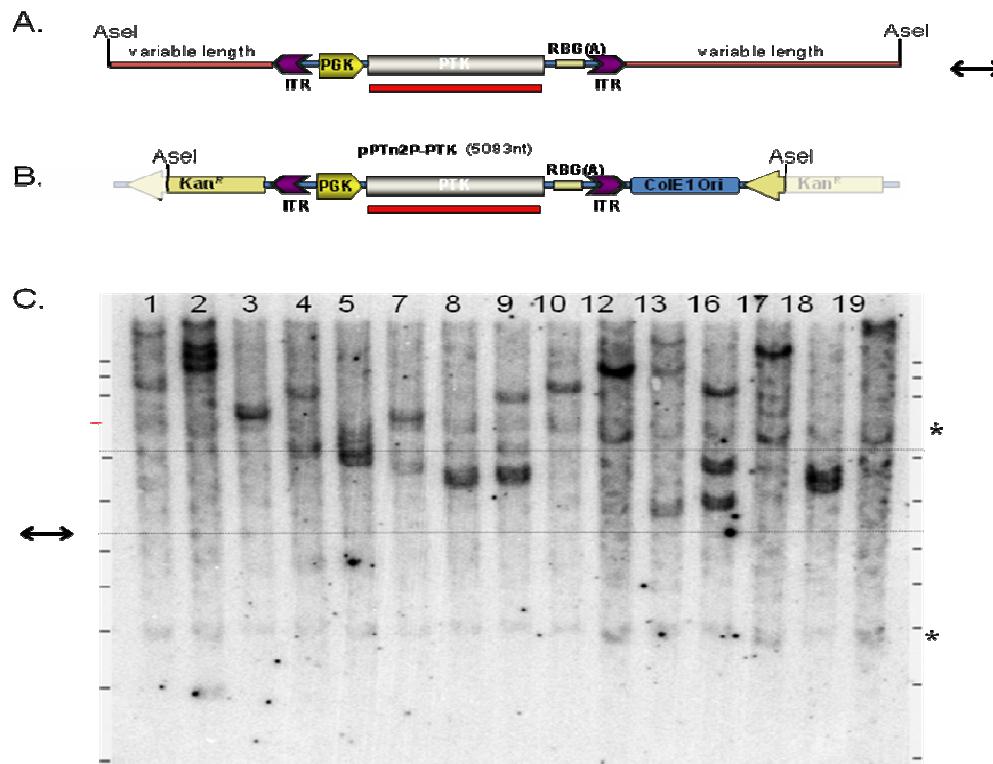


Figure 3. Evaluation of diversity and number of Passport genomic integrations. (A) Transposase-mediated recombination into the genome should result in transposon fragments of variable length after digestion with AseI. The sizes of the fragments are dependent on the proximity of AseI recognition sites in the neighboring chromatin, and can be observed following Southern hybridization. (B) Commonly, when DNA integrates without the enzymatic activity of transposase, head-to-tail concatemers of variable length are formed and integrate into the genome by non-homologous end-joining. In this case, the size of this internal high-representative fragment (~5.1 kb) is predictable based on the location of AseI sites within the transposon donor plasmid. (C) An image of our Southern hybridization of 15 independent HT1080 clones is shown. The paired head-to-tail arrows indicate the expected position of pPTnP-PTK concatemers that could potentially form during integration by non-homologous end joining. The line with outward facing arrowheads represents the size of the transposon and therefore the minimal expected size of a hybridizing fragment integrated by TnT. The asterisks mark two bands present in the HT1080 DNA that hybridize weakly with the PTK probe used here.

Passport-like transposons are present in other fish and amphibian genomes

The availability of sequenced genomes provides an opportunity to compare and categorize all transposons within a species and derive consensus sequences with a minimum of experimental bias. While *Passport* elements were originally isolated from plaice, nearly identical elements have been identified in other flatfish, including flounder and turbot (99% and 98% DNA identity). A recent search of ENSEMBL revealed the presence of additional related transposases with high nucleotide identity (>80%)

