

piggyBac internal sequences are necessary for efficient transformation of target genomes

X. Li*, R. A. Harrell†, A. M. Handler‡, T. Beam‡,
K. Hennessy* and M. J. Fraser Jr*

*Department of Biological Sciences, and Center for Tropical Diseases Research and Training, University of Notre Dame, Notre Dame, IN, USA; †USDA-ARS, Center for Medical, Agricultural & Veterinary Entomology, Gainesville, FL, USA; and ‡Department of Biological Sciences, University of Saint Francis, Fort Wayne, IN, USA

Abstract

A previously reported *piggyBac* minimal sequence cartridge, which is capable of efficient transposition in embryo interplasmid transposition assays, failed to produce transformants at a significant frequency in *Drosophila melanogaster* compared with full-length or less extensive internal deletion constructs. We have re-examined the importance of these internal domain (ID) sequences for germline transformation using a PCR strategy that effectively adds increasing lengths of ID sequences to each terminus. A series of these *piggyBac* ID synthetic deletion plasmids containing the 3xP3-ECFP marker gene are compared for germline transformation of *D. melanogaster*. Our analyses identify a minimal sequence configuration that is sufficient for movement of *piggyBac* vectored sequences from plasmids into the insect genome. Southern hybridizations confirm the presence of the *piggyBac* transposon sequences, and insertion site analyses confirm these integrations target TTAA sites. The results verify that ID sequences adjacent to the 5' and 3' terminal repeat domains are crucial for effective germline transformation with *piggyBac* even though they are not required for excision or interplasmid transposition. Using this information we reconstructed an inverted repeat cartridge, ITR1.1k, and a minimal *piggyBac* transposon vector, pXL-BacII-ECFP, each of which contains these identified ID sequences in addition to the terminal repeat configura-

tion previously described as essential for mobility. We confirm in independent experiments that these new minimal constructs yield transformation frequencies similar to the control *piggyBac* vector. Sequencing analyses of our constructs verify the position and the source of a point mutation within the 3' internal repeat sequence of our vectors that has no apparent effect on transformation efficiency.

Keywords: *piggyBac*, transposon, transformation.

Introduction

The *piggyBac* transposable element was originally isolated from the Lepidopteran cell line TN-368 as a gene-disrupting insertion within spontaneously arising baculovirus plaque morphology mutants (Fraser *et al.*, 1985; Cary *et al.*, 1989). *piggyBac* encodes a transposase function that operates using a precise cut-and-paste mechanism, targeting and duplicating TTA tetranucleotide sequences upon insertion, and reforming a single target site upon excision (Cary *et al.*, 1989; Fraser *et al.*, 1995; Elick *et al.*, 1996a,b; Fraser *et al.*, 1996). Structurally, the *piggyBac* element is a 2472 bp short inverted terminal repeat transposable element composed of a 2374 bp transposase-encoding internal domain (ID) flanked by asymmetric terminal repeat domains (TRD). The 5' TRD consists of a 19 bp internal repeat sequence separated from the 13 bp terminal repeat sequence by 3 bp, while the 3' TRD has a 31 bp spacer separating the internal and terminal repeat sequences (Cary *et al.*, 1989; Fraser *et al.*, 1995; Elick *et al.*, 1996b).

piggyBac-derived vectors are capable of mediating germline transformation in a wide variety of insect species including *Ceratitis capitata* (Handler *et al.*, 1998), *Drosophila melanogaster* (Handler & Harrell, 1999; Berghammer *et al.*, 1999; Horn & Wimmer, 2000), *Tribolium castaneum* (Berghammer *et al.*, 1999), *Bactrocera dorsalis* (Handler & McCombs, 2000), *Bombyx mori* (Tamura *et al.*, 2000), *Pectinophora gossypiella* (Peloquin *et al.*, 2000), *Aedes aegypti* (Kokoza *et al.*, 2001; Lobo *et al.*, 2002), *Anastrepha suspensa* (Handler & Harrell 2001b), *Anopheles gambiae* (Grossman *et al.*, 2001), *Musca domestica* (Hediger *et al.*, 2001), *Anopheles albimanus* (Perera *et al.*, 2002), *Anopheles stephensi* (Nolan *et al.*, 2002) and *Lucilia cuprina* (Heinrich *et al.*, 2002). Most

doi: 10.1111/j.1365-2583.2004.00525.x

Received 10 September 2003; accepted after revision 29 July 2004. Correspondence: Dr Malcolm Fraser, Department of Biological Sciences, and Center for Tropical Diseases Research and Training, University of Notre Dame, Notre Dame, IN 46556, USA. Tel.: +1 574 631 6209; fax: +1 574 631 7413; e-mail: Malcolm.J.Fraser.1@nd.edu

recently, a non-insect species, the planarian *Girardia tigrina*, has also been successfully transformed using *piggyBac*, suggesting a wider range of utility for this element (Gonzalez-Estevez *et al.*, 2003). A common feature of all of these transformations has been the use of either a complete or a partial internally deleted (0.8 kb deletion in the transposase coding sequence; Handler & Harrell, 1999; Horn & Wimmer, 2000) *piggyBac* transposon as the transformation vector.

Using plasmid-based mobility assays, Li *et al.* (2001b) had previously reported that the 5' and 3' *piggyBac* TRD sequences were sufficient for excision or transposition in injected insect embryos. However, personal communications from several labs attempting to utilize the previously identified minimal *piggyBac* sequences in transformation experiments indicated significant difficulty in obtaining germline transformants (John Peloquin, UC Riverside; Richard Beeman, Kansas State University). Our own attempts to transform *D. melanogaster* with one of our minimal *piggyBac* constructs also yielded an extremely low frequency of 0.6% transformants, significantly lower than the previously reported 26% transformation frequency for *piggyBac* in this species (Handler & Harrell, 1999).

These results imply some *piggyBac* internal sequences essential for efficient germline transformation of the insect genome are absent in our previous internal deletion constructs, even though they are apparently not required for interplasmid transposition. Similar observations have been reported for other Class II transposons such as the *Mos1 mariner* element. *Mariner* transposons lacking any internal sequences are inactive in transposition assays, while retention of a few additional nucleotides of the internal sequence restores full activity (Tosi & Beverley, 2000). However, efficient germline transformation also requires the presence and proper spacing of multiple sequences internal to the element as well as the inverted repeats (Lohe & Hartl, 2001; Lozovsky *et al.*, 2002).

In this report, we identify ID sequences adjacent to the 5' and 3' TRD of the *piggyBac* transposon that contribute to a high frequency of germline transformants in *D. melanogaster*. We analyse a series of PCR synthesized deletion vectors constructed with the 3xP3-ECFP gene as a transformation marker (Horn & Wimmer, 2000). These vectors define a minimal ID configuration that is sufficient to effect efficient germline transformation of *D. melanogaster*. Using this information we construct a new ITR cartridge, called ITR1.1k, and verify its utility in converting an existing plasmid into a mobilizable *piggyBac* vector that enables efficient germline transformation. We also construct a transposon-based cloning vector, pXL-BacII, for insertion of sequences within a minimal *piggyBac* transposon and verify its capabilities in germline transformations. Sequence analysis of these constructs reveals an introduced point mutation within the internal repeat sequence of the 3' TRD that has no apparent effect on mobility of the element.

Results

Transformation experiments with synthetic deletion constructs

Initial attempts to transform *D. melanogaster* with plasmids having only TRD sequences as specified in our previous report (Li *et al.*, 2001b) yielded transformation frequencies far less than full-length *piggyBac* constructs. The p(PZ)-BacEYFP construct contains the ITR cartridge of Li *et al.* (2001b) composed of the 5' and 3' TRD and the spacer sequence, while the pBS-pBac/DsRed retains only 2 bp of 5' ID and 36 bp of 3' ID sequences in addition to the 5' and 3' TRD. Neither of these constructs was able to generate germline transformants at the frequencies previously reported for full-length vectors (Handler & Harrell, 1999) or the less extensive internal deletion construct pBac{3xP3-ECFPafm} (Horn & Wimmer, 2000). This forced us to re-examine the potential involvement of *piggyBac* ID sequences in generating germline transformations.

We examined the requirements for TRD adjacent ID sequences of the *piggyBac* transposon using a synthesized cartridge strategy based upon construction of the previously reported ITR cartridge (Li *et al.*, 2001b) rather than digesting with an endonuclease and selecting clones representing an internal deletion series. Each of the *piggyBac* synthetic internal deletion plasmids was formed from the pIAO-P/L-589 plasmid (Li *et al.*, 2001b) by PCR amplification across the facing TRDs and spacer sequences with primers that recognize 5' or 3' ID sequences adjacent to the respective TRDs (Fig. 1). The fragments generated were cloned into a pBSII-3xP3-ECFP plasmid and sequenced (see Experimental procedures).

Each of the synthetic deletion series plasmids and the control plasmid, pBac{3xP3-ECFPafm}, were co-injected with the *hsp70*-regulated transposase helper into *w*¹¹¹⁸ embryos, with surviving adults backcrossed, and G1 adult progeny screened for fluorescence. Positive transformants exhibited fluorescent eyes with cyan and green filter sets but not with the yellow filter set. Transformation frequencies from all injections are listed in Table 1. The p(PZ)-Bac-EYFP plasmid, which was constructed using the ITR cartridge previously described (Li *et al.*, 2001b), yielded a relatively low transformation frequency of 0.6% compared with the control plasmid, pBac{3xP3-ECFPafm}, frequency of 12.9% (Table 1).

