5/30/03

for *Methods in Molecular Biology*
Sarah Melville, Editor; Humana Press

in vitro shuttle mutagenesis using engineered *Mariner* transposons

Kelly A. Robinson[a], Sophie Goyard[b], and Stephen M. Beverley[a]*

[a] Washington University Medical School

Department of Molecular Microbiology

660 S. Euclid Ave.

St. Louis, MO 63110 USA

[b] present address : Immunophysiologie et Parasitisme Intracellulaire, Institut Pasteur, 25 rue du Dr. Roux 75724 Paris cedex 15 France.

* Corresponding author; beverley@borcim.wustl.edu

Telephone: 314-747-2630, FAX: 314-747-2634

Key words: protozoan parasite, shuttle mutagenesis, *in vitro* transposition, Tc1/*mariner* transposon family, *Mos1* transposon

Running Title: *in vitro mariner* transposition.

Abstract

Advances in our understanding of the protozoan parasite *Leishmania* have been facilitated by the development of molecular and genetic tools. One powerful approach for gene identification and analysis is transposon mutagenesis. This can be performed directly *in vivo*, but often it is more convenient to generate transpositions *in vitro* for subsequent analysis *in vivo*, in a process termed 'shuttle mutagenesis'. The *Drosophila* element *mariner* is well-suited for application by either route. Minimal *mariner* elements containing *cis*-acting elements required for transposition have been generated, which can be further modified to suit the needs of the experimenter. Additional genetic markers and/or reporters can be introduced which are useful for procedures such as insertional mutagenesis, shotgun sequencing, or the generation of protein and transcriptional fusions for subsequent analysis. Active transposase can readily be generated following expression in *E. coli*, and efficiencies of 10^{-3}/target can be obtained, allowing the generation of large transposon insertion libraries suitable for subsequent screening *in vivo*. In this article, we will explain the steps necessary to purify active *Mos1* transposase and conduct an *in vitro* transposition reaction. We will also discuss some of the considerations relevant to the design and application of functional *mariner*-elements (donor plasmids) relevant to studies in *Leishmania* and other organisms.

1. Introduction

Protozoan parasites like *Leishmania* are responsible for a number of illnesses that cause significant mortality and morbidity throughout the world (1). Genetic and genomic tools now available for the study of the parasite promise to greatly increase our understanding of how this

parasite survives and causes disease, and ultimately lead to improved methods for overcoming this disease by immunization or chemotherapy (2-5).

Two common tasks in parasite genetics are first the identification of genes mediating interesting functions, and secondly dissection of the role, regulation and localization of encoded proteins. A powerful tool suitable for this task in many organisms is transposon-based mutagenesis (6). This can be performed directly *in vivo*, where both the transposon and active transposase are introduced or expressed in the parasite, or *in vitro*, where transposition is performed *in vitro* and the products introduced into the parasite for subsequent analysis (Fig. 1). *In vivo* strategies are especially powerful when incorporated into forward genetic approaches, as mutants generated are simultaneously tagged by the transposon, which can then be used to recover the affected gene. Unfortunately, *Leishmania* is an asexual diploid in the laboratory, and for most loci recovery of loss of function mutations requires at least two genetic events (7). However this approach is widely used in haploid organisms, or ones in which homozygosity can be readily attained in some manner (for example by sexual crossing).

In vivo transposition systems can be challenging to set up, due to the constraints inherent in engineering transposase expression and controlling transposition. For many purposes, *in vitro* transposition is more convenient and as powerful. In a process termed shuttle mutagenesis, transfectable molecular constructs (for example, Leishmania DNA cloned in the shuttle E. coli-Leishmania vector cLHYG; 8) are subjected to transposition *in vitro*, and then the population of independent insertions scored for phenotypes following transfection back into *Leishmania* (Fig. 1). The transposon insertion library also can be used for rapid and systematic DNA sequencing if necessary, using primers situated within the transposon. We have found this approach

especially useful in mapping the active gene within cosmids recovered in a variety of functional genetic screens in *Leishmania*.

Beyond their role as insertional mutagens, transposons can be engineered to contain reporters such as the green fluorescent protein (GFP), β-galactosidase, β-glucosidase or β-lactamase, or selectable markers such as *NEO*, *HYG* or *PHLEO*, which mediate resistance to G418/geneticin, hygromycin B or phleomycin/zeocin, respectively. Following transposition, activation of the reporter or marker can then be used to identify and/or select for transcribed or translated regions of the genome, a procedure commonly referred to as gene/protein 'trapping'. By studying expression and/or localization of the reporter proteins, one can then conveniently (albeit indirectly) monitor gene expression and protein localization (9, 10).

Several transposon systems have been engineered to the point that they are readily incorporated into shuttle transposon mutagenesis strategies; these include Tn7 (11), Tn5 (12), Ty1 (13), Mu (14) and several Tc1/*mariner* family elements (15-17). Relevant factors include the availability, cost and/or ease of purification of transposase, the randomness with which a given transposon inserts into target DNA, the requirements for specific *cis*-acting elements required within the transposon itself (for example the size or properties of the flanking inverted repeats), the ability of the transposon to carry 'cargo' of sufficient size and with the desired properties, and ease of use. We have found the *Drosophila mariner* element *Mos1* to be satisfactory in these respects, and here we describe the basic elements of the transposon relevant to its application *in vitro*.

1.1. *mariner* and *in vitro* transposition reactions. The *Drosophila* element *Mos1* is a member of the *mariner*/Tc1 family, which occurs in most kingdoms of living organisms (18, 19). Typically *mariner*/Tc1 elements are small, encoding only the transposase and *cis*-acting elements

required for transposition such as the flanking inverted repeats (20, 21). Transposition occurs through a cut and paste mechanism (21), in which recognition of the inverted repeats (IR) by the transposase results in excision of the donor element. This is subsequently inserted into a TA dinucleotide of the target molecule, and accompanied by duplication of the TA flanking the insertion site (22). *In vitro* experiments have shown that the transposition reaction requires only transposase and transposon *cis*-elements, without the need for cellular factors (15, 17, 21, 23, 24). For *Drosophila Mos1*, the *cis*-acting elements required for transposition include the 28 bp 5' and 3' inverted repeats, along with some internal nucleotides (no greater than 38 and 5 additional internal nucleotides on the 5' and 3' sides, respectively; 17, 23). While here we focus on *in vitro* applications, it is notable that the *mariner* system has been shown to function *in vivo* in a variety of different organisms including *Leishmania*, insects and vertebrates (25-29). Most of the *mariner* transposon derivatives described below can be used *in vivo* as well, in any species.

Our understanding of the mechanism of *Mos1* transposition has lead to the development of a minimal, "empty" transposon donor, pELHY6Δ-0 (Fig. 2), into which a variety of transposon 'cargos' have been inserted previously (23). New cargos may be designed and rapidly introduced into this vector as desired by the experimenter. A variety of transposons have been created previously, and potential applications such as insertional mutagenesis and gene trapping are summarized in Table 1. A typical *in vitro Mos1* transposition reaction consists of the donor plasmid (ex. /GEP3/, Fig. 2, Table 1), target DNA (a cLHYG-based cosmid for example; Fig. 2) and purified transposase. The properties and requirements for these elements are discussed below.

