

# New *Mos1 mariner* transposons suitable for the recovery of gene fusions in vivo and in vitro

Sophie Goyard<sup>a</sup>, Luiz R.O. Tosi<sup>a,1</sup>, Julia Gouzova<sup>a</sup>, John Majors<sup>b</sup>, Stephen M. Beverley<sup>a,\*</sup>

<sup>a</sup>Department of Molecular Microbiology, Washington University Medical School, Box 8230, 660 South Euclid Avenue, St. Louis, MO 63110, USA

<sup>b</sup>Department of Biochemistry and Molecular Biophysics, Washington University Medical School, 660 South Euclid Avenue, St. Louis, MO 63110, USA

Received by D. Finnegan

## Abstract

The *Drosophila Mos1* element can be mobilized in species ranging from prokaryotes to protozoans and vertebrates, and the purified transposase can be used for in vitro transposition assays. In this report we developed a 'mini-*Mos1*' element and describe a number of useful derivatives suitable for transposon mutagenesis in vivo or in vitro. Several of these allow the creation and/or selection of tripartite protein fusions to a green fluorescent protein–phleomycin resistance (GFP-PHLEO) reporter/selectable marker. Such X-GFP-PHLEO-X fusions have the advantage of retaining 5' and 3' regulatory information and N- and C-terminal protein targeting domains. A *Mos1* derivative suitable for use in transposon-insertion mediated linker insertion (TIMLI) mutagenesis is described, and transposons bearing selectable markers suitable for use in the protozoan parasite *Leishmania* were made and tested. A novel 'negative selection' approach was developed which permits in vitro assays of transposons lacking bacterial selectable markers. Application of this assay to several *Mos1* elements developed for use in insects suggests that the large *mariner* pM[*en*] element used previously in vivo is poorly active in vitro, while the *Mos1*-Act-EGFP transposon is highly active. © 2001 Elsevier Science B.V. All rights reserved.

**Keywords:** In vitro transposition; GFP + ; Reporter gene; Gene inactivation; Linker-insertion mutagenesis

## 1. Introduction

Transposon mutagenesis has been widely used as a tool for genetic mapping, DNA sequencing, insertional inactivation, and promoter strength assays (Berg et al., 1989). Recently, methods for carrying out transposition in vitro with purified transposase and engineered transposons have been developed for several systems, including Tn5 (Goryshin and Reznikoff, 1998), Ty1 (Merkulov and Boeke, 1998), Tn7 (Biery et al., 2000), and Tc1/*mariner* elements including *Himar1* (Lampe et al., 1996), Sleeping Beauty (Fischer et al., 2001), and *Mos1* (Tosi and Beverley, 2000). These have facilitated rapid generation of large transposon insertion libraries into a variety of molecular cloning

vectors and targets, and have been proven to be robust and valuable tools in different genetics studies.

Our laboratory has developed in vitro as well as in vivo approaches using the *Drosophila mauritiana* element *Mos1* (Gueiros-Filho and Beverley, 1997; Coates et al., 1998; Tosi and Beverley, 2000; Beverley et al., 2002). *Mos1* is one of the defining members of the Tc1/*mariner* transposon family (Robertson, 1995), whose activity was first detected in *Drosophila* in vivo. As typical for members of this family, *Mos1* is small (1.3 kb) and contains the transposase gene flanked by terminal inverted repeats (Plasterk, 1996; Hartl et al., 1997). Transposition is mediated by a cut and paste mechanism following recognition of the terminal inverted repeats by the transposase, and insertion into a TA target site dinucleotide (Plasterk, 1996). An efficient system for in vitro transposition of *Mos1*-derived elements has been developed (Tosi and Beverley, 2000). In vivo, transposition of *mariner*-based elements has been engineered into organisms ranging from prokaryotes to protozoans to vertebrates (Gueiros-Filho and Beverley, 1997; Coates et al., 1998; Fadool et al., 1998; Sherman et al., 1998; Rubin et al., 1999; Zhang et al., 2000; Fischer et al., 2001).

The ability to use the same transposon in vivo and in vitro is a powerful advantage, permitting investigators to use both

Abbreviations: IR, inverted repeat; GFP, green fluorescent protein; PHLEO, phleomycin/zeocin resistance marker; NEO, Tn5 neomycin resistance marker; Km, Tn903 kanamycin resistance marker; HYG, hygromycin resistance marker; SAT, nourseothricin resistance marker; AG3, *Leishmania* splice acceptor; PCR, polymerase chain reaction; bp, base pair

\* Corresponding author. Tel.: +1-314-747-2630; fax: +1-314-747-2634.

E-mail address: beverley@borcim.wustl.edu (S.M. Beverley).

<sup>1</sup> Present address: Departamento de Biologia Celular e Molecular e Biotecnologia Patogênicos, Faculdade Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto – SP, Brazil.

approaches as most appropriate. In this work we describe a variety of new *Mos1* mariner derivatives that are suitable for a diverse range of transposon applications such as insertional mutagenesis, identification and recovery of gene fusions and transposon-mediated linker insertion (TIMLI) mutagenesis (Hayes and Hallet, 2000).

One goal was the design of transposons suitable for the generation of protein fusions to the green fluorescent protein (GFP) and the universal selectable marker for phleomycin resistance (PHLEO). Additionally, we describe a useful assay for transposition that does not utilize transposon-borne genetic markers, instead relying upon negative selection to detect transposition events. This allows testing of transposons lacking selectable markers or suicide replication origins required for the in vitro transposition approach, due to constraints of size or functionality in their target organism or applications.

These new transposons should be applicable in studies of many species, and expand the utility of the *Mos1* mariner system as an experimental tool. A convenient and highly active ‘mini-mariner’ element has also been generated which enables rapid generation of new transposons for future applications. We have tested these transposons in the parasitic protozoan *Leishmania*, and several of the transposons described in this work contain *Leishmania*-specific regulatory sequences that enhance their utility for work in this species.

## 2. Materials and methods

### 2.1. Bacteria and growth conditions

*Escherichia coli* strains DH10B (lab strain B2192; Gibco BRL), DH5 $\alpha$ - $\lambda$ pir (B3684; (Garraway et al., 1997), Top10 (B 4532) and BLR (DE3) pLysS (B3842; Novagen) were used in this study. DH5 $\alpha$ - $\lambda$ pir was used to provide the *pir* gene function for plasmids bearing the R6K origin of replication. All strains were grown in standard LB media supplemented when necessary with the appropriate antibiotics (ampicillin 100  $\mu$ g/ml; zeocin 50  $\mu$ g/ml; kanamycin 50  $\mu$ g/ml; hygromycin 125  $\mu$ g/ml; nourseothricin 50  $\mu$ g/ml; chloramphenicol 30  $\mu$ g/ml). Transposons pEL-Apr (B3780) and pM[*cn*] (B3779) were described previously (Coates et al., 1998; Gehring et al., 2000) and were provided by A. Gehring and A. James, respectively. Transposon pMos1-Act-EGFP (B3758) was provided by H. Zieler (unpublished work) and contains a *Drosophila* actin promoter driving a GFP + reporter gene inserted within a *mariner* element. Plasmid pBSCm (B1930; Cm<sup>R</sup>) was obtained from Stratagene.