Eight of the eleven synthetic ID deletion plasmids yielded positive transformants at an acceptable frequency compared with the control. The 5' ID deletion constructs pBSII-ECFP-R1/L5, pBSII-ECFP-R2/L5, pBSII-ECFP-R3/L5 and pBSII-ECFP-R4/L5 had variable deletions of the *piggyBac* 5' ID, retaining sequences from 66 bp (nucleotides 36–101; GenBank accession number AR307779) to 542 bp (nucleotides 36–567) of the *piggyBac* sequence. Each of these 5' ID deletions yielded ECFP-positive germline transformants at frequencies from 8.9% ($\pm 1.0\%$) to 15.0% ($\pm 0.6\%$)

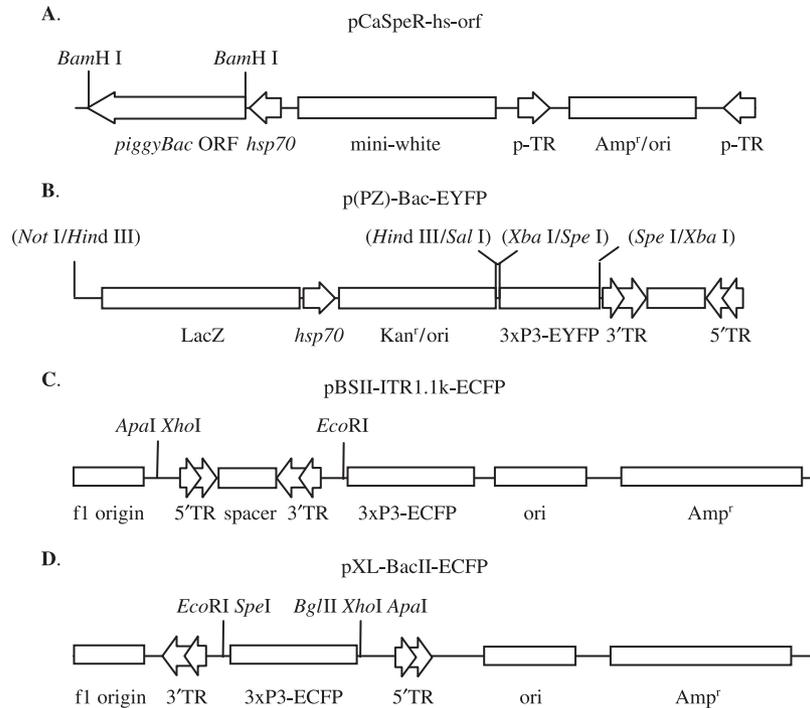


Figure 1. Construction of important plasmids developed in this study. (A) Diagram of the pCaSpeR-hs-orf helper used for the transformation assays. The *piggyBac* ORF *Bam*HI cassette was cloned as a PCR product into the *Bam*HI site of the pCaSpeR-hs adjacent to the *hsp70* promoter. (B) Diagram of the p(PZ)-Bac-EYFP construct demonstrating the inefficiency of the ITR cartridge (Li *et al.*, 2001b) for transformation. A 7 kb *Hind*III fragment containing *LacZ*, *hsp70* and *Kan*/ori sequences was excised from plasmid p(PZ) (Rubin & Spradling, 1983), and ligated to form a p(PZ)-7kb intermediate plasmid. The ITR cartridge was excised from pBSII-ITR (Li *et al.*, 2001b) using *Not*I and *Sal*I, blunt ended and inserted into the blunt-ended *Hind*III site of the p(PZ)-7kb plasmid. A 3xP3-EYFP *Spe*I fragment excised from pBac{3xP3-EYFPafm} (Horn & Wimmer, 2000) was then inserted into the *Xba*I site to form p(PZ)-Bac-EYFP. (C) Diagram of the pBSII-ITR1.1k-ECFP minimal *piggyBac* vector constructed by PCR amplification from the pAO-P/L 589 plasmid (Li *et al.*, 2001b), which contains a minimal *piggyBac* cartridge with inverted 5' and 3' TRDs separated by a 589 bp λ DNA spacer sequence, and incorporate additional subterminal ID sequences necessary for efficient transformation. This construct is tagged by the addition of the 3xP3-ECFP marker gene excised as a *Spe*I fragment from the plasmid pBac{3xP3-ECFPafm} (Horn & Wimmer, 2000). (D) Diagram of the *piggyBac* minimal vector pXL-BacII-ECFP, constructed from the pBSII-ITR1.1k plasmid essentially as previously described (Li *et al.*, 2001b), with the addition of the 3xP3-ECFP *Spe*I fragment from pBac{3xP3-ECFPafm}.

(Table 1) when paired with 1 kb of the 3' ID sequence (nucleotides 1454–2409). These results demonstrated a minimal sequence of no more than 66 bp of the 5' ID is essential for effective germline transposition.

The R4 minimum 5' ID sequence primer was then used in combination with a series of 3' ID deletion primers to generate the constructs pBSII-ECFP-R4/L4, pBSII-ECFP-R4/L3, pBSII-ECFP-R4/L2 and pBSII-ECFP-R4/L1. Of these four constructs, only pBSII-ECFP-R4/L1, which represented the greatest deletion of 3' ID sequence (2284–2409 of the *piggyBac* sequence), failed to yield transformants. Again, frequencies for the constructs that yielded positive transformants compared favourably with the control (Table 1). We therefore deduced that the minimal 3' ID sequence requirement for efficient germline transformation was between 125 bp (L1) and 378 bp (L2) of the 3' TRD adjacent ID sequence.

Construction of the ITR1.1k minimal sequence *piggyBac* cartridge

To construct a minimal sequence cartridge using the information gained from the synthetic deletion analysis we first

assembled combinations of 5' and 3' minimal sequences and tested their transformation capabilities. The pBSII-ECFP-R-TR/L construct is composed of a 35 bp 5' TRD lacking any 5' ID sequence, coupled to a fragment containing the 63 bp 3' TRD and 172 bp of the adjacent 3' ID sequence. This combination did not yield any transformants, confirming the necessity for having 5' ID sequences in combination with 3' ID sequences for efficient transformation.

Unexpectedly, addition of 66 bp of the 5' ID sequences to the 5' TRD sequences in the construct pBSII-ECFP-R1/L was not sufficient to recover transformation capacity when paired with the 172 bp 3' ID sequences, even though the lower limit of essential 5' ID sequences had been previously defined as 66 bp using pBSII-ECFP-R1/L5 (Table 2). Increasing the 5' ID sequences to 276 bp in the pBSII-ITR1.1k-ECFP plasmid recovered the full transformation capability when paired with the 172 bp 3' ID sequence (Table 2). The minimal operational sequence requirement for 5' ID sequences is therefore between 276 and 66 bp when coupled to a minimal 3' ID sequence of 172 bp.

Table 1. Transformation of *Drosophila melanogaster*

Plasmid	Experiment	Embryos injected	Embryos hatched	Adults mated	Transformed lines	Frequency	Overall frequency	SD	SE
p(PZ)-Bac-EYFP	1	920	136	55	1	1.8%	0.6%	1.0%	±0.6%
	2	910	120	56	0	0.0%			
	3	900	120	55	0	0.0%			
pBSII-ECFP-R1/L5	1	350	86	21	2	9.5%	8.9%	1.8%	±1.0%
	2	280	70	16	1	6.3%			
	3	360	84	33	3	9.1%			
pBSII-ECFP-R2/L5	1	320	37	11	1	9.1%	12.5%	7.7%	±5.4%
	2	300	38	5	1	20.0%			
pBSII-ECFP-R3/L5	1	220	39	7	1	14.3%	15.0%	0.8%	±0.6%
	2	430	88	13	2	15.4%			
pBSII-ECFP-R4/L5	1	220	59	12	1	8.3%	12.9%	5.3%	±3.7%
	2	510	123	19	3	15.8%			
pBSII-ECFP-R4/L4	1	340	108	21	1	4.8%	10.7%	16.8%	±11.9%
	2	330	61	7	2	28.6%			
pBSII-ECFP-R4/L3	1	220	39	9	0	0.0%	9.7%	12.9%	±7.4%
	2	240	53	14	1	7.1%			
	3	250	55	8	2	25.0%			
pBSII-ECFP-R4/L2	1	320	43	11	1	9.1%	10.8%	4.9%	±3.5%
	2	530	148	25	4	16.0%			
pBSII-ECFP-R4/L1	1	350	89	30	0	0.0%	0.0%	N/A	N/A
	2	160	33	16	0	0.0%			
	3	330	78	25	0	0.0%			
	4	150	31	15	0	0.0%			
pBSII-ECFP-R-TR/L	1	280	73	31	0	0.0%	0.0%	N/A	N/A
	2	330	96	40	0	0.0%			
pBSII-ECFP-R1/L	1	220	63	19	0	0.0%	0.0%	N/A	N/A
	2	290	80	23	0	0.0%			
	3	330	104	27	0	0.0%			
pBac{3xP3-ECFPafm}	1	300	45	14	2	14.3%	12.9%	1.8%	±1.3%
	2	350	59	17	2	11.8%			
pBSII-ITR1.1k-ECFP	1	530	128	36	5	13.9%	13.9%	N/A	N/A
pXL-BacII-ECFP	1	500	80	14	3	21.4%	22.2%	0.9%	±0.6%
	2	520	101	22	5	22.7%			
pBSII-ITR1.1k-ECFP*	1	515	120	22	8	36.4%	36.4%	N/A	N/A
pXL-BacII-ECFP*	1	533	199	88	22	25.0%	25.0%	N/A	N/A

*These injections were done independently (Handler lab) using a 0.4 : 0.2 µg/µl vector–helper concentration ratio of DNA. Statistical analysis of the data revealed no significant difference between frequencies obtained with any of the synthetic deletion mutants that yielded detectable numbers of transformants and the control plasmid pBac{3xP3-ECFPafm}. The assay cannot be interpreted to represent relative efficiencies of transformation among these constructs, but only whether a particular construct was able to generate transformants at a detectable frequency with the number of surviving injected flies analysed.

Two independent verifications of the pBSII-ITR1.1 k-ECFP plasmid transforming capabilities were conducted for transformation of *D. melanogaster*. These transformation experiments resulted in calculated frequencies of 13.9% (Fig. 1) and 36% (Table 1). We attribute the rather large discrepancy in frequencies to differences in injection protocols between labs. Unless otherwise indicated, the transformation frequencies presented in Table 1 and Fig. 1 were obtained with injections of 0.6 : 0.6 µg/µl vector–helper concentration ratios. The increased efficiency of transformation for pBSII-ITR1.1k-ECFP observed in the second independent trial seems to be

related to a decreased vector–helper concentration in *D. melanogaster*.