The donor plasmid pELHY6Δ-0 contains outside of the minimal transposon a bacterial/*Leishmania* selectable marker (*HYG* here) and origin of replication (OriR6K) (Fig. 2). For propagation in *E. coli*, the OriR6K origin requires the *pir* gene product, which is provided by the use of appropriate *E. coli* host strains when growing this plasmid (often harboring a λ*pir* lysogen). This allows one to select against the donor plasmid following transformation of the *in vitro* transposition reaction mix into *pir- E. coli* (this comprises virtually all common *E. coli* recipients). In this particular donor plasmid, the presence of a *Leishmania* marker permits introduction into parasites by *HYG* selection and incorporation into *in vivo* transposition approaches. Note that the signals for replication, transcription and/or mRNA processing differ considerably between *E. coli* and *Leishmania*; briefly, *E. coli* markers require promoters and plasmids require origins of replication for episomal maintenance. In contrast, in *Leishmania* all that is required is a *trans*-splice acceptor site upstream of the marker ORF, as transcriptional and replication origin requirements are quite relaxed (30). These differences must be taken into account when designing new transposons, and it can also be used to the researcher's advantage in various ways.

As a target plasmid, most common laboratory plasmid, cosmid and BAC vectors can be used; the specific requirements are that the target DNA should not contain an OriR6K for replication. The target marker should not be the same as ones borne within the transposon or in its donor background (note that in the cLHYG example shown, the *HYG* gene lacks an *E. coli* promoter and thus does not confer resistance in bacteria). Transposition efficiencies are highest if the target DNA is supercoiled and its "quality" is high (17).

Mos1 transposase is purified and stored as described below. Typically transposons require Mg^{+2} for activity, however Mn^{+2} can be used as this relaxes the requirement for insertion

into TA dinucleotides (17). While some *mariner* family transposases show a phenomenon called overproduction inhibition (15), *Mos1* shows simple saturation kinetics, and increasing transposase yields increasing transpositions until a plateau is reached (17).

Following incubation, the *in vitro* transposition mix (which contains both donor and target plasmids as well as the desired transpositions) is transformed into *pir- E. coli* and plated on drugs which select for both the transposon (phleomycin in the example of /GEP3/ here) and the target (ampicillin here). Transposition efficiencies can be calculated by comparing platings on ampicillin alone vs. ampicillin + phleomycin, and can approach 10^{-3}/target.

The number of individual transpositions required depends upon the particular application. While *mariner* demands TA residues for insertion under standard conditions, these are sufficiently abundant even in the GC-rich *Leishmania* genome to provide plenty of potential target sites, and the requirement for TA can be relaxed if transposition is performed in the presence of Mn^{+2} (17). We have found that for cosmid targets, several hundred independent insertions are usually sufficient for sequencing and inactivation of most potential target genes. For the recovery of specific gene fusions, larger libraries may be required, since one has the additional constraints of inserting into TAs in the appropriate strand and reading frame. For these purposes one thousand independent insertions into a cosmid target should suffice. Note that with current *in vitro* transposition efficiencies and *Leishmania* transfection efficiencies (31) one may contemplate scoring libraries in excess of 10^5 independent insertions.

1.2 The *mariner* toolkit. Table 1 describes some of the *mariner* derivatives that have been developed and used successfully in our laboratory. A variety of applications and transposons can be envisaged, and the ones below provide some perspective on the factors relevant to their design and utilization. The transposons are described briefly below; many can be used for the recovery

and/or selection of gene fusions in *Leishmania* and other organisms, as they lack species-specific regulatory elements (the GEP transposon series, for example). All transposons can be used for primer-island sequencing, and insertional inactivation. Most transposons contain autonomous bacterial selectable markers and can be used in the *in vitro* system, except for pELHY6TK-PG (thereby restricting it to *in vivo* applications). Since *Leishmania* uses a polycistronic transcriptional mechanism to generate mRNAs and relies heavily on post-transcriptional regulatory mechanisms to control protein expression, we have given the most attention to transposons that facilitate the recovery of protein fusions.

Transposons /GEP3*, /GEP3/ and /GEP2/ contain a GFP-PHLEO fusion protein (Table 1); the linker peptide between the GFP and PHLEO additionally functions as an *E. coli* promoter, and in bacteria this cassette confers phleomycin resistance constitutively (as required for use in the *in vitro* system). Importantly, the GFP lacks an initiating ATG codon, and thus in eukaryotes GFP-PHLEO expression can only be obtained following insertion of the /GEP transposons into an open reading frame expressed by the target DNA (Fig. 3). Such insertions can be selected for by phleomycin resistance (only in eukaryotes), or screened for by GFP expression. Note that phleomycin resistance can be affected by compartmentalization of the fusion protein; if the PHLEO protein domain is restricted to a compartment such as the glycosome which is segregated from the nucleus (the site of action of phleomycin), resistance will be abrogated (32).

/GEP3* differs from /GEP3/ and /GEP2/ in that it contains a stop codon following the GFP-PHLEO fusion protein. Thus, the protein fusions recovered bear the N but not C terminus of the trapped protein (Fig. 4A,C). In /GEP3/ and /GEP2/ the stop codon has been eliminated so that an intact reading frame is maintained across the entire transposon, enabling the recovering of

protein fusions which bear both the N and C terminus of the trapped protein; this type of transposon is referred to as a 'sandwich' transposon (Fig. 4B,C). Since *mariner* elements must insert into TAs which can occur in any reading frame, /GEP3/ differs from /GEP2/ in which frame can be trapped (due to the sequence of the 5' mariner inverted repeats it is not possible to make a "/GEP1/" transposon for protein trapping; Fig. 3). While all GEP transposons can be used to study translational regulation, the ability of sandwich transposon to retain both N and C terminal sequences which often contain important protein targeting information (for example, membrane anchoring domains; Fig. 4C) is important for some purposes.

Transposon /NEO*ELSAT (Table 1) creates a translational fusion to the *NEO* selectable marker. It should be noted that the ability of the NEO protein to tolerate N-terminal fusions varies considerably amongst different protein targets, in contrast to GFP, PHLEO and β-galactosidase which are more permissive. An additional element in this transposon is the selectable marker *SAT* (streptothricin resistance), which contains both a *Leishmania* splice acceptor site and an *E. coli* promoter, allowing for selection for the transposon with *SAT* and protein fusions with *NEO* in *Leishmania*.

Transposon GFP*K (Table 1) can used to generate GFP fusions in a manner similar to the /GEP transposons, as its GFP also lacks an ATG start codon. Additionally, it bears a rare restriction endonuclease (I-*Ppo*I) which is helpful in mapping transposon insertion sites in large targets or the genome *in vivo*.

The transposon carried in **pELHY6TK-PG** (Table 1) contains a PHLEO-GUS translational fusion containing its own start codon. This transposon lacks a constitutive bacterial selectable marker, and contains an *E. coli* oriC replication origin; thus it cannot be used in the *in vitro* system, and can only be used *in vivo*. There, the oriC replication origin facilitates the

recovery of candidate transpositions back from *Leishmania* into *E. coli* if desired. This transposon donor is carried on pELHY6TK, which is a modified version of pELHY6Δ-0; it additionally bears a conditionally negative selectable marker (herpesvirus Thymidine Kinase) active in *Leishmania* (33).

