

Short communication

# Highly conserved *piggyBac* elements in noctuid species of Lepidoptera

Grazyna J. Zimowska, Alfred M. Handler\*

Center for Medical, Agricultural, and Veterinary Entomology, Agricultural Research Service, US Department of Agriculture,  
1700 SW 23rd Drive, Gainesville, FL 32608, USA

Received 18 October 2005; received in revised form 14 February 2006; accepted 1 March 2006

## Abstract

The *piggyBac* transposable element was originally discovered in a *Trichoplusia ni* cell line and nearly identical elements were subsequently discovered in the tephritid fly, *Bactrocera dorsalis*. This suggested the existence of *piggyBac* in additional insects and this study shows highly conserved, though not identical, *piggyBac* sequences in the noctuid species *Helicoverpa armigera*, *H. zea*, and *Spodoptera frugiperda*, as well as new *piggyBac* sequences from the *T. ni* organismal genome. Genomic *piggyBac* elements could not be unambiguously identified in several other moth species indicating a discontinuous presence of *piggyBac* in the Lepidoptera. Most sequences have greater than 95% nucleotide identity to the original IFP2 *piggyBac*, except for a more diverged sequence in *S. frugiperda*, having ~78% identity. Variants of 1.3 and 0.8 kb sequences found in both *H. armigera* and *H. zea* most likely became established by interbreeding, supporting the notion that the species are conspecific. None of the independent *piggyBac* sequences isolated from *T. ni* larval genomes are identical to IFP2, though all have an uninterrupted reading frame with the potential for encoding a functional transposase. The *piggyBac* sequences from *T. ni* and the *Helicoverpa* species, as well as those previously reported from *B. dorsalis*, all share three common nucleotide substitutions resulting in a single amino acid substitution in the transposase. This suggests that the original IFP2 *piggyBac* is a related variant of a predecessor element that became widespread. The existence of conserved *piggyBac* elements, some of which may have been transmitted horizontally between lepidopteran species, raises important considerations for the stability and practical use of *piggyBac* transformation vectors.

© 2006 Elsevier Ltd. All rights reserved.

**Keywords:** *piggyBac*; Lepidoptera; *Helicoverpa*; Transposable elements; Horizontal transfer

## 1. Introduction

The *piggyBac* transposable element is one of the most widely used vectors for gene transfer in insects, and the only functional class II element to be identified in lepidopterans (see Handler, 2002). It was originally discovered as the causative agent of few polyhedra (FP) mutations in baculoviruses passed through the *Trichoplusia ni* TN-368 cell line (Fraser et al., 1983). The isolated element, originally known as IFP2, is 2472 bp in length having 13 bp short inverted terminal repeat (ITR) sequences, 19 bp sub-terminal ITRs and a 1.8 kb open reading frame (ORF) (Cary et al., 1989). It is part of a subclass of elements that insert exclusively in TTAA target sites and excise only in a precise fashion (Fraser et al., 1983;

Elick et al., 1995). Although the functional IFP2 transposon was discovered as a repetitive element in several *T. ni* cell lines (Fraser et al., 1996), the same element was not immediately evident in the organismal genome. Until the discovery of nearly identical elements in the oriental fruit fly, *Bactrocera dorsalis*, *piggyBac* was not known to exist in any other species (Handler and McCombs, 2000). The transposon was identified in several wild and mutant strains of *B. dorsalis* by Southern hybridization, that was verified by PCR analysis in this and other *Bactrocera* species. *piggyBac* was found to exist throughout the *B. dorsalis* species complex, with most sequences having greater than 95% nucleotide identity to the *T. ni* element (G.J. Zimowska, A.M. Handler, unpublished).

The finding of nearly identical elements in phylogenetically and geographically distinct insect species suggested that *piggyBac* or related elements may exist in other species as well, consistent with other class II elements that are

\*Corresponding author. Tel.: +1 352 374 5793; fax: +1 352 374 5794.  
E-mail address: [handler@nersp.nerdc.ufl.edu](mailto:handler@nersp.nerdc.ufl.edu) (A.M. Handler).

members of large families, such as *hAT* and *mariner/Tc* that exist in diverse organisms (Kempken and Windhofer, 2001; Avancini et al., 1996). A database search for *piggyBac* elements revealed amino acid sequence similarities in a variety of organisms with very distant elements present in plants and animals (Sarkar et al., 2003). Three clades of elements were found in insects and vertebrates including a complete related transposon in the *Anopheles gambiae* genome, and the most closely related element, *yabusame-1*, from *Bombyx mori* having 36% amino acid sequence identity to IFP2 (Omuro, N. and Shimada, T., GenBank accession no. AB162707). Thus, similar to other transposon families the known *piggyBac*-related elements remain highly divergent. A less common finding is the presence of nearly identical elements in distantly related species, which is thus far limited to *piggyBac* and several *mariner* subfamilies (Lampe et al., 2003).

In an effort to identify *piggyBac* sequences in other lepidopteran species and in the *T. ni* organismal genome, we embarked on a survey for *piggyBac* by Southern blot-hybridization and PCR. *piggyBac* sequences very closely related to IFP2 were discovered in several noctuid moth species, which contributes to our understanding of the distribution and evolution of *piggyBac*, and has implications for its use as a transformation vector in these species.