to *Passport* in the genomes and EST collections of *Xenopus tropicalis*, and pufferfish (*Takifugu rubripes*), stickleback (*Gasterosteus aculeatus*), medaka (*Oryzias latipes*), Atlantic salmon (*Salmo salar*) and rainbow trout (*Oncorhynchus mykiss*). *Passport*-like transposons were absent from all other ENSEMBL genomes, including those of the zebrafish (*Danio rerio*), despite the wide range and high copy number of other *Tc1*-like elements in this species. Comparison of the encoded transposase amino acid sequences show that relatives of *Passport* form a distinct family of *Tc1*-like transposons that is further divided into two subfamilies, including *Eagle/Glan* and *Barb/SSTN/RTTN* (Figure 4). The salmonids (salmon and rainbow trout) contain members of both subfamilies, whilst *X. tropicalis*, pufferfish, stickleback and medaka contain only the *Eagle/Glan* subfamily. The structure of *Passport* is somewhat intermediate between that of *Eagle/Glan* and *Barb/SSTN/RTTN*, in that its ITRs bear a strong resemblance to *Barb/SSTN/RTTN* (Figure 5A) whereas its transposase-coding region seems to bear more resemblance to the *Eagle/Glan* subfamily. Intriguingly, alignment of the DNA-binding domains of the transposases demonstrates a distinction between *Eagle/Glan* and *Passport/Barb/SSTN/RTTN* (Figure 5B), a difference that may be functionally connected to the ITRs of these elements.

Table 1. *Passport* junctions from integration into the human genome

(L)	Integration Site	(R)	Chrm	Pos	Gene ID
ATGATGCAGCTGGATCCGAT	TA	ATCGGTACCATTTAAATCTG	--		VECTOR
CTACCCAGACTCATTGATT	TA	actggaaagtctcttggta	2	p21	Intergenic
TTTCAATTCTTTGAATGTA	TA	cctacgaatagaattgtcg	2	q24	PLA2R1
gttggaaacttaacttgaac	TA	GTATAGAAAGGATGTCCGAA	2	q37	Intergenic
GTCCAGAAGTGAGTCAGAT	TA	gatcaattctgttagcacct	3	q25	Intergenic
GTTTTTATTATCTTGAGTA	TA	taccatgaattggcactgct	4	q32	(hmm3072864)
gatggttgcattaaacaatt	TA	TGTCCTAACATTATGCACAAT	5	p13	Intergenic
agacatagatgttacatata	TA	GATTAGTGTATTGTAGATA	6	p21	SUPT3H
tacatggtagttaaaatta	TA	CATCACTTTGTATATGGAGC	6	q14	Intergenic
catcttttatattgttagg	TA	GTAAGTGTATATTCAAACC	6	q21	FYN
GCAGAGGCCCTGTGTCAAGGTT	TA	aatgtgagctgcaggcagag	6	q26	TULP4
TCAAAGCAAGAAAGATTAT	TA	gctcgagtctctgcaacaaa	7	q32	PODXL
gagtggctaagtaggatatt	TA	GGTTCTAAAGCTAATAGAG	9	p24	PCD1LG2
TGTTGTCAAGTTATTGATA	TA	catccttaataatgccttt	9	q22	FANCC
CGCACCAAGTCGATAGTATT	TA	tgctaaagtctctctgaaat	11	q24	ETS1
GTACGTATAGATTGACTGG	TA	tacaacccctctggggcggc	12	p11	PPFIB1
gatgctagagaatcaacttt	TA	ATTCCAAAACCTGGTACATT	12	p12	PLEKHA5
TAATAGTGTGAGTGGTATC	TA	tctccactcaagaaaaatgg	12	p13	(hmm15010263)
gcatccccacagacacacct	TA	CCTGTTCAAGTGCAGGCACCT	12	p13	Intergenic
CAGCTCTCCCTCTGCCTCCC	TA	ttataagaacactgatgatt	12	q13	Intergenic
TCTATCATTACCCATGGCC	TA	gatcatgaaactgagtccta	12	q24.1	(hmm1230274)
AGAGGAGAGAAGGGAGCTT	TA	atacagttcgtcaaaag	13	q14	RCBTB1
TCCCTATAAGCTCTACCATG	TA	cctacagtccctaggcaga	14	q22	Intergenic
GCAAAACCACCATGGCATATG	TA	tagctatgtacaacatgc	15	q11	Intergenic
ttggtttgaataactggttt	TA	GTTCAATGTCAACCCTGCAA	15	q26	ALDH1A3
CCCAGAAACCAGGCCATAATC	TA	ctcatttagcaaaaatcatg	17	q21	NBR1
TAGTTATTTATACTAAGGTG	TA	aatgattgctgtcccactca	18	p11	(hmm1912534)
ttacacatatgatgccatgc	TA	TCTTATTGTTCTGTAGCTT	22	q13	Intergenic

