Insect Molecular Biology

Insect Molecular Biology (2009) **18**(6), 747–757 doi: 10.1111/j.1365-2583.2009.00917.x

Molecular analysis of a mutant *Anticarsia gemmatalis multiple nucleopolyhedrovirus* (AgMNPV) shows an interruption of an inhibitor of apoptosis gene (*iap-3*) by a new class-II *piggyBac*-related insect transposon

M. P. Carpes*, J. F. Nunes*, T. L. Sampaio*, M. E. B. Castro†, P. M. A. Zanotto‡ and B. M. Ribeiro*

**Laboratório de Microscopia Eletrônica, Departamento de Biologia Celular, Universidade de Brasília, Brasília, DF, Brazil; †Embrapa Recursos Genéticos e Biotecnologia-Núcleo Temático de Controle Biológico (NTCB), Brasília, DF, Brazil; and ‡Laboratório de Evolução Molecular e Bionformática, Departamento de Microbiologia, Instituto de Ciências Biomédicas, Universidade de São Paulo, SP, Brazil*

Abstract

A new *piggyBac*–related transposable element (TE) was found in the genome of a mutant *Anticarsia gemmatalis multiple nucleopolyhedrovirus* interrupting an inhibitor of apoptosis gene. This mutant virus induces apoptosis upon infection of an *Anticarsia gemmatalis* cell line, but not in a *Trichoplusia ni* cell line. The sequence of the new TE (which was named IDT for *iap disruptor transposon*) has 2531 bp with two DNA sequences flanking a putative Transposase (Tpase) ORF of 1719 bp coding for a protein with 572 amino acids. These structural features are similar to the *piggyBac* TE, also reported for the first time in the genome of a baculovirus. We have also isolated variants of this new TE from different lepidopteran insect cells and compared their Tpase sequences.

Keywords: *piggyBac*, *Anticarsia gemmatalis*, vApAg, transposon, AgMNPV.

Introduction

Insect virology has been an active area of research, mainly because of the baculoviruses. This large family of double-stranded DNA viruses has several useful applications in biotechnology. In Brazil, the baculovirus *Anticarsia gemmatalis multiple nucleopolyhedrovirus* (AgMNPV) has been used for the biological control of the velvet bean caterpillar, *Anticarsia gemmatalis* (Lepidotera: Noctuidae), since the beginning of the 80s and is the world's most successful biological control program using a virus, with more than one million soy bean hectares being applied with this virus (Moscardi, 1999; Oliveira *et al*., 2006). Baculoviruses are also used as expression vectors for heterologous protein production (Kost *et al*., 2005) and more recently, they have shown a potential to be used in gene therapy, since they have been shown to enter mammalian cells, but do not replicate in them (Hu, 2006; Gao *et al*., 2007).

Upon viral infection, an insect host cell can activate a complex cell death mechanism known as apoptosis. Cell death is carried out by a class of proteases known as caspases (cysteine-aspartate proteases). Initiator caspases cleave and trigger the activation of effector caspases that effectively kill the cell by means of a protein complex called apoptosome (Zimmermann *et al*., 2001).

Since apoptosis is an important antiviral response in insects (Zhang *et al*., 2002; Clarke & Clem, 2003a,b; Silveira *et al*., 2005, 2007), viruses must make use of countermeasures against it. The baculoviruses can have two types of apoptosis inhibitor proteins, the P35 and IAP (inhibitor of apoptosis) proteins (Clem, 2005). The pan caspase inhibitor P35 is found only in baculoviruses and recently a homolog was found in an entomopoxvirus (Means *et al*., 2007) and it has been shown to block apoptosis triggered by several stimuli, including viral infections (Clem, 2005). The *iap* gene family comprises inhibitors of apoptosis proteins that are found not only in baculoviruses, but also in *Drosophila*, nematodes, and even in mammals.

In a previous work, we have isolated a mutant baculovirus (derived from AgMNPV) that caused apoptosis in an *Anticarsia gemmatalis* cell line (UFL-AG-286), but not in a

First published online 25 September 2009.

