ELSEVIER

Gene 198 (1997) 27–35

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

Insertional mutagenesis by a modified in vitro Ty1 transposition system

Levi A. Garraway, Luiz R.O. Tosi, Yixin Wang, Jeffrey B. Moore, Deborah E. Dobson, Stephen M. Beverley *

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA, USA

Received 24 January 1997; accepted 18 April 1997; Received by C.M. Kane

Abstract

Transposable elements are useful tools for insertional mutagenesis and have many potential applications in the characterization of complex genomes. Here we describe a system which facilitates the construction of large transposon insertion libraries useful for genome sequencing and functional genomic analysis. We developed two transposons, TyK and TyK′GFP+, which can be introduced into target DNAs by Ty1-mediated transposition in vitro, and several modifications which decrease the frequency of false transposition events and direct the recovery of transpositions into passenger rather than vector DNA. Insertions of TyK′GFP+ additionally may yield fusions to the *Aequorea* green fluorescent protein (GFP), useful in studies of gene expression and protein targeting. Transposition in vitro was obtained into target DNAs of up to 50 kb in size, restriction mapping showed insertion to be relatively random, and the sequence of 55 insertion sites showed neither strong site nor base compositional preference. Our data suggest that TyK-based artificial transposons will be suitable for a variety of genetic applications in many organisms. © 1997 Elsevier Science B.V.

Keywords: Transposon; DNA sequencing; Green fluorescent protein; Gene fusions

1. Introduction

The characterization of complex eukaryotic genomes by determination of their complete sequence has resulted in a shift in emphasis by genome scientists (Johnston, 1996; Pennisi, 1996). Now, the major task is to gain an understanding of the cellular function of genes and their utilization under different environmental or developmental conditions. Transposable elements offer many potential advantages suitable for incorporation into genome-wide studies, as modified transposons are popular tools for insertional mutagenesis, ''primer island'' sequencing and the construction of gene fusions (Casadaban et al., 1983; Silhavy et al., 1984; Seifert et al., 1986; Berg et al., 1989; Burns et al., 1994; Sherratt, 1995; Morgan et al., 1996). One powerful approach for genome scientists is the method of shuttle transposon mutagenesis (Seifert et al., 1986). Burns et al. (1994) used this approach to generate nearly 3000 independent *Saccharomyces cerevisiae* protein fusions to *Escherichia coli* β-galactosidase, facilitating studies of cellular localization and function. Similarly, Chun and Goebl (1996) used a related approach to identify essential genes affecting bud morphogenesis in *S. cerevisiae*. Both of these studies required the generation of large, random transposon insertion libraries, and here we describe a convenient system which facilitates their construction for use in any organism.

In selecting transposition systems, important factors are the randomness of insertion site and the ability to design transposons for specific applications. Often, specialized bacterial strains or manipulations are required for transposon delivery, which can prove daunting to novices. A yeast Ty1-based transposition system was recently described which offers several advantages (Devine and Boeke, 1994). The transposition reaction is performed in vitro, employing only a linear Ty1 donor fragment, a target plasmid and transposase in the form of Ty1 VLPs which are readily prepared (Boeke et al., 1985). Although endogenous Ty1 itself shows a preference for certain locations within the yeast genome such as tRNA loci or other transposon LTRs (Ji et al., 1993),

* Corresponding author. Present address: Department of Molecular Microbiology, Washington University School of Medicine, 660 S. Euclid Avenue, St. Louis, MO 63105, USA. Tel.: +1 314 7472630; Fax: +1 314 747-2634; e-mail: beverley@borcim.wustl.edu

Abbreviations: GFP, green fluorescent protein; LTR, long terminal repeat; PEG, poly(ethylene glycol); RT, reverse transcriptase; VLP, virus-like particle.

PII S0378-1119(97)00288-6

in vitro Ty1 transposition appears to occur much more randomly into the target (Boeke et al., 1985; this work). The structural requirements for efficient in vitro transposition of a donor Ty1 transposon fragment are minimal, requiring only four specific bases at the termini (Boeke et al., 1985).

By virtue of its randomness and versatility, Ty1-mediated in vitro transposition could prove widely applicable in studies of gene function and regulation. We have introduced several modifications facilitating the generation of large, random "transposon libraries" useful in DNA sequencing, functional gene inactivation and the construction of gene fusions. Two new Ty1-based transposons have been developed: TyK, which uses the Tn903 kanamycin resistance marker and contains rare-cutting endonuclease sites helpful in mapping insertions; and TyK′GFP+, a TyK derivative capable of generating fusions to the *Aequorea* GFP (Chalfie et al., 1994). The transposons are borne on a donor plasmid chosen to decrease the incidence of "false" positive transposition events and we developed a target cloning vector which directs virtually all recoverable transposition events into the passenger DNA. With this system, we show that both TyK and TyK′GFP+ integrate randomly into several target DNAs derived from the trypanosomatid protozoan parasite *Leishmania major*. We have used this approach to generate several transposon insertion libraries which have proved useful in genetic studies involving parasite drug resistance and virulence in our laboratory.

2. Materials and methods

2.1. Strains and plasmid constructions

Yeast strain JB224 (Boeke et al., 1985; Beverley lab strain B2089) was provided by G. Fink (Massachusetts Institute of Technology). JB224 contains the plasmid pTyH3, which carries the VLP-encoding *TYA* and *TYB* genes downstream of the inducible GAL4 promoter.

The construction of the transposon donor plasmid pTyK (Fig. 2; lab strain B2074) and the sequence of the mini-transposon TyK (Fig. 1), are provided in GenBank entry U63147 or are available from the authors. pTyK contains a tetracycline-resistance gene and R6K origin of replication, which requires the *pir* gene product, and was propagated in strain DH5$\alpha\lambda$*pir* (lab strain B2046; provided by S. Chiang). pTyK′GFP+ (lab strain B2798) contains a modified GFP cassette from plasmid pSynGFPS65T (Haas et al., 1996; lab strain B2798), inserted into the *Not*I site of pTyK (Fig. 1). The sequence of pTyK′GFP+ is provided in GenBank entry U84737.