### 2.2. Construction of *Mos1* derivatives

Plasmid DNA preparations, restriction enzyme digests, and ligations were carried out using standard methods (Maniatis et al., 1982). DH5 $\alpha$ - $\lambda$ pir was transformed using

CaCl<sub>2</sub> procedures, while electroporation was used to transform DH10B and Top 10. Typically blunt-ended DNAs were generated by the action of T4 DNA polymerase in the presence of dNTPs, except when indicated. In the following descriptions, a slash (/) signifies a gene cassette that lacks an initiating ATG codon (for example /GFP) or stop codon (/GEP3/), while an asterisk (\*\*) signifies a cassette bearing a stop codon separating a reporter gene and selectable marker (such as /GFP\*K). The plasmid backbone bearing most of the transposons (pELHY6 $\Delta$ -0) is shown in Fig. 1 and a summary of the transposons is shown in Fig. 2.

#### 2.2.1. pELHYGmos1 (B3520)

pELHYGmos1 was created by inserting *PacI*–*HindIII* fragment (blunt) containing the wild type *Mos1* element from pNEB-*Mos1* (B3082) into the *Bam*HI site (blunt) of pEL-HYG (B2766; Garraway et al., 1997).

#### 2.2.2. pHM6K (B3545)

pHM6K was created by inserting a *Bam*HI–*Sma*I fragment (containing the R6K origin of replication of plasmid pGP704 (Miller and Mekalanos, 1988) into pELHYGmos1 digested with *Bgl*II–*Bsa*AI, thereby exchanging the OriC with the R6K origin of replication.

#### 2.2.3. pELHY6-0 (B3546)

pELHY6-0 was created in two steps. First the chloramphenicol

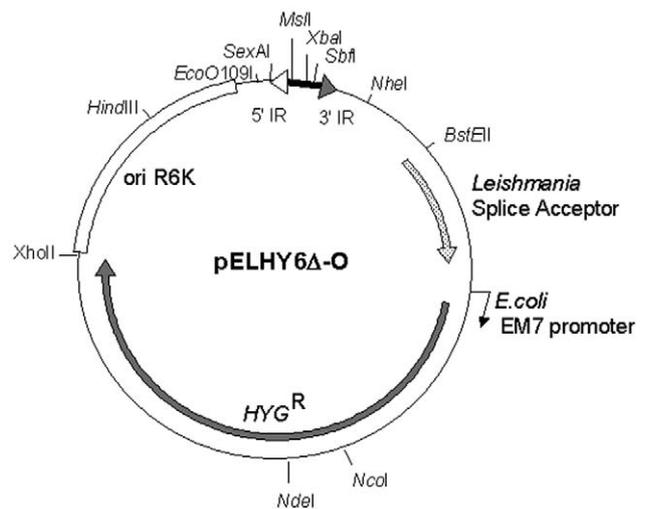


Fig. 1. Structure of the ‘empty’ minimal *Mos1* pELHY6 $\Delta$ -0 transposon donor plasmid. This plasmid contains the R6K origin of replication (oriR6K) and a hygromycin resistance marker (HYG<sup>R</sup>), allowing it to be used conveniently as a donor in the in vitro transposition reaction (Tosi and Beverley, 2000), and an ‘empty’ *Mos1* cassette consisting of 5' and 3' *Mos1* IRs (open and filled triangles, respectively). The unique restriction sites *Msl*I, *Xba*I and *Sbf*I separating the IRs are convenient for the introduction of selectable markers and reporters as summarized in Figs. 2 and 3. Relevant restriction sites are indicated. The HYG<sup>R</sup> marker is active in *E. coli* and *Leishmania*, due respectively to the presence of the bacterial EM7 promoter and *Leishmania* AG3 splice acceptor upstream of the resistance marker (Garraway et al., 1997).