Transposon /-2x5 (Table 1) was designed for transposon-mediated linker-insertional mutagenesis (TIMLI; 34). It contains an *E. coli* kanamycin resistance marker, flanked by a "symmetric" mariner element in which the 5' IR was duplicated. Importantly, this IR contains two sites that occur relatively infrequently in *Leishmania*, *Sex*AI and *Bsr*GI (and most importantly should not occur in targets where they are to be used). In TIMLI mutagenesis, one first generates a large library of transposition events into the target. Then, this transposition pool is collected en mass, DNA prepared, digested with *Sex*AI or *Bsr*GI, diluted and self-ligated, and transformed back into *E. coli*. This yields excision of the transposon, leaving behind only an insertion of 12 or 18 nucleotides (encoding 4 or 6 amino acids respectively). Thus one can generate a library of short-peptide insertions for subsequent functional analysis, such as the mapping of protein domains and activities.

In this chapter we will describe how to express and purify active *Mos1* transposase, and carry out *in vitro* transposition reactions.

2. Materials

2.1 Expression of *Mos1* transposase

1. *E. coli* strain expressing T7 polymerase (BLR (DE3)) from Novagen.
2. Vector expressing His_6-tagged *Mos1* transposase (pET19-Tpase, Beverley lab strain B4289; 23)

3. 1 M IPTG (isopropyl-β-D-thio-galactopyranoside) stock solution
4. Resuspension buffer (20 mM Tris-HCl pH 7.6, 2 mM $MgCl_2$, 25 % sucrose, 0.6 mM phenylmethylsulfonyl fluoride (PMSF), 1 mM benzamidine (BZA) and 1 mM dithiothreitol (DTT)
5. Liquid nitrogen or Sonicator

2.2 Purification of transposase

1. Lysis buffer (20 mM Tris-HCl pH 7.6, 4 mM EDTA, 200 mM NaCl, 1 % deoxycholate, 1 % nonylphenoxy polyethoxy ethanol (NP-40), 0.6 mM PMSF, 1 mM BZA, 1 mM DTT)
2. DNAseI
3. $MgCl_2$
4. Lysozyme
5. Buffer A (20 mM Tris-HCl pH 8, 500 mM NaCl, 6 M guanidine HCl, 1 % NP-40, 70 mM imidazole)
6. Wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 100 mM imidazole, adjusted to pH 8.0 using NaOH)
7. Elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 500 mM imidazole, adjusted to pH 8.0 using NaOH)
8. Ni-NTA agarose (Qiagen, # 30210)
9. Purification column (Qiagen, # 34964)
10. SDS-PAGE gel
11. Dialysis slide (Slide-A-Lyzer® dialysis cassette, 10,000 MW cutoff ; Pierce catalog number 66425)

12. Dialysis buffer A (10 % glycerol, 25 mM Tris-HCl pH 7.6, 50 mM NaCl, 5 mM $MgCl_2$, 2 mM DTT)

13. Dialysis buffer B (10 % glycerol, 25 mM Tris-HCl pH 7.6, 50 mM NaCl, 5 mM $MgCl_2$, 0.5 mM DTT)

14. 100 % Glycerol

2.3 *In vitro* transposition assay

1. Reaction buffer (25 mM HEPES pH 7.9, 1 mM DTT, 5 mM $MgCl_2$, 100 mM $NaCl_2$)
2. 100 % Glycerol
3. Purified BSA (10 mg/ml) (New England Biolabs)
4. Stop buffer (50 mM Tris-HCl pH 7.6, 0.5 mg/ml proteinase K, 10 mM EDTA, 250 mg/ml yeast tRNA)
5. 3 M NaAcetate
6. 100 % ethanol
7. 70 % ethanol
8. Bacterial electoporator
9. *pir- E. coli* electrocompetent cells like DH10B
10. 25:24:1 Phenol/chloroform/isoamyl alcohol
11. 10 mM Tris pH 7.5
12. Ampicillin, hygromycin, nourseothricin (Dr.Walter Werner; WeBioAge@aol.com) and Zeocin (Invitrogen)
13. Lb medium (with appropriate drug)/plates

3. Methods

3.1 Expression of *Mos1* transposase

1. Transform BLR (DE3) with plasmid expressing *Mos1* transposase; plate on ampicillin (100 µg/ml).

2. The following day, pick a number of colonies and resuspend into LB with ampicillin (100 µg/ml) .

3. Use the above solution to inoculate 100 ml of media. Incubate culture at 37°C to an OD600 of 0.6.

4. Induce expression of transposase by adding IPTG to a final concentration of 1 mM. After 5 hours of induction harvest cells by centrifugation at 1,303g for 10 min at 4°C.

5. Resuspend in 0.5 ml of resuspension buffer and flash freeze in liquid nitrogen and store at -80 c. (*see* **Note 1**)

3.2 Purification of *Mos1* transposase

1. Thaw cells at room temperature.

2. Add 1 mg/ml of lysozyme and incubate for 5 min at room temperature. Next, add 1 ml of lysis buffer and incubate at room temperature for 15 minutes and then for 20 min after adding 60 µg of DNAseI and $MgCl_2$ to 10 mM.

3. All subsequent steps are conducted at 4°C.

4. Pellet inclusion bodies at 14,000g in a microcentrifuge and wash three times with 1 ml of 100 mM Tris-HCL pH 7.6. Resuspend inclusion bodies to a final volume of 4 ml in Buffer A. Add 1 ml of 50 % Ni-NTA agarose and gently shake solution for 1 hour.

5. Load mixture onto column (Qiagen) and collect flow-through.

6. Wash twice with 4 ml of Wash buffer and collect wash fractions.

7. Elute four times with 0.5 ml of Elution buffer.

8. Run 20 µl of eluted protein on an SDS-PAGE gel. On a 12 % SDS-PAGE gel the purified His-tagged *Mos1* transposase runs at approximately 50 kDa (*see* **Note 2**). Also load an aliquot of the non-induced culture, column flow-through and wash flow-through.

9. Pool the fractions containing the most transposase. Place the transposase into dialysis slide, ensuring not to overfill. Conduct dialysis in 1 liter of dialysis buffer A for 6 to 8 hours at 4°C. Replace dialysis buffer with 1 liter of dialysis buffer B and incubate overnight at 4°C. (*see* **Note 3**)

10. Centrifuge solution at 10,000g for 20 min at 4°C to remove precipitate.

11. Add glycerol to a final concentration of 50 %. Store at -20°C.

3.3 *In vitro* transposition

1. Set up a standard transposition reaction in a 0.6 ml microcentrifuge tube to a final volume of 20 µl.

   2 µl of 10x transposition buffer

   2 µl of 100 % glycerol (Warm glycerol at 65 C to ease pipetting)

   0.5 µl of BSA at 10 mg/ml

   1 µl of donor plasmid (32 fmol)

   6 µl of target plasmid (10 fmol, *see* **Note 4**)

   5 µl of transposase (100 nM, *see* **Note 5**)

   3.5 µl of sdH2O

2. Incubate reaction at 30 C for 1 hour to overnight.

3. Add 80 µl of stop buffer and incubate for 30 minutes at 37 °C

4. Add 100 µl of 25:24:1 phenyl/chloroform/isoamyl alcohol and vortex. Separate phases by centrifugation at 14,000g for 15 min. Remove approximately 90 µl of the upper layer into a 1 ml microcentrifuge tube. Add 10 µl of 3 M NaAcetate and 250 µl of EtOH and incubate at -80°C for at least 1 hour.

5. Precipitate DNA by centrifugation at 14,000g for 30 minutes at 4°C. Wash precipitated DNA with 1 ml of 70 % EtOH and centrifuge at 14,000g for 15 minutes at 4°C. Resuspend the pellet in 10 µl of 10 mM Tris pH 7.5.