Correspondence: Professor Bergmann Morais Ribeiro, Laboratório de Microscopia Eletrônica, Departamento de Biologia Celular, Universidade de Brasília, Brasília, DF, CEP 70910-900, Brazil. Tel.: +55 61 3307 2424; fax: +55 61 3347 6533; e-mail: bergmann@unb.br

Trichoplusia ni cell line (BTI-Tn5B1-4) (Silveira *et al.*, 1999). This virus was also shown to induce apoptosis *in vivo* (Silveira *et al.*, 2007). We and others have sequenced the complete AgMNPV genome recently (Oliveira *et al.*, 2006), GenBank accession number DQ813662) and shown that AgMNPV possess three *iap* genes (*iap-1*, *-2* and *-3*).

Transposable elements (TEs) represent a type of mobile genetic element that can move into genomes, being found in almost all organisms, and acting as a powerful agent of evolution (Feschotte & Pritham, 2007). TEs can be classified into two major classes: The retrotransposons, which use an RNA intermediate for transposition like a retrovirus, and DNA transposons, which mobilize themselves in a classic 'cut-and–paste' mechanism, in which the TE is removed from the donor site and reinserted into another DNA region (Feschotte & Pritham, 2007). One simple type of TE is known as insertion sequence (IS), a small DNA element that catalyses its own movement or recombination by a multifunctional enzyme known as Transposase (Tpase).

It is very common to find TEs in any organism genome, but less common is to find an active TE. One example is a widely studied and known insect TE, called *piggyBac*. This lepidopteran TE was found in a mutant baculovirus called FP (few polyhedra) in 1983 and was derived from the cabbage looper *T. ni* insect cell line TN-368 (Fraser *et al.*, 1983). Initially called IFP-2 at that time, this TE has an internal open reading frame (ORF) of around 1.8 kb coding for a Tpase, which is needed for the transposition event. This TE has been used in genetic manipulation, such as the production of genetically modified insects (Handler, 2002). Several types of elements that resemble *piggyBac* have been found in the genome of different insects, and even in other organisms (Handler & McCombs, 2000), characterizing a superfamily of TEs, known today as the *piggyBac* superfamily (Zimowska & Handler, 2006; Feschotte & Pritham, 2007; Sun *et al.*, 2008). Baculoviruses may play an important role in the spread of these elements in different insect genomes (Sarkar *et al.*, 2003), since these viruses can, in some cases, infect more than one lepidopteran species, and by doing so they can contribute to the horizontal spread of these TEs in nature (Handler, 2002).

In this work, we present evidence that a genetic alteration in a mutant *Anticarsia gemmatalis multiple nucleopolyhedrovirus* (called vApAg) making it incapable of blocking apoptosis in *A. gemmatalis* cells, resulted from the insertion of a novel class II insect TE into the *iap-3* gene. We have also isolated variants of this new TE from different lepidopteran insect cells and compared their Tpase sequences.

Results

Analysis of the iap *genes from the mutant virus*

In order to verify if the vApAg *iap* genes had any modification, we performed PCR reactions using the AgMNPV and vApAg DNAs as templates and specific oligonucleotides, originally designed for the *iap* genes of the wild type AgMNPV virus (Fig. 1). All PCR reactions amplified expected fragments of the predicted size, except for the *iap-3 gene* (Fig. 1). The PCR reaction for the *iap-3* gene of the wild-type AgMNPV virus amplified an expected fragment of around 900 bp (Carpes *et al.*, 2005). However, the PCR reaction with the vApAg DNA showed a fragment of around 3500 bp. The sequence analysis of this unexpected PCR fragment showed an insertion of 2531 bp (Fig. 2) that interrupted the *iap-3* gene after a TTAA tetranucleotide at the position + 59 related to the translation start codon of the *iap-3* ORF (Fig. 3). This insertion caused the duplication of the *iap*-3 tetranucleotide TTAA, a very common event in transposition of the TTAA family of TEs. The 2531 bp TE that interrupted the *iap*-3 gene had two DNA sequences flanking a putative Tpase ORF of