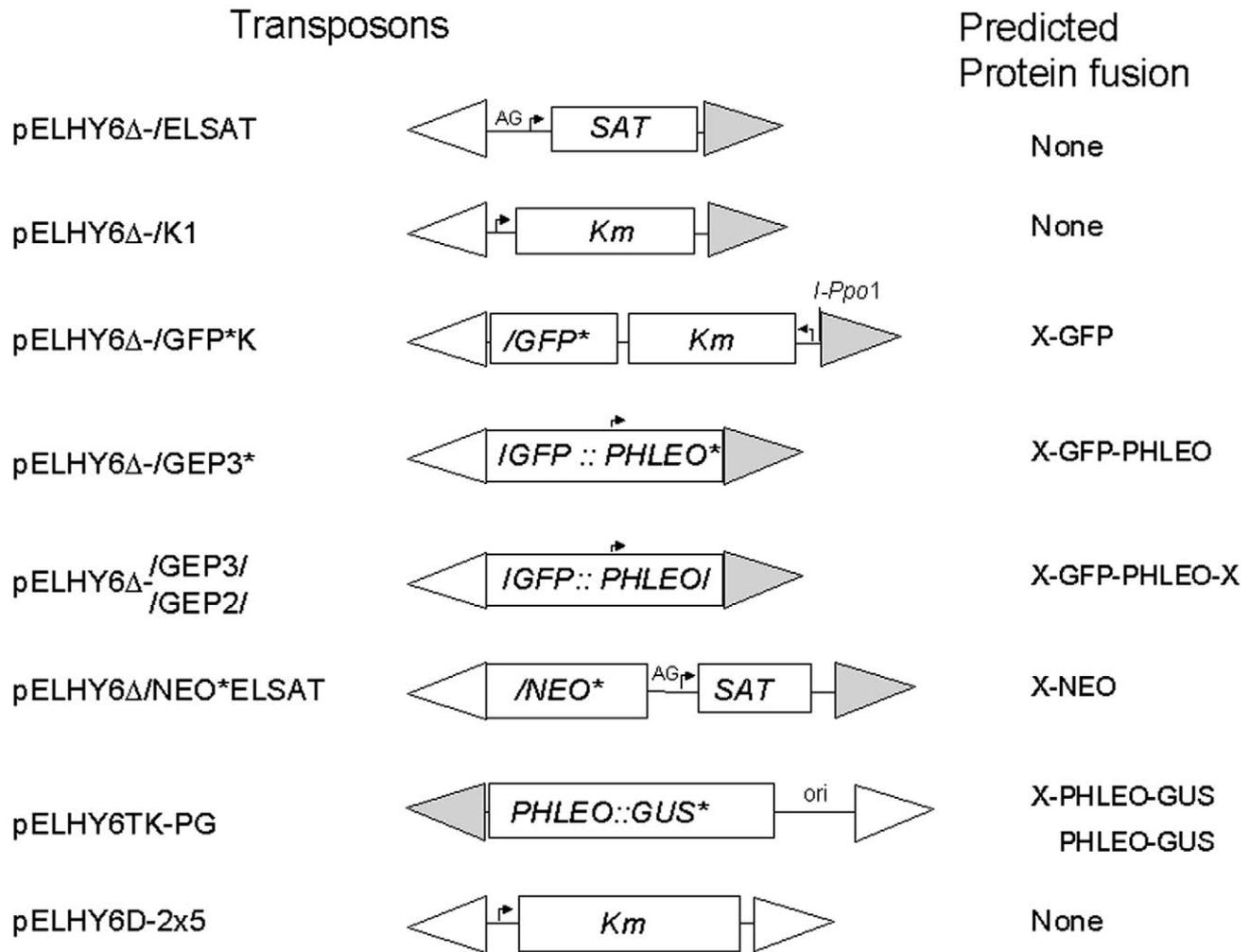


Fig. 2. Summary structures of MosI-based transposons and their potential applications. All transposons shown can be used for applications such as gene inactivation/disruption and primer island sequencing. Symbols are as follows: Open triangle, 5' IR; shaded triangle, 3' IR; AG, *Leishmania* trans-splice acceptor site; arrow, *E. coli* promoter; SAT, nourseothricin resistance marker; Km, kanamycin resistance marker; GFP, modified green fluorescent protein GFP<sup>+</sup>; PHLEO, phleomycin/zeocin resistance marker; NEO, Tn5-derived G418 resistance marker; ori, oriC. Naming conventions for fusion protein reporter or marker genes are '/', the indicated gene lacks an ATG start codon or stop codon, and '\*', the indicated gene contains terminal stop codon.

nicol gene (*SphI*–*SacI* fragment made blunt) from pUC18CM (B1935; Schweizer, 1990) was inserted into pELHYGmosI cut with *DraI*–*SspI*, yielding pHMCm (B4287). Then a *BamHI* (blunt)–*NdeI* fragment of pHMCm containing the *SbfI*–*XbaI* linker portion and the 3' IR and the 5' end of HYG marker was ligated with a *SspI*–*NdeI* fragment of pHM6K containing the 5' IR–Ori R6K–3'HYG marker.

#### 2.2.4. pELHY6TK-0 (B3653)

pELHY6TK-0 was made by the insertion of the HSV thymidine kinase gene (*AflII*–*SacII* fragment made blunt; from pXG-TK; B1317; (LeBowitz et al., 1992) into pELHY6-0 cut with *NheI* (blunt).

#### 2.2.5. pELHY6Δ-0 (B3653)

This construct was made by self-ligation of pELHY6-0, after digestion with *MamI* and *AocI* and made blunt. A restriction map is shown in Fig. 1.

#### 2.2.6. pELHY6Δ-K1 (B3682)

This construct contains the *Km* gene from pUC4K (Pharmacia; strain B4534), obtained by digestion with *HincII*, inserted into the *MspI* site of pELHY6Δ-0.

#### 2.2.7. pELHY6TK-PG (B3615)

This plasmid contains a *NcoI*–*BsaI* fragment (blunt) bearing a *phleo::gus* fusion and OriC of pUT90 (B2000; Cayla, France) inserted into the *XbaI* site (blunt) of pELHY6TK-0.

#### 2.2.8. pELHY6Δ-/GFP\*K (B3677)

A */GFP<sup>+</sup>*–*Km* fragment was obtained from TyKGFP (Garraway et al., 1997; B2798) by digestion with *NcoI*, partially filling-in with dATP + dCTP, digestion with *XmnI*, and made blunt with mung bean nuclease. This fragment was inserted into pELHY6Δ-0 previously cut with *MspI*. It should be noted that the *Km* gene is oriented in the opposite direction to the */GFP* cassette.

### 2.2.9. pELHY6Δ-/GEP3\* (B3692)

This construct was obtained in two steps: first, oligonucleotides SMB 860 (5'-gggatatctgtacaattGGCATAGTATATCGGCAaAGTATAAATACG; lower case signifies bases not present in the target DNA used for PCR amplification) and SMB 861 (5'-gggtctagaCTAGGATCCGTCCTGCTCCTCGGCCAC) were used to amplify the phleomycin resistance gene and promoter (*PHLEO*) from pUT90 by PCR using the Taq polymerase (one cycle of 5 min at 95°C, five cycles of 1 min at 95°C, 1 min at 55°C, 1 min at 70°C; 20 cycles of 1 min at 95°C, 1 min at 65°C, 1 min at 70°C. Primer B860 introduces a point mutation in the EM7 promoter, creating an open reading frame across this region. The PCR product was digested by *Bsr*GI and *Xba*I and ligated into pBS-GFP + (B2798) cut with *Bsr*GI and *Xba*I, yielding pBS-GEP (B4516). pBS-GFP + contains a gene for a modified GFP, with a S56T mutation and modified to contain GC-rich codons; this GFP is distributed by Clontech as EGFP. pBS-GEP was used as a template for PCR with the oligonucleotides SMB 862 (5'-gggatcGGT-GAGCAAGGGCGAGGAGCTG) and SMB 861, using the Taq polymerase (one cycle of 5 min at 95°C; five cycles of 1 min at 95°C, 1 min at 55°C, 1 min at 70°C; 15 cycles of 1 min at 95°C, 1 min at 65°C, 1 min at 70°C). The PCR product which corresponds to the /GFP + with no ATG fused to the promoter-*PHLEO* gene was digested by *Eco*RV and *Xba*I, and inserted into pELHY6Δ-0 digested by *Msp*II and *Xba*I.

### 2.2.10. pELHY6Δ-/GEP3/(B3851)

This construct was made by insertion of a double-stranded adaptor formed by annealing of the oligonucleotides SMB 1323 (5'-GATCGGCTCGAGCTGCA) and SMB 1324 (5'-GCTCGCC) into pELHY6Δ-/GEP3\* which had been digested by *Bam*HI and *Sbf*I.

### 2.2.11. pELHY6Δ-/GEP2/(B3860)

This construct was obtained by the insertion of a PCR product containing 'GFP + (with no ATG) fused to the promoter and *PHLEO* gene, synthesized using oligonucleotides SMB 945 (5' gggatcGGGTGAGCAAGGGCGAG) and SMB946 (5'-ccctgcagGTCCTGCCTCCTCGGCCAC), pBS-GEP DNA as a template, and the PCR protocol described for pELHY6Δ-/GEP3\* above. The PCR product was digested by *Eco*RV and *Pst*I, and inserted into of pELHY6Δ-0 digested by *Msp*II and *Sbf*I.

### 2.2.12. pELHY6Δ-/NEO\*ELSAT (B4537)

This transposon was constructed in two steps. First, a cassette containing the *SAT* marker preceded by a *Leishmania* splice acceptor and an *E. coli* promoter from pELSAT (B2764) was obtained by digestion with *Bam*HI and *Bgl*II, and inserted into the *Bam*HI site of pHM3 (Kaestner et al., 1994) (B482) to create pHM3SAT2 (B4538). This has a *lacZ::NEO* fusion gene in the same orientation as the *SAT* marker. A 1.7 kb fragment containing the /NEO and *SAT*

markers was excised with *Xho*I and *Xba*I, made blunt-ended with mung bean nuclease and cloned into the *Msp*II site of pELHY6Δ-0, yielding pELHY6Δ-/NEO\*ELSAT.