6. Electroporate 2 µl of the purified transposition reaction into DH10B electrocompetent cells (*see* **Note 6**). Add 1 ml of Lb media and incubate for 1 hour at 37°C. Plate transfectants onto selective plates (*see* **Note 7 and 8**) and incubate at 37°C overnight.

7. Depending on the purposes, colonies may be picked individually or large pools made for DNA preparations and subsequent analysis *en mass*, for example following transfection into *Leishmania* or other organisms (*see* **Note 9**; 31).

4. NOTES

1. Alternatively, one can sonicate cells using a microtip at 40-50 % power for 20 bursts. The cells must be kept on ice during sonication. Afterwards one can proceed directly to step 3.2.2 excluding the addition of lysozyme.

2. The predicted molecular weight of the His-tagged *Mos1* transposase is 43.6 kDa. The altered mobility of the transposase during electrophoresis may result from the presence of the basic histidine residues. Antibodies to the poly-histidine residues specifically recognize the 50 kDa band by western blot hybridization.

3. The recovery of active, properly folded Tpase is very sensitive to the refolding conditions. Previous work has demonstrated that rapid dilution at low pH or dialysis of detergent solubilized proteins results in no active enzyme (17). Omission of the column purification step also results in no enzyme activity likely due to the presence of an unknown inhibitory factor. However, rapid dilution at pH 8 or refolding on a column using a linear urea gradient (8-0 M urea) has been shown to yield Tpase activity (17, 24). Thus, while the refolding protocol described in this article has worked effectively in this laboratory ultimately the optimal refolding conditions needs to be qualitatively determined by each investigator.

4. The quality of DNA is very important to transposition efficiencies. The preparation of donor and recipient DNA containing a high proportion of supercoiled DNA results in high transposition efficiencies. Qiagen midi preparations are generally suitable for this goal. However, when preparing cosmid DNA one should take special care to avoid shearing the DNA. Transposition efficiency reaches a maximum at around 150 ng of donor plasmid.

5. The concentration of transposase is determined by the micro-BCA method (Pierce). Concentration may also be determined by UV absorbance (ε_{280nm} = 76, 989 $M^{-1}cm^{-1}$); note that these measurements lead to differences in estimation of transposase concentration by a factor of 3 (17). Different batches of purified transposase can have different transposition efficiencies. This difference probably arises from batches of *Mos1* transposase containing different amounts of correctly folded transposase. Therefore, each batch should be tested before conducting large scale transposition reactions. The transposition efficiency reaches a plateau at around 100 nM of transposase and remains at this level at higher concentrations (17).

6. The use of high efficiency electrocompetent cells work best. Invitrogen GeneHogs® electrocompetant cells can yield 1 x 10^{10} transformants per μg of pUC vector.

7. Plate 10 μl of a 1/100 dilution of the transformation onto medium containing the appropriate antibiotic for the resistance marker found (ex. Ap, 100 μg/ml, Fig. 2) on the recipient plasmid. The number of Ap^R colonies multiplied by the dilution factor (in this case by 10000) is the transformation efficiency. The remaining 990 μl of cells are plated onto medium containing antibiotics for the resistance markers found on the recipient plasmid and the transposon cassette (such as Ap and Phleo in the example shown in Fig, 2; 100 μg/ml and 50 μg/ml respectively). Transposition efficiency is determined by dividing the number of Ap^R,$Phleo^R$ colonies by the transformation efficiency (obtained in previous step). Control transposition efficiencies should range from 10^{-4} to 10^{-3}.

8. One can also estimate the transposition efficiency of a vector like pELHY6TK-PG which contains no bacterial selectable marker. To accomplish this, a negative selectable marker like the product of the *ccdB* (control of cell death) gene in placed in the bacterial plasmid, which additionally containing positive selectable marker (Kanamycin, Km^R). Examples of this are the pZERO system available from Invitrogen. Transformation of this plasmid into bacteria lacking the gyrase gene results in Km+ colonies. However, cell death occurs when transfected into strains containing gyrase like TOP10. An *in vitro* transposition reaction is performed with equimolar amounts of this plasmid (Km^R) and a standard plasmid target (Chloramphenicol, Cm^R) in a strain lacking gyrase. The transposition reaction is subsequently transformed into TOP10 bacteria. One half of the transformation is plated onto chloramphenicol plates (30 μg/ml) and the other half is plated onto kanamycin plates (50 μg/ml). The number of colonies obtained when plated on Kanamycin represents transposition events into the *ccdB* gene. The ratio of Km^R / Cm^R represents the transposition efficiency of this transposon.

9. If one is planning transfections into *Leishmania* it is critical to first determine the sensitivity of your specific strain under the exact circumstances you plan to use. Drug sensitivities vary greatly amongst different strains and species, in different media, and interactions can occur if two drugs are used simultaneously. First determine the EC_{50} in liquid media; then, carry out a 'mock' transfection followed by inoculation or plating onto media containing drug concentrations ranging upwards from 3-4 of the liquid media EC_{50}. Since high drug concentrations inhibit the recovery of *bona fide* transfectants, the goal is to identify the minimal drug concentration that kills untransfected/drug sensitive cells. In *Leishmania* drug concentrations typically are 15-30 μg/ml for G418/geneticin, 25-40 μg/ml for phleomycin, 15-30 μg/ml for hygromycin B and 100 μg/ml for nourseothricin, but exceptions are common.

Acknowledgements

The authors thank Lon-Fye Lye for assistance with the transposon protocols and Rosa Reguera for comments on the manuscript. This work is supported by NIH research fellowship (AI50339) to KAR and NIH grants AI 29646 to SMB.

Table and Figure Legends.

Table 1. Examples of *Mos1* transposons and their properties and applications. All transposons listed can be used for insertional mutagenesis and sequencing. Symbols are as described: AG, *Leishmania* splice acceptor; *SAT*, nourseothricin resistance marker; *Km*, kanamycin resistance marker, GFP, modified green fluorescent protein; *PHLEO*, phleomycin/zeocin resistance marker; "/"indicates a gene lacking a start or stop codon respectively; "*" indicates an in-frame stop codon; black arrows, *E. coli* promoter; open triangle, 5' inverted repeat; grey triangle, 3' inverted repeat; oriC, oriC origin of replication; and *Bsr*GI and *Sex*AI, unique restrictions sites used in TIMLI mutagenesis..

Fig. 1. Transposon mutagenesis strategies. (A) Shuttle mutagenesis begins with *in vitro* transposition into target DNA (plasmid or cosmid) to create an insertion library. The transposon library is then transfected into *Leishmania*, and recovered by selection on the drug resistance marker found on the transposon therein in various ways. (B) *In vivo* mutagensis requires the establishment of an active transposition system in the parasite itself. This could be accomplished in several ways; stable expression of transposase and stable introduction of the transposon has been successful (29). Ideally, one would prefer transient introduction of the transposition system, for example using transient or regulated expression of transposase, and/or transient introduction of donor transposons. Alternatively, one could form a transposase-transposon complex *in vivo*, and then introduce this for subsequent transposition *in vitro*. This has not yet been demonstrated in the *mariner* system but it works well with transposon Tn5 in a variety of eukaryotes (35, 36). Reprinted with permission from (37).