## 2. Materials and methods

### 2.1. Insects

Lepidopteran species used to prepare genomic DNA samples were obtained from laboratory insect colonies maintained routinely at USDA-ARS-CMAVE, except for *H. armigera* samples that were kindly provided by Dr. Rod Mahon (University of Sydney).

### 2.2. Southern DNA hybridization

Genomic DNAs were prepared from larvae using DNazol (Molecular Research Center) with the addition of RNase. DNA samples (~5 µg) were digested with indicated restriction enzymes, electrophoresed in 0.8% agarose gels, transferred to nylon membranes and immobilized by ultraviolet irradiation. Hybridizations were performed under moderate stringency conditions in phosphate buffer pH 7.5; 1% BSA; 7% SDS at 65 °C with an initial wash in 2 × SSC; 0.2% SDS at room temperature and two washes in 1 × SSC; 0.1% SDS at 55 °C for 30 min. Autoradiography was performed by exposure on Kodak X-Omat film at –80 °C. Various exposures were used to determine fragment identities and numbers.

### 2.3. PCR analysis

Direct PCR amplification was performed on genomic DNA samples prepared with DNazol (Molecular Research Center) using the Expand High Fidelity PCR

System (Roche) having proofreading capability under the following cycling conditions: 94 °C for 2 min, 94 °C for 30 s, 60 °C for 30 s, 68 °C for 2 min for 35 cycles with a final extension at 72 °C for 10 min. Amplified DNA was separated and visualized on 0.8% agarose gels. Isolated products were subcloned into TOPO TA cloning vectors (Invitrogen) and sequenced using M13 forward and reverse primers, and internal *piggyBac* primers, with most PCR products sequenced on both strands. Forward (F) and reverse (R) primers used for amplification of *piggyBac* sequences are given below. All positions are based on IFP2 sequence except for the forward primer Tn-gDNA which is located in the 5' insertion site genomic DNA of Tn-pBac5 (see Supplementary data 1). Inverse PCR for determining insertion site sequences was performed as described previously (Handler and Harrell, 2001).

1F: 5'-CCCTAGAAAGATAGTCTGCG-3'

445F: 5'-CCAGAGCGATACAGAAGAAGC-3'

1293R: 5'-ACTCCGTTGGTCTGTGTTCC-3'

2039R: 5'-TTGCCTTTCGCCTTATTTTAG-3'

2125R: 5'-CTTATTAGTCAGTCAGAAACAAC-3'

2443R: 5'-CGTTAAAGATAATCATGCGTAAAATT-GAC-3'

Tn-gDNA: 5'-TATGAGTTAAATCTTAAAAGTCACG-3'

### 2.4. Sequence analysis and comparison

Nucleotide and amino acid sequence analysis and comparisons were performed using MegAlign (DNAS-TAR, Inc.) and GeneWorks 2.5 (Oxford Molecular Group) software, and BLASTX and BLASTP (Altschul et al., 1997). Sequence pair distances were determined from a ClustalV multiple sequence alignment (Higgins et al., 1992), and parsimony analysis of ClustalX alignments (Thompson et al., 1997) was performed with PAUP\* v.4.0b10 (Swofford, 2001). Amino acid sequences were aligned by pairwise alignment to IFP2 with manual introduction of gaps and frameshifts to maintain a consensus sequence.

## 3. Results

### 3.1. Southern DNA hybridization

To determine the existence of *piggyBac* or *piggyBac*-related sequences with significant identity in moth species, Southern hybridization with *piggyBac* probes was performed on genomic DNA from *Ephestia kühniella*, *Galleria mellonella*, *Helicoverpa armigera*, *Helicoverpa zea*, *Heliothis virescens*, *Plodia interpunctella*, *Spodoptera frugiperda*, and *T. ni*. Digests were performed with *HindIII*, which does not cut within the *T. ni* IFP2 *piggyBac* element, and *BglII* which digests at nt 675 and *PstI* which digests at nt 1704 (Fig. 1). Using probe for the 5' and 3' regions of *piggyBac*, a single band for each *piggyBac* element was expected with *HindIII* digestion, with two bands resulting