The integration sites show the sequence outside the left ITR (L), the TA that is duplicated upon integration, and the sequence outside the right ITR (R). The first sequence indicates the sequence found in the donor plasmid (red), while the remaining

represent 27 Passport integrations sites all of which occurred by transposition as indicated by the exact junction at the ITR with a TA dinucleotide from the genome. In each case the sequence represented in CAPS was cloned by blocked LM-PCR and the sequence in lower case was derived from genome sequence data. In many cases, the Passport transposon integrated into known or (predicted) genes (Locus). The transposon integrations targeted a wide variety of chromosomal positions (Chrm Pos).

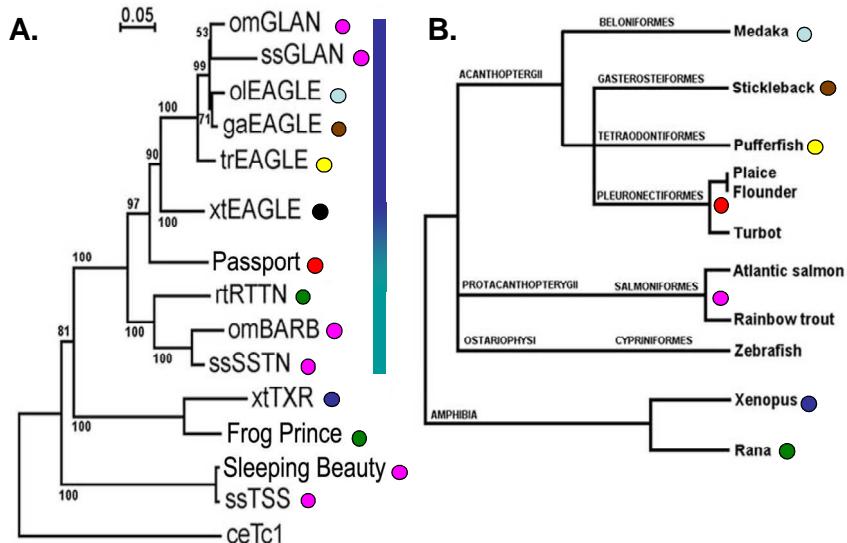


Figure 4. Phylogeny of Passport-like transposons and their hosts. (A) Neighbor-joining plot of multiply aligned transposase consensus amino acid sequences. Sequences were aligned with ClustalW, and plotted with NJplot. Numbers represent the percentage frequencies with which the tree topology was returned after 1000 iterations. The tree is rooted to Tc1 from *C. elegans*. Transposon designation is prefixed by host species identifier; om, rainbow trout; ol, medaka; ga, stickleback; tr, pufferfish; ss, Atlantic salmon; ce, *C. elegans*; rt, *Rana temporaria* (frog); xt, *Xenopus tropicalis*. Passport, Frog Prince and Sleeping Beauty were isolated from *Pleuronectes platessa*, *Rana sylvestris*, and a variety of salmonid species, respectively. (B) Phylogeny of host species adapted from Nelson (50). The colored dots assist in pairing the transposons shown in A with the species from B.

DISCUSSION

We have shown for the first time that a vertebrate-derived transposon from the *Tc1* family is active in its native form. *Passport* promotes TnT into vertebrate cells, with up to a 40-fold increase in HeLa cells and a 20-fold increase in HT1080 cells when compared to transgenesis without transposase. This corresponds to a rate of transposition up-to half the level we observed for SB11, itself a hyperactive mutant that is about 3-fold more active than the originally reanimated SB10 (27). Reanimation of SB, *Frog Prince* and *HsMar1* relied on phylogeny-informed reconstruction of extinct elements (11–13). Efforts to develop hyperactive transposases for application to TnT and gene therapy have applied both structure-based (33,34) and phylogenetics-informed (27,35,36) approaches. Indeed, the native *Passport* transposase sequence has been considered in improvements to SB and it contains several residues that have been synthetically introduced to generate hyperactive SB mutants, including; L205 and VR207/8 (35), and R130 and Q243 (27). Changes have also been made in the cis-acting ITR (36,37), as well as the spacer sequence between the ITRs of the SB transposon (26,36), resulting in the development of improved transposons, and evidence that only flanking IR/DR are required to constitute an effective transposon, a finding recapitulated in our study.