### 2.2.13. pELHY6Δ-ELSAT (B3683)

A cassette containing a *SAT* marker preceded by a *Leishmania* splice acceptor and an *E. coli* promoter from pELSAT (B2764) was obtained by digestion with *Sca*I and *Stu*I, and inserted into pELHY6Δ-0 cut with *Msp*II.

### 2.2.14. pELHY6Δ-2x5 (B3746)

This transposon is flanked symmetrically by identical 5' *mariner* IRs. It was obtained by ligation of a 1.3 kb fragment bearing a 5' IR-*Km* cassette from pELHY6Δ-K1 (B3682; made by digestion with *Eco*O109I, filling-in, and then digestion with *Xba*I), to a 2.1 kb fragment from plasmid pELHY6-0 (B3546) bearing the 5' IR/ori R6K/EM7/AG3/HYG regions (made by digestion with *Bsa*BI, filling-in, and digestion with *Xba*I).

The sequences of all PCR products and critical regions of the transposons were confirmed by automated DNA sequencing with appropriate primers.

## 2.3. Transposase purification and transposition assay

We engineered an N-terminal His<sub>6</sub>-tagged *Mos*I transposase by insertion of the appropriate *Nde*I–*Bam*HI fragment of pET3a-Tpase (B3297; (Tosi and Beverley, 2000) into pET19B (B4536) (Novagen) digested with the same enzymes, yielding pET19-Tpase (B4289). The recombinant His<sub>6</sub>-transposase protein was purified from *E. coli* BLR(DE3) (B3842) bearing pET19-TPASE. Expression was induced with IPTG, and inclusion bodies were recovered and solubilized in buffer A (20 mM Tris–HCl (pH 8), 500 mM NaCl, 6 M guanidine HCl, 1% NP-40) containing 70 mM imidazole as described (Tosi and Beverley, 2000). Transposase was recovered by binding to a NiSO<sub>4</sub> chelation column according to the manufacturer's instructions (Qiagen). After elution with 500 mM imidazole in buffer A, the protein was renatured as described (Tosi and Beverley, 2000). The purified His<sub>6</sub>-transposase was kept at –20°C in 50% glycerol. The properties of the tagged transposase are very similar to the native transposase, however its solubility following concentration is much better (data not shown).

Standard transposition reactions contained 0.3 μg (plasmid sizes 3–10 kb) or 0.5–1 μg target DNA (cosmid sizes 45 kb), 100 ng of transposon donor, and 100 ng of purified *Mos*I transposase, and were performed as previously described (Tosi and Beverley, 2000). Transpositions were recovered following transformation into *E. coli* lacking the *pir* gene product, thereby selecting against the R6K origin of replication present on the transposon donor plasmid.

### 3. Results

#### 3.1. Construction of a mini-*Mos1* element suitable to create translational fusions

We first created a ‘mini-*Mos1*’ element that would be a suitable platform for subsequent derivatives. Previously, it has been shown in an in vitro system that a modified *Mos1* element containing only the 28 bp terminal IRs was inactive (Fig. 3A), while the smallest active derivative developed retained 38 and 33 bp inside of the 5′ and 3′ IRs, respectively (Tosi and Beverley, 2000). Similar findings were reported for a mini *Himar1* element, containing 69 + 36 bp within the IRs (Lampe et al., 1999). In order to generate transpositional protein fusions, we needed to eliminate or avoid stop codons within the essential *cis*-acting sequences of the planned ‘mini-*Mos1*’ element. Additionally, the absence of an in-frame start codon was preferable, since potentially these could eliminate the dependence of fusion protein expression on acquisition of an upstream (target) start codon. Conceptual translation across the *Mos1* 5′ IR revealed a stop and start codon in frame 1, but uninterrupted reading frames in frames 2 and 3; these two reading frames also lacked ATG codons prior to the *MspI* site (Fig. 3A). The 3′ IR contained a stop codon in frame 4, and ATG codon in frames 5 and 6 (Fig. 3A). Thus only fusions in frames 2 and 3 across the 5′ IR are suitable for the exclusive recovery of protein fusions without the possibility of initiation at an IR-internal ATG.

With these factors in mind, we generated an ‘empty’ mini-*Mos1* derivative, carried on a small donor plasmid suitable for use in vivo or in vitro (Fig. 1). pELHY6Δ-0 contains an R6K origin of replication used in the in vitro transposition assay (Tosi and Beverley, 2000), a hygromycin B resistance marker active in both *E. coli* and *Leishmania* (Garraway et al., 1997), and a ‘mini-*Mos1*’ element containing the 5′ and 3′ IRs separated by a 61 bp region with unique restriction sites such as *MspI* suitable for the insertion of reporter genes or markers.

The activity of the *cis*-acting sequences within the pELHY6Δ mini-*Mos1* element was tested following insertion of two transposons bearing *Km* or *SAT* resistance markers (pELHY6Δ-K1 and pELHY6Δ-ELSAT, respectively), inserted into the transposon-internal *MspI* site. The in vitro transposition efficiencies were  $5\text{--}8 \times 10^{-4}$  (transpositions/target) for pELHY6Δ-K1, and  $1\text{--}2 \times 10^{-3}$  for pELHY6Δ-ELSAT. These efficiencies are comparable to the most active *Mos1* derivatives studied previously (Tosi and Beverley, 2000).

The pELHY6Δ-K1 or pELHY6Δ-ELSAT transposons can be used in many general applications such as gene disruption or primer island sequencing. In *E. coli*, the neomycin resistance marker *SAT* can be used as an alternative to kanamycin selection. Additionally, the *SAT* marker in pELHY6Δ-ELSAT is functional in *Leishmania*, as it has *Leishmania* specific processing signals for gene expression (a *trans*-splice acceptor site upstream of the *SAT* ORF; Garraway et al., 1997).