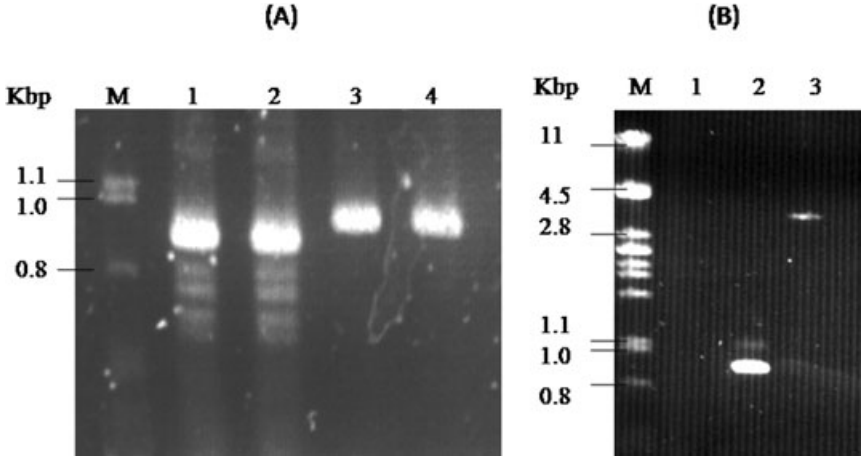

Figure 1. Inhibitor of apoptosis (*iap*) genes PCR amplifications from wild-type AgMNPV and mutant vApAg viruses. (A) Agarose gel (0.8%) electrophoresis of PCR products using oligonucleotides specific for *iap-1* (lanes 1 and 2) and *iap-2* (lanes 3 and 4) genes and DNA from wild-type AgMNPV (lanes 1 and 3) and mutant vApAg (lanes 2 and 4) viruses. (B) Agarose gel (0.8%) electrophoresis of PCR products using oligonucleotides specific for the *iap-3* (lanes 2 and 3) and DNA from wild-type AgMNPV (lanes 2) and mutant vApAg (lane 3) viruses. Lane 1 in (B) is the result of a PCR reaction without DNA. Lane M in (A) and (B) is the molecular mass marker lambda phage DNA digested with *Pst*I restriction enzyme.