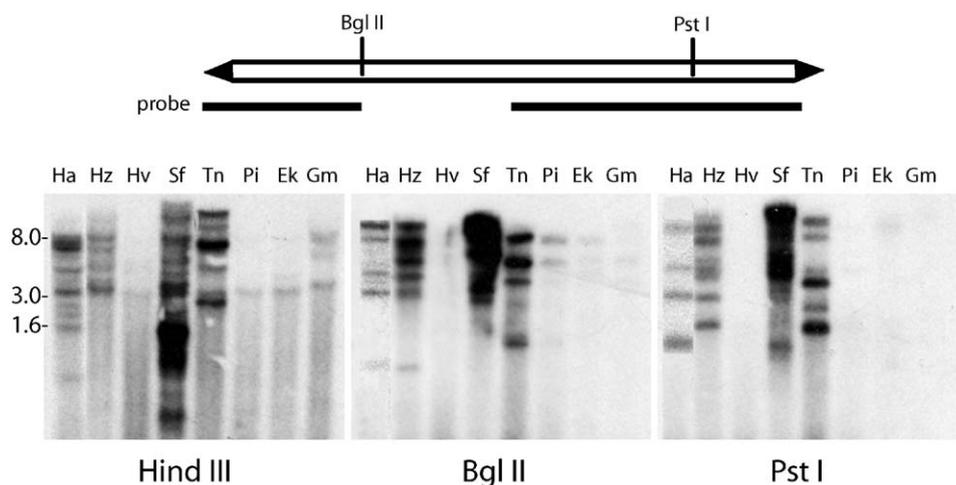


Fig. 1. Southern DNA hybridization analysis of *piggyBac* in genomic DNA from *H. armigera* (Ha), *H. zea* (Hz), *H. virescens* (Hv), *S. frugiperda* (Sf), *T. ni* (Tn), *P. interpunctella* (Pi), *E. kühniella* (Ek) and *G. mellonella* (Gm). For each blot, the restriction enzyme used for digestion is indicated with DNA size markers shown to the left of the autoradiograms. The schematic above shows relative *Bgl*II and *Pst*I restriction site positions within *piggyBac* IFP2, which does not contain a *Hind*III site. Fragments of *piggyBac* IFP2 used as 5' and 3' probe are indicated below.

from digests with the latter two enzymes. While the total digested DNA was nearly the same for all samples (data not shown), strong hybridization signals were obtained primarily from the *H. armigera*, *H. zea*, *S. frugiperda* and *T. ni* samples in all three blots (Fig. 1), with significantly weaker signals having similar hybridization profiles to one another observed in the other four species. Similar results were obtained from three or more independent hybridizations for each species (data not shown). From the *Hind*III digestion approximately eight bands were observed in *H. armigera*, five bands in *H. zea*, fifteen or more bands in *S. frugiperda* (observed in shorter blot exposures) and six bands in *T. ni*. These are minimum numbers since bands with higher intensity probably represent overlapping elements.

For the *H. virescens*, *P. interpunctella*, *E. kühniella*, and *G. mellonella* samples faint signals of two to three bands with *Hind*III, three to four bands for *Bgl*II, and one to two bands for *Pst*I were observed (some requiring several-fold longer blot exposures). For these hybridizations a majority of the fragments were of similar size. Faint signals were also observed from these species in additional hybridizations using different DNA preparations and *piggyBac* probes (data not shown).

### 3.2. Sequence analysis and comparisons

Total genomic DNA was assayed for the presence of *piggyBac* sequences using several sets of *piggyBac* primers for PCR amplification. The nucleotide positions for non-degenerate primers based on the original IFP2 *piggyBac* sequence, and a list of generated sequences are given in Table 1. While *piggyBac*-related sequences were generated from almost all species, the Southern hybridization data indicated an unambiguous source of *piggyBac* only from the genomes of *H. armigera*, *H. zea*, *S. frugiperda* and *T. ni*.

For the purpose of this analysis, we will focus on sequences from these species.

Primers pairs were first designed to generate products that include separate but overlapping segments of the *piggyBac* transposase coding region. PCR derived products with different combinations of primers were subcloned, sequenced and initially compared directly to the corresponding *piggyBac* IFP2 sequence (Table 1, Fig. 2). Selected representative sequences from each species and consensus sequences from their conceptual translation were compared to each other by multiple nucleotide and amino acid sequence alignments using ClustalV (Table 2).

#### 3.2.1. *T. ni* elements

Using *T. ni* genomic DNA as a template, PCR products were generated with almost every primer pair combination. From multiple products of different lengths, five or more unique *T. ni* sequences of approximately 1.6–2.3 kb, spanning most of the 1.8 kb transposase ORF are listed in Table 1. More unique sequences were generated (10 in total) but at present they have not been unambiguously distinguished from those listed (requiring distinct insertion sites).

The five *piggyBac* sequences from *T. ni* (Tn) had a small number of randomly distributed point mutations, no inserted nucleotides, and one sequence (Tn-pBac4) having a 6-bp deletion (Table 1). The identity between each of these elements and with IFP2 is greater than 99% (Table 2), and all have an uninterrupted ORF with the potential to generate functional transposase-products. Since IFP2 was originally discovered in infectious AcNPV (Fraser et al., 1983), two independent 5' integration sites in *T. ni* (one from Tn-pBac5) were identified by inverse PCR and their sequences subjected to BLASTX (Altschul et al., 1997), confirming that these *piggyBac* sequences were not derived from a contaminating baculovirus (see Supplementary

Table 1  
Summary of unique *piggyBac* sequences isolated by PCR

Sequence	GenBank accession nos.	Primers	Length <sup>a</sup>	Sequence comparison to IFP2 (acc no. J04364)		
				Single substitutions	Deletions <sup>b</sup>	Insertions
Tn-pBac1	DQ236231	445F-2039R	1595	3	—	—
Tn-pBac2	DQ236235	445F-2125R	1681	5	—	—
Tn-pBac3	DQ236237	445F-2125R	1681	4	—	—
Tn-pBac4	DQ236238	445F-2125R	1675	4	(6)	—
Tn-pBac5	DQ236240	Tn-gDNA-2443R	2443	10	—	—
Ha-pBac1	DQ236220	1F-1293R	1192	38	(3), (104)	2 × (1), (4)
Ha-pBac2	DQ236221	1F-1293R	1192	38	(3), (104)	2 × (1), (4)
Ha-pBac3	DQ236222	1F-1293R	1295	36	(3)	(1), (4)
Ha-pBac4	DQ236223	1F-2443R	777	149	(1667), 13 × (1), (2)	6 × (1), 5 × (2)
Ha-pBac5	DQ236224	1F-2443R	764	148	(1667), 14 × (1), (2)	7 × (1), 5 × (2)
Hz-pBac1	DQ236225	1F-1293R	1190	35	(3), (104)	2 × (1), (2)
Hz-pBac2	DQ236226	1F-1293R	1191	35	(3), (104)	(1), (4)
Hz-pBac3	DQ236227	1F-1293R	1260	36	(3), (28), (8)	2 × (1), (4)
Hz-pBac4	DQ236228	1F-2443R	781	140	(1667), 13 × (1), (2)	7 × (1), 6 × (2)
Hz-pBac5	DQ236229	1F-2443R	768	138	(1667), 13 × (1), (2), (4), (9)	7 × (1), 6 × (2)
Sf-pBac1	DQ236230	445F-2039R	1587	262	3 × (1), (2), (3)	—

<sup>a</sup>Length and sequence comparisons given in nucleotides.