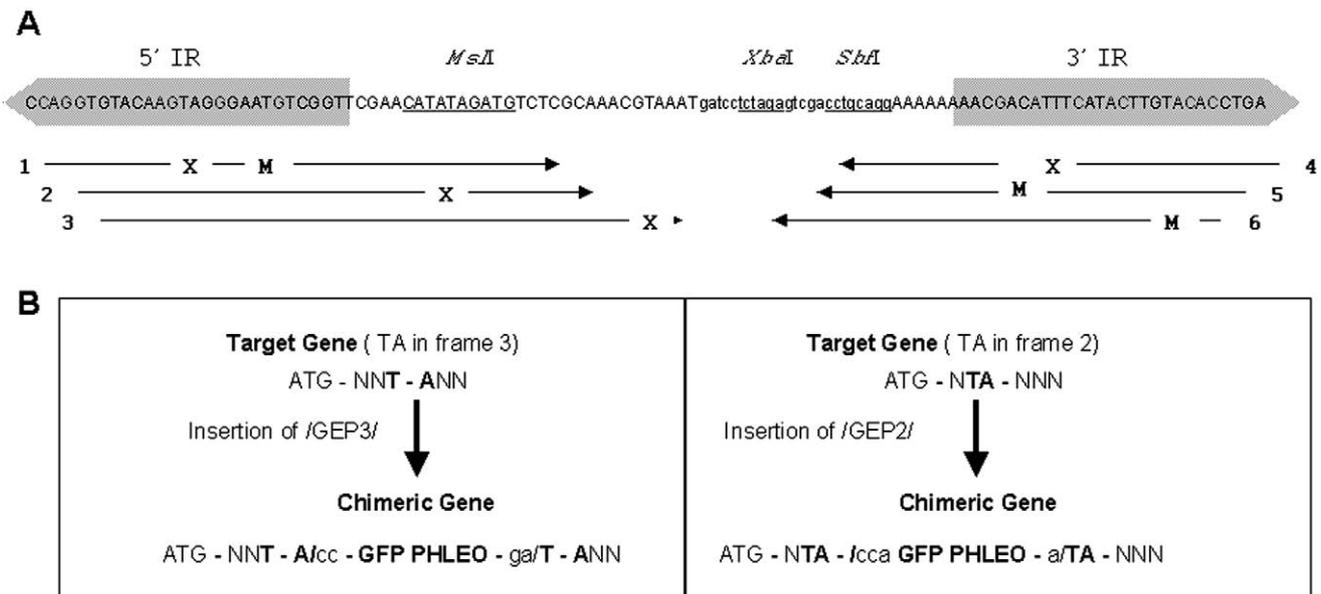


Fig. 3. Properties of the minimal *Mos1* element in pELHY6-0 or pELHY6Δ-0 and the design of translational gene fusions. (A) Nucleotide sequence of the IRs and intervening region in pELHY6-0. Shaded regions indicate the 5′ and 3′ IRs, and the unique *MspI*, *SbfI* and *XbaI* sites are shown. Nucleotides in lower case are not found in the original *Drosophila Mos1* element. Beneath the sequence, conceptual translations across all three reading frames from the left (5′ IR) or right (3′ IR) sides are shown; potential initiating ATG and stop codons are indicated by M and X, respectively. (B) The diagram shows the productive targets for /GEP3/ and /GEP2/ transposon insertions into TA dinucleotides located in target protein frames 3 and 2, respectively, and the results of transpositions therein.

### 3.2. Transposons suitable for recovery of translational gene fusions to GFP

The intrinsically fluorescent green fluorescent protein (GFP) has been widely used as a reporter gene for protein expression studies. Fluorescence can be observed with proteins targeted to all cellular compartments, and GFP-positive cells can be quantitated and recovered by flow cytometry. We developed several transposons with the GFP+ protein as a reporter; this modified GFP contains the S65T mutation and a GC-rich codon bias (Haas et al., 1996).

#### 3.2.1. pELHY6D-/GFP\*K

This construct contains the *E. coli* Km marker and a GFP+ derivative lacking an initiating ATG codon, engineered into the reading frame 2 across the 5' IR (/GFP+; Figs. 2 and 3A). The first in-frame ATG within the transposon 5' IR–GFP+ fusion protein is located past the GFP+ chromophore. Hence, this construct can yield active GFP+ only if it acquires an upstream in-frame start codon following transposition, yielding protein fusions with the GFP+ tag at the C terminus. Additionally, /GFP\*K contains a rare I-PpoI restriction site, whose asymmetric location facilitates mapping of transposon insertions (Fig. 2). Transposition efficiencies of  $5 \times 10^{-4}$  (transpositions/target) were obtained for pELHY6Δ-/GFP\*K. This transposon has been used successfully for primer-island sequencing of *Leishmania* cosmids (Pedrosa et al., 2001).

#### 3.2.2. pELHYG6Δ-/GEP3\*

In this transposon, the GFP+ cassette was fused to the phleomycin resistance marker (PHLEO), allowing selection for either phleomycin/zeocin resistance or GFP expression. The linker separating the PHLEO and GFP cassettes consists of a modified EM7 *E. coli* promoter, designed and inserted in a way that yields an uninterrupted reading frame across the /GFP+–EM7–PHLEO (Fig. 2). In the /GEP3\* transposon the fusion protein is expressed through the second reading frame of the 5' IR (Fig. 3A,B).

/GEP3\* is a compact 1.2 kb transposon that contains an *E. coli* selectable marker (PHLEO), and a GFP+–PHLEO marker whose expression depends of the provision of an upstream ATG following transposition. The PHLEO marker encodes a 124-amino-acid bleomycin–phleomycin binding protein, which mediates resistance in virtually all species (Drocourt et al., 1990). After the in vitro reaction, transposition pools are obtained by selection in *E. coli* for phleomycin resistance. These transposon pools can then be introduced into other organisms for selection for X-GFP-PHLEO protein fusion expression (through GFP by flow cytometry or by phleomycin selection), or for transcriptional activation of PHLEO (since there is a suitable initiating ATG codon upstream of PHLEO but not GFP+).

To check the functionality of the GFP fusion within pELHY6Δ-/GEP3\*, we generated GFP fusion proteins in

*E. coli*. The target plasmid pTZ18 (Pharmacia) expresses the lacZ α-peptide, which is commonly used for α-complementation of 5' truncated β-galactosidase genes in many *E. coli* strains. In vitro transposition was carried out with the pELHY6Δ-/GEP3\* donor and the pTZ18 target, the products were transformed into *E. coli* strain DH10B (lacZΔM15), and transposition events were recovered by plating in the presence of phleomycin and ampicillin with a frequency of  $10^{-3}$ . Of these colonies, 30% showed a lacZ phenotype, due to insertion of the transposon into the α-peptide gene (the remaining 70% represented insertions into nonessential regions of pTZ18, which is consistent with the relative target sizes). We analyzed 21 LacZ colonies and found that seven of these exhibited strong green fluorescence; this is the approximate frequency expected, since insertion into only 1 reading frame should give GFP expression (Fig. 3).

#### 3.2.3. pELHYG6Δ-/GEP3/ and /GEP2/

These transposons are similar to pELHYG6Δ-/GEP3\*, except that the stop codon of PHLEO was deleted and its coding region fused to the 3' MosI IR in a manner which yielded an ORF that spans the entire 1.2 kb transposon (Figs. 2 and 3B). This ORF was additionally designed so that following insertion of /GEP3/ or /GEP2/ (which yields a duplication of the target TA dinucleotide), the ORF would continue into the C-terminus of the target protein. Thus, these transposons yield 'sandwich' protein fusions of the type X-GFP-PHLEO-X. Transposons /GEP3/ and /GEP2/ differ in which frame the obligatory MosI TA target can occur and still yield active fusions (Fig. 3B); /GEP3/ yields fusions only when the target TA is in the third reading frame of the target protein, while /GEP2/ yields fusions only into the second reading frame (fusions to TAs in frame 1 cannot be recovered, due to the presence of a stop codon). The frequencies of transposition obtained for these two construct is similar to the one obtained for pELHY6Δ-/GEP3\*, indicating that the absence of a stop codon at the end of the phleomycin resistance gene does not alter the selection in *E. coli*. As for /GEP3\*, transposition events are recovered first by selection for phleomycin resistance in *E. coli*, and then gene fusions identified by transfection into the target organism followed by selection in vivo for GFP-PHLEO 'sandwich' protein fusions, or PHLEO protein or transcriptional fusions.