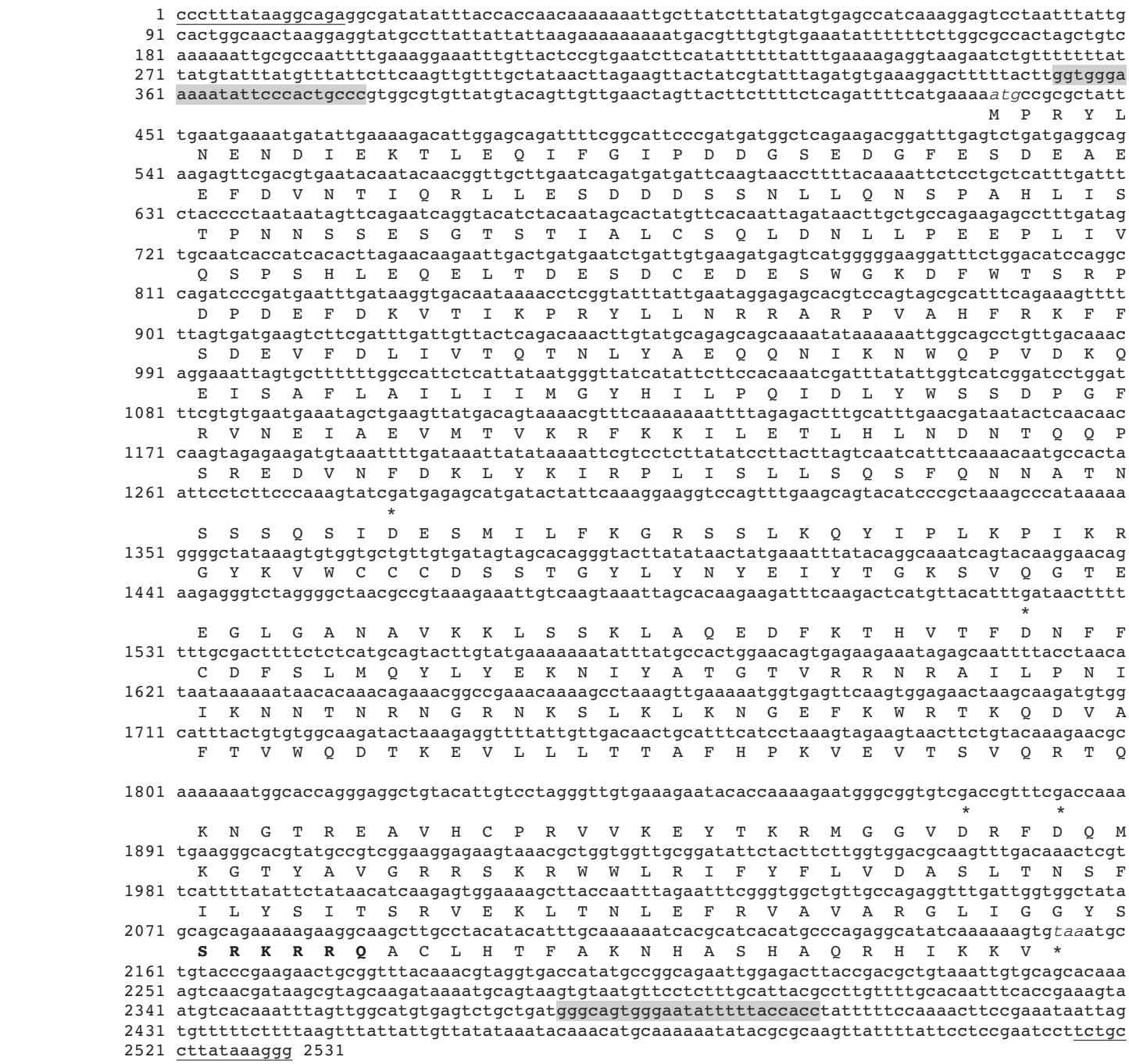

Figure 2. Complete nucleotide and amino acid sequences of the *iap disruptor transposon* (IDT). A flanking DNA region upstream of the Tpase ORF has 437 nucleotides, followed by the Tpase ORF (nucleotides 438–2157), and a downstream region of 373 nucleotides. The start and stop codons of the Tpase ORF are shown in italics. The ITRs (inverted terminal repeats) in the upstream (nucleotides 1–16) and downstream (nucleotides 2515–2531) regions are underlined. The IIRs (inverted internal repeats) are also shown in grey (nucleotides 354–379 and nucleotides 2377–2399). The Tpase ORF has 1179 coding for a protein with 572 amino acids. Four aspartate (D) residues, thought to be part of a putative DDE/DDD motif, are shown with an asterisk (*) above the sequence. A putative nuclear localization signal was detected at amino acids 546–551 (bold letters SRKRRQ) by analyzing the protein sequence at the PredictNLS server (prediction and analysis of nuclear localization signals) at the web site: http://cubic.bioc.columbia.edu/predictnls/.

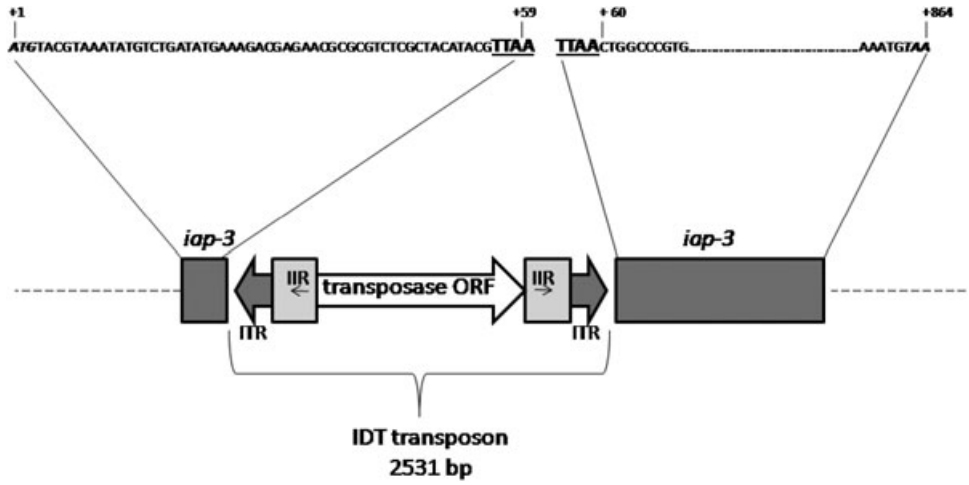

Figure 3. Diagram of the IDT TE insertion into the *iap-3* ORF of vApAg mutant virus. Sequence analysis of the *iap-3* gene of the mutant vApAg virus showed the insertion of a 2531 pb fragment into the *iap-3* ORF after nucleotide +59 relative to the start codon of the *iap*-3 gene. The tetranucleotide TTAA was duplicated after insertion which is characteristic of a transposition event by a *piggyBac*-related TE. ITR and IIR represent the terminal inverted repeats and internal inverted repeats found in the IDT sequence. The positions of selected nucleotides in the *iap-3* ORF are shown above the diagram.