The construction of pELHYG and pELSAT is described in GenBank entries U84738 and U84739. pELHYG (2128 bp, lab strain B2766; Fig. 2) contains a *Leishmania* splice acceptor (that located 3′ of the *DHFR-TS* gene, Kapler et al., 1990), followed by a synthetic T7/*E.coli* promoter from plasmid pUT100 (Cayla, Toulouse, France), a hygromycin resistance gene (*Hy*) and a minimal ColE1 origin of replication. *Hy* was replaced with the streptothricin resistance gene (*SAT*) in pELSAT (1650 bp, lab strain B2764). The T7/*E. coli* promoter drives expression of the *Hy* (or *SAT*) gene.

2.2. Preparation of VLPs

VLPs were prepared according to a modified version of protocols described previously (Eichinger and Boeke, 1988; Braiterman et al., 1994). JB224 was cultured with shaking in 500 ml case/-ura medium (1.8 g/l yeast nitrogen base, 6.6 g/l autoclaved casamino acids, 20 mg/L-tryptophan) with 1% raffinose at room temperature for 1–2 days. Galactose was added to 2% to induce expression of VLPs, and the incubation continued another 24 h. Cells were collected by centrifugation (Sorvall GSA, Newtown, CT, USA; 10 min at 6000 rpm; 4°C), washed with 10 ml water, and suspended in 5 ml of ice-cold buffer B/Mg (10 mM Hepes–KOH pH 7.8, 15 mM KCl, 5 mM $MgCl_2$, 3 mM DTT and 10 μg/ml aprotinin). Pre-chilled nitric acid-washed glass beads (Sigma, St. Louis, MO, uSA) were added until no free liquid remained, and the cells were lysed by vortexing vigorously for 3 min, followed by 2 min on ice; this was repeated four to five times or until lysis was complete, as monitored by phase-contrast microscopy. The lysate was collected (2–3 ml) and the beads were washed three times with 2 ml buffer B/Mg. The lysate and washes were combined (7–8 ml) and centrifuged (Sorvall SA600, 10 min at 10 000 rpm; 4°C). The supernatant was layered onto a linear 20–70% sucrose gradient in buffer B/EDTA (10 mM EDTA replaces the $MgCl_2$ and aprotinin in buffer B/Mg), followed by centrifugation for 24 h at 25 000 rpm and 4°C in a Beckman SW28 rotor. One milliliter fractions were collected through an 18-gauge needle. In some experiments, VLP-containing fractions were purified further by centrifugation through two-step gradients consisting of the following in a SW28 polyallomer tube: 5 ml of 70% sucrose, 5 ml of 30% sucrose and 20 ml of 20% sucrose in buffer B/EDTA. These were centrifuged for 3 h at 25 000 rpm in a SW28 rotor at 4°C.

2.3. Reverse transcriptase assay

VLPs were monitored by their RT activity (Garfinkel et al., 1985). The assay mixture consisted of 6 μl of gradient fraction and 24 μl of RT assay mix (50 mM Tris–Cl pH 8.0, 10 mM $MgCl_2$, 15 μM dGTP, 10 μg/ml poly(C), 0.7 μg/ml oligo(dG), 2% β-mercaptoethanol, 33 nCi [α-^{32}P]dGTP). Samples were incubated for 1 h at 37°C. An aliquot of 25 μl of the assay mix was spotted onto a Whatman DE81 disc. Discs were washed three times for 20 min each in 300 ml of 2 × SSPE (0.3 M NaCl, 2 mM EDTA, 20 mM NaH_2PO_4, pH 7.0) at room

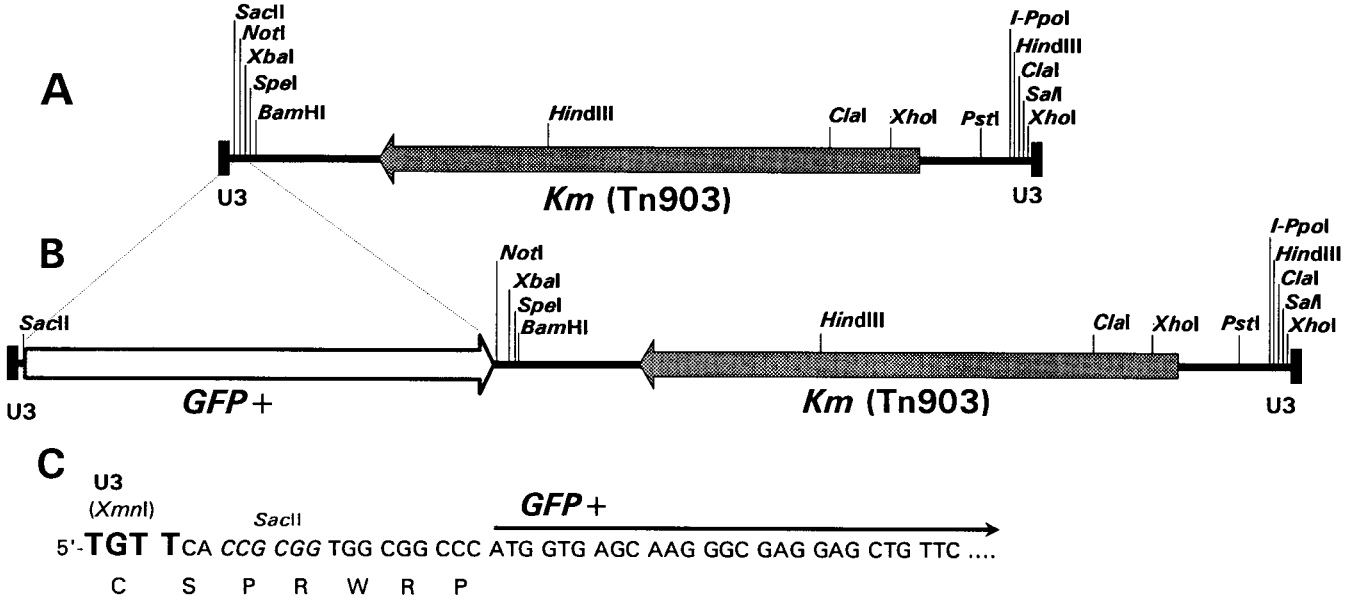

Fig. 1. Properties of TyK and TyK′GFP+. Restriction map and features of TyK (panel A) and TyK′GFP+ (panel B). The Tn903 kanamycin-resistance gene (shaded arrow), GFP (open arrow), U3 termini (small filled rectangles) and relevant restriction sites are shown. *Not*I, *Xba*I, *Spe*I, *Pst*I and *I-Ppo*I are unique within both pTyK and pTyK′GFP+. Panel C shows the left end of TyK′GFP+, with translation into protein through the start of GFP+ (marked by overlining arrow). (*Xmn*I) denotes the 5 bp half-site created following *Xmn*I digestion to expose the U3 termini. The sequences of TyK and TyK′GFP+ can be found in GenBank entries U63147 and U84737.