### 3.3. A transposon for recovery of translational gene fusions to NEO

pELHY6Δ-/NEO\*ELSAT (Fig. 2) contains the G418-resistance marker NEO, lacking its own ATG and fused across the second reading frame of the 5' IR. This transposon additionally carries a SAT marker active in both *Leishmania* and *E. coli* as described in Section 3.1. This 1.7 kb transposon has an in vitro efficiency of transposition efficiency of  $5 \times 10^{-4}$ . Transpositions can be recovered in *E.*

*coli* following by selection for SAT, and gene fusions to NEO are identified by selection for G418 resistance in the target organism.

#### 3.4. A transposon suitable for the recovery of transcriptional gene fusions in vivo

Transposon pELHY6TK-PG contains a PHLEO–GUS fusion protein bearing its own initiation codon, inserted into the second reading frame of the 3' IR, and a ColE1 origin of replication (Fig. 2). GUS refers to *E. coli*  $\beta$ -glucuronidase, for which there are several sensitive colorimetric and fluorimetric assays. Thus the PHLEO–GUS reporter can be used as a selectable marker and reporter enzyme. Since it lacks an *E. coli* selectable marker, this transposon can only be used for in vivo applications, selecting for PHLEO–GUS transcriptional or translational gene fusions (Fig. 2). Additionally, the ColE1 replication origin hinders the use of this transposon with in vitro transposition assays. However, following transposition in the target organism in vivo, the transposon-internal ColE1 origin can facilitate the recovery of the transposon plus flanking sequences by transformation into *E. coli*.

This transposon is carried on the delivery plasmid pELHY6TK, which bears the same elements as pELHY6 (*E. coli/Leishmania* selectable HYG marker, R6K origin of replication) and additionally the eukaryotic negative selectable marker HSV thymidine kinase gene (TK) under the control of *Leishmania* regulatory elements (LeBowitz et al., 1992). This allows the investigator to select against the episomal delivery plasmid following its use in *Leishmania* by selection with ganciclovir (LeBowitz et al., 1992).

#### 3.5. Transposition for TIMLI mutagenesis: pELHY6 $\Delta$ -2x5

A useful approach for the study of protein and gene functions has been the linker-insertion method, and an elegant transposon-mediated linker-insertional (TIMLI) mutagenesis method has been described which enables rapid and broad application of this method (Hayes and Hallet, 2000). First, a pool of transpositions into the desired target protein is generated; then, the transposon is excised en masse, leaving behind a collection of proteins bearing small residual transposon 'footprints'. In TIMLI mutagenesis, these are engineered to be short, 4–6 amino acid insertions.

We developed a *mariner* derivative suitable for TIMLI mutagenesis. The 5' IR of *Mos1* contains *BsrGI* and *SexAI* restriction sites that could be used; however, the 3' IR lacks the *SexAI* site. We created a 'symmetrical' *Mos1* derivative, in which the 3' IR had been replaced with the 5' IR, and the *Km* marker was inserted between the flanking 5' IRs (pELHY6 $\Delta$ -2x5; Fig. 2). The transposition efficiency for pELHY6 $\Delta$ -2x5 ( $1-2 \times 10^{-3}$ ) was slightly higher than that of pELHY6 $\Delta$ -K1, which bears the normal 5' and 3' IRs ( $5-8 \times 10^{-4}$ ). Following generation of pools into appropriate targets with pELHY6 $\Delta$ -2x5 by in vitro transposition and

selection for *Km*<sup>R</sup>, the transposons can be excised en masse by digestion with either *BsrGI* or *SexAI* followed by self-ligation, leaving respectively behind a 18 or 12 bp (6 or 4 amino acid) 'footprint' lacking stop codons in all three frames. Significantly, both *BsrGI* and *SexAI* sites are relatively infrequent in both target genomes and vectors. We have successfully tested in TIMLI mutagenesis using the *BsrGI* site (data not shown).

#### 3.6. A 'Negative' in vitro transposition assay not requiring transposon-internal *E. coli*-selectable marker or suicide origins or replication

The ability to assay transposition of engineered *Mos1* derivatives in vitro allows one to rapidly test their activity prior to their use in more laborious circumstances, such as in eukaryotes. However, occasionally one wishes to utilize transposons in vivo whose properties are not appropriate for the in vitro system. For example, they may lack an appropriate *E. coli* selectable marker (such as transposons pM[*cn*] and pMos1-Act-EGFP), or contain replication origins other than the *pir*-dependent R6K origin (an example being pELHY6TK-PG; Section 3.4). While these elements could be engineered into the transposon, sometimes this is incompatible with function in vitro or in vivo, often simply due to the increase in size. To overcome this limitation, we developed an approach in which transposition was scored by inactivation of a negative selectable marker, removing the requirement for particular origins of replication on the delivery vector or transposon-internal *E. coli* markers. As a negative marker, we chose the lethal gene *ccdB* (control of cell death), expressed under the control of the lactose promoter in the plasmid pZero-2 (Invitrogen). In this assay, donor transposons were incubated with pZero-2-*ccdB*, which bears the *Km* marker. When expressed following induction with IPTG, *ccdB* is lethal and *Km*<sup>R</sup> colonies are not obtained, thus allowing the recovery of transposition events into *ccdB*.

We tested this approach in a standard in vitro transposition assay containing equimolar amounts of two plasmid targets, one bearing *ccdB* (pZero-2, *Km*<sup>R</sup>, *ccdB*<sup>+</sup>) and a standard plasmid target (pBSCm, *Cm*<sup>R</sup>). The transposon donor was pELHY6 $\Delta$ -ELSAT, which contains the SAT marker under control of an *E. coli* promoter. Following transposition in vitro, the products were transformed into the *E. coli* strain Top10 under conditions where *ccdB* expression is lethal. Transposition of ELSAT into pBSCm was scored as SAT<sup>R</sup> + *Cm*<sup>R</sup> colonies and its efficiency as SAT<sup>R</sup> + *Cm*<sup>R</sup>/*Cm*<sup>R</sup>. Transposition into the *ccdB* gene was scored as *Km*<sup>R</sup> colonies and the efficiency of transposition into *ccdB* was estimated as *Km*<sup>R</sup>/*Cm*<sup>R</sup> (since the wild-type pZero-2 plasmid is lethal). In these experiments, the transposition efficiency into pBSCm was  $5-10 \times 10^{-4}$  (Table 1). There was a background of *Km*<sup>R</sup> colonies arising from spontaneous mutations which reduce *ccdB* expression, of about  $1 \times 10^{-5}$ , which corresponds to a background relative to the