Five recovered pBSII-ITR1.1k-ECFP transformed strains were used to perform genetic mapping to identify their chromosome locations. Several of the strains had insertions on the second and third chromosomes (including strain 1), while strain 3 had an insertion on the X chromosome. We chose strains 1 and 3 for further analyses.

Direct PCR analysis of integrations

Genomic DNAs from each of the transformed strains obtained with the synthetic deletion constructs in Fig. 1, as

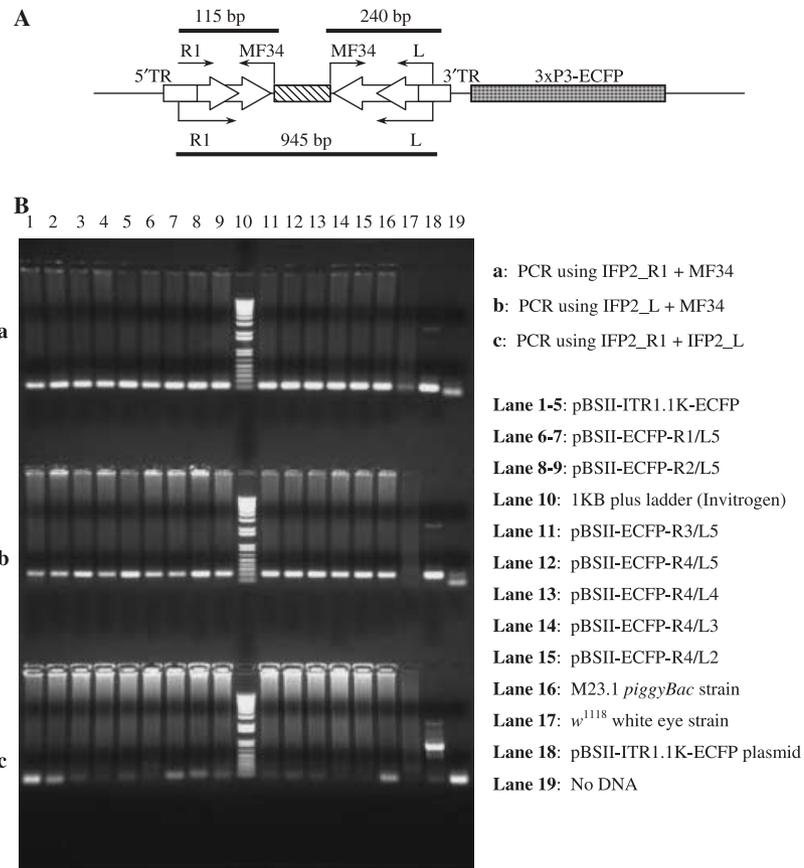


Figure 2. Direct PCR analysis of transformed flies. (A) Diagram of a generalized synthetic deletion construct indicating the position of primers and expected fragment. Three sets of PCR primers were used to verify the *piggyBac* insertion. The first primer set (IFP2_R1 + MF34) detects the 5' terminal region (115bp), the second primer set (IFP2_L + MF34) detects the 3' terminal region (240bp) and the third primer set (IFP2_R1 + IFP2_L) detects the presence of the external spacer sequence (945bp). (B) Direct PCR results. (a) The first primer set yields a 115 bp fragment in all transformed strains confirming the 5' terminal region. A less effectively amplified 115 bp fragment is also evident in the w^{1118} strain, reflecting the probable presence of *piggyBac*-like sequences in the *D. melanogaster* genome. (b) The second primer set yields the expected 240 bp fragment in all transformed strains, confirming the 3' terminal region, whereas this fragment is absent in the w^{1118} strain. (c) The external spacer primer set failed to amplify a sequence in any of the transformed strains or the control w^{1118} .

well as the *piggyBac*-positive strain M23.1 and the negative white eye strain w^{1118} , were used to perform two sets of PCRs to verify the presence of the *piggyBac* 5' and 3' terminal repeat regions. An additional negative control PCR was performed on all transformants to show the absence of the external lambda phage DNA stuffer sequence (Fig. 2).

The first set of PCRs utilized the IFP2_R1 and MF34 primers to amplify the 5' terminal repeat regions, and the second set of PCRs used the IFP2_L and MF34 primers to amplify the 3' terminal repeat regions. All of the synthetic deletion transformed strains, the M23.1 control strain and the plasmid control yielded a strong PCR product of the correct size for each of the primer sets, confirming the presence of both of the *piggyBac* terminal repeat regions in all of the transformed strains. Interestingly, the white eye strain w^{1118} yielded a very weak product of the correct size with the 5' terminal repeat PCR amplification, but failed to generate a product with the 3' terminal specific primer set.

A third set of PCRs was performed using the IFP2_R1 and IFP2_L primers in an attempt to amplify the external lambda phage DNA stuffer sequence, which would be present if an insertion resulted from recombination of the entire plasmid sequence rather than transposition. The control product from this PCR reaction is a 925 bp fragment,

and no such corresponding fragments were generated with any of the transformed strain genomic DNAs.

Southern hybridization analysis

Southern hybridization analysis was performed to verify the copy number and further confirm transposition of the *piggyBac* deletion plasmids into the *Drosophila* genome (Figs 3 and 4). Genomic DNAs from two of the pBSII-ITR1.1k-ECFP strains (strains 1 and 3) and one of each of the other strains were digested with *HindIII*, with the pBSII-ITR1.1k-ECFP plasmid *HindIII* digest as a plasmid control. The *HindIII* digestion of all transformed strains is expected to generate four fragments after transpositional insertion: the pBSII plasmid backbone fragment (2960 bp), the 3xP3-ECFP marker fragment (1158 bp), the *piggyBac* 5' terminus fragment and the *piggyBac* 3' terminus fragment. Using the pBSII-ITR1.1k-ECFP plasmid as probe, all four fragments generated by the *HindIII* digestion may be detected.

The diagnostic 2960 bp pBSII backbone and 1158 bp ECFP marker fragments were present in all of the transformed strains examined. All of these strains also exhibited at least two additional bands corresponding to the *piggyBac* termini and adjacent sequences at the integration site (Fig. 3). These

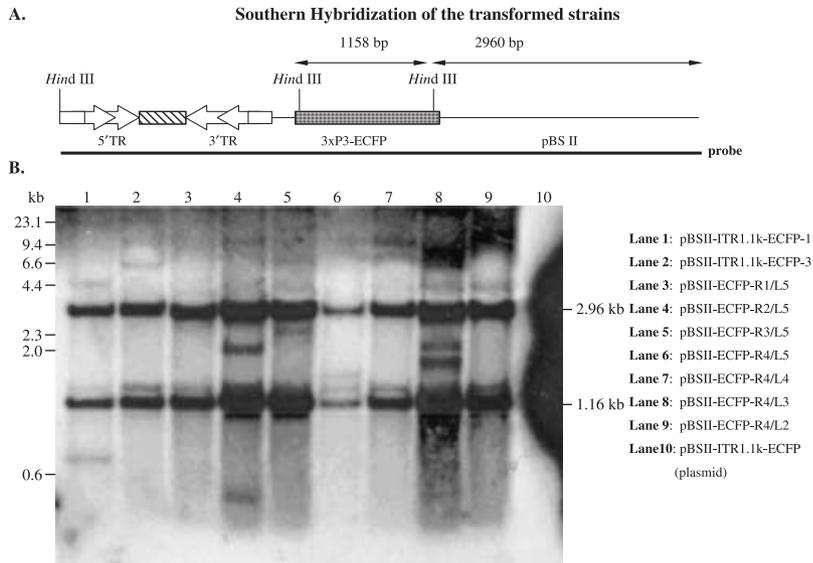


Figure 3. Southern hybridization analysis of synthetic deletion plasmid transformed strains. Genomic DNAs from selected strains and the pBSII-ITR1.1k-ECFP plasmid control were digested with *Hind*III and hybridized to the pBSII-ITR1.1k-ECFP plasmid probe. (A) Map of the pBSII-ITR1.1k-ECFP plasmid showing the size of expected diagnostic fragments. (B) All transformed strains exhibit the two diagnostic bands (2.96 and 1.16 kb) and at least two additional bands reflecting the *piggyBac* terminal adjacent sequences at the site of integration. We also observe a weak 1.3 kb band in all strains that probably represents a *piggyBac*-like sequence in the w^{1118} genome. The reduced intensity of the two additional bands representing joining sequences between the *piggyBac* termini and adjacent genomic DNA in each of the transformed strains is probably due to weaker hybridization of the 200–300 bp of AT-rich sequences of this region of the probe.

p(PZ)-Bac-EYFP Southern Hybridization

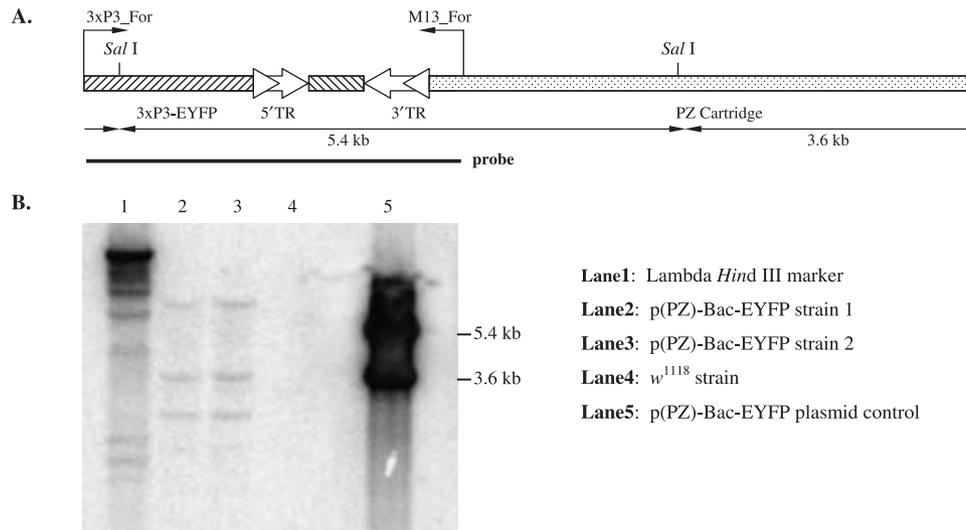


Figure 4. Southern hybridization analysis of the single p(PZ)-Bac-EYFP transformant. Genomic DNA from the p(PZ)-Bac-EYFP strain and the w^{1118} white eye strain were digested with *Sal*I, with a *Sal*I digest of the p(PZ)-Bac-EYFP plasmid serving as control. The probe was PCR amplified from p(PZ)-Bac-EYFP using the primers 3xP3_for and M13_For. (A) Map of the p(PZ)-Bac-EYFP plasmid illustrating the position of *Sal*I sites, the region used as the probe and expected size (3.6 kb) for the diagnostic hybridizing fragment. (B) The two p(PZ)-Bac-EYFP transgenic sublines (lanes 2 and 3) exhibit the diagnostic 3.6 kb band and two additional bands representing junction fragments containing genomic sequences and *piggyBac* ends at the single insertion site.

results confirmed that the observed frequencies were the result of transpositional integrations.