A

gAEAGLE	TACAGTGCCTTGC AAA GTATTGGGCCCCCTTG AA ACTTTG AC CTTTGCnnn..
omGLAN	TACAGTGCCTTGC AAA GTATTGGGCCCCCTTG AA ACTTTG AC CTTTGCnnn..
oleEAGLE	TACAGTGCCTTGC AAA GTATTGGGCCCCCTTG AA ACTTTCAACCCTTTGCnnn..
xteEAGLE	TACAGTGCCTTGC AAA GTATTGGGCCCCCTTG AA ACTTTCCACATTTGTnnn..
ombARB	TACAGTGCCTTGC AAA GTATTGGGCCCCCTTG AA CGTTTTTCCTATTTGTnnn..
ssSSTN	TACAGTGCCTTGC AAA GTATTGGGCCCCCTTG AA CGCTTTCCCTATTTGTnnn..
rtrRTN	TACAGTGCCTTGC AAA GTATTGGGCCCCCTTG AA CGCTTTACCTATTTGTnnn..
Passport	TACAGTGCCTTGC AAA GTATTGGGCCCCCTTG AA CGCTTTACATTTGTnnn..

* * * *

gAEAGLE	..nnn AAAA ..CTG AAA AAATTGGGCGTGCAA AA ATTATTCA GCCCC .TTTACTTTCA
omGLAN	..nnn AAAA ..CTG AAA AAATTGGGCGTGCAA AA ATTATTCA GCCCC .TTTACTTTCA
oleEAGLE	..nnn AAAA ..CTG AAA AAATTGGGCGTGCAA AA ATTATTCA GCCCC .TTTACTTTCA
xteEAGLE	..nnn ATAA ..CTG CAAA GTGGGGTGTGCGTAATTATTCA GCCCC TTTGGTCTGA
ombARB	..nnn AACTGGTGCCTGCATATGTATTCA CCCC .TTTGTATGA
ssSSTN	..nnn AAACTMCAGAAA ACTGGTGCCTGCATATGTATTCA CCCC .TTTGTATGA
rtrRTN	..nnn AA ..ACTGATAATTGGCATGTGCGTATGTATTCA CCCC .TTTGTATGA
Passport	..nnn AAAT ..CTG AAA ACTGTGACTGCATATGTATTCA CCCC .TTTACTGTGA

* *

B

gAEAGLE	MKNKEH I RQVRDTVV E KFK G FGFYKK I SQ L N I P R STVQ A IILK W KEYQTTANLQRPGRP
omGLAN	MKNKEH I RQVRDTVV E KFK G FGFYKK I SQ L N I P R STVQ A IILK W KEYQTTANLQRPGRP
oleEAGLE	MKNKEH I RQVRDTVV E KFK G FGFYKK I SQ L N I P R STVQ A IILK W KEYQTTANLQRPGRP
xteEAGLE	MKSKEH I RQVRD V E E KFK G GLGYKK I S K ALN I P R STVQ A IOKW E Y G ITTVNLPRQGRP
ombARB	MKTKE L SKQVRD V E E KFK G RSGLGYKK I S E TLN I P Q STIK S I I KKW E Y G ITTNLPKEGRP
ssSSTN	MKTKE L SKQVRD V E E KFK G RSGLGYKK I S E TLN I P Q STIK S I I KKW E H G TTANLPRDGRT
rtrRTN	MKTKE L SKQVRD V E E KFK G GLGYKK I S R LN I S T IK S I R KW E Y G TTANLPRGGRP
Passport	MKTKE L SKQVRD V E E KFK G GLGYKK I S R LN I S T IK S I R KW E Y G TTANLPRGGRP