temperature, and bound radioactivity was counted by Cerénkov emission. 22 units of M-MuLV Reverse Transcriptase (Boehringer-Mannheim, Indianapolis, IN, USA) provided a quantitative positive control. One RT unit was defined as that which incorporates 1 nmol of dGMP/10 min. Fractions containing high RT activity (>180 U/ml) were aliquoted and stored at −80°C, where they were stable for several months. Typically, enough VLPs for 400–500 transposition reactions could be obtained from a single preparation.

2.4. Nuclease assays

VLP samples (0.5 RT units) were incubated with 1 μg of a large target DNA (cosmid cHP; 51 kb) in the in vitro transposition reaction buffer (described below). Reactions were terminated by the addition of 1 μl 0.5 M EDTA (pH 8.0), incubation at 65°C for 20 min, extraction with phenol/chloroform/isoamyl alcohol (25:24:1), and the DNAs were precipitated with ethanol and suspended in 10 mM Tris, 1 mM EDTA (pH 8.0). DNA integrity was first assessed by gel electrophoresis, and samples exhibiting minimal degradation were evaluated by quantitative DNA transformation. An aliquot of 0.1 μg DNA of experimental and control DNAs were transformed into *E. coli*, and 1% of the total was plated on LB-ampicillin plates (50 μg/ml ampicillin).

2.5. Preparation of DNAs

The purity and integrity of DNAs used for in vitro transposition is very important. *E. coli* DH5α*λpir* (Herrero et al., 1990) transformed with pTyK or pTyK′GFP+ was cultured in 500 ml 2× YT medium (16 g/l tryptone, 10 g/l yeast extract and 5 g/l NaCl) supplemented with 10 μg/ml kanamycin and 10 μg/ml tetracycline at 37°C for 24 h with vigorous aeration. DNA was prepared by the alkaline lysis and PEG precipitation procedure (Ausubel et al., 1990), which gave the best yield. Following digestion with *Xmn*I, the transposon fragment was purified by gel electrophoresis using Ultrapure Low Melt Agarose (Gibco–BRL, Gaithersburg, MD, USA). The 1.2 kb (TyK) or 2 kb band (TyK′GFP+) was recovered by electroelution (Ausubel et al., 1990), extracted with phenol/chloroform/isoamyl alcohol (25:24:1), precipitated with ethanol, and suspended in 10 mM Tris, 1 mM EDTA (pH 8.0) at a concentration of 300 μg/ml.

Transposition targets are listed in the legends to Tables 1 and 2. DNAs for transposition and sequencing were prepared by alkaline lysis and PEG precipitation (clones pSN7-1 and pSN6-1.1; Ausubel et al., 1990) or by Qiagen Plasmid Maxi Kits (all other targets). The sequences of insertions were obtained using Taquenase (Amersham USB, Arlington Heights, IL, USA) and the oligonucleotide primers SMB216 (5′CATTTGGCTACC-TTAAGAGA) and SMB217 (5′CTACAACAAAGCTC-TCATCA) for TyK, or primers SMB216 and SMB458 (5′TTCACGTCGCCGTCCAG) for TyK′GFP+.

2.6. In vitro transposition reactions

Transposition reactions (Devine and Boeke, 1994) contained: 2 μl 10× transposition buffer (100 mM Tris