Table 1  
Use of a ‘negative’ selection transpositional assay into the lethal *E. coli* *ccdB* gene to compare in vitro transposition efficiencies<sup>a</sup>

Donor transposon	Relative activity (%)
pELHY6Δ-SAT	100
No transposon	1–2.9
pEL-Apr	37–150
pMos1-Act-EGFP	22–27
pM[ <i>cn</i> ]	1–2.9

<sup>a</sup> In vitro transposition reactions with the indicated transposons were performed with an equimolar mixture of the target plasmids pZero-2 (bearing the *ccdB*<sup>+</sup> and Km<sup>R</sup> markers) and pBSCm (bearing the Cm<sup>R</sup> marker), as described in Section 2. Typically, 3–5 × 10<sup>6</sup> Cm<sup>R</sup> colonies were obtained (this is the plasmid transformation efficiency) and ~90 Km<sup>R</sup> colonies were obtained (this is the background arising from spontaneous loss of *ccdB*). The number of Cm<sup>R</sup> colonies was used to normalize for transformation efficiency, and the efficiency of transposition obtained with the control transposon pELHY6Δ-SAT (SAT<sup>R</sup> + Cm<sup>R</sup>/Cm<sup>R</sup>) was 5–10 × 10<sup>-4</sup>. The relative transposition activity of test transposons was calculated as (Km<sup>R</sup><sub>test</sub>/Km<sup>R</sup><sub>control</sub>) × (Cm<sup>R</sup><sub>test</sub>/Cm<sup>R</sup><sub>control</sub>) × 100.

fully active transposon pELHY6Δ-SAT activity of about 1–3% (Table 1). These data confirmed the utility of the negative selection assay for comparing the in vitro activity of a given transposon against a highly active transposon control.

We then used the ‘negative transposition’ assay to test several transposons that have been used successfully in vivo, but whose transposition activity in vitro had not previously been measured. Transposon pEL-Apr has been used successfully in vivo in *Streptomyces coelicolor* (Gehring et al., 2000), and its efficiency was comparable to the control transposon pELHY6Δ-ELSAT (Table 1). We also tested two transposons that have been developed previously for use in insects. Transposon pMos1-Act-EGFP showed good activity, ranging from 22 to 27% that of pELHY6Δ-ELSAT (Table 1). In contrast, the pM[*cn*] element, which has been used successfully in transposition assays in mosquitoes (Coates et al., 1998), showed no activity above background (Table 1). This 6 kb element contains a 4.7 kb *Drosophila melanogaster cn* gene, inserted within an intact *Mos1* element. Possibly, the size of the pM[*cn*] element reduces its transposition efficiency in vitro; studies of the related element *Himar* shows a significant decrease in transposition efficiency with increasing size (Lampe et al., 1998).

#### 4. Discussion

In this work we have described a collection of *Mos1*-based transposons, delivery plasmids and ‘negative’ transposition assays that together facilitate the design, testing and application of *mariner* transposon mutagenesis in many organisms. Most of the new transposons have been designed to be efficiently mobilized by the in vitro transposition system, and following the generation of suitable insertion libraries, to allow the recovery of a variety of translational

and transcriptional gene fusions following transfection into target cells or species. Our collection enables fusions to be made to useful reporters such as the GFP, β-glucuronidase or β-galactosidase, or eukaryotic selectable markers such as *NEO* or *PHLEO*. These allow a variety of selective or sensitive screening strategies to be incorporated into one’s experimental strategy. All of the mini-*Mos1* transposons are additionally suitable for approaches such as primer-island sequencing or insertional inactivation. As necessary, new elements can be designed and inserted into the mini-*Mos1* element carried on pELHY6Δ-0 (Fig. 1).

Of particular note are transposons /GEP2/ and /GEP3/, which yield tri-partite or ‘sandwich’ protein fusions of the form X-GFP-PHLEO-X (Fig. 2). Such fusions retain 5’ and 3’ flanking sequence as N- and C-terminal protein elements, increasing the likelihood of correct expression and cellular localization of the resultant fusion proteins. We have used these transposons successfully in a shuttle mutagenesis approach to recover the expected protein fusions from a number of *Leishmania* genes (unpublished data; Beverley et al., 2002). Additionally, we developed a *mariner*-based system for TIMLI mutagenesis (Hayes and Hallet, 2000), which enables the rapid generation of a series of small insertions across the protein of interest for subsequent functional studies (pELHY6Δ-2x5, Fig. 2).

A number of the transposons bear *Leishmania* selectable markers, such as the *HYG* element in the pELHY6 derivatives (which is also active in *E. coli*) or the *SAT* element in pELHY6Δ-ELSAT (Fig. 2). Amongst the applications that could be envisioned for these transposons, one that is particularly useful is their ability to rapidly generate constructs suitable for gene inactivation by homologous gene replacement. In this approach, in vitro mutagenesis is performed with a transposon bearing a *Leishmania* marker (such as pELHY6Δ-ELSAT), and insertions inactivating a target gene of interest identified rapidly by PCR. This provides a convenient gene targeting fragment suitable for homologous gene replacement following transfection into *Leishmania*. This approach could be used in many organisms by using mini-*Mos1* elements with appropriate species-specific markers.

As part of this work, we also developed a novel ‘negative’ transposition assay, in which transposition is scored by inactivation of a lethal target gene (*ccdB*<sup>+</sup>; Table 1). This assay has the advantage of not requiring specific origins of replication or transposon-internal selectable *E. coli* markers. While these elements are essential for the in vitro transposition assay, they can compromise transposon function in eukaryotes, arising from factors such as increased size (Lampe et al., 1998). The negative selection approach eliminates this problem while allowing the investigator to test the functionality of prospective transposons (Table 1). Remarkably, we found that the *Mos1*-based pM[*cn*] transposon used previously with mosquitoes to recover a limited number of transposition events in vivo (Coates et al., 1998) did not show good activity in vitro (Table 1). In contrast, the

pMos1-Act-EGFP transposon showed strong activity, perhaps due to its smaller size. Possibly, the efficiency measured in vitro may not be strictly correlated with that observed in vivo (Fischer et al., 2001). Nonetheless, these findings suggest that in vivo transposition in mosquitoes with the *mariner* system could be improved substantially by using more active transposons, as judged by the in vitro transposition assay.

In summary, this collection of engineered mariner elements provides investigators studying many species a useful ‘toolkit’ for systematically applying transposon mutagenesis in vitro and in vivo.

## Acknowledgements

We thank Jenni Lovett for the construction of pELHY6TK-0 and D. Dobson, J. Revollo, K. Zhang and H. Zieler for comments. This work was supported by grants from the NIH (AI29646), FAPESP (98/09805-0) and WHO/TDR (990881). We thank H. Zieler, A. James and A. Gehring for providing the transposons tested in Table 1.