<sup>b</sup>Contiguous nucleotides deleted or inserted given in parentheses; number of multiples indicated by '×'

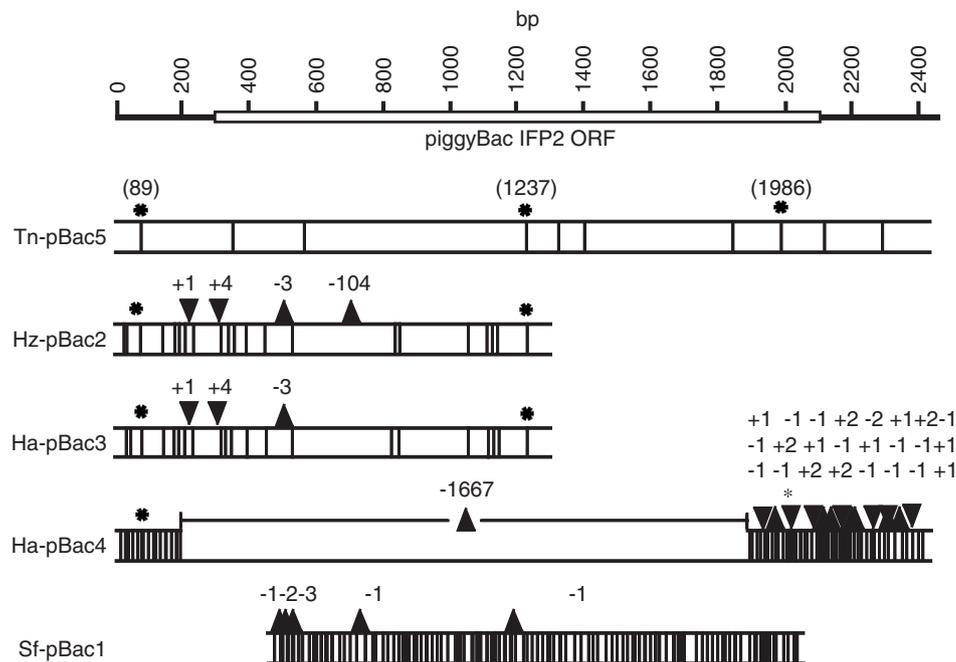


Fig. 2. Representative *piggyBac* nucleotide sequences from *T. ni* (Tn), *H. zea* (Hz), *H. armigera* (Ha) and *S. frugiperda* (Sf) compared to the analogous IFP2 sequence. Above is a schematic of the IFP2 *piggyBac* element with length in base pairs (bp) and its ORF (open bar). Sequences are presented as horizontal open bars with positions of single base pair substitutions (vertical lines), deletions (upward triangles), and insertions (downward triangles) indicated. Nucleotide numbers are shown above the deletions (–) and insertions (+). Asterisks mark signature base pair substitutions present in all sequences except those from *S. frugiperda*, with the nucleotide positions for Tn-pBac5 (and IFP2) in parentheses.

data 1). The 5' terminal and sub-terminal repeat sequences from these elements are identical to IFP2 (data not shown).

Two nucleotide substitutions from the IFP2 element were present in all *T. ni* sequences at IFP2 positions 89 (A to G) in the non-transcribed 5' region and 1237 (G to A) yielding a synonymous substitution (Thr303) (Fig. 2). A

third substitution present in Tn-pBac5 at position 1986 (T to C) is non-synonymous resulting in a Met553Tyr substitution. Despite greater overall degeneracy, the substitutions were also found in all elements from *H. armigera* and *H. zea* where the corresponding sites were sequenced, but not in *S. frugiperda* (Fig. 2).

Table 2  
Sequence pair distance table of *piggyBac* nucleotide and amino acid sequences aligned by ClustalV

		Percent nucleotide identity <sup>a</sup>										
		1	2	3	4	5	6	7	8	9	10	11
1	IFP2	—	99.6	99.9	99.8	95.4	95.3	96.8	95.8	95.5	95.1	78.5
2	Tn-pBac1	99.6	—	99.8	99.6	95.3	95.1	96.7	95.7	95.4	95.7	78.3
3	Tn-pBac3	100.0	99.6	—	99.9	95.7	95.5	97.0	96.1	95.8	96.0	78.6
4	Tn-pBac5	99.6	99.3	99.6	—	95.5	95.4	96.9	96.0	95.7	95.9	78.4
5	Ha-pBac1	93.5	93.1	93.5	93.1	—	99.1	96.9	99.3	99.2	91.4	75.8
6	Ha-pBac2	93.1	92.7	93.1	92.7	98.8	—	96.8	99.2	99.1	91.2	75.7
7	Ha-pBac3	95.0	94.7	95.0	94.7	94.7	94.3	—	97.3	97.2	97.4	77.3
8	Hz-pBac1	93.9	93.5	93.9	93.5	99.6	99.2	95.1	—	99.3	91.8	76.2
9	Hz-pBac2	93.5	93.1	93.5	93.1	98.4	98.0	94.7	98.8	—	91.6	76.1
10	Hz-pBac3	93.3	92.9	93.3	92.9	88.6	88.2	94.4	89.0	88.6	—	75.2
11	Sf-pBac1	79.6	79.3	79.6	79.3	75.6	75.2	77.1	76.0	75.6	74.0	—