Figure 5. Comparison of repeat sequences and transposase DNA-binding domains of Passport-like transposons. (A) Comparison of terminal and internal 50 repeats of the left ITR. Host species identifier as for Figure 4 prefixes transposon designation. Gray-shaded residues are common to the majority of sequences. Bars below the line indicate the conserved repeat units from within the inverted repeats. Pink highlighted sequences delineate differences between the Eagle/Glan and SSTN/Barb families. For the ITR sequences, Passport clearly aligns more closely with the SSTN/Barb families within the direct repeats, but evidence of convergence towards Eagle/Glan sequences are observed just outside of these direct repeats as indicated by the asterisks. (B) Comparison of the putative DNA-binding domains of Passport and related transposases. Pink-shaded residues indicate the amino acids that distinguish members of the Eagle/Glan family from SSTN/Barb. Residues shaded red indicate the convergence of the active Passport sequence towards the Eagle/Glan family to which it is more closely related over the length of the entire protein, whereas blue residues show some convergence of the *X. tropicalis* Eagle element towards the SSTN/Barb subfamily. Residues shaded in yellow seem to be unique within Passport and may be important for its activity.

We demonstrated that *Passport* transposons display overproduction inhibition analogous to other Tc1/mariner elements (5). Interestingly, despite using identical promoters in the SB and *Passport* transposase expression constructs, optimal transposition and the emergence of overproduction inhibition for *Passport* occurred under conditions expected to correspond to significantly higher levels of transposase expression. We can estimate that optimal transposition for *Passport* requires more than double the amount of transposase expression as SB, since their maximal transposition occurred at Tn:Ts molar equivalents of 1:5 and 1:2, respectively. This could result from differences in the translational efficiency or stability of the encoded transposases, differences in the affinities of the transposases for their corresponding transposons, or from innate variance in transpositional activity (disparities not unexpected when comparing native and hyperactive transposon systems).

We conducted a limited survey of *Passport* integration sites and observed a significant preference for integration into genes (likelihood ratio >5000:1). This characteristic has also been observed for the piggyBac transposon system, a non-Tc1 element (38), but contrasts sharply with the random

integration site preferences for the SB transposon system (39), suggesting *Passport* may be especially suitable for functional genomics applications. Indeed, there are a variety of transposons useful in vertebrate cells, including; *SB* (11), *Frog Prince* (12), *Tol2* (40), *minos* (41,42), *piggyBac* (43,44), *Ac/Ds* (45,46), *TolI* (47), *HsMar1* (13), *Harbinger* (48) and now *Passport*. The unique transposon ITRs and preferences for target sites provides a toolbox that can be implemented in response to a variety of needs. Particularly relevant applications include the serial introduction of multiple transgenes into cells/animals without disturbing previously integrated transposons, and enhanced insertional mutagenesis screens that capitalize on differences in integration site preferences (49).

An examination of genome-sequence data for diverse organisms shows that *Tc1* elements related to *Passport* are also present in *X. tropicalis* and in other fish species. In *X. tropicalis* these transposons have been termed *Eagle* (20) and in rainbow trout *Glan* and *Barb* (19). Members of the *Eagle/Glan* family are phylogenetically widespread and their distribution is generally in agreement with the accepted phylogeny for these species (50), indicating vertical transmission. On the other hand, *Barb/SSTN/RTTN* family members are restricted to salmonid fish and *Rana* frogs (14), likely representing horizontal transmission.

Passport transposons appear to be an intermediate between the *Eagle/Glan* group and the *Barb/SSTN/RTTN* group of transposons, with ITRs bearing a strong resemblance to *Barb/SSTN/RTTN* but a transposase more akin to the *Eagle/Glan* subfamily. Perhaps transposon ‘hybridization’ has resulted in the genesis of *Passport* in Pleuronectid genomes. Regardless, the correspondence of differences in the DNA-binding domains of these closely related transposases and their cognate ITRs provides a rationale for future investigations focused on the evolution and engineering of transposase specificity.

Passport represents the first opportunity to study the biology of a native functionally intact vertebrate *Tc1* element. As such, it may provide important information about transposon function and regulation that could have been lost or modified during synthetic reanimation or subsequent hyperactivity mutagenesis. Provided suitable culture conditions can be developed for plaice cells, *Passport* may also provide the seminal opportunity to study evolved regulatory or accommodating interactions between vertebrate *Tc1* transposons and the host genome, perhaps missing when elements are artificially introduced into a naive genome.