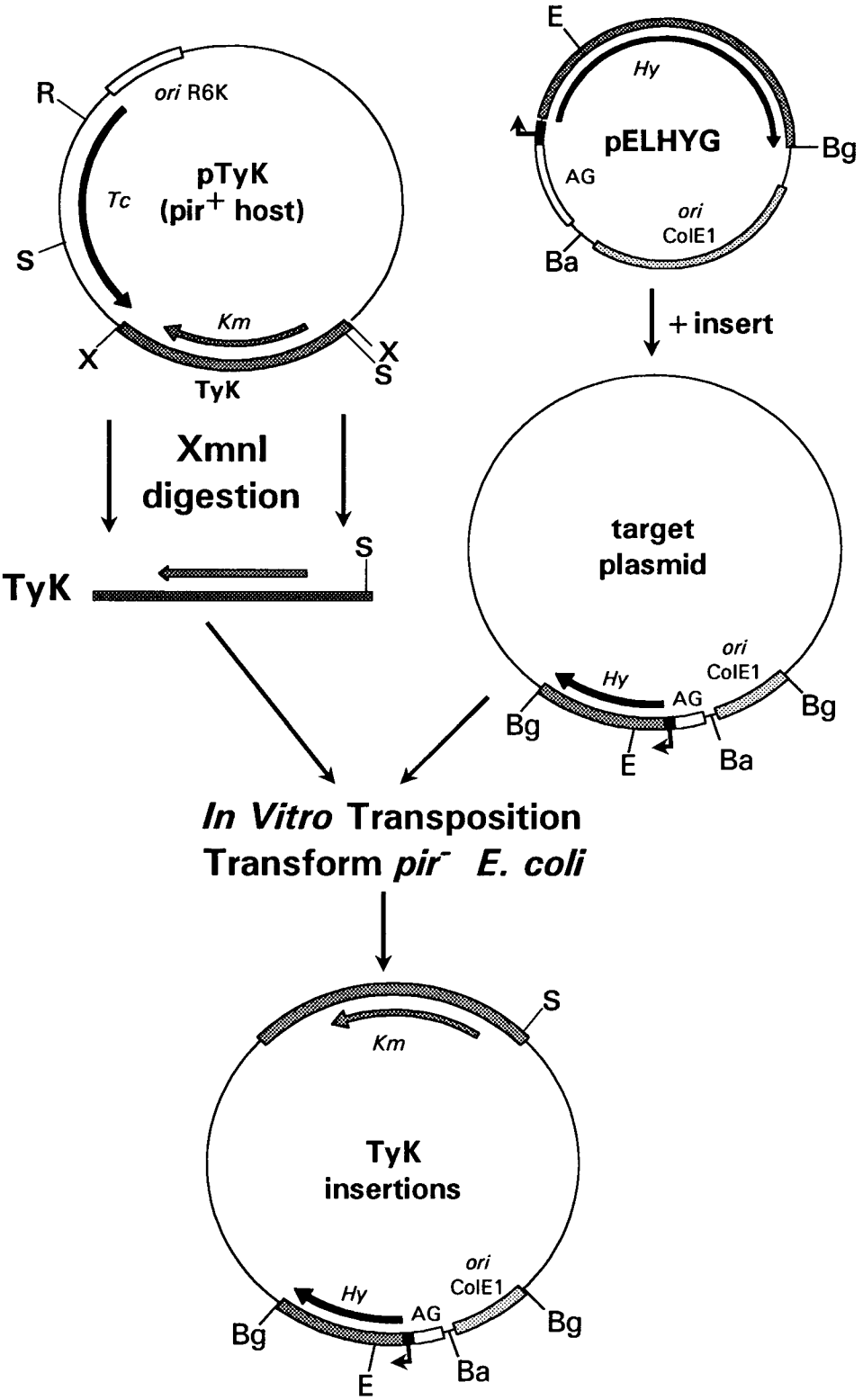

Fig. 2. Transposon mutagenesis using TyK. The suicide vector pTyK contains TyK (bounded by *Xmn*I sites), a tetracycline resistance marker (*Tc*) and an R6K replication origin (*ori* R6K) which requires the *pir* gene product (Kolter et al., 1978). Ba, Bg, E, R, S and X denote *Bam*HI, *Bgl*II, *Eco*RI, *Eco*RV, *Sal*I and *Xmn*I sites, respectively. TyK is released by digestion with *Xmn*I, purified, and used for in vitro transposition into the target plasmid. pELHYG contains a hygromycin resistance marker (*Hy*; shaded box), a *Leishmania trans*-splice acceptor site (AG; open box), a T7/*E. coli* promoter (bent arrow) and a minimal colE1 replication origin (ori ColE1; shaded box); pELSAT (not shown) contains a streptothricin resistance marker (*SAT*) replacing *Hy*. After insertion of passenger DNA (narrow line in target plasmid) and Ty1-mediated in vitro transposition, insertions are detected by transformation into *pir*$^{-}$ *E. coli*, which selects against contaminating donor plasmid and plating on drugs selecting for both transposon (kanamycin) and target (hygromycin in this example). Plasmids are not drawn to scale.

Table 1
In vitro transposition using TyK or TyK′GFP+

Target[a]	Size (kb)	Transposon	Total colonies[b]	Km^R colonies	Transposition efficiency[c]	% Single insertion[d]
pELSAT-A3[e]	10	TyK′GFP+	1.7×10^7	2200	1.3×10^{-4}	100% (15/15)
pBS-VIR1[f]	11	TyK	ND	176	ND	95% (167/176)
pSN7-1[g]	12	TyK	ND	5000	ND	98% (106/108)
pELHYG-H1[h]	14	TyK′GFP+	4.6×10^6	275	5.6×10^{-5}	94% (51/54)
pELHYG-H2[i]	15	TyK′GFP+	1.9×10^7	365	1.8×10^{-5}	90% (61/65)
pSN6-1.1[j]	21	TyK	(A) 3.1×10^7	450	1.3×10^{-5}	89% (16/18)
			(B) 1.6×10^8	340	2.1×10^{-6}	98% (54/55)
ID-2[k]	22	TyK	(A) ND	400	ND	86% (50/58)
			(B) 1.3×10^7	45	3.5×10^{-6}	100% (9/9)
cHP[l]	51	TyK	3.7×10^6	57	2.2×10^{-5}	77% (44/57)

[a]All passenger DNAs were from *L. major* genomic DNA.
[b]Transformants obtained by plating on ampicillin, hygromycin or nourseothricin as appropriate.
[c]Km^R colonies containing single insertions divided by total transformants.
[d]Clones without single insertions either contained multiple insertions or more complex rearrangements.
[e]An 8.5 kb *Sau*3A partial fragment inserted into the *Bam*HI site of pELSAT (strain B2765).
[f]An 8.5 kb *Bgl*II fragment inserted in the *Bam*HI site of pBluescript (strain B2150).
[g]Two 3 kb partial *Sau*3A fragments inserted in the *Bam*HI site of pSNBR (Callahan and Beverley, 1991; strain B3068).
[h]A 12 kb *Eco*RI fragment of the H region inserted into the *Bgl*II site of pELHYG (strain B3026).
[i]A 13 kb *Eco*RI fragment of the H region inserted into the *Bgl*II site of pELHYG (strain B3107).
[j]A 14 kb partial *Sau*3A fragment inserted in the *Bam*HI site of pSNBR (strain B3070).
[k]A 12 kb *L. major* fragment in cLHYG (Ryan et al., 1993) (strain B1520).
[l]A 41 kb genomic DNA fragment in cLHYG (strain B2856).
ND, not determined.
(A) and (B) refer to two independent experiments.