Analysis of insertion site sequences

To further verify that *piggyBac*-mediated transposition of the synthetic deletion constructs occurred in these transformants, individual insertion sites were examined by isolating joining regions between the transposon and genomic sequences using either universal PCR or inverse PCR. Subsequent sequencing analysis of these joining regions

demonstrated that all of the insertions occurred exclusively at single TTAA target sites that were duplicated upon insertion, and all insertion sites had adjacent sequences that were unrelated to the vector (Table 2). The two pBSII-ITR1.1k-ECFP strains 1 and 3 have a single insertion on the third and X chromosome, respectively. These data are consistent with the information obtained from genetic crosses with balancer strains (data not shown).

During sequence analysis of the integration sites a reported point mutation in our constructs was confirmed

Table 2. Transformed *Drosophila* insertion sites

Strain name	Chromosome location	Insertion site sequence 3' junction	5' junction
p(PZ)-Bac-EYFP	3R	CCAAACTTCGGCGATGTTTCTTAA	-- piggyBac --
pBSII-ITR1.1k-ECFP-1	3R	TAGAATTCATGTTTCCAATTTTAA	-- piggyBac --
pBSII-ITR1.1k-ECFP-3	X		-- piggyBac --
pBSII-ECFP-R1/L5	3L	TGGGTGGCACGTTGTGGATTTAA	-- piggyBac --
pBSII-ECFP-R2/L5	2L	AAATACGTCACCTCCCTTAA	-- piggyBac --
pBSII-ECFP-R3/L5	2R	AGCTGCACCTCACCGGATGCCTTAA	-- piggyBac --
pBSII-ECFP-R4/L5	2R	CCCAAAGTATAGTTAAATAGCTTAA	-- piggyBac --
pBSII-ECFP-R4/L4	2R	GT T TAT T TATGAT TAGAGCCTTAA	-- piggyBac --
pBSII-ECFP-R4/L3	2R	TGT TGT T T T T TGTCCCCACGTTAA	-- piggyBac --
pBSII-ECFP-R4/L2	2L	CTGCCTCTAGCCGCTGCT T TAT TAA	-- piggyBac --

The 5' and 3' flanking sequences for the inserted *piggyBac* sequences in each strain were obtained using end-specific inverse PCR (see Experimental procedures) followed by cloning and sequencing of the recovered fragments. The chromosomal locations were determined from the sequences using the Blast search program against the available *Drosophila* sequence in the GenBank library.

Table 3. Percentage of each nucleotide at *piggyBac* insertion sites flanking sequences from position -10 to +10

Nucleotide	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	TTAA	+1	+2	+3	+4	+5	+6	+7	+8	+9	+10
A	22	31	38	33	26	27	16	18	18	29		41	28	43	41	42	43	28	34	33	40
C	20	19	22	17	17	23	15	20	26	15		11	20	18	20	15	17	21	23	16	11
G	28	19	17	16	24	8	24	16	19	12		18	29	22	13	20	12	23	6	15	11
T	30	31	23	34	33	42	45	46	37	44		30	23	17	26	23	28	28	37	36	38

The available *piggyBac* insertion sites include insertion sites in transformed insect genomes (Elick *et al.*, 1996a,b; Handler *et al.*, 1998, 1999; Handler & Harrell, 1999; Peloquin *et al.*, 2000; Tamura *et al.*, 2000; Grossman *et al.*, 2001; Hediger *et al.*, 2001; Kokoza *et al.*, 2001; Li *et al.*, 2001a; Heinrich *et al.*, 2002; Lobo *et al.*, 2002; Mandrioli & Wimmer, 2002; Nolan *et al.*, 2002; Perera *et al.*, 2002; Thomas *et al.*, 2002; Sumitani *et al.*, 2003; data from this report), insertion sites in baculoviruses (Cary *et al.*, 1989; Fraser *et al.*, 1995) and insertion sites in transposition assay target plasmid pGDV1 (Thibault *et al.*, 1999; Grossman *et al.*, 2000; Li *et al.*, 2001a; Lobo, Li and Fraser, unpublished data). No consensus aside from the TTAA target site is apparent among these insertion sites. However, the *piggyBac* transposable element does have a preference of inserting in the TA-rich region with 4–5 Ts before and 5–6 As after the TTAA target site.

that occurs at position 2426 in the *piggyBac* sequence, within the 3' TRD at the boundary of the 31 bp spacer and the internal repeat sequence. This point mutation was apparently generated in constructing the pIAO-P/L plasmid (Li *et al.*, 2001b) and was therefore present in all of the constructs generated by the PCR syntheses employed in this study. This point mutation had no apparent effect on the transformation frequencies, as evidenced by the efficiency of transformation obtained with pBSII-ITR1.1k-ECFP.

The available *piggyBac* insertion site data from previous reports and these studies were compiled and aligned using ClustalX to identify a potential common insertion site motif (Table 3). No apparent consensus motif arose from the comparison of these sequences outside of the required TTAA target site.

Discussion

Attempts to transform insects using plasmids containing a previously reported *piggyBac* ITR minimal sequence cartridge (Li *et al.*, 2001b), which has facing 5' and 3' TRDs with their respective TTAA target sites and is completely devoid of ID sequences, failed to produce a transformation frequency that was comparable with frequencies obtained

with full-length or conservative ID deletion constructs (Handler & Harrell, 1999; Horn & Wimmer, 2000). As is illustrated in this report, frequencies of transposition obtained for the ITR-based p(PZ)-Bac-EYFP and the similarly constructed pBS-pBac/DsRed were far less than expected. Although Southern hybridization and inverse PCR analyses did confirm that the single transformant recovered with p(PZ)-Bac-EYFP had resulted from transpositional insertion, the efficient transposition of *piggyBac* minimal vectors evidenced in interplasmid transposition assays (Li *et al.*, 2001b) did not necessarily predict the properties of *piggyBac* transposon movement in germline transformations.

The fact that germline transposition involves distinctly different cell populations than interplasmid transposition in injected embryos may explain these discrepancies. Similar discrepancies between transformation results and artificial transposition assays have been reported with other Class II transposons (Tosi & Beverley, 2000; Lohe & Hartl, 2001; Lozovsky *et al.*, 2002). In addition, the *Hermes* transposable element undergoes normal cut-and-paste transposition in plasmid-based transposition assays (Sarkar *et al.*, 1997b), but germline integrations in *Ae. aegypti* seem to occur either through general recombination or through a partial replicative transposition mechanism (Jasinskiene *et al.*, 2000).

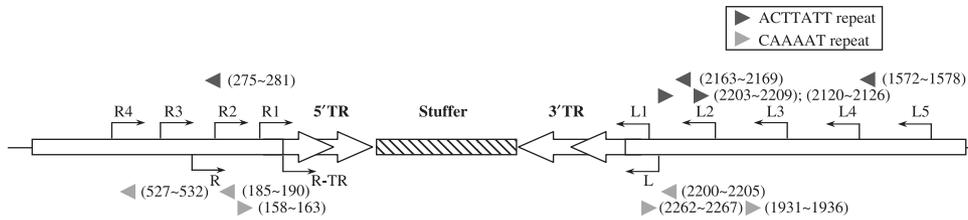


Figure 5. Schematic illustration of the locations of the two short repeat sequence motifs identified in the TRD adjacent ID sequences of *piggyBac*. Several of these repeat motifs are within the regions between R and R1, or L and L2, which appear to be the critical regions based on our transformation results. These repeats are also found in other positions of the *piggyBac* ID sequence.

The synthetic cartridge approach that we used to examine the role of ID sequences in effecting efficient germline transposition has the advantage of examining the involvement of sequences through reconstruction rather than by analysis of successive internal deletions. The main disadvantage of this approach in analysing *piggyBac* is the high AT content of the transposon, which limits the position of useful primers. As a result, our present analyses cannot define the exact limits of the requisite sequences. However, we are able to delimit the critical sequences to a relatively narrow 250 bp of TRD adjacent ID sequences.

Transformation results from synthetic unidirectional deletion plasmids suggest that no more than 66 bp (nt 36–101) of the *piggyBac* 5' ID sequence and 378 bp (nt 2031–2409) of the *piggyBac* 3' ID sequence are necessary for efficient transformation when these deletions are paired with long (378 or 311 bp, respectively, or longer) ID sequences from the opposite end of the transposon. The transformation data from the pBSII-ITR1.1k-ECFP plasmid further defines the 3' ID essential sequence as 172 bp (nt 2237–2409). Combining this same 172 bp 3' ID sequence with only the 5' TRD in the pBSII-ECFP-R-TR/L plasmid yielded no transformants, demonstrating that the 3' ID sequence alone was insufficient for full mobility. Unexpectedly, adding the 66 bp 5' ID sequence in pBSII-ECFP-R1/L also does not allow recovery of full transformation capability while the same 66 bp does allow full transformation capability when coupled to the larger (955 bp) 3' ID sequence in pBSII-ECFP-R1/L5. This result cannot be explained by size alone, because the ITR cartridge strategy used to test these deletion sequence constructs effectively replaces the rest of the *piggyBac* ID with the 2961 bp pBSII plasmid sequence.