Fig. 2. *Mos1* vectors and the *in vitro* transposase reaction. **Panel A:** The 'empty'donor plasmid pELHY6Δ-0 contains the minimal *cis*-element (open arrow heads) with the 5' and 3' IRs and some internal nucleotides (Fig. 3A). The vector contains the *E. coli* OriR6K origin (striped box) for propagation in a *λpir*+ strain and a *HYG* selectable marker (internal black line); this particular marker contains a *Leishmania* splice acceptor site (AG) for expression in *Leishmania* and an *E. coli* promoter (black arrow). For use one can insert a variety of 'cargo' within the inverted repeats, at the unique *Msl*I, *Xba*I or *Sbf*I sites (Fig. 3A). In this figure the donor element is pELHY6Δ-/GEP3/, created by insertion of the /GEP3/ (Table 1). For simplicity only the *E. coli PHLEO* resistance marker is shown in /GEP3/. **Panel B:** The basic *in vitro* transposase reaction contains a donor element, target DNA and transposase (shaded area). In this example, the transposon target is a cosmid DNA which contains an *E. coli* ampicillin resistance marker (Ap) and an OriC origin of replication (open box). After *in vitro* transposition into a TA dinucleotide, the DNA is transformed into a *λpir*- strain (like DH10B) to select against the donor plasmid. Bacterial transformants are plated on Ap/PHLEO to select for transposition. Target DNAs containing transpositions can then be transfected into *Leishmania*, and fusion proteins identified by selecting or screening for PHLEO or GFP expression respectively.

Fig. 3. Nucleotide sequences of *Mos1 cis*-element and pELHY6Δ-/GEP3/, pELHY6Δ-/GEP2/ chimeric genes. (A) The minimal *Mos1 element* used in our work contains essential *cis*-elements consisting of the 5' and 3' inverted repeats (shaded gray arrows) and the internal 38 and 5 internal nucleotides (non-shaded capital letters). The six potential *Mos1* reading frames coming in from flanking DNAs across the IRs are shown with arrows (labeled 1-6) while start and stop codons in each frame is shown by M or X respectively. Unique restriction sites found

within the empty transposon suitable for the addition of 'cargo' are shown. (B) Putative chimeric genes created by insertion of the /GEP3/ and GEP2/ transposons into target TA dinucleotides in the third or second reading frame respectively. Reprinted with permission from (23).

Fig. 4. Diagram of potential translational fusions obtained with transposons /GEP3* and /GEP3/. Both transposons yield translation fusions when inserted in frame into target ORFs, which express a GFP-PHLEO resistance fusion protein domain (Table 1). Note that phleomycin resistance can be affected by compartmentalization of the fusion protein; if the PHLEO protein domain is restricted to a compartment such as the glycosome which is segregated from the nucleus (the site of action of phleomycin), resistance will be abrogated (32). (A) The pELHY6Δ-/GEP3* transposon encodes a bifunctional GFP-PHLEO protein with a stop codon after the PHLEO domain. Thus fusion proteins contain only N-terminal sequence information from the target. (B) The pELHY6Δ-/GEP3/ transposon contains an ORF across the entire transposon (both IRs and the bifunctional GFP-PHLEO protein). C) Comparison of the use of 'terminator' vs 'sandwich' protein fusions. In this example the properties of fusion proteins generated by /GEP3* and /GEP3/ are compared following in frame insertion into a typical membrane surface protein, which bears an N terminal signal peptide and C-terminal membrane anchor. With /GEP3* the C-terminal segment is lost, resulting in secretion of the fusion protein from the cell, while with /GEP3/ retention of the C-terminal segment results in the formation of a surface membrane anchored fusion protein.

References

1. **Herwaldt, B. L.** (1999). Leishmaniasis. *Lancet* **354:**1191-9.
2. **Beverley, S. M.** 2003. Genetic and genomic approaches to the analysis of *Leishmania* virulence. *In* J. M. Marr, T. Nilsen, and R. Komuniecki (ed.), Molecular & Medical Parasitology. Academic Press, New York.
3. **Beverley, S. M.** (2003). Protozomics: trypanosomatid parasite genetics comes of age. *Nat Rev Genet* **4:**11-9.
4. **Myler, P. J., and Stuart, K. D.** (2000). Recent developments from the *Leishmania* genome project. *Curr Opin Microbiol* **3:**412-6.
5. **Swindle, J., and Tait, A.** 1996. Trypanosomatid genetics, p. 6-34. *In* D. F. Smith and M. Parsons (ed.), Molecular Biology of Parasitic Protozoa. IRL Press, Oxford.
6. **Hamer, L., DeZwaan, T. M., Montenegro-Chamorro, M. V., Frank, S. A., and Hamer, J. E.** (2001). Recent advances in large-scale transposon mutagenesis. *Curr Opin Chem Biol* **5:**67-73.
7. **Gueiros-Filho, F. J., and Beverley, S. M.** (1996). Selection against the dihydrofolate reductase-thymidylate synthase (*DHFR-TS*) locus as a probe of genetic alterations in *Leishmania major*. *Molec. Cell. Biol.* **16:**5655-5663.
8. **Ryan, K. A., Dasgupta, S., and Beverley, S. M.** (1993). Shuttle cosmid vectors for the trypanosomatid parasite *Leishmania*. *Gene* **131:**145-50.
9. **Cecconi, F., and Meyer, B. I.** (2000). Gene trap: a way to identify novel genes and unravel their biological function. *FEBS Lett* **480:**63-71.
10. **Stanford, W. L., Cohn, J. B., and Cordes, S. P.** (2001). Gene-trap mutagenesis: past, present and beyond. *Nat Rev Genet* **2:**756-68.
11. **Biery, M. C., Lopata, M., and 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.
12. **Goryshin, I. Y., and Reznikoff, W. S.** (1998). Tn5 in vitro transposition. *J Biol Chem* **273:**7367-74.
13. **Merkulov, G. V., and Boeke, J. D.** (1998). Libraries of green fluorescent protein fusions generated by transposition *in vitro*. *Gene* **222:**213-22.
14. **Lavoie, B. D., and Chaconas, G.** (1996). Transposition of phage Mu DNA. *Curr Top Microbiol Immunol* **204:**83-102.