		Percent amino acid identity <sup>b</sup>										
--	--	--	--	--	--	--	--	--	--	--	--	--

<sup>a</sup>Based on alignment to sequences corresponding to IFP2 nucleotides 445-1293.

<sup>b</sup>After conceptual translation and introduction of frameshifts to maintain optimized alignment with IFP2.

### 3.2.2. *Helicoverpa* elements

*H. zea* and *H. armigera* DNA gave rise to five PCR products each with the 1F-1293R and 1F-2443R primer pairs that span the 5' half, and nearly the complete IFP2 element, respectively (Table 1, Fig. 2). Sequences from both species were closely related to each other and IFP2, but more highly diverged than the *T. ni* sequences (Table 2, Fig. 2). They were placed into two groups based on structure, with representatives in both species. The 1F-1293R primer pair resulted in one group (Ha-pBac3/Hz-pBac3) of moderately degenerate sequences spanning ~1300 nucleotides, with a ~1200 nucleotide sub-group having a 104-bp deletion (Ha-pBac1,2/Hz-pBac1,2). A second group (Ha-pBac4,5/Hz-pBac4,5) had more highly degenerate ~0.8 kb sequences with a long common 1667 bp deletion spanning most of the ORF.

Only the 1.2–1.3 kb sequences were aligned with nucleotide identities >95% to IFP2 and 91–99% identity with each other (Table 2). Amino acid identities for these sequences were >93% to IFP2 and between 88 and 99% to each other. As noted, the common nucleotide substitutions identified in *T. ni* were also present in the *Helicoverpa* sequences. The existence of nearly identical sequences in both species that are significantly diverged from the functional IFP2 element, Ha-pBac4,5 and Hz-pBac4,5 in particular, indicates either interbreeding or a very recent divergence between the species, or recent horizontal transmission of the elements.

### 3.2.3. *S. frugiperda* elements

Repeated PCR of *S. frugiperda* genomic DNA with the primer sets yielded only three identical 1587 bp sequences with the 445F-2040R primer pair, that was most highly diverged relative to the other elements (Table 1, Fig. 2). Comparison of the common 848 bp sequence to IFP2 showed 78% identity with similar identities to the other

sequences, with amino acid identities ranging from 74% to 79% (Table 2). Unlike the *Helicoverpa* and *T. ni piggyBac* sequences, the conserved nucleotide substitutions were not evident in the *S. frugiperda* sequence. Southern hybridization in *S. frugiperda* indicated the greatest number of related elements, and thus the inability to amplify these sequences is likely due to degeneracy or deletions at the priming sites.

### 3.3. Phylogenetic analysis

To determine the phylogenetic relationship of the common overlapping 848 bp amplified sequence in the different lepidopteran species and to IFP2 *piggyBac* (nts 445-1293), unique nucleotide and amino acid consensus sequences from each species were compared by parsimony analysis of ClustalX multiple sequence alignments (see Supplementary data 2 for amino acid alignments). These relationships are presented graphically using the Neighbor-Joining method with bootstrap (1000 replications) and plotted as unrooted phylograms (Fig. 3). The clade structure of the trees is consistent, with the *H. armigera* and *H. zea* elements forming common clades independent from the *T. ni* clade. *S. frugiperda* sequences show the most distant relationships based on branch length. Direct comparisons of the *H. armigera* and *H. zea* nucleotide sequences, and their conceptual translation within the reading frame, suggest a common ancestor for the two types of elements. Comparison of the predicted amino acids in the consensus ORFs of representative sequences from the different species indicate that only the *T. ni* transposases have the potential for function (see Supplementary data 2). These sequences contain only infrequent single amino acid changes with uninterrupted open reading frames. Amino acid sequences from the other species are highly related to *piggyBac* transposase, however, multiple