The functional integrity and phylogenetic restriction of *Passport* sequences suggest contemporaneous activity. Further study of *Passport* integration sites by high throughput sequencing could provide important information about the history of transposition in flatfish, contributing to our understanding of the roles of DNA transposons in vertebrate genome evolution (51). The prevalence and genomic location of *Passport* transposons among geographically diverse plaice and related natural fish populations may be helpful in determining the time course and extent of *Passport* infiltration. Of particular interest is the potential to address the hypothesis that transposon activity may increase in response to stress (52), a phenomenon recently supported by observations in other teleosts (19). Although speculative, perhaps *Passport* activity could prove a useful indicator for managing geographically disperse flatfish fisheries in the face of ecological and fishing pressures.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING This work and funding for open access publication was supported by grants from the USDA National Research Initiative (2008-35205-18852 to S.C.F.).

Conflict of interest statement. None declared.

REFERENCES

1. Girard,L. and Freeling,M. (1999) Regulatory changes as a consequence of transposon insertion. *Dev. Genet.*, 25, 291–296.
2. Lander,E.S., Linton,L.M., Birren,B., Nusbaum,C., Zody,M.C., Baldwin,J., Devon,K., Dewar,K., Doyle,M., FitzHugh,W. et al. (2001) Initial sequencing and analysis of the human genome. *Nature*, 409, 860–921.
3. Plasterk,R.H., Izsvák,Z. and Ivics,Z. (1999) Resident aliens: the Tc1/mariner superfamily of transposable elements. *Trends Genet.*, 15, 326–332.
4. Robertson,H.M. (1995) The Tc1-mariner superfamily of transposons in animals. *J. Insect Physiol.*, 41, 99–105.
5. Hartl,D.L., Lozovskaya,E.R., Nurminsky,D.I. and Lohe,A.R. (1997) What restricts the activity of mariner-like transposable elements. *Trends Genet.*, 13, 197–201.
6. Jacobson,J.W., Medhora,M.M. and Hartl,D.L. (1986) Molecular structure of a somatically unstable transposable element in *Drosophila*. *Proc. Natl Acad. Sci. U.S.A.*, 83, 8684–8688.
7. Collins,J., Forbes,E. and Anderson,P. (1989) The Tc3 family of transposable genetic elements in *Caenorhabditis elegans*. *Genetics*, 121, 47–55.
8. Moerman,D.G. and Waterston,R.H. (1984) Spontaneous unstable unc-22 IV mutations in *C. elegans* var. Bergerac. *Genetics*, 108, 859–877.
9. Barry,E.G., Witherspoon,D.J. and Lampe,D.J. (2004) A bacterial genetic screen identifies functional coding sequences of the insect mariner transposable element Famar1 amplified from the genome of the earwig, *Forficula auricularia*. *Genetics*, 166, 823–833.
10. Munoz-Lopez,M., Siddique,A., Bischerour,J., Lorite,P., Chalmers,R. and Palomeque,T. (2008) Transposition of Mboumar-9: identification of a new naturally active mariner-family transposon. *J. Mol. Biol.*, 382, 567–572.
11. Ivics,Z., Hackett,P.B., Plasterk,R.H. and Izsvák,Z. (1997) Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. *Cell*, 91, 501–510.
12. Miskey,C., Izsvák,Z., Plasterk,R.H. and Ivics,Z. (2003) The Frog Prince: a reconstructed transposon from *Rana pipiens* with high transpositional activity in vertebrate cells. *Nucleic Acids Res.*, 31, 6873–6881.
13. Miskey,C., Papp,B., Mates,L., Sinzelle,L., Keller,H., Izsvák,Z. and Ivics,Z. (2007) The ancient mariner sails again: transposition of the human Hsmar1 element by a reconstructed transposase and activities of the SETMAR protein on transposon ends. *Mol. Cell. Biol.*, 27, 4589–4600.
14. Leaver,M.J. (2001) A family of Tc1-like transposons from the genomes of fishes and frogs: evidence for horizontal transmission. *Gene*, 271, 203–214.
15. Leaver,M.J., Wright,J. and George,S.G. (1997) Structure and expression of a cluster of glutathione S-transferase genes from a marine fish, the plaice (*Pleuronectes platessa*). *Biochem. J.*, 321 (Pt 2), 405–412.
16. Izsvák,Z., Ivics,Z. and Hackett,P.B. (1995) Characterization of a Tc1-like transposable element in zebrafish (*Danio rerio*). *Mol. Gen. Genet.*, 247, 312–322.
17. Robertson,H.M. (1993) The mariner transposable element is widespread in insects. *Nature*, 362, 241–245.
18. Pocwierz-Kotus,A., Burzynski,A. and Wenne,R. (2007) Family of Tc1-like elements from fish genomes and horizontal transfer. *Gene*, 390, 243–251.
19. Krasnov,A., Koskinen,H., Afanasyev,S. and Molsa,H. (2005) Transcribed Tc1-like transposons in salmonid fish. *BMC Genomics*, 6, 107.