Although the frequencies obtained for a given construct may be higher or lower relative to the control, these frequency estimates cannot be considered to have statistical significance relative to one another because the actual number of transformed lines generated for each construct is too low to permit meaningful analyses. In addition, variation between injection experiments can be quite high, as evidenced by the control injections performed in the Handler laboratory vs. those performed in the Fraser laboratory. Ultimately, our experiments were designed merely to detect

the limits of ID sequence that yielded acceptable transformation frequencies and not to evaluate the effectiveness of the deleted regions relative to one another.

Taken together, our results indicate the presence of critical sequences between nucleotides 66 and 311 of the 5' ID sequence used for construction of the pBSII-ITR1.1k-ECFP, because this construct exhibits full transforming capability when matched with the L 3' ID sequence. We also infer that compensating sequences must be present in 3' ID sequences longer than 172 bp, because the 955 bp 3' ID sequence included with primer L5 is able to compensate for the 66 bp 5' ID sequence (construct pBSII-ECFP-R1/L5). Exactly what these signals are cannot be defined in this study, but we do note the presence of small repeats in the 5' ID sequence of pBSII-ITR1.1k-ECFP that are matched by similar sequences in the 3' ID sequences included in construct pBSII-ECFP-R1/L5. These small repeats (Fig. 5) occur in direct or opposite orientations and are also found in several other locations within the *piggyBac* ID. Although we cannot say at this point whether the number, combination and/or orientation of these repeats is significant in conferring efficient transpositional capability, there does seem to be a correlation between efficient transgenesis and the presence of at least one CAAAAT repeat in the 3' ID sequence combined with at least one in the 5' ID sequence, or the compensating presence of two or three sequence repeats in the 3' ID. Further analyses are necessary to confirm the importance, if any, of this small repeat in facilitating transpositional activity of *piggyBac* constructs.

Previous observations of efficient interplasmid transposition for the *piggyBac* ITR construct, completely devoid of *piggyBac* ID sequences, support a mechanism for movement in which the *piggyBac* transposase binds at the terminal repeat regions (IR, spacer and TR) to effect transposition (Li *et al.*, 2001b). Because the cut-and-paste reactions of excision and transposition do not appear to require ID sequences, the relatively unsuccessful application of the previously constructed ITR cartridge for germline transformation suggests the required ID sequences may be involved in other aspects of the transformation process than the mechanics of cut-and-paste. These other aspects seem to be linked to differential movement in germline cells.

The presence of sequences important for full transforming capability within internal domains of transposons is not without precedent. Transposase binding to target sequences at or near the ends of the element is necessary to generate a synaptic complex that brings the ends of the element together for subsequent DNA cleavage (reviewed by Saedler & Gierl, 1996), but the efficiency of this interaction can be influenced by other sequences in the transposon.

Multiple transposase binding sites or accessory factor binding sites are identified in other Class II transposon systems. Efficient transposition of *mariner* requires the continuity of several internal regions of this element and their proper spacing with respect to the terminal repeats (Lohe & Hartl, 2001; Lozovsky *et al.*, 2002), although they are not essential for *in vitro* transposition (Tosi & Beverley, 2000). The P element transposase binding occurs at 10 bp subterminal sequences present at both 5' and 3' ends, while the 31 bp terminal inverted repeat is recognized by a *Drosophila* host protein, IRBP (inverted repeat binding protein), and an internally located 11 bp inverted repeat is shown to act as a transpositional enhancer *in vivo* (Rio & Rubin, 1988; Kaufman *et al.*, 1989; Mullins *et al.*, 1989). The maize *Ac* transposase binds specifically and co-operatively to repetitive ACG and TCG trinucleotides, which are found in more than twenty copies in both 5' and 3' subterminal regions, although the *Ac* transposase also weakly interacts with the terminal repeats (Kunze & Starlinger, 1989; Becker & Kunze, 1997). The *TNPA* transposase of the *En/Spm* element binds a 12 bp sequence found in multiple copies within the 5' and 3' 300 bp subterminal repeat regions (Gierl *et al.*, 1988; Trentmann *et al.*, 1993). The *Arabidopsis* transposon *Tag1* also requires minimal subterminal sequences and a minimal internal spacer between 238 bp and 325 bp for efficient transposition (Liu *et al.*, 2000). The *Sleeping Beauty* (SB) transposable element contains two transposase binding sites (DRs) at the end of the ~230 bp terminal inverted repeats (Ivics *et al.*, 1997). The DNA-bending protein HMGB1, a cellular co-factor, was found to interact with the SB transposase *in vivo* to stimulate preferential binding of the transposase to the DR further from the cleavage site, and promoted bending of DNA fragments containing the transposon IR (Zayed *et al.*, 2003).

These examples suggest that the *piggyBac* transposase or some host accessory factors could be binding to the identified critical TRD adjacent ID regions to promote efficient transposition in germline cells. These subterminal ID sequences may serve as additional *piggyBac* transposase binding sites, thus increasing the efficiency of movement by co-operative binding of the transposase. Alternatively, these sequences may serve as some accessory factor binding site(s) responsible for efficient alignment of the termini or facilitating association of the transposon with chromatin-complexed genomic sequences.

Our results force a reassessment of the reliability of plasmid-based transposition assays in predicting *piggyBac* movement for transgenesis. Plasmid-based transposition assays, although facilitating mutational analyses of the transposon, are likely to be reliable predictors of *in vivo* movement only when alterations lead to a loss of movement. This difference is probably due to the fact that plasmid-based assays indicate the activity of the transposon in somatic cells whereas transformation assays assess movement in germline cells.

Chromatin in the primordial germ cells is structured and regulated differently than that of blastoderm cells (reviewed by Wolffe, 1996). This difference could contribute to different results in the two types of assays. Interplasmid transposition assays utilize purified supercoiled DNA as the target, whereas transformation assays target chromatin. Nucleosome formation on negatively supercoiled DNA occurs virtually instantaneously *in vitro* (Pfaffle & Jackson, 1990), and target plasmid DNA introduced into the embryo cells would most likely form nucleosome structures, but there will be a significant difference in complexity compared with chromatin. This difference in complexity could be the cause of different transposition results. Alternatively, the absence of additional transposase or accessory factor binding sites on the transposon could result in less efficient translocation of the DNA to the nucleus, or lessened affinity of the transposon/transposase complex for the genomic DNA.

Sequence analysis of integrated constructs and subsequent detailed analysis of all the constructs confirms a point mutation first reported to us by Dr Peter Clyne (UC San Francisco; pers. comm.) in the TRD of all constructs used in this study. This mutation is a C→A transversion in the 19 bp internal repeat sequence of the 3' TRD (Fig. 6). This point mutation originated during construction of the pIAO plasmid (Li *et al.*, 2001b), and is most likely the result of misincorporation during PCR amplification. However, our results confirm that this mutation has no significant effect on the transformation efficiency.

Under the conditions of our direct PCR amplification using *piggyBac* 5' terminus-specific primers we generated a weak band of the same size as the expected *piggyBac* band from control *w*¹¹⁸ flies. The Southern hybridizations detected a 1.3 kb band in all of the transformants that was distinct from the pBSII backbone fragment (2.96 kb) and 3xP3-ECFP (1.16 kb) marker bands. *piggyBac*-like sequences have been detected in many species by PCR and Southern hybridization analysis using probes derived from the *piggyBac* 5' terminal region, including moths, flies and beetles (reviewed by Handler, 2002). A homology search against the available sequence database has identified the existence of the *piggyBac*-like sequences in the *D. melanogaster* genome (Sarkar *et al.*, 2003). We interpret our results as reflecting the presence of one of these degenerate *piggyBac*-like sequences in the *Drosophila* genome.

The insertion sites in the transformed fly strains were identified by either universal PCR or inverse PCR techniques.

	13 bp TR	31 bp spacer	19 bp IR
Wide-type	CCCTAGAAAGATA	ATCATATTGTGACGTACGTTAAAGATAATCA	TGCGTAAAAATTGACGCATG
Mutation	CCCTAGAAAGATA	ATCATATTGTGACGTACGTTAAAGATAATCA	TG <u><u>A</u></u> GTAAAAATTGACGCATG

Figure 6. Identified point mutation in the 3' internal repeat sequence. A point mutation was discovered in the 19 bp internal repeat sequence (IR) of the 3' TRD in all of the constructs derived from the pIAO-P/L 589 plasmid (Li *et al.*, 2001b). This nucleotide substitution from C to A (bold and underlined) had no apparent effect on the transposition frequency of any of these constructs relative to the pBac{3xP3-EYFP} control plasmid.

All insertions occurred exclusively at TAA sites, verifying that these insertions were due to a specific *piggyBac* transposase-mediated mechanism (Fraser *et al.*, 1995). A ClustalX alignment of all *piggyBac* insertion sites identified in this and previous reports, including insertion sites in the transposition assay target plasmid pGDV1 (Sarkar *et al.*, 1997a), baculoviruses and transformed insect genomes, does not reveal any further significant similarities (Table 3). The proposed existence of a larger *piggyBac* insertion consensus sequence YYT T T T T /AARTAAAYAG (Y = pyrimidine, R = purine, / = insertion point) by Cary *et al.* (1989) and Grossman *et al.* (2000), and a short 8 bp consensus sequence A/TNA/TTTAAA/T proposed by Li *et al.* (2001a) seem to be contradicted by the accumulated insertion site data. We do note a decided preference for *piggyBac* insertion within TAA target sites flanked by 4–5 Ts on the 5' side and 5–6 As on the 3' side (Table 3).