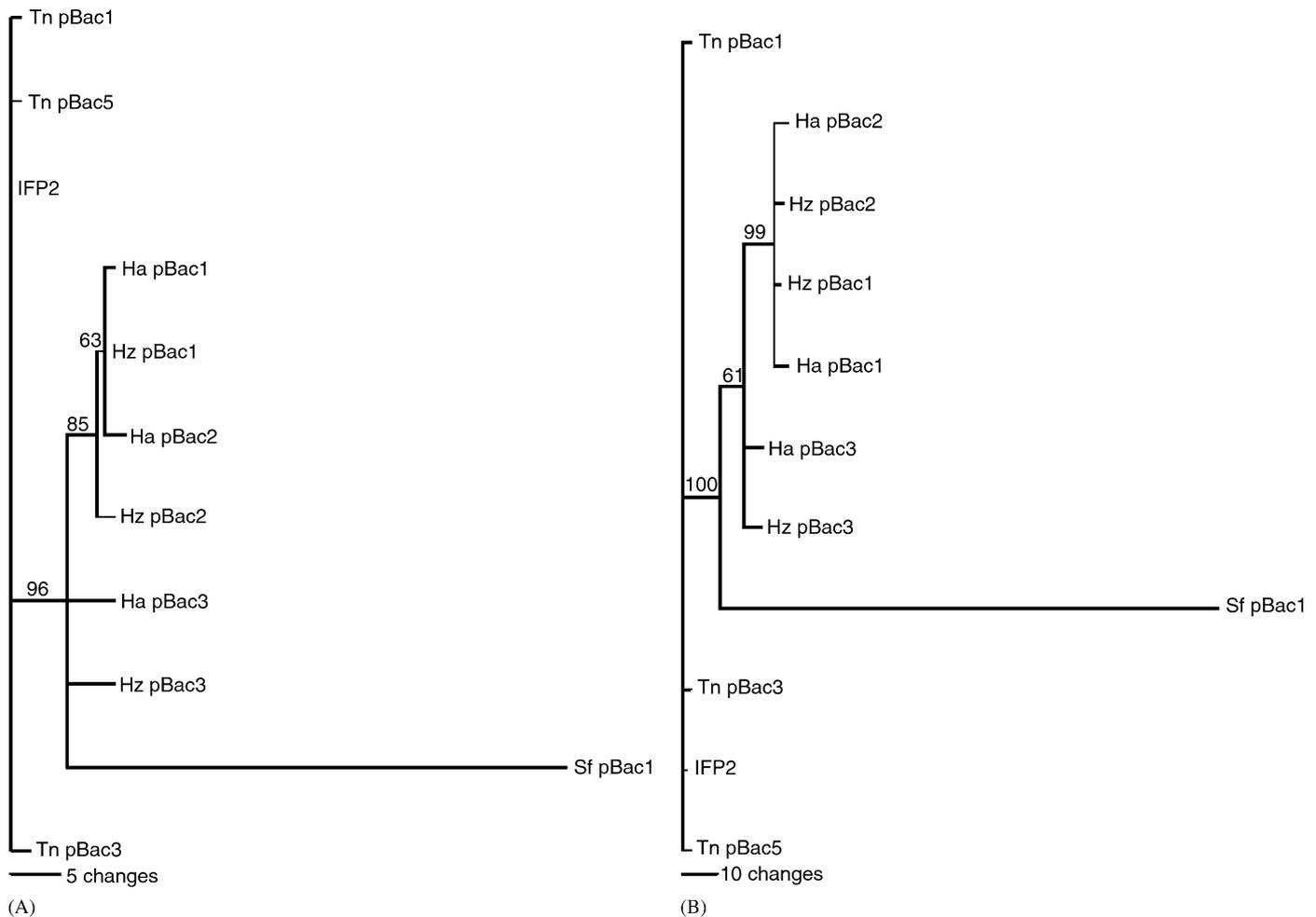


Fig. 3. Phylogenetic relationships of *piggyBac* nucleotide (A) and amino acid (B) sequences derived from the genomic DNA of lepidopteran species. Unrooted neighbor-joining (NJ) phylograms were generated from ClustalX multiple sequence alignments of the analogous sequences spanning nucleotides 445-1293 in IFP2, and the conceptual translation of these sequences after introduction of frameshifts to maintain a consensus sequence (see Supplementary data 2). Numbers at the nodes are bootstrap values (> 50%) based on 1000 replications of the data set calculated using PAUP\* v.4.0b10 (Swofford, 2001).

deletions and frame shifts required for the consensus transcript are unlikely to generate functional products.

#### 4. Discussion

A survey of several lepidopteran species for the presence of *piggyBac*, or closely related sequences, yielded positive hybridization evidence for four noctuid species, including *T. ni* in which the original IFP2 element was discovered (Fraser et al., 1983). Two of the species were the closely related *Helicoverpa* species, *H. zea* and *H. armigera* (Mitter et al., 1993) in addition to *S. frugiperda*. Fewer than 10 elements per genome could be estimated for the *Helicoverpa* species and *T. ni*, with potentially twice as many or more estimated for *S. frugiperda*. Faint hybridization signals sharing similar fragment size profiles were detected in multiple analyses of the noctuid *H. virescens*, and the pyralid species *P. interpunctella*, *E. kühniella*, and *G. mellonella*. While highly conserved *piggyBac* sequences were generated by PCR from some DNA preparations, a

genomic presence of closely related *piggyBac* elements in these species remains ambiguous. Potentially, the hybridization signals are to more highly diverged related elements, or to low copy elements that inhabit co-existing organisms, and this is a subject for further study. This survey for *piggyBac* elements is far from exhaustive, and thus it is likely that *piggyBac*-related elements of varying similarity exist in other Lepidoptera, and other insect species. The hybridization data, however, indicates that the presence of elements closely related to *piggyBac* in the Lepidoptera is discontinuous.

Nucleotide sequence analysis of amplified *piggyBac* DNA in the noctuids, and the conceptual translation of these sequences, supports the close relationship of these elements, with all sequence identities to *piggyBac* IFP2 above 95%, except for the more diverged *piggyBac*-related sequence from *S. frugiperda* having 78% identity. The most conserved sequences (>99% identity) were from *T. ni*, but notably, none of the unique sequences identified in organismal DNA were identical to the IFP2 *piggyBac*

isolated from the *T. ni* TN-368 cell line (Fraser et al., 1983; Cary et al., 1989). The several base substitutions in the *T. ni* sequences do not disrupt the transposase reading frame, and for at least two elements the 5' terminal sequence is identical to IFP2.