Table 2
TyK and TyK′GFP+ insertion sites into *Leishmania* DNA

GC content	Site sequence			
5/5	CGCGG	GGGGG		
4/5	TCCCG	GCAGG		
	GCTCC	GCGAC		
3/5	GAGTC	GGAAC	GTGTG	
	CTTGC	AGCTC	GATGG	
	ACGTC	GCAAG	GAGTG	
	GTCAC	GGGAA	GACAC	
	GCTCT	CCACT		
2/5	GAAGT	CACTT	GTATG	AGATC
	CGTAT	TACAC	GAAAG	CTGAA
	GAAAC	TCTTC	GATAG	AGTTC
	TTCGT	AAGAC	CTTTC	CTTAG
	CTTTG			
1/5	ACATA	ATACA	TATAG	GAAAT
	AATGT	TGATT	CTATA	GAAAT
	AAATC	AATCT	ATATG	AAAAC
	GATTT	GAATA	CAAAA	TTTAG
	TTTTC			
0/5	AAATA			

Underlining denotes sites obtained from sequencing both ends of the TyK insertion. Sequences were obtained from transpositions into pSN7-1, pBS-VIR1 and pELHYG-H1 (Table 1), or pSC1 (lab strain B3069), a 6 kb partial *Sau*3A fragment from *L. major* inserted into the *Bam*HI site of pSNBR (Callahan and Beverley, 1991).

pH 7.5, 150 mM $MgCl_2$, 100 mM KCl, 10 mM DTT); 5 μl 20% PEG 8000; 500 ng purified TyK or TyK′GFP+ donor fragment; 1–2 μg of target plasmid DNA; VLPs (0.5 RT units) and sterile water to 20 μl. Important control reactions lacked either the transposon DNA fragment, target plasmid, or VLPs. Reactions were mixed gently, incubated for 1 h at 30°C and terminated by the addition of 1 μl 0.5 M EDTA (pH 8.0), or 5 μl of a stop solution containing 5 mg/ml Proteinase K in 0.25 M EDTA, followed by incubation at 65°C for 20 min. DNAs were then extracted with phenol/chloroform/isoamyl alcohol (25:24:1), ethanol precipitated, and resuspended in 10 μl TE (10 mM Tris, 1 mM EDTA, pH 8.0). A sample of 2 μl were used for electrotransformation of *E. coli* strain DH10B. Aerosol-resistant pipette tips and freshly prepared reagents were used.

2.7. E. coli *electrotransformation*

Electrocompetent DH10B cells were prepared by a modification of previous protocols (Dower et al., 1988). Cells were grown at 37°C in 500 ml salt-free medium (20 g/l tryptone, 5 g/l yeast extract) to an absorbance of 0.6. Cultures were chilled to 0°C, and washed once with 500 ml water and once with 500 ml 10% glycerol at 0°C. Cells were resuspended in 1 ml ice-cold 10% glycerol, aliquoted, frozen in a dry ice/ethanol bath, and stored at −80°C. An aliquot of 35 μl was taken for each electroporation, using a BTX Electro Cell Manipulator at 12.5 kV/cm and 129Ω. Transformants recovered for 1 h in 1 ml SOC (Ausubel et al., 1990), and were plated on LB-agar plates containing kanamy-

cin (30–50 μg/ml) and either hygromycin (50 μg/ml, pELHYG targets), nourseothricin sulfate (50 μg/ml, pELSAT targets) or ampicillin (50 μg/ml, all other targets). Transformation efficiencies were determined by plating in the absence of kanamycin.

3. Results

3.1. Properties of TyK, TyK′GFP+ and pTyK

The features of the mini-Ty1 element TyK are shown in Fig. 1A. TyK contains the Tn903 kanamycin-resistance gene (*Km*), which we found to be superior to and more convenient than the trimethoprim-resistance marker in the original Ty1 donor plasmid pAT-2 (Devine and Boeke, 1994). TyK also contains several unique and/or rare restriction sites, including that for *I-Ppo*I, an intron-encoded endonuclease (Muscarella et al., 1990), which facilitate mapping and simplify the construction of more specialized artificial transposons. TyK′GFP+ (Fig. 1B) contains the coding region for the *Aequorea* GFP, modified to contain a S65T mutation and GC-rich codons in order to increase the fluorescence signal (Haas et al., 1996; Fig. 1B). In-frame insertion of TyK′GFP+ into protein coding genes is a convenient way to create fusions to this useful reporter protein (Fig. 1C).

TyK transposon fragments are prepared by digestion of the donor plasmid with *Xmn*I to expose the termini required for in vitro transposition (Fig. 1, Fig. 2). In early studies we found that trace contamination of the donor transposon fragment with parental plasmid frequently led to the recovery of "false" transpositions, which were colonies containing both the target and donor plasmids. We inserted TyK into a "suicide" vector containing a tetracycline-resistance marker and an R6K replication origin which functions only in the presence of the *pir* gene (Kolter et al., 1978; Fig. 2). By transforming in vitro transposition products into *pir*$^-$ *E. coli*, the donor pTyK plasmid was selected against (Fig. 2). This dramatically reduced the occurrence of false positives from values ranging up to 100% with early Ty1 donor plasmids, to less than 2% with pTyK-derived DNA (data not shown).

3.2. Preparation and quality of VLPs

Purified VLPs were obtained following induction of yeast strain JB224, which contains an integrated Ty1 element under the control of a GAL4 promoter (Boeke et al., 1985). VLPs were purified by centrifugation over linear sucrose gradients and identified by RT activity. Often, preparations contained a nuclease activity, as revealed by losses in transformation efficiency of the target DNA and/or its degradation, observed following gel electrophoresis; this activity could not be removed by additional rounds of purification (not shown). Co-sedimentation of VLPs with a Mg^{+2}-activated nuclease has been described (Braiterman and Boeke, 1994), and it may be an intrinsic component of the VLPs. Heterogeneity in VLP particle structure has been reported (Brookman et al., 1995) and we noticed that during purification, the ratio of nuclease to RT activity was not constant across the sucrose gradient fractions. In the VLP assay shown in Fig. 3, fractions 5 and 6 exhibited comparable RT activity; however, the frequency of transposition was 5-fold higher for fraction 6, due to the presence of nuclease activity in fraction 5. This empirical approach permitted identification of VLP fractions showing maximal transposition activity and was especially useful for transposition into large targets (cHP, 51 kb; Table 1).