Based on the minimal *piggyBac* vector pBSII-ITR1.1k-ECFP, we constructed a plasmid minimal vector, pXL-BacII-ECFP, which also yields a high frequency of transformation in *D. melanogaster* (Table 1). Our results confirm that both the pBSII-ITR1.1k-ECFP and the pXL-BacII-ECFP minimal vectors can serve as highly efficient *piggyBac* transformation vectors. For more information on these minimal vectors, please visit the *piggyBac* website at <http://piggyBac.bio.nd.edu>.

Experimental procedures

Plasmids

The pCaSpeR-hs-orf helper plasmid was constructed by PCR amplifying the *piggyBac* open reading frame using IFP2orf_For and IFP2orf_Rev primers, cloning into the pCRII vector (Invitrogen), excising with *Bam*HI and inserting into the *Bam*HI site of the P element vector, pCaSpeR-hs (Thummel & Pirrotta, 1992). A single clone with the correct orientation and sequence was identified and named pCaSpeR-hs-orf (Fig. 1).

The p(PZ)-Bac-EYFP plasmid (Fig. 1) was constructed from the p(PZ) plasmid (Rubin & Spradling, 1983) by digesting with *Hind*III and recircularizing the 7 kb fragment containing *LacZ*, *hsp70* and Kan/ori sequences to form the p(PZ)-7kb plasmid. The ITR cartridge was excised from pBSII-ITR (Li *et al.*, 2001b) using *Not*I and *Sal*I and blunt-end cloned into the *Hind*III site of the p(PZ)-7kb plasmid. A 3xP3-EYFP marker gene was PCR amplified from pBac{3xP3-EYFPafm} (Horn & Wimmer, 2000), digested with *Spe*I and inserted into the *Xba*I site to form p(PZ)-Bac-EYFP.

The pBSII-3xP3-ECFP plasmid was constructed by PCR amplifying the 3xP3-ECFP marker gene from pBac{3xP3-ECFPafm} (Horn & Wimmer, 2000) using the primer pair ExFP_For and ExFP_Rev

(Table 3), digesting the amplified fragment with *Spe*I, and cloning it into the *Xba*I site of pBlueScript II plasmid (Stratagene).

The *piggyBac* synthetic internal deletion plasmids were constructed by PCR amplification from the pIAO-P/L-589 bp plasmid (Li *et al.*, 2001b) using a series of primers (Table 4). Nine PCR products were generated using the combination of IFP2_R4 against all five IFP2_L primers and IFP2_L5 against all four IFP2_R primers. Two additional PCR products were also obtained using the IPF2_R-TR + IFP2_L and IFP2_R1 + IFP2_L primer pairs. These PCR products were then cloned into the pCR II vector (Invitrogen), excised by *Spe*I digestion and cloned into the *Spe*I site of the pBSII-3xP3-ECFP plasmid to form the *piggyBac* internal deletion series (Fig. 7). The pBSII-ITR1.1k-ECFP plasmid (Fig. 1) was constructed by cloning the *Eco*R V/*Dra*I fragment from pIAO-P/L-589 bp, which contained both *piggyBac* terminal repeats, into the *Eco*RV site of pBSII-3xP3-ECFP. The pXL-BacII-ECFP plasmid (Fig. 1) was constructed essentially as described previously (Li *et al.*, 2001b) by PCR amplifying the ITR1.1k cartridge from pBSII-ITR1.1k-ECFP plasmid using MCS_For and MCS_Rev primers, each containing flanking *Bg*III sites, cutting with *Bg*III, re-ligating and cutting again with *Bss*HIII, then inserting into the *Bss*HIII sites of the pBSII plasmid.

The pBS-pBac/DsRed1 plasmid was constructed by excising the 731 bp *Asel* fragment from p3E1.2, including 99 bp of 3' *piggyBac* terminal sequence and adjacent NPV insertion site sequence, and ligating it as a blunt fragment into a unique *Kpn*I-blunted site in pBS-KS (Stratagene). The resulting plasmid was digested with *Sac*I and blunted, digested with *Pst*I and ligated to a 173 bp *Hinc*II-*Nsi*I fragment from p3E1.2, including 38 bp of 5' *piggyBac* terminal sequence. The pBS-pBac minimal vector was marked with the polyubiquitin-regulated DsRed1 digested from pB[PUbDsRed1] (Handler & Harrell, 2001a) and inserted into an *Eco*RI-*Hind*III deletion in the internal cloning site within the terminal sequences.

Transformation of *D. melanogaster*

The *D. melanogaster* w^{1118} white eye strain was used for all micro-injections employing a modification of the standard procedure described by Rubin & Spradling (1982) in which the dechlorination step was eliminated. Equal concentrations (0.5 µg/µl) of each of the internal deletion plasmids or the control plasmid pBac{3xP3-ECFPafm} were injected along with an equal amount of the pCaSpeR-hs-orf helper plasmid into embryos followed by a 1 h heat shock at 37 °C and recovery overnight at room temperature. Emerging adults were individually mated with w^{1118} flies, and progeny were screened as larvae using an Olympus SZX12 fluorescent dissecting microscope equipped with GFP (480 nm excitation/510 nm barrier), CFP (436 nm excitation/480 nm barrier) and YFP (500 nm excitation/530 barrier) filter sets. Two positive adults from each of the vials were crossed with w^{1118} to establish germline-transformed strains. The pBS-pBac/DsRed1 minimal vector was also injected and screened using a HQ Texas Red® filter no. 41004 (Handler & Harrell, 2001a).

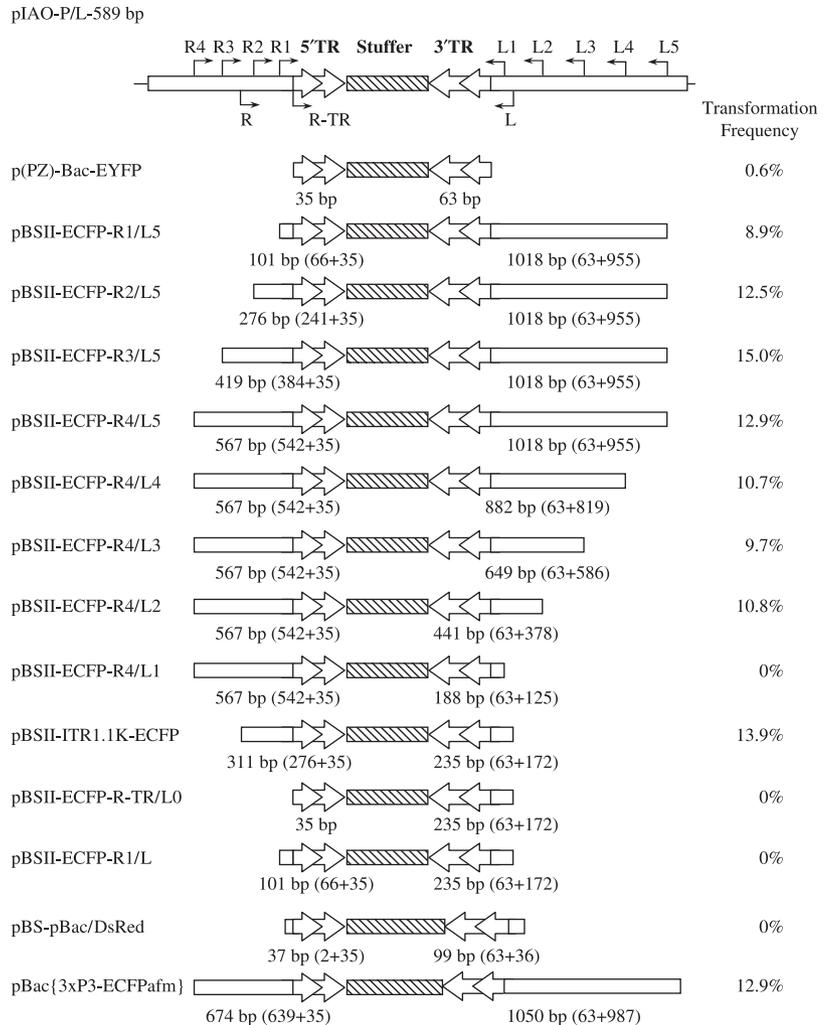


Figure 7. Schematic illustration of TRD and adjacent ID regions present in plasmids and synthetic *piggyBac* internal deletion series constructs tested for transformation efficiency. The plasmids p(PZ)-Bac-EYFP and all pBSII-ECFP synthetic deletions are based on sequences amplified from the pIAO-P/L-589 construct of Li *et al.* (2001b). All plasmids have the 35 bp 5' TRD and 63 bp 3' TRD, and include variable lengths of 5' and 3' adjacent ID sequence. The relative transformation frequency for each plasmid is listed to the right for convenience.

Direct PCR analysis

Genomic DNAs from each of the transformed stains, the w^{1118} wild-type strain and a *piggyBac*-positive strain M23.1 (Handler & Harrell, 1999) were prepared using a modified DNAzol procedure. About sixty flies from each strain were combined with 150 μ l of DNAzol (Molecular Research Center, Inc.) in a 1.5 ml eppendorf tube. The flies were homogenized, an additional 450 μ l of DNAzol was added and the homogenates were incubated at room temperature for 1 h. The DNAs were extracted twice with phenol-chloroform (1 : 1 ratio), and the aqueous fractions were transferred to new tubes for precipitation of the DNA with an equal volume of 2-propanol. The DNA pellets were washed with 70% ethanol, air dried and resuspended in 150 μ l of dH_2O containing 10 μ g of RNase A.

Two sets of direct PCRs were performed to identify the presence of *piggyBac* sequences in transformed fly genomes. Primers MF34 and IFP2_L were used to identify the presence of the *piggyBac* 3' terminal repeat, and MF34 and IFP2_R1 were used to identify the *piggyBac* 5' terminal repeat. To exclude the possibility of recombination, a second PCR was also performed using the IFP2_R1 and IFP2_L primers to amplify the external stuffer fragment (Li *et al.*, 2001b) between the terminal repeat regions.