20. Sinzelle,L., Pollet,N., Bigot,Y. and Mazabraud,A. (2005) Characterization of multiple lineages of Tc1-like elements within the genome of the amphibian *Xenopus tropicalis*. *Gene*, 349, 187–196.
21. Clark,K.J., Carlson,D.F., Foster,L.K., Kong,B.W., Foster,D.N. and Fahrenkrug,S.C. (2007) Enzymatic engineering of the porcine genome with transposons and recombinases. *BMC Biotechnol.*, 7, 42.
22. Ohlfest,J.R., Frandsen,J.L., Fritz,S., Lobitz,P.D., Perkinson,S.G., Clark,K.J., Nelsestuen,G., Key,N.S., McIvor,R.S., Hackett,P.B. et al. (2005) Phenotypic correction and long-term expression of factor VIII in hemophilic mice by immunotolerization and nonviral gene transfer using the Sleeping Beauty transposon system. *Blood*, 105, 2691–2698.
23. Schaefer-Klein,J., Givol,I., Barsov,E.V., Whitcomb,J.M., VanBrocklin,M., Foster,D.N., Federspiel,M.J. and Hughes,S.H. (1998) The EV-O-derived cell line DF-1 supports the efficient replication of avian leukosis-sarcoma viruses and vectors. *Virology*, 248, 305–311.
24. Kong,B.W., Foster,L.K. and Foster,D.N. (2007) Establishment of an immortal turkey turbinate cell line suitable for avian metapneumovirus propagation. *Virus Res.*, 127, 106–115.
25. Abuin,A. and Bradley,A. (1996) Recycling selectable markers in mouse embryonic stem cells. *Mol. Cell. Biol.*, 16, 1851–1856.
26. Izsvak,Z., Ivics,Z. and Plasterk,R.H. (2000) Sleeping Beauty, a wide host-range transposon vector for genetic transformation in vertebrates. *J. Mol. Biol.*, 302, 93–102.
27. Geurts,A.M., Yang,Y., Clark,K.J., Liu,G., Cui,Z., Dupuy,A.J., Bell,J.B., Largaespada,D.A. and Hackett,P.B. (2003) Gene transfer into genomes of human cells by the Sleeping Beauty transposon system. *Mol. Ther.*, 8, 108–117.
28. Mikkelsen,J.G., Yant,S.R., Meuse,L., Huang,Z., Xu,H. and Kay,M.A. (2003) Helper-independent Sleeping Beauty transposontransposase vectors for efficient nonviral gene delivery and persistent gene expression in vivo. *Mol. Ther.*, 8, 654–665.
29. Gill,D.R., Smyth,S.E., Goddard,C.A., Pringle,I.A., Higgins,C.F., Colledge,W.H. and Hyde,S.C. (2001) Increased persistence of lung gene expression using plasmids containing the ubiquitin C or elongation factor 1alpha promoter. *Gene Therapy*, 8, 1539–1546.
30. Niwa,H., Yamamura,K. and Miyazaki,J. (1991) Efficient selection for high-expression transfecants with a novel eukaryotic vector. *Gene*, 108, 193–199.
31. Garrick,D., Fiering,S., Martin,D.I. and Whitelaw,E. (1998) Repeatinduced gene silencing in mammals. *Nat. Genet.*, 18, 56–59.
32. Altschul,S.F., Gish,W., Miller,W., Myers,E.W. and Lipman,D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.*, 215, 403–410.
33. Yant,S.R., Park,J., Huang,Y., Mikkelsen,J.G. and Kay,M.A. (2004) Mutational analysis of the N-terminal DNA-binding domain of Sleeping Beauty transposase: critical residues for DNA binding and hyperactivity in mammalian cells. *Mol. Cell. Biol.*, 24, 9239–9247.
34. Zayed,H., Izsvak,Z., Khare,D., Heinemann,U. and Ivics,Z. (2003) The DNA-bending protein HMGB1 is a cellular cofactor of Sleeping Beauty transposition. *Nucleic Acids Res.*, 31, 2313–2322.
35. Baus,J., Liu,L., Heggestad,A.D., Sanz,S. and Fletcher,B.S. (2005) Hyperactive transposase mutants of the Sleeping Beauty transposon. *Mol. Ther.*, 12, 1148–1156.
36. Zayed,H., Izsvak,Z., Walisko,O. and Ivics,Z. (2004) Development of hyperactive Sleeping Beauty transposon vectors by mutational analysis. *Mol. Ther.*, 9, 292–304.
37. Cui,Z., Geurts,A.M., Liu,G., Kaufman,C.D. and Hackett,P.B. (2002) Structure-function analysis of the inverted terminal repeats of the Sleeping Beauty transposon. *J. Mol. Biol.*, 318, 1221–1235.