15. **Lampe, D. J., Churchill, M. E., and Robertson, H. M.** (1996). A purified *mariner* transposase is sufficient to mediate transposition in vitro. *EMBO J* **15:**5470-9.
16. **Fischer, S. E., Wienholds, E., and Plasterk, R. H.** (2001). Regulated transposition of a fish transposon in the mouse germ line. *Proc Natl Acad Sci U S A* **98:**6759-64.
17. **Tosi, L. R., and 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:**784-90.
18. **Doak, T. G., Doerder, F. P., Jahn, C. L., and Herrick, G.** (1994). A proposed superfamily of transposase genes: transposon-like elements in ciliated protozoa and a common "D35E" motif. *Proc Natl Acad Sci U S A* **91:**942-6.
19. **Robertson, H. M., and Lampe, D. J.** (1995). Recent horizontal transfer of a *mariner* transposable element among and between Diptera and Neuroptera. *Mol Biol Evol* **12:**850-62.
20. **Hartl, D. L., Lohe, A. R., and Lozovskaya, E. R.** (1997). Modern thoughts on an ancyent marinere: function, evolution, regulation. *Annu Rev Genet* **31:**337-58.
21. **Vos, J. C., De Baere, I., and Plasterk, R. H.** (1996). Transposase is the only nematode protein required for *in vitro* transposition of Tc1. *Genes Dev* **10:**755-61.
22. **Plasterk, R. H.** (1996). The Tc1/*mariner* transposon family. *Curr Top Microbiol Immunol* **204:**125-43.
23. **Goyard, S., Tosi, L. R., Gouzova, J., Majors, J., and Beverley, S. M.** (2001). New *Mos1 mariner* transposons suitable for the recovery of gene fusions *in vivo* and *in vitro*. *Gene* **280:**97-105.
24. **Zhang, L., Dawson, A., and Finnegan, D. J.** (2001). DNA-binding activity and subunit interaction of the *mariner* transposase. *Nucleic Acids Res* **29:**3566-75.
25. **Coates, C. J., Jasinskiene, N., Miyashiro, L., and James, A. A.** (1998). *Mariner* transposition and transformation of the yellow fever mosquito, *Aedes aegypti*. *Proc Natl Acad Sci U S A* **95:**3748-51.
26. **Lohe, A. R., and Hartl, D. L.** (1996). Reduced germline mobility of a *mariner* vector containing exogenous DNA: effect of size or site? *Genetics* **143:**1299-306.
27. **Garza, D., Medhora, M., Koga, A., and Hartl, D. L.** (1991). Introduction of the transposable element *mariner* into the germline of *Drosophila melanogaster*. *Genetics* **128:**303-10.

28. **Fadool, J. M., Hartl, D. L., and Dowling, J. E.** (1998). Transposition of the *mariner* element from *Drosophila mauritiana* in zebrafish. *Proc Natl Acad Sci U S A* **95:**5182-6.
29. **Gueiros-Filho, F. J., and Beverley, S. M.** (1997). Trans-kingdom transposition of the *Drosophila* element *mariner* within the protozoan *Leishmania*. *Science* **276:**1716-9.
30. **Papadopoulou, B., Roy, G., and Ouellette, M.** (1994). Autonomous replication of bacterial DNA plasmid oligomers in *Leishmania*. *Mol. Biochem. Parasitol.* **65:**39-49.
31. **Robinson, K., and Beverley, S. M.** (2003). Improved methods for transfection and tests of RNA interference (RNAi) activity in *Leishmania*. *submitted*.
32. **Flaspohler, J. A., Rickoll, W. L., Beverley, S. M., and Parsons, M.** (1997). Functional identification of a *Leishmania* gene related to the peroxin 2 gene reveals common ancestry of glycosomes and peroxisomes. *Molec. Cell. Biol.* **17:**1093-101.
33. **LeBowitz, J. H., Cruz, A., and Beverley, S. M.** (1992). Thymidine kinase as a negative selectable marker in *Leishmania major*. *Mol Biochem Parasitol* **51:**321-5.
34. **Hayes, F., and Hallet, B.** (2000). Pentapeptide scanning mutagenesis: encouraging old proteins to execute unusual tricks. *Trends Microbiol* **8:**571-7.
35. **Shi, H., Wormsley, S., Tschudi, C., and Ullu, E.** (2002). Efficient transposition of preformed synaptic Tn5 complexes in *Trypanosoma brucei*. *Mol Biochem Parasitol* **121:**141-4.
36. **Goryshin, I. Y., Jendrisak, J., Hoffman, L. M., Meis, R., and Reznikoff, W. S.** (2000). Insertional transposon mutagenesis by electroporation of released Tn5 transposition complexes. *Nat Biotechnol* **18:**97-100.
37. **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, F., MacMillan, S., Muo, J. N., Schwertman, L. A., Smeds, M. R., and Wang, Y.** (2002). Putting the *Leishmania* genome to work: functional genomics by transposon trapping and expression profiling. *Philos Trans R Soc Lond B Biol Sci* **357:**47-53.

Table 1

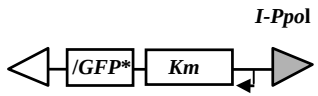
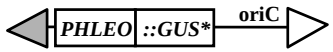
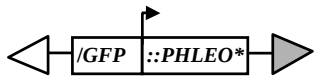
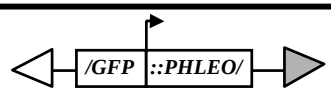
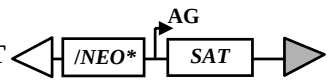
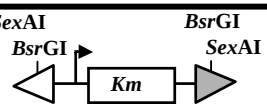
Transposon	Transposon Elements: Bacterial marker	Transposon Elements: Eukaryotic marker	Transposon Elements: Reporter	Application: Translational fusions	Application: Transcriptional fusions	Application: In vivo	Application: In vitro	Application: TIMLI
pELHY6Δ-/GFP*K 	Km	—	—		+	—	+	+	—
pELHY6TK-PG 	—	PHLEO	GUS		PHLEO GUS	PHLEO GUS	+	—	—
pELHY6Δ-/GEP3* 	PHLEO	PHLEO	GFP		GFP PHLEO	PHLEO	+	+	—
pELHY6Δ-/GEP3/ /GEP2/ 	PHLEO	PHLEO	GFP		GFP PHLEO	PHLEO	+	+	—
pELHY6Δ-/NEO*ELSAT 	SAT	NEO SAT	—		NEO	SAT	+	+	—
pELHY6Δ-/-2x5 	Km	—	—		—	—	—	+	+