Southern hybridization analysis

Southern hybridization analysis was performed using a standard procedure with minor modifications (Ausubel *et al.*, 1994). Approximately 8 μ g of genomic DNA (isolated as above) from each of the transformed fly strains was digested with 40 units of *Hind*III for 4 h, followed by agarose gel electrophoresis. The gel was then denatured, neutralized, transferred to nylon membranes and baked at 80 $^{\circ}C$ for 4 h, and the membranes were prehybridized overnight. A synthetic probe was prepared by nick translation (Invitrogen kit) using ^{32}P -labelled dGTP against the pBSII-ITR1.1K-ECFP plasmid template. Purified probe was hybridized at 65 $^{\circ}C$ overnight followed by several washes, and the membranes were first exposed on phosphor screens (Kodak) overnight for scanning with a Storm phosphor Scanner (Molecular Dynamics System), and then exposed on X-ray film (Kodak).

Universal PCR and inverse PCR analysis

The *piggyBac* insertion sites in the transformed fly strains were identified using either universal PCR (Beeman & Stauth, 1997) or inverse PCR techniques (Ochman *et al.*, 1988). For the universal PCR, the IFP2_L (3' TR) or IPR2_R1 (5' TR) primer was combined

Table 4. A listing of the synthetic oligonucleotide primers used in this study

Primer	Sequence
Internal deletion primers	
IFP2_R1	ACTTCTAGAGTCCTAAATTGCAACAGCGAC
IFP2_R2	ACTTCTAGACACGTAAGTAGAACATGAAATAAC
IFP2_R3	ACTTCTAGATCACTGTCTAGAATCCTCACCAAC
IFP2_R4	ACTTCTAGAGAAGCAATGAAGAACCTGG
IFP2_L1	ACTTCTAGAAATAAATAAATAAATAAATAAATTG
IFP2_L2	ACTTCTAGAGAAAAGCAAAATGCATCGTGC
IFP2_L3	ACTTCTAGACGCAAAAAATTTATGAGAAACC
IFP2_L4	ACTTCTAGAGATGAGGATGCTTCTATCAACG
IFP2_L5	ACTTCTAGACGCGAGATACCGGAAGTACTG
IFP2_L	ACTTCTAGACTCGAGAGAGAATGTTTAAAGTTTTGTT
IFP2_R-TR	ACTTCTAGACATGCGTCAATTTTACGCAGACTATCTTTC TAGGG TTAATCTAGCTGCATCAGG
Other primers	
ExFP_For	ACGACTAGTGTCCCAACAATGGTTAATTCG
ExFP_Rev	ACGACTAGTGCCGTACGCGTATCGATAAGC
IFP2orf_For	GGATCCTATA TAATAAAATG GGTAGTTCTT
IFP2orf_Rev	GGATCCAAATCAACAACAATTTATTTATG
MF34	GGATCCTCTAGATTAACCCTAGAAAGATA
Univ-1	TAATACGACTCACTATAGGGNNNNNNNNNNCTAT
Univ-2	TAATACGACTCACTATAGGGNNNNNNNNNNAGTGC
Univ-3	TAATACGACTCACTATAGGGNNNNNNNNNNGAATTC
Univ-4	TAATACGACTCACTATAGGGNNNNNNNNNNAGTACT
Univ-5	TAATACGACTCACTATAGGGNNNNNNNNNNNAAGCTT
Univ-6	TAATACGACTCACTATAGGGNNNNNNNNNNGGATCC
Univ-7	TAATACGACTCACTATAGGGNNNNNNNNNNCTAG
iPCR_R1	ATTTTACGCAGACTATCTTCTTA
T7	TTAATACGACTCACTAT
MF14	GGATCCGCGGTAAGTGTCACTGA
JF3	GGATCCTCGATATACAGACCGATAAAAACACATG
IFP2_Lb	ACTGGGCCACTAATAATAAATCAACAACAAAC
iPCR_6	TTATTTTCATGTTCTACTTACGTG
iPCR_L1	TGATTATCTTTAACGTCACGTCAC
MF04	GTCAGTCCAGAAACAACCTTTGGC
IFP2_L-R+	CTAGAAATTTATTTATGTTTATTTATTTATTA
MCS_For	ACGCGTAGATCTAATACGACTCACTATAGGG
MCS_Rev	ACGCGTAGATCTAATTAACCTCACTAAAGGG

with one of seven universal primers (Table 4) during the first round of PCR (94 °C 1 min, 40 °C 1 min, 72 °C 2 min, thirty-five cycles). Two microlitres of the reaction mix from the first-round PCR was then used for a second round of PCR (94 °C 1 min, 50 °C 1 min, 72 °C 2 min, thirty-five cycles) using IFP2_L1 (3' TR) or iPCR_R1 (5' TR) together with a T7 primer (nested on the universal primer).

Inverse PCRs were performed by digesting 5 µg of the genomic DNAs from each of the transformed strains completely with *Hin*P1I for the 3' end or *Taq*I for the 5' end, followed by purification using the GeneClean kit (Q-Biogene) and self-ligation in a 100 µl volume overnight. The self-ligated DNAs were precipitated and resuspended in 30 µl ddH₂O. A 5 µl portion of each ligation was used for first-round PCR (94 °C 1 min, 40 °C 1 min, 72 °C 2 min, thirty-five cycles) with primer pairs IFP2_R1 + MF14 for the 5' end, and JF3 + IFP2_Lb for the 3' end (Table 4). Two microlitres of the first-round PCR products were used as templates for the second-round PCR (94 °C 1 min, 50 °C 1 min, 72 °C 2 min, thirty-five cycles) using primer pairs iPCR_R1 + iPCR_6 for the 5' end and iPCR_L1 + MF04 for the 3' end. The primer pair iPCR_L1 + IFP2_L-R was used for the second-round PCR of the 3' end of pBSII-ITR1.1k-ECFP strains. All the PCR products were cloned into the pCRII vector (Invitrogen) and sequenced. Sequences were subjected to a Blast search of the NCBI database to identify

the locations of the insertions. MacVector 6.5.3 (Oxford Molecular Group) and ClustalX (Jeanmougin *et al.*, 1998) were used for sequence alignments.

Acknowledgements

We would like to thank Ms Anna Wysinski for helping with injections and maintaining fly stocks, and Dr Neil Lobo for helping with sequencing analysis. This research was supported by National Institutes of Health grant AI48561 to M.J.F.

References

- Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1994) *Current Protocols in Molecular Biology*. John Wiley & Sons, Inc., New York.
- Becker, H.A. and Kunze, R. (1997) Maize Activator transposase has a bipartite DNA binding domain that recognizes subterminal sequences and the terminal inverted repeats. *Mol Gen Genet* **254**: 219–230.
- Beeman, R.W. and Stauth, D.M. (1997) Rapid cloning of insect transposon insertion junctions using 'universal' PCR. *Insect Mol Biol* **6**: 83–88.
- Berghammer, A.J., Klingler, M. and Wimmer, E.A. (1999) A universal marker for transgenic insects. *Nature* **402**: 370–371.
- Cary, L.C., Goebel, M., Corsaro, B.G., Wang, H.G., Rosen, E. and Fraser, M.J. (1989) Transposon mutagenesis of baculoviruses: analysis of *Trichoplusia ni* transposon IFP2 insertions within the FP-locus of nuclear polyhedrosis viruses. *Virology* **172**: 156–169.
- Elick, T.A., Bauser, C.A. and Fraser, M.J. (1996a) Excision of the *piggyBac* transposable element *in vitro* is a precise event that is enhanced by the expression of its encoded transposase. *Genetica* **98**: 33–41.
- Elick, T.A., Bauser, C.A., Principe, N.M. and Fraser, M.J. Jr (1996b) PCR analysis of insertion site specificity, transcription, and structural uniformity of the Lepidopteran transposable element IFP2 in the TN-368 cell genome. *Genetica* **97**: 127–139.
- Fraser, M.J., Brusca, J.S., Smith, G.E. and Summers, M.D. (1985) Transposon-mediated mutagenesis of a baculovirus. *Virology* **145**: 356–361.
- Fraser, M.J., Cary, L., Boonvisudhi, K. and Wang, H.G. (1995) Assay for movement of Lepidopteran transposon IFP2 in insect cells using a baculovirus genome as a target DNA. *Virology* **211**: 397–407.
- Fraser, M.J., Ciszczon, T., Elick, T. and Bauser, C. (1996) Precise excision of TTAAspecific lepidopteran transposons *piggyBac* (IFP2) and *tagalong* (TFP3) from the baculovirus genome in cell lines from two species of Lepidoptera. *Insect Mol Biol* **5**: 141–151.
- Gierl, A., Lutticke, S. and Saedler, H. (1988) *TrnA* product encoded by the transposable element *En-1* of *Zea mays* is a DNA binding protein. *EMBO J* **7**: 4045–4053.
- Gonzalez-Estevez, C., Momose, T., Gehring, W.J. and Salo, E. (2003) Transgenic planarian lines obtained by electroporation using transposon-derived vectors and an eye-specific GFP marker. *Proc Natl Acad Sci USA* **100**: 14046–14051.
- Grossman, G.L., Rafferty, C.S., Fraser, M.J. and Benedict, M.Q. (2000) The *piggyBac* element is capable of precise excision and transposition in cells and embryos of the mosquito, *Anopheles gambiae*. *Insect Biochem Mol Biol* **30**: 909–914.