3.3. In vitro transposition using TyK and TyK′GFP+

Table 1 lists the results of several in vitro transposition reactions in which either TyK or TyK′GFP+ were inserted into target plasmids ranging from 10 to 51 kb in size. Depending upon the target and VLP preparation, transposition efficiencies ranged from 10^{-4} to 10^{-6}, comparable to that reported previously with the mini-Ty1 element AT-2 (Devine and Boeke, 1994). Often, higher yields of transposition products were obtained

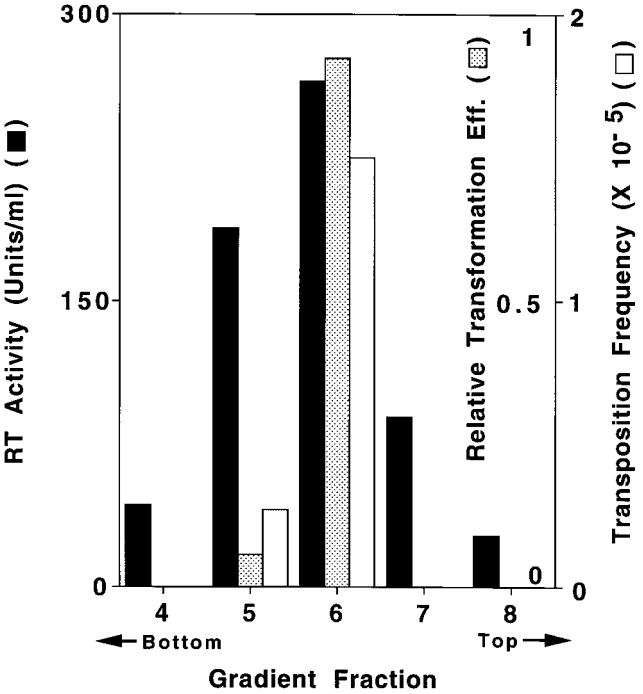

Fig. 3. Activity and quality of VLPs. Sucrose gradient fractions from a representative purification were assayed for VLP reverse transcriptase activity (black bars), transposition efficiency into the 15 kb target pSN6-1.1 (stippled bars) and transformation efficiency (a measure of nuclease activity; grey bars).

with smaller targets (Table 1), which probably reflects the lower transformation efficiency of large plasmid targets and their increased sensitivity to nuclease contamination. Typically, 90–100% of kanamycin-resistant colonies contained plasmids bearing a single insertion (determined by restriction digestion; Table 1), although this declined to 77% for cHP (51 kb). Clones without single insertions either contained multiple insertions or had undergone more complex rearrangements, as noted previously (Devine and Boeke, 1994).

3.4. Location and sequence of TyK insertion sites

Analysis of numerous transpositions into several targets showed that the orientation and distribution of TyK insertions was seemingly random with respect to site and orientation (Fig. 4). No statistical difference was found between insertion into the coding versus non-coding strand. The insertion site was determined for 31 TyK and 24 TyK′GFP+ transpositions (Table 2). Where both flanking sequences were determined, the TyK insertions were flanked by the expected 5 bp duplication of target sequences (Boeke, 1989). No obvious target site bias was evident in the events analysed here, although there was a preference for AT-rich sites (Fig. 4C). The average base composition for the TyK and TyK′GFP+ insertion sites analysed here was 43% GC (45 and 42%, respectively), which may be compared with the base composition of *S. cerevisiae* (39% GC; Dujon, 1996) and *L. major* (61% GC, Chance, 1972). Effectively, TyK and TyK′GFP+ insert randomly within the target DNA.

3.5. Preferential recovery of transpositions into passenger DNA

Many popular cloning vectors contain non-essential DNA, thereby allowing recovery of transpositions into vector rather than passenger DNA sequences. For example, approx. 15% of the TyK insertions into pBS-VIR1 (Fig. 4A) occurred in the pBluescript vector. To diminish the recovery of background transpositions, we constructed two specialized target vectors, pELHYG and pELSAT (Fig. 2, Fig. 4B), which are only 2128 and 1650 bp in size, respectively. They contain a minimal ColE1 replication origin, a small *Leishmania* splice acceptor, a T7/*E. coli* promoter and hygromycin (*Hy*) or streptothricin/nourseothricin (*SAT*) resistance markers. Transposition into all but the *Leishmania* splice acceptor should functionally inactivate either pEL vector. This was confirmed with pELHYG-H1 (Table 1), which consists of a 12 kb *L. major Eco*RI fragment; none of the 45 TyK′GFP+ insertions analysed occurred in pELHYG vector sequences (Fig. 4B).

4. Discussion

We have tested several modifications of the in vitro Ty1 transposition system which increase its usefulness in genomic applications. The TyK-based transposons contain several unique and rare restriction sites (including one for the *I-Ppo*I intron-encoded nuclease) which facilitate mapping. They are carried by a suicide donor plasmid which significantly reduces the incidence of "false positive" events, and we have developed a transposition target vector where insertions into the passenger DNA are preferentially recovered. Together, these modifications comprise a highly efficient system for the creation of large transposon libraries and simplify the application of Ty1-based transposition strategies to a variety of functional genetic analyses.

An important concern for transposon mutagenesis is the randomness of insertion, as many transposons show a biased site preference. Our data show that transposition using TyK and TyK′GFP+ is effectively random with respect to *Leishmania* or vector DNA, strand, or coding versus non-coding regions. This is particularly notable as the base composition of targets such as *L. major* DNA differs considerably from that of the yeast host of Ty1, *S. cerevisiae*. Interestingly, native Ty1 elements are 5.9 kb in size, significantly larger than either TyK (1.2 kb) or TyK′GFP+ (2 kb). Modified Ty1 elements which are larger than TyK or TyK′GFP+ may thus also function in vitro.

The randomness of insertion, efficiency of transposition and relatively low background attainable with this system has permitted the generation of large, random transposon insertion libraries for many target plasmids in our laboratory. Several of these have been used successfully for DNA sequencing within plasmids and cosmids. For several cosmids bearing *Leishmania* genes involved in drug resistance or the parasite infectious cycle, transposon mutagenesis has been used to identify the relevant gene by insertional inactivation.