Three signature base substitutions from the IFP2 sequence were identified in all *T. ni* and *Helicoverpa* sequences, except those from *S. frugiperda*. Notably, these substitutions also exist in the *piggyBac* sequences amplified in the oriental fruit fly and other *Bactrocera* species (Handler and McCombs, 2000; G.J. Zimowska, A.M. Handler, unpublished). Only the third nucleotide substitution causes a non-synonymous amino acid change, and if this does not affect transposase function, then it is likely that IFP2 is a variant of a more ancestral sequence that became widespread.

Similar to other transposable elements, significantly diverged *piggyBac*-related elements have been found throughout the animal and plant kingdoms (Sarkar et al., 2003) which is consistent with the wide function of *piggyBac* in insects, as well as a mammalian species (Handler, 2002; Ding et al., 2005). To date, however, *piggyBac* is only one of three known class II transposons in which nearly identical elements are found in separate orders of insects. The *mariner*-related irritans subfamily *Himar* element is found in *Haemotobia irritans* and in *Chrysoperla plorabunda* (Robertson and Lampe, 1995), and mellifera subfamily elements are found in four different insect orders (Lampe et al., 2003). The phylogenetic distance between these insects is a strong indication that these elements traversed species horizontally, and it is proposed that these transfers are integral to the *mariner* life cycle (Lampe et al., 2003). Similarly, the existence of nearly identical *piggyBac* sequences in *T. ni* and *B. dorsalis* (Handler and McCombs, 2000) also indicated recent horizontal transmission. A systematic determination of horizontal transfer of *piggyBac* between noctuids is beyond the scope of this analysis, and is not straightforward given the relatively close, yet imprecise phylogenetic relationship of the taxa. Noctuoidea phylogeny based on morphological characteristics place Spodoptera and *Helicoverpa* in adjacent taxa (Noctuidae and Heliothinae, respectively) within the same sub-family clade, with Trichoplusia (Plutellinae) separated by 11 taxa in a different sub-family (Speidel et al., 1996). In contrast, molecular phylogenetic analysis of *Elongation factor-1 $\alpha$*  (Mitchell et al., 1997) and *Dopa decarboxylase* (Fang et al., 2000) support a closer relationship between Trichoplusia and Spodoptera than with *Helicoverpa*, with the latter study having Trichoplusia and Spodoptera in adjacent taxa with *Helicoverpa* most distantly related, separated by nine taxa. Nevertheless, none of these phylogenies are consistent with neutral evolution of the *piggyBac* sequences found in *T. ni* and *Helicoverpa*, that are much more closely related to the functional IFP2 element than the *S. frugiperda* sequence. If *T. ni* harbors only functional elements, or elements nearly identical to functional elements, this would strongly

suggest that these elements were recently introduced into the species, or that species-specific purifying selection has occurred.

The 0.8 and 1.3 kb *piggyBac* sequences in *Helicoverpa* are nearly identical to one another in the two species. If their terminal sequences remained intact, they may have been horizontally transmitted between species, but a simpler explanation is that both types of elements were established by interbreeding. Evidence exists for mating compatibility between particular strains of *H. zea* and *H. armigera* (Laster and Hardee, 1995; Laster and Sheng, 1995), and if the elements share the same genomic insertion sites in both species, this would support the notion that the *Helicoverpa* species are actually conspecific, as proposed in early studies (see Pogue, 2004).

The discovery of closely related *piggyBac* elements in noctuid species has implications for the use of *piggyBac* transformation vectors in these species, and for insects in general. In species where functional elements exist, such as *T. ni*, transgene stability is a primary concern, though if initial vector integration can be achieved, new vectors are available that can be stabilized post-integration by deletion of a terminal sequence (Handler et al., 2004). This will also limit concerns for remobilization of vectors where the potential source of functional transposase is less obvious. For species having multiple defective non-autonomous elements, such as *S. frugiperda*, achieving successful transformation may be impeded by potential repression of transposon movement similar to P cytotype repression, or dysgenic effects where mobilization of defective elements causes lethality or sterility (see Engels, 1996). While the influence of defective elements have only been clearly associated with P in *Drosophila*, and may depend on a relatively large number (~50 or more), concerted efforts to genetically transform *S. frugiperda* with *piggyBac* have thus far been unsuccessful (R. Harrell and A. Handler, unpublished; R. Nagoshi, personal communication). It is therefore important to be cognizant of caveats related to the use of *piggyBac* vectors in species containing autonomous or non-autonomous elements, and to further investigate potential negative interactions.

## Acknowledgements

Grateful appreciation is extended to Dr. Adam Williams, University of Melbourne, for sharing preliminary data on *piggyBac* in *Helicoverpa armigera*, and to Dr. Rod Mahon, University of Sydney, for providing *H. armigera* DNA samples. Funding from the USDA National Research Initiative—Competitive Grants Program is gratefully acknowledged.

## Appendix A. Supplementary materials

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ibmb.2006.03.001.