- Grossman, G.L., Rafferty, C.S., Clayton, J.R., Stevens, T.K., Mukabayire, O. and Benedict, M. (2001) Germline transformation of the malaria vector, *Anopheles gambiae*, with the *piggyBac* transposable element. *Insect Mol Biol* **10**: 597–604.
- Handler, A.M. (2002) Use of the *piggyBac* transposon for germline transformation of insects. *Insect Biochem Mol Biol* **32**: 1211–1220.
- Handler, A.M. and Harrell, R.A., 2nd (1999) Germline transformation of *Drosophila melanogaster* with the *piggyBac* transposon vector. *Insect Mol Biol* **8**: 449–457.
- Handler, A.M. and Harrell, R.A., 2nd (2001a) Polyubiquitin-regulated DsRed marker for transgenic insects. *Biotechniques* **31**: 824–828.
- Handler, A.M. and Harrell, R.A., 2nd (2001b) Transformation of the Caribbean fruit fly, *Anastrepha suspensa*, with a *piggyBac* vector marked with polyubiquitin-regulated GFP. *Insect Biochem Mol Biol* **31**: 199–205.
- Handler, A.M. and McCombs, S.D. (2000) The *piggyBac* transposon mediates germ-line transformation in the Oriental fruit fly and closely related elements exist in its genome. *Insect Mol Biol* **9**: 605–612.
- Handler, A.M., McCombs, S.D., Fraser, M.J. and Saul, S.H. (1998) The lepidopteran transposon vector, *piggyBac*, mediates germ-line transformation in the Mediterranean fruit fly. *Proc Natl Acad Sci USA* **95**: 7520–7525.
- Hediger, M., Niessen, M., Wimmer, E.A., Dubendorfer, A. and Bopp, D. (2001) Genetic transformation of the housefly *Musca domestica* with the lepidopteran derived transposon *piggyBac*. *Insect Mol Biol* **10**: 113–119.
- Heinrich, J.C., Li, X., Henry, R.A., Haack, N., Stringfellow, L., Heath, A.C. and Scott, M.J. (2002) Germ-line transformation of the Australian sheep blowfly *Lucilia cuprina*. *Insect Mol Biol* **11**: 1–10.
- Horn, C. and Wimmer, E.A. (2000) A versatile vector set for animal transgenesis. *Dev Genes Evol* **210**: 630–637.
- Ivics, Z., Hackett, P.B., Plasterk, R.H. and Izsvak, Z. (1997) Molecular reconstruction of *Sleeping Beauty*, a *Tc1*-like transposon from fish, and its transposition in human cells. *Cell* **91**: 501–510.
- Jasinskiene, N., Coates, C.J. and James, A.A. (2000) Structure of *hermes* integrations in the germline of the yellow fever mosquito, *Aedes aegypti*. *Insect Mol Biol* **9**: 11–18.
- Jeanmougin, F., Thompson, J.D., Gouy, M., Higgins, D.G. and Gibson, T.J. (1998) Multiple sequence alignment with Clustal X. *Trends Biochem Sci* **23**: 403–405.
- Kaufman, P.D., Doll, R.F. and Rio, D.C. (1989) *Drosophila* P element transposase recognizes internal P element DNA sequences. *Cell* **59**: 359–371.
- Kokoza, V., Ahmed, A., Wimmer, E.A. and Raikhel, A.S. (2001) Efficient transformation of the yellow fever mosquito *Aedes aegypti* using the *piggyBac* transposable element vector pBac[3xP3-EGFP afm]. *Insect Biochem Mol Biol* **31**: 1137–1143.
- Kunze, R. and Starlinger, P. (1989) The putative transposase of transposable element *Ac* from *Zea mays* L. interacts with subterminal sequences of *Ac*. *EMBO J* **8**: 3177–3185.
- Li, X., Heinrich, J.C. and Scott, M.J. (2001a) *piggyBac*-mediated transposition in *Drosophila melanogaster*: an evaluation of the use of constitutive promoters to control transposase gene expression. *Insect Mol Biol* **10**: 447–455.
- Li, X., Lobo, N., Bauser, C.A. and Fraser, M.J. (2001b) The minimum internal and external sequence requirements for transposition of the eukaryotic transformation vector *piggyBac*. *Mol Genet Gen* **266**: 190–198.
- Liu, D., Mack, A., Wang, R., Galli, M., Belk, J., Ketpura, N.I. and Crawford, N.M. (2000) Functional dissection of the *cis*-acting sequences of the Arabidopsis transposable element *Tag1* reveals dissimilar subterminal sequence and minimal spacing requirements for transposition. *Genetics* **157**: 817–830.
- Lobo, N.F., Hua-Van, A., Li, X., Nolen, B.M. and Fraser, M.J. Jr (2002) Germ line transformation of the yellow fever mosquito, *Aedes aegypti*, mediated by transpositional insertion of a *piggyBac* vector. *Insect Mol Biol* **11**: 133–139.
- Lohe, A.R. and Hartl, D.L. (2001) Efficient mobilization of *mariner* *in vivo* requires multiple internal sequences. *Genetics* **160**: 519–526.
- Lozovsky, E.R., Nurminsky, D., Wimmer, E.A. and Hartl, D.L. (2002) Unexpected stability of *mariner* transgenes in *Drosophila*. *Genetics* **160**: 527–535.
- Mandrioli, M. and Wimmer, E.A. (2002) Stable transformation of a *Mamestra brassicae* (lepidoptera) cell line with the lepidopteran-derived transposon *piggyBac*. *Insect Biochem Mol Biol* **33**: 1–5.
- Mullins, M.C., Rio, D.C. and Rubin, G.M. (1989) *cis*-acting DNA sequence requirements for P-element transposition. *Genes Dev* **3**: 729–738.
- Nolan, T., Bower, T.M., Brown, A.E., Crisanti, A. and Catteruccia, F. (2002) *piggyBac*-mediated germline transformation of the malaria mosquito *Anopheles stephensi* using the red fluorescent protein dsRED as a selectable marker. *J Biol Chem* **277**: 8759–8762.
- Ochman, H., Gerber, A.S. and Hartl, D.L. (1988) Genetic applications of an inverse polymerase chain reaction. *Genetics* **120**: 621–623.
- Peloquin, J.J., Thibault, S.T., Staten, R. and Miller, T.A. (2000) Germ-line transformation of pink bollworm (Lepidoptera: *gelechiidae*) mediated by the *piggyBac* transposable element. *Insect Mol Biol* **9**: 323–333.
- Perera, O.P., Harrell, I.I.R.A. and Handler, A.M. (2002) Germ-line transformation of the South American malaria vector, *Anopheles albimanus*, with a *piggyBac*/EGFP transposon vector is routine and highly efficient. *Insect Mol Biol* **11**: 291–297.
- Pfaffle, P. and Jackson, V. (1990) Studies on rates of nucleosome formation with DNA under stress. *J Biol Chem* **265**: 16821–16829.
- Rio, D.C. and Rubin, G.M. (1988) Identification and purification of a *Drosophila* protein that binds to the terminal 31-base-pair inverted repeats of P transposable element. *Proc Natl Acad Sci USA* **85**: 8929–8933.
- Rubin, G.M. and Spradling, A.C. (1982) Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218**: 348–353.
- Rubin, G.M. and Spradling, A.C. (1983) Vectors for P element-mediated gene transfer in *Drosophila*. *Nucleic Acids Res* **11**: 6341–6351.
- Saedler, H. and Gierl, A. (eds) (1996) *Transposable Elements*. Springer-Verlag, Berlin.
- Sarkar, A., Coates, C.J., Whyard, S., Willhoelt, U., Atkinson, P.W. and O'Brochta, D.A. (1997a) The *Hermes* element from *Musca domestica* can transpose in four families of *cyclorrhaphan* flies. *Genetica* **99**: 15–29.
- Sarkar, A., Yardley, K., Atkinson, P.W., James, A.A. and O'Brochta, D.A. (1997b) Transposition of the *Hermes* element in embryos of the

- vector mosquito, *Aedes aegypti*. *Insect Biochem Mol Biol* **27**: 359–363.
- Sarkar, A., Sim, C., Hong, Y.S., Hogan, J.R., Fraser, M.J., Robertson, H.M. and Collins, F.H. (2003) Molecular evolutionary analysis of the widespread *piggyBac* transposon family and related 'domesticated' sequences. *Mol Genet Genomics* **270**: 173–180.
- Sumitani, M., Yamamoto, D.S., Oishi, K., Lee, J.M. and Hatakeyama, M. (2003) Germline transformation of the sawfly, *Athalia rosae* (Hymenoptera: Symphyta), mediated by a *piggyBac*-derived vector. *Insect Biochem Mol Biol* **33**: 449–458.
- Tamura, T., Thibert, C., Royer, C., Kanda, T., Abraham, E., Kamba, M., Komoto, N., Thomas, J.L., Mauchamp, B., Chavancy, G., Shirk, P., Fraser, M., Prudhomme, J.C. and Couble, P. (2000) Germline transformation of the silkworm *Bombyx mori* L. using a *piggyBac* transposon-derived vector. *Nat Biotechnol* **18**: 81–84.
- Thibault, S.T., Luu, H.T., Vann, N. and Miller, T.A. (1999) Precise excision and transposition of *piggyBac* in pink bollworm embryos. *Insect Mol Biol* **8**: 119–123.
- Thomas, J.L., Da Rocha, M., Besse, A., Mauchamp, B. and Chavancy, G. (2002) 3xP3-EGFP marker facilitates screening for transgenic silkworm *Bombyx mori* L. from the embryonic stage onwards. *Insect Biochem Mol Biol* **32**: 247–253.
- Thummel, C.S. and Pirrotta, V. (1992) New pCaSpeR P element vectors. *Dros Info Service* **71**: 150–150.
- 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**: 784–790.
- Trentmann, S.M., Saedler, H. and Gierl, A. (1993) The transposable element *En/Spm*-encoded *TNPA* protein contains a DNA binding and a dimerization domain. *Mol Gen Genet* **238**: 201–208.
- Wolffe, A.P. (1996) Histone deacetylase: a regulator of transcription. *Science* **272**: 371–372.
- Zayed, H., Izsvak, Z., Khare, D., Heinemann, U. and Ivics, Z. (2003) The DNA-bending protein HMGB1 is a cellular cofactor of *Sleeping Beauty* transposition. *Nucleic Acids Res* **31**: 2313–2322.