## References

- Berg, C.M., Berg, D.E., Groisman, E.A., 1989. Transposable elements and the genetic engineering of bacteria. In: Berg, D.E., Howe, M.M. (Eds.). *Mobile DNA*, American Society for Microbiology, Washington, DC.
- Beverley, S.M., Akopyants, N.S., Goyard, S., Matlib, R.S., Gordon, J.L., Brownstein, B.H., Stormo, G.D., Bukanova, E.N., Hott, C.T., Li, T., MacMillan, S., Muo, J.N., Schwertman, L.A., Smeds, M.R., Wang, Y., 2002. Putting the *Leishmania* genome to work: functional genomics by transposon trapping and expression profiling. *Philos. Trans. R. Soc. Biol. Sci.* 357 (1417).
- Biery, M.C., Lopata, M., Craig, N.L., 2000. A minimal system for Tn7 transposition: the transposon-encoded proteins TnsA and TnsB can execute DNA breakage and joining reactions that generate circularized Tn7 species. *J. Mol. Biol.* 297, 25–37.
- Coates, C.J., Jasinskiene, N., Miyashiro, L., James, A.A., 1998. *Mariner* transposition and transformation of the yellow fever mosquito, *Aedes aegypti*. *Proc. Natl. Acad. Sci. USA* 95, 3748–3751.
- Drocourt, D., Calmels, T., Reynes, J.P., Baron, M., Tiraby, G., 1990. Cassettes of the *Streptoalloteichus hindustanus ble* gene for transformation of lower and higher eukaryotes to phleomycin resistance. *Nucleic Acids Res.* 18, 4009.
- Fadool, J.M., Hartl, D.L., Dowling, J.E., 1998. Transposition of the *mariner* element from *Drosophila mauritiana* in zebrafish. *Proc. Natl. Acad. Sci. USA* 95, 5182–5186.
- Fischer, S.E., Wienholds, E., Plasterk, R.H., 2001. Regulated transposition of a fish transposon in the mouse germ line. *Proc. Natl. Acad. Sci. USA* 98, 6759–6764.
- Garraway, L.A., Tosi, L.R., Wang, Y., Moore, J.B., Dobson, D.E., Beverley, S.M., 1997. Insertional mutagenesis by a modified in vitro Tyl transposition system. *Gene* 198, 27–35.
- Gehring, A.M., Nodwell, J.R., Beverley, S.M., Losick, R., 2000. Genome-wide insertional mutagenesis in *Streptomyces coelicolor* reveals additional genes involved in morphological differentiation. *Proc. Natl. Acad. Sci. USA* 97, 9642–9647.
- Goryshin, I.Y., Reznikoff, W.S., 1998. Tn5 in vitro transposition. *J. Biol. Chem.* 273, 7367–7374.
- Gueiros-Filho, F.J., Beverley, S.M., 1997. *Trans*-kingdom transposition of the *Drosophila* element *mariner* within the protozoan *Leishmania*. *Science* 276, 1716–1719.
- Haas, J., Park, E.C., Seed, B., 1996. Codon usage limitation in the expression of HIV-1 envelope glycoprotein. *Curr. Biol.* 6, 315–324.
- Hartl, D.L., Lohe, A.R., Lozovskaya, E.R., 1997. Modern thoughts on an ancient *marinere*: function, evolution, regulation. *Annu. Rev. Genet.* 31, 337–358.
- Hayes, F., Hallet, B., 2000. Pentapeptide scanning mutagenesis: encouraging old proteins to execute unusual tricks. *Trends Microbiol.* 8, 571–577.
- Kaestner, K.H., Montoliu, L., Kern, H., Thulke, M., Schutz, G., 1994. Universal  $\beta$ -galactosidase cloning vectors for promoter analysis and gene targeting. *Gene* 148, 67–70.
- Lampe, D.J., Churchill, M.E., Robertson, H.M., 1996. A purified *mariner* transposase is sufficient to mediate transposition in vitro. *EMBO J.* 15, 5470–5479.
- Lampe, D.J., Grant, T.E., Robertson, H.M., 1998. Factors affecting transposition of the *Himar1 mariner* transposon in vitro. *Genetics* 149, 179–187.
- Lampe, D.J., Akerley, B.J., Rubin, E.J., Mekalanos, J.J., Robertson, H.M., 1999. Hyperactive transposase mutants of the *Himar1 mariner* transposon. *Proc. Natl. Acad. Sci. USA* 96, 11428–11433.
- LeBowitz, J.H., Cruz, A., Beverley, S.M., 1992. Thymidine kinase as a negative selectable marker in *Leishmania major*. *Mol. Biochem. Parasitol.* 51, 321–325.
- Maniatis, T., Fritsch, E.F., Sambrook, J., 1982. *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Merkulov, G.V., Boeke, J.D., 1998. Libraries of green fluorescent protein fusions generated by transposition in vitro. *Gene* 222, 213–222.
- Miller, V.L., Mekalanos, J.J., 1988. A novel suicide vector and its use in construction of insertion mutations: osmoregulation of outer membrane proteins and virulence determinants in *Vibrio cholerae* requires *toxR*. *J. Bacteriol.* 170, 2575–2583.
- Pedrosa, A.L., Ruiz, J.C., Tosi, L.R., Cruz, A.K., 2001. Characterization of three chromosomal ends of *Leishmania major* reveals transcriptional activity across arrays of reiterated and unique sequences. *Mol. Biochem. Parasitol.* 114, 71–80.
- Plasterk, R.H., 1996. The Tc1/*mariner* transposon family. *Curr. Top. Microbiol. Immunol.* 204, 125–143.
- Robertson, H.M., 1995. The Tc1-*mariner* Superfamily of Transposons in Animals. *J. Insect Physiol.* 41, 99–105.
- Rubin, E.J., Akerley, B.J., Novik, V.N., Lampe, D.J., Husson, R.N., Mekalanos, J.J., 1999. In vivo transposition of *mariner*-based elements in enteric bacteria and mycobacteria. *Proc. Natl. Acad. Sci. USA* 96, 1645–1650.
- Schweizer, H.P., 1990. The pUC18CM plasmids: a chloramphenicol resistance gene cassette for site-directed insertion and deletion mutagenesis in *Escherichia coli*. *Biotechniques* 8, 612–616.
- Sherman, A., Dawson, A., Mather, C., Gilhooley, H., Li, Y., Mitchell, R., Finnegan, D., Sang, H., 1998. Transposition of the *Drosophila* element *mariner* into the chicken germ line. *Nat. Biotechnol.* 16, 1050–1053.
- Tosi, L.R., Beverley, S.M., 2000. *cis* and *trans* factors affecting *Mos1 mariner* evolution and transposition in vitro, and its potential for functional genomics. *Nucleic Acids Res.* 28, 784–790.
- Zhang, J.K., Pritchett, M.A., Lampe, D.J., Robertson, H.M., Metcalf, W.W., 2000. In vivo transposon mutagenesis of the methanogenic archaeon *Methanosarcina acetivorans* C2A using a modified version of the insect *mariner*-family transposable element *Himar1*. *Proc. Natl. Acad. Sci. USA* 97, 9665–9670.