38. Wilson,M.H., Coates,C.J. and George,A.L. Jr (2007) piggyBac transposon-mediated gene transfer in human cells. *Mol. Ther.*, 15, 139–145.
39. Yant,S.R., Wu,X., Huang,Y., Garrison,B., Burgess,S.M. and Kay,M.A. (2005) High-resolution genome-wide mapping of transposon integration in mammals. *Mol. Cell. Biol.*, 25, 2085–2094.
40. Hori,H., Suzuki,M., Inagaki,H., Oshima,T. and Koga,A. (1998) An active Ac-like transposable element in teleost fish. *J. Marine Biotechnol.*, 6, 206–207.
41. Klinakis,A.G., Zagoraiou,L., Vassilatis,D.K. and Savakis,C. (2000) Genome-wide insertional mutagenesis in human cells by the Drosophila mobile element Minos. *EMBO Rep.*, 1, 416–421.
42. Franz,G. and Savakis,C. (1991) Minos, a new transposable element from *Drosophila hydei*, is a member of the Tc1-like family of transposons. *Nucleic Acids Res.*, 19, 6646.
43. Ding,S., Wu,X., Li,G., Han,M., Zhuang,Y. and Xu,T. (2005) Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. *Cell*, 122, 473–483.
44. Fraser,M.J., Ciszczon,T., Elick,T. and Bauser,C. (1996) 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.*, 5, 141–151.
45. Emelyanov,A., Gao,Y., Naqvi,N.I. and Parinov,S. (2006) Transkingdom transposition of the maize dissociation element. *Genetics*, 174, 1095–1104.
46. McClintock,B. (1950) The origin and behavior of mutable loci in maize. *Proc. Natl Acad. Sci. U.S.A.*, 36, 344–355.
47. Koga,A., Shimada,A., Kuroki,T., Hori,H., Kusumi,J., Kyono-Hamaguchi,Y. and Hamaguchi,S. (2007) The Tol1 transposable element of the medaka fish moves in human and mouse cells. *J. Human Genet.*, 52, 628–635.
48. Sinzelle,L., Kapitonov,V.V., Grzela,D.P., Jursch,T., Jurka,J., Izsvak,Z. and Ivics,Z. (2008) Transposition of a reconstructed Harbinger element in human cells and functional homology with two transposon-derived cellular genes. *Proc. Natl Acad. Sci. U.S.A.*, 105, 4715–4720.
49. Thibault,S.T., Singer,M.A., Miyazaki,W.Y., Milash,B., Dompe,N.A., Singh,C.M., Buchholz,R., Demsky,M., Fawcett,R., Francis-Lang,H.L. et al. (2004) A complementary transposon tool kit for *Drosophila melanogaster* using P and piggyBac. *Nat. Genet.*, 36, 283–287.
50. Nelson,J.S. (2006) Fishes of the World, 4th edn., John Wiley, Hoboken, NJ.
51. Wessler,S.R. (2006) Transposable elements and the evolution of eukaryotic genomes. *Proc. Natl Acad. Sci. U.S.A.*, 103, 17600–17601.
52. McClintock,B. (1984) The significance of responses of the genome to challenge. *Science (New York, N.Y.)*, 226, 792–801.