## References

- Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res.* 25, 3389–3402.
- Avancini, R.M., Walden, K.K., Robertson, H.M., 1996. The genomes of most animals have multiple members of the Tc1 family of transposable elements. *Genetica* 98, 131–140.
- Cary, L.C., Goebel, M., Corsaro, H.H., Wang, H.H., Rosen, E., Fraser, M.J., 1989. Transposon mutagenesis of baculoviruses: analysis of *Trichoplusia ni* transposon IFP2 insertions within the FP-Locus of nuclear polyhedrosis viruses. *Virology* 161, 8–17.
- Ding, S., Wu, X., Li, G., Han, M., Zhuang, Y., Xu, T., 2005. Efficient transposition of the *piggyBac* (PB) transposon in mammalian cells and mice. *Cell* 122, 473–483.
- Elick, T.A., Bauser, C.A., Principe, N.M., Fraser, M.J., 1995. 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.
- Engels, W.R., 1996. P Elements in *Drosophila*. In: Saedler, H., Gierl, A. (Eds.), *Transposable Elements*. Springer, Berlin, pp. 103–123.
- Fang, Q.Q., Mitchell, A., Regier, J.C., Mitter, C., Friedlander, T.P., Poole, R.W., 2000. Phylogenetic utility of the nuclear gene dopa decarboxylase in noctuid moths (Insecta: Lepidoptera: noctuoidea). *Mol. Phylogenet. Evol.* 15, 473–486.
- Fraser, M.J., Smith, G.E., Summers, M.D., 1983. Acquisition of host-cell DNA-sequences by baculoviruses—relationship between host DNA insertions and FP mutants of *Autographa-californica* and *Galleria-mellonella* nuclear polyhedrosis viruses. *J. Virol.* 47, 287–300.
- Fraser, M.J., Clszczon, Elick, T., Bauser, C., 1996. Precise excision of TTAA-specific lepidopteran transposons *piggyBac* (IFP2) and *tagalong* (TFP3) from the baculovirus genome in cell lines from two species of Lepidoptera. *Insect. Mol. Biol.* 5, 141–151.
- Handler, A.M., 2002. Use of the *piggyBac* transposon for germ-line transformation of insects. *Insect. Biochem. Mol. Biol.* 32, 1211–1220.
- Handler, A.M., Harrell, R.A., 2001. Transformation of the Caribbean fruit fly with a *piggyBac* transposon vector marked with polyubiquitin-regulated GFP. *Insect. Biochem. Mol. Biol.* 31, 199–205.
- Handler, A.M., 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., Zimowska, G.J., Horn, C., 2004. Post-integration stabilization of a transposon vector by terminal sequence deletion in *Drosophila melanogaster*. *Nat. Biotechnol.* 22, 1150–1154.
- Higgins, D.G., Bleasby, A.J., Fuchs, R., 1992. CLUSTAL V: improved software for multiple sequence alignment. *Comput. Appl. Biosci.* 8, 189–191.
- Kempken, F., Windhofer, F., 2001. The *hAT* family: a versatile transposon group common to plants, fungi, animals, and man. *Chromosoma* 110, 1–9.
- Lampe, D.J., Witherspoon, D.J., Soto-Adames, F.N., Robertson, H.M., 2003. Recent horizontal transfer of mellifera subfamily mariner transposons into insect lineages representing four different orders shows that selection acts only during horizontal transfer. *Mol. Biol. Evol.* 20, 554–562.
- Laster, M.L., Hardee, D.D., 1995. Intermating compatibility between North-American *Helicoverpa-zea* and *Heliothis-armigera* (Lepidoptera, noctuidae) from Russia. *J. Econ. Entomol.* 88, 77–80.
- Laster, M.L., Sheng, C.F., 1995. Search for hybrid sterility for *Helicoverpa-zea* in crosses between the North-American *Heliothis-zea* and *Helicoverpa-armigera* (Lepidoptera, noctuidae) from China. *J. Econ. Entomol.* 88, 1288–1291.
- Mitchell, A., Cho, S., Regier, J.C., Mitter, C., Poole, R.W., Matthews, M., 1997. Phylogenetic utility of elongation factor-1 alpha in noctuoidea (Insecta: Lepidoptera): the limits of synonymous substitution. *Mol. Biol. Evol.* 14, 381–390.
- Mitter, C., Poole, R.W., Matthews, M., 1993. Biosystematics of the Heliothinae (Lepidoptera: Noctuidae). *Annu. Rev. Entomol.* 38, 207–225.
- Pogue, M.G., 2004. A new synonym of *Helicoverpa zea* (Boddie) and differentiation of adult males of *H. zea* and *H. armigera* (Hübner) (Lepidoptera: Noctuidae: heliothinae). *Ann. Entomol. Soc. Am.* 97, 1222–1226.
- Robertson, H.M., Lampe, D.J., 1995. Recent horizontal transfer of a mariner transposable element among and between Diptera and Neuroptera. *Mol. Biol. Evol.* 12, 850–862.
- Sarkar, A., Sim, C., Hong, Y.S., Hogan, J.R., Fraser, M.J., Robertson, H.M., Collins, F.H., 2003. Molecular evolutionary analysis of the widespread *piggyBac* transposon family and related “domesticated” sequences. *Mol. Genet. Genomics* 270, 173–180.
- Speidel, W., Fänger, H., Naumann, C.M., 1996. The phylogeny of the Noctuidae (Lepidoptera). *Systematic Entomol.* 21, 219–252.
- Swofford, D.L., 2001. PAUP\*: phylogenetic analysis using parsimony (\*and other methods). Sinauer Associates, Sunderland, MA.
- Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucl. Acids Res.* 25, 4876–4882.