Fig. 1

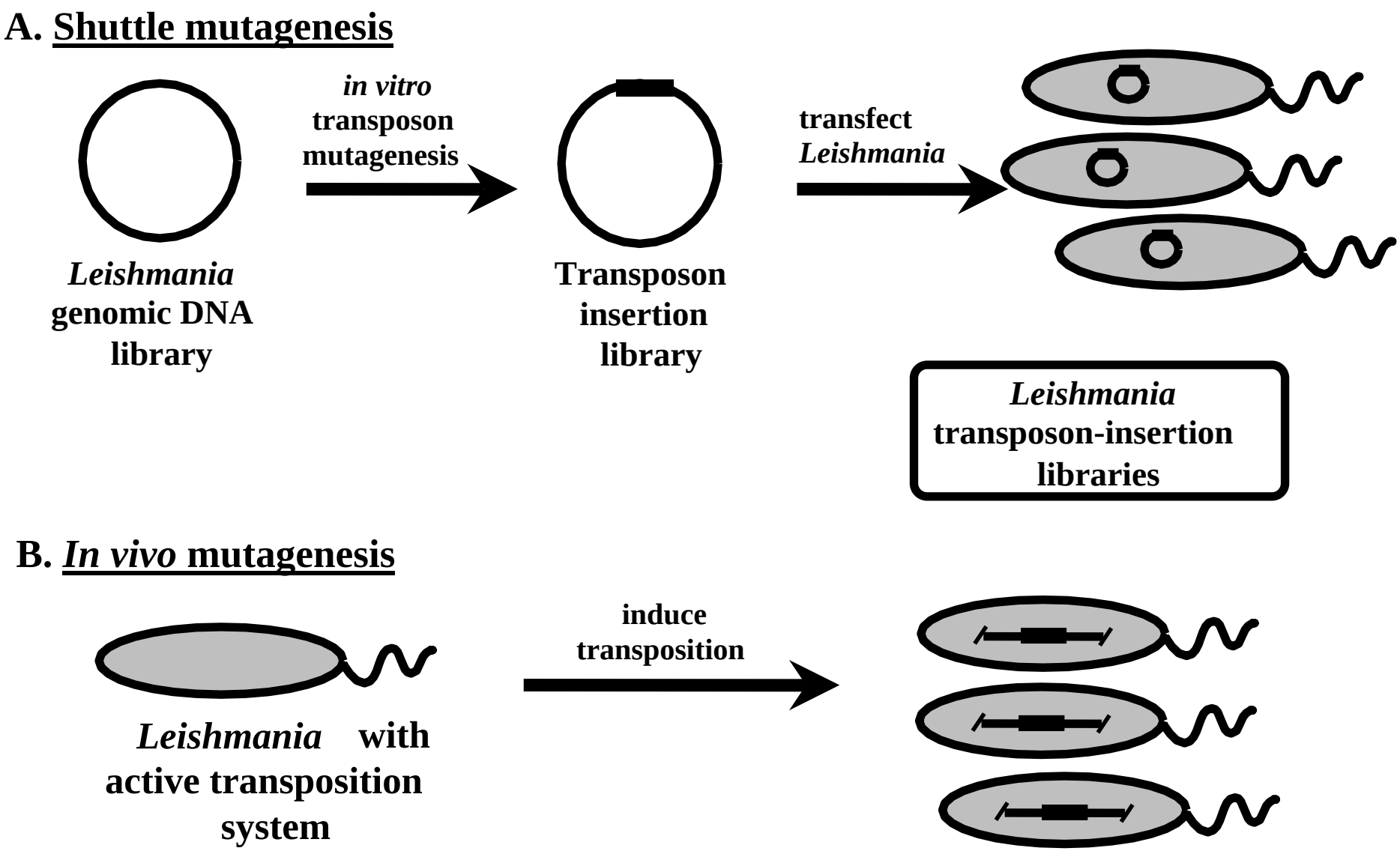

A. Shuttle mutagenesis
in vitro transposon mutagenesis
transfect Leishmania
Leishmania genomic DNA library
Transposon insertion library
Leishmania transposon-insertion libraries
B. In vivo mutagenesis
induce transposition
Leishmania with active transposition system

Fig. 2

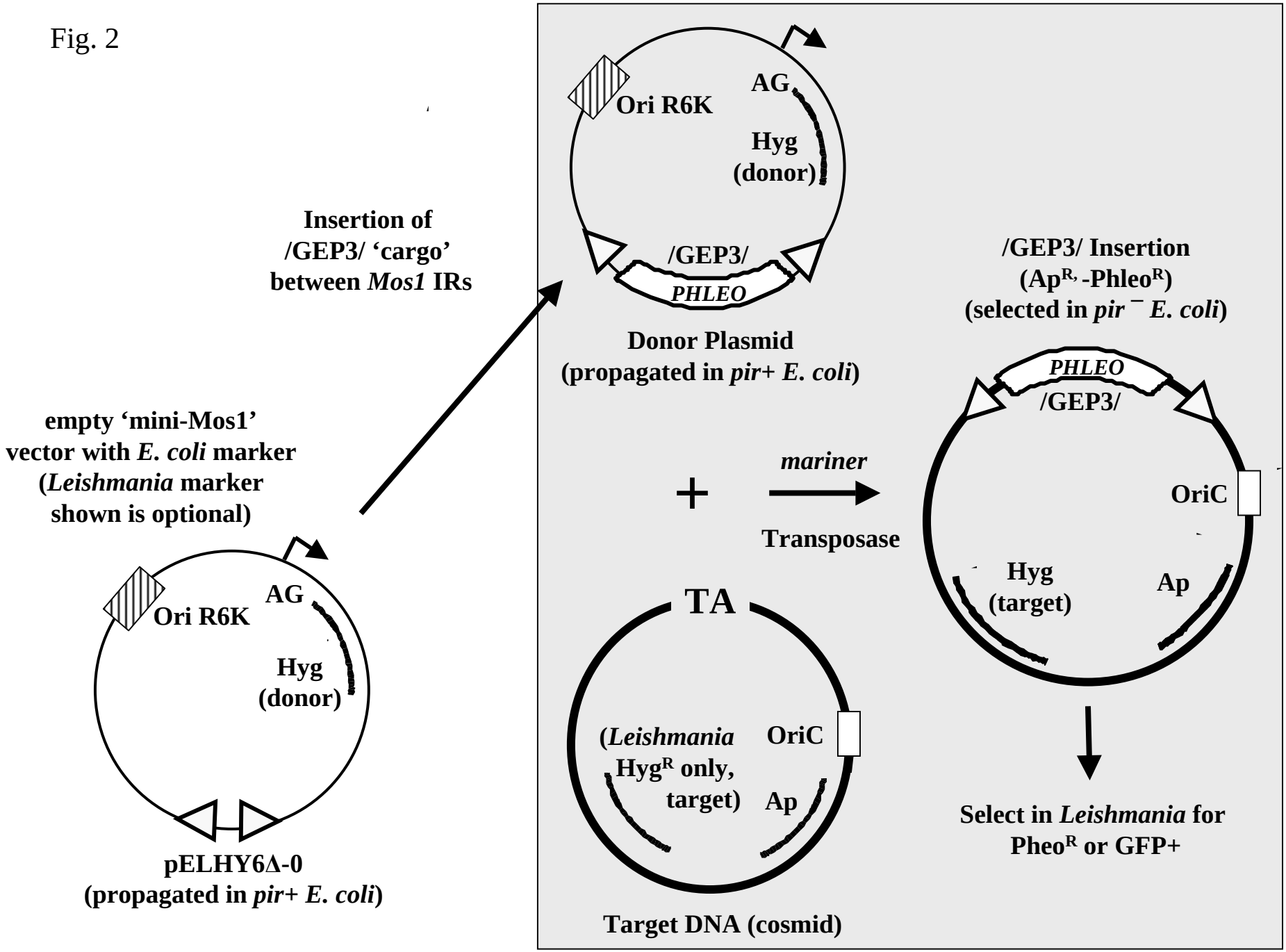
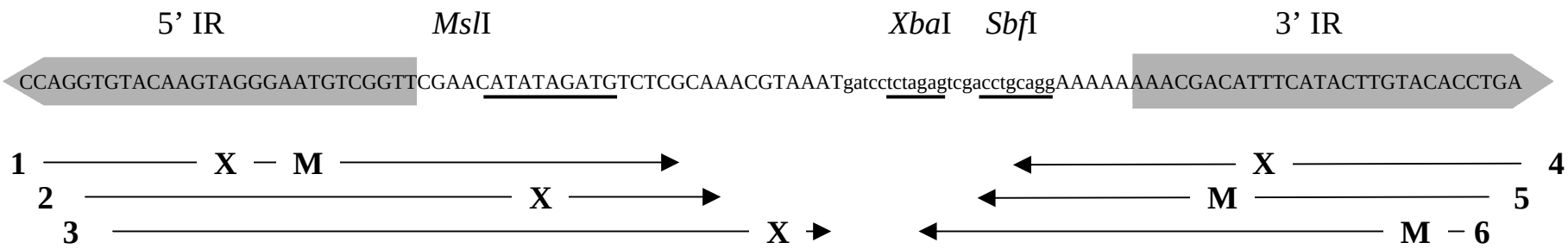