Insertion of the specialized transposon TyK′GFP into protein-coding genes could be used to generate fusions to GFP, because TyK′GFP+ was shown to transpose as predicted (Fig. 4B, Table 2). GFP has many advantages as a reporter protein, including the ease of screening by microscopy or fluorescence-activated cell sorter technology (Chalfie et al., 1994; Ha et al., 1996). One application of this vector is in the identification of cells bearing GFP fusions showing interesting patterns of cellular localization and/or regulation. The potential of this approach is illustrated by a recent report employing yeast expressing a modified inducible GFP-containing Ty1 element, which yielded an interesting panel of protein fusions exhibiting nuclear localization (Sawin and Nurse, 1996). The TyK′GFP+ system described here would permit a similar study to be performed in any organism, by generating libraries of TyK′GFP+

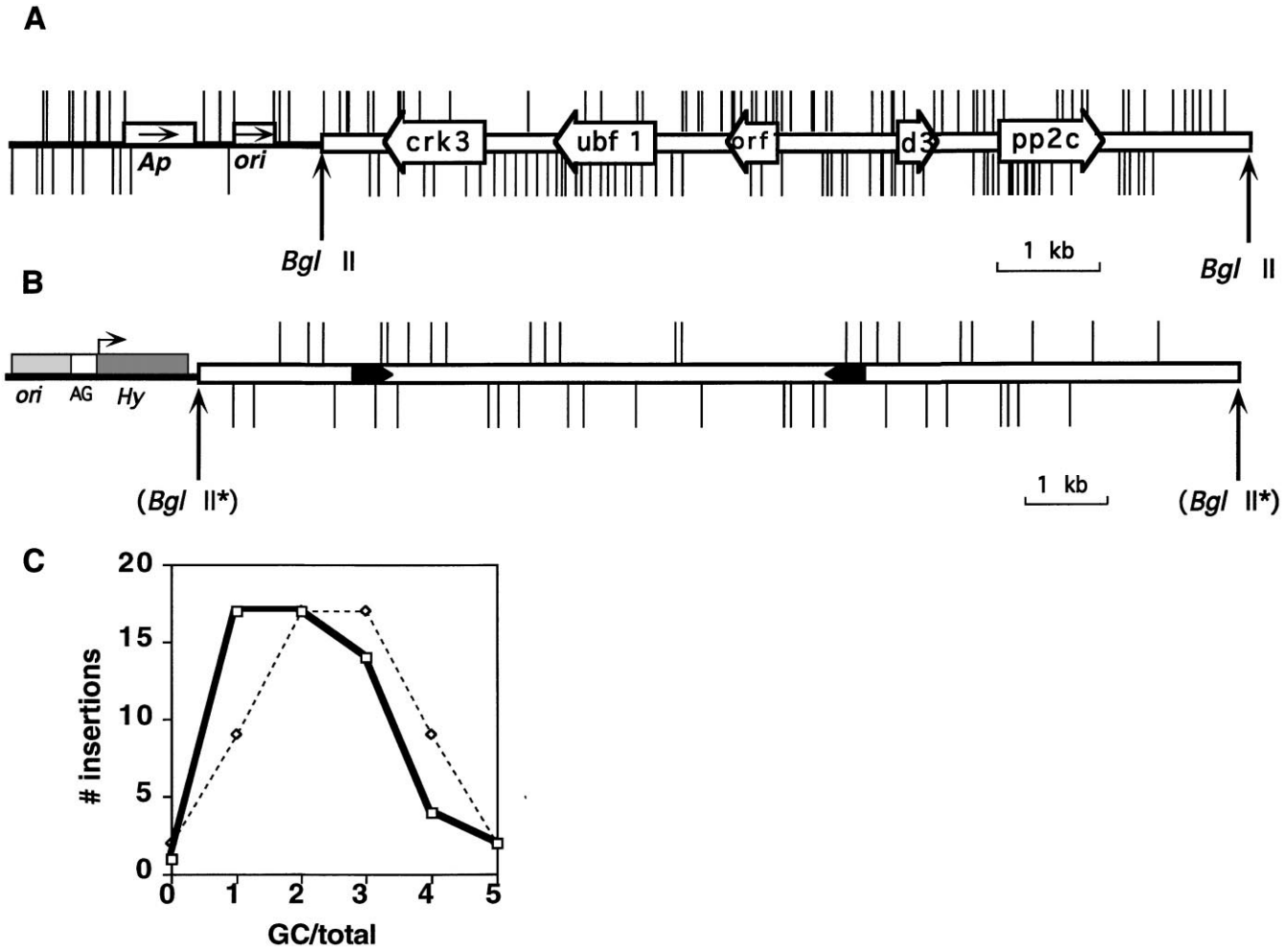

Fig. 4. Randomness of TyK insertion. A. Insertion into pBS-VIR1. The pBLUESCRIPT vector is shown by a dark line, the 8.5 kb insert of *L. major* genomic DNA by a thicker open line and known *Leishmania* or vector genes by larger open arrows or boxes. TyK insertions with *Km* in the forward or reverse orientation are shown by marks above or below the map, respectively. B. Insertion into pELHYG-H1. The pELHYG vector is denoted by a line; its elements are described in Fig. 2. The 12 kb insert is shown by a thick open line; (*Bgl*II*) denotes the site in pELHYG used to insert the genomic DNA, destroyed after end repair. Blackened arrows within the 12 kb insert denote inverted repeats from the *L. major* H region (unpublished data). TyK′GFP+ insertions and orientation are denoted as in panel A. C. Frequency of TyK and TyK′GFP+ insertion sites (Table 2) as a function of GC content. The dashed line shows the expectation for a random binomial distribution.

insertions into random genomic DNAs carried by appropriate shuttle vectors.

In the future, we envisage the use of TyK-based vectors for a variety of other applications. For example, mobile drug resistance cassettes can readily be designed which, after insertion into the desired target gene, could be used to generate selectable markers required for homologous gene replacement procedures. The ease and features of the TyK system should facilitate similar studies in target DNAs from any organism.

Acknowledgement

We thank G. Fink for yeast strain JB224, E.I. Vivas and M. Dziejman for plasmids used in construction of pTyK, S. Chiang for DH5$\alpha\lambda$*pir*, S. Cilmi for pSC1 and Lynne Garrity for ID-2; C. Andre, J. Boeke, J. Brookman and members of K. Struhl's laboratory for advice and discussions; and F. Gueiros-Filho for comments on this manuscript. This work was supported by NRSA Fellowships HG00092 (L.A.G.) and AI09382 (Y.W.), a NSERC Canada Fellowship (J.B.M.) and NIH grants to D.E.D. and S.M.B. L.R.O.T. is a PEW Latin American Fellow.

References

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K.S., 1990. Current Protocols in Molecular Biology. Wiley, New York, NY.