Insertion of
/GEP3/ 'cargo'
between *Mos1* IRs
empty 'mini-Mos1'
vector with *E. coli* marker
(*Leishmania* marker
shown is optional)
Ori R6K
AG
Hyg
(donor)
pELHY6Δ-0
(propagated in *pir*+ *E. coli*)
Ori R6K
AG
Hyg
(donor)
/GEP3/
PHLEO
Donor Plasmid
(propagated in *pir*+ *E. coli*)
+
mariner
Transposase
TA
(*Leishmania*
Hyg^R only,
target)
OriC
Ap
Target DNA (cosmid)
/GEP3/ Insertion
(Ap^R, -Phleo^R)
(selected in *pir* $^-$ *E. coli*)
PHLEO
/GEP3/
OriC
Hyg
(target)
Ap
Select in *Leishmania* for
Pheo^R or GFP+

Fig. 3

A


5’ IR
MslI
XbaI
SbfI
3’ IR
CCAGGTGTACAAGTAGGGAATGTCGGTTCGAACATATAGATGTCTCGCAAACGTAAATgatcctctagagtcgacctgcaggAAAAAAAAACGACATTTCATACTTGTACACCTGA
1 X M
2 X
3 X
X 4
M 5
M 6

B

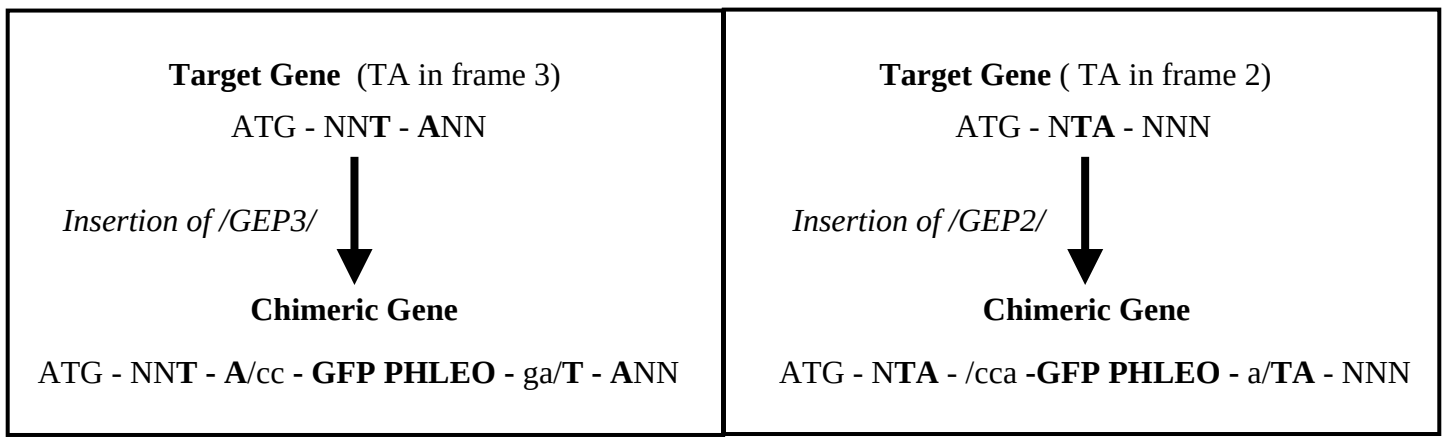
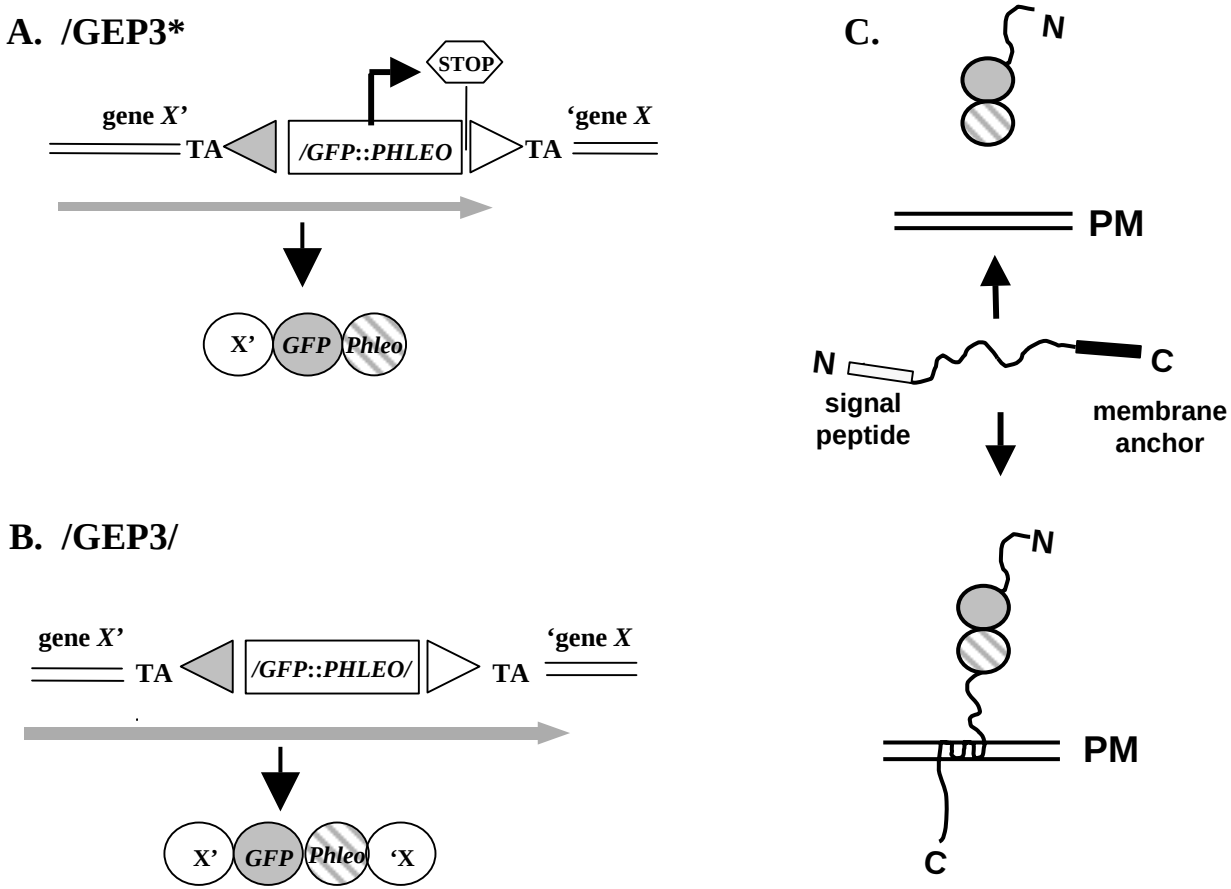

Target Gene (TA in frame 3)
ATG - NNT - ANN
Insertion of /GEP3/
Chimeric Gene
ATG - NNT - A/cc - GFP PHLEO - ga/T - ANN
Target Gene (TA in frame 2)
ATG - NTA - NNN
Insertion of /GEP2/
Chimeric Gene
ATG - NTA - /cca -GFP PHLEO - a/TA - NNN

Fig. 4


A. /GEP3*
STOP
gene X'
TA
/GFP::PHLEO
TA
'gene X
X'
GFP
Phleo
B. /GEP3/
gene X'
TA
/GFP::PHLEO/
TA
'gene X
X'
GFP
Phleo
'X
C.
N
PM
N
C
signal peptide
membrane anchor
N
PM
C