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.
Boeke, J.D., 1989. Transposable elements in *Saccharomyces cerevisiae*. In: Berg, D.E., Howe, M.M. (Eds.), Mobile DNA. American Society for Microbiology, Washington, DC, pp. 335–374.
Boeke, J.D., Garfinkel, D.J., Styles, C.A., Fink, G.R., 1985. Ty elements transpose through an RNA intermediate. Cell 40, 491–500.
Braiterman, L.T., Boeke, J.D., 1994. In vitro integration of retrotransposon Ty1: a direct physical assay. Mol. Cell. Biol. 14, 5719–5730.
Braiterman, L.T., Monokian, G.M., Eichinger, D.J., Merbs, S.L., Gabriel, A., Boeke, J.D., 1994. In-frame linker insertion mutagenesis of yeast transposon Ty1: phenotypic analysis. Gene 139, 19–26.
Brookman, J.L., Stott, A.J., Chessman, P.J., Burns, N.R., Adams, S.E., Kingsman, A.J., Gull, K., 1995. An immunological analysis of Ty1 virus-like particle structure. Virology 207, 59–67.
Burns, N., Grimwade, B., Ross-Macdonald, P.B., E, Y., Choi, K.F., Roeder, G.S., Snyder, M., 1994. Large-scale analysis of gene expression, protein localization, and gene disruption in *Saccharomyces cerevisiae*. Genes Dev. 8, 1087–1105.
Callahan, H.L., Beverley, S.M., 1991. Heavy metal resistance: a new role for P-glycoproteins in *Leishmania*. J. Biol. Chem. 266, 18427–18430.
Casadaban, M.J., Martinez-Arias, A., Shapira, S.D., Chow, J., 1983. *β*-Galactosidase gene fusions for analyzing gene expression in *Escherichia coli* and yeast. Methods Enzymol. 100, 293–308.
Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W., Prasher, D.C., 1994. Green fluorescent protein as a marker for gene expression. Science 263, 802–805.
Chance, M.L., 1972. DNA base composition differences between species of *Leishmania*. Trans. R. Soc. Trop. Med. Hyg. 66, 352
Chun, K.T., Goebl, M.G., 1996. The identification of transposon-tagged mutations in essential genes that affect cell morphology in *Saccharomyces cerevisiae*. Genetics 142, 39–50.
Devine, S.E., Boeke, J.D., 1994. Efficient integration of artificial transposons into plasmid targets in vitro: a useful tool for DNA mapping, sequencing, and genetic analysis. Nucleic Acids Res. 22, 3765–3772.
Dower, W.J., Miller, J.F., Ragsdale, C.W., 1988. High efficiency transformation of *E. coli* by high voltage electroporation. Nucleic Acids Res. 16, 6127–6145.
Dujon, B., 1996. The yeast genome project: what did we learn? Trends Genet. 12, 263–270.
Eichinger, D.J., Boeke, J.D., 1988. The DNA intermediate in yeast Ty1 element transposition copurifies with virus-like particles: cell-free Ty1 transposition. Cell 54, 955–966.
Garfinkel, D.J., Boeke, J.D., Fink, G.R., 1985. Ty element transposition: reverse transcriptase and virus-like particles. Cell 42, 507–517.
Ha, D.S., Schwarz, J.K., Turco, S.J., Beverley, S.M., 1996. Use of the green fluorescent protein as a marker in transfected *Leishmania*. Mol. Biochem. Parasitol. 77, 57–64.
Haas, J., Park, E.C., Seed, B., 1996. Codon usage limitation in the expression of HIV-1 envelope glycoprotein. Curr. Biol. 6, 315–324.
Herrero, M., de Lorenzo, V., Timmis, K.N., 1990. Transposon vectors containing non-antibiotic resistance selection markers for cloning and stable chromosomal insertion of foreign genes in Gram-negative bacteria. J. Bacteriol. 172, 6557–6567.
Ji, H., Moore, D.P., Blomberg, M.A., Braiterman, L.T., Voytas, D.F., Natsoulis, G., Boeke, J.D., 1993. Hotspots for unselected Ty1 transposition events on yeast chromosome III are near tRNA genes and LTR sequences. Cell 73, 1007–1018.
Johnston, M., 1996. Towards a complete understanding of how a simple eukaryotic cell works. Trends Genet. 242, 242–243.
Kapler, G.M., Zhang, K., Beverley, S.M., 1990. Nuclease mapping and DNA sequence analysis of transcripts from the dihydrofolate reductase-thymidylate synthase (R) region of *Leishmania major*. Nucleic Acids Res. 18, 6399–6408.
Kolter, R., Inuzuka, M., Helinski, D.R., 1978. Trans-complementation-dependent replication of a low molecular weight origin fragment from plasmid R6K. Cell 15, 1199–1208.
Morgan, B.A., Conlon, F.L., Manzanares, M., Millar, J.B., Kanuga, N., Sharpe, J., Krumlauf, R., Smith, J.C., Sedgwick, S.G., 1996. Transposon tools for recombinant DNA manipulation: characterization of transcriptional regulators from yeast, *Xenopus* and mouse. Proc. Natl. Acad. Sci. USA 93, 2801–2806.
Muscarella, D.E., Ellison, E.L., Ruoff, B.M., Vogt, V.M., 1990. Characterization of *I-Ppo*I, an intron-encoded endonuclease that mediates homing of a group I intron in the ribosomal DNA of *Physarum polycephalum*. Mol. Cell. Biol. 10, 3386–3396.
Pennisi, E., 1996. From genes to genome biology. Science 272, 1736–1738.
Ryan, K.A., Dasgupta, S., Beverley, S.M., 1993. Shuttle cosmid vectors for the trypanosomatid parasite *Leishmania*. Gene 131, 145–150.
Sawin, S.E., Nurse, P., 1996. Identification of fission yeast nuclear markers using random polypeptide fusions with green flourescent protein. Proc. Natl. Acad. Sci. USA 93, 15146–15151.
Seifert, H.S., Chen, E.Y., So, M., Heffron, F., 1986. Shuttle mutagenesis: a method of transposon mutagenesis for *Saccharomyces cerevisae*. Proc. Natl. Acad. Sci. USA 83, 735–739.
Sherratt, D.J., 1995. Mobile Genetic Elements. IRL Press, Oxford.
Silhavy, T.J., Berman, M.L., Enquist, L.W., 1984. Experiments with gene fusions. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.