1719 bp coding for a protein with 572 amino acids (Fig. 2). The right and left flanking DNA sequences (Fig. 2) had two perfect inverted terminal repeats (ITRs) of 16 nucleotides (Fig. 2). We have also found two almost perfect internal inverted repeats (IIRs) of 25 bp in length on the left and right sequences of the TE (Fig. 2). These structural features of two inverted terminal repeats and internal repeats within two sequences flanking a putative Tpase ORF is normally found in TEs known as insertion sequences (IS), like the *piggyBac* TE. Besides these structural features, the insertion site (TTAA) of the IDT TE is identical to the *piggyBac* TE, also reported for the first time in the genome of a baculovirus (Fraser *et al*., 1983). This new TE found in the mutant baculovirus vApAg was named IDT (*iap disruptor transposon*).

Amplification, cloning and sequencing of the IDT-related Tpase from different insect cell lines

IDT-related sequences in the genome of different lepidopteran insects were searched by PCR using oligonucleotides specific for the inverted repeat (ITR) flanking the IDT TE and DNA from three insect cell lines (UFL-AG-286, BTI-Tn5B1-4 and IPLB-SF21-AE). All amplification reactions produced 2500 bp fragments, indicating the presence of this type of TE in at least three lepidopteran insects (data not shown). Primers were then designed for the amplification of the putative Tpase ORFs from the three TEs. DNA fragments (around 1700 bp) of the putative Tpases from the three cell lines were amplified, cloned and sequenced (not shown). The sequence analysis showed that the Tpase ORF from the UFL-AG-286 cell line had a 100% identity with the IDT Tpase isolated from the mutant vApAg. This result was expected, since the mutant virus probably acquired the TE from this cell line. The UFL-AG-286 Tpase ORF was shown to have 1719 bp coding for a putative protein of 572 amino acids (Fig. 2) with a predicted mass of 66.06 kDa. On the other hand, the BTI-Tn5B1-4 and IPLB-SF21-AE Tpases were shown to have ORFs of 1707 bp coding for proteins with 568 amino acids and predicted molecular mass of 65.43 and 65.42 kDa, respectively (data not shown).

Comparative sequence analysis of the Tpases

We compared the Tpase sequences of the IDT TE found in the mutant vApAg virus, with other lepidopteran-selected Tpases from *piggyBac*-related TEs by BLASTP analysis (Table 1).

The comparative analysis among different Tpases (Table 1) showed that the IDT Tpase was 22 amino acids shorter than the *piggyBac* Tpase, and with 26–46% amino acid identity and similarity, respectively. As expected, the IDT Tpase was identical to the UFL-AG-286 Tpase (not shown) and was shown to have greater than 91% identity with the *T. ni* (TN5B) and *Spodoptera frugiperda* (SF21) Tpases described in this work. These insect cell Tpases have 25% and 44–45% amino acid identity and similarity, respectively, with the *piggyBac* Tpase (Table 1).

Multiple alignments of the selected Tpases showed that although the different transposases have considerable sequence variability, they share several blocks of conserved amino acids (Fig. 4). A putative DDE/DDD motif was found in all Tpases analysed, and a putative nuclear localization signal was detected at amino acids 546-551 (SRKRRQ) only in the IDT, TN5B and SF21 Tpases by

Table 1. Percentage of identity and similarity at the amino acid level of *piggyBac*-related transposases (Tpases) found in lepidopteran insects. The data presented were obtained using the BLASTP program at the NCBI home page (http://www.ncbi.nlm.nih.gov/)

	Amino acid identity/similarity (%)							
Transposases (Tpase)	IDT. Tpase	TN5B. Tpase	SF21. Tpase	*piggyBac*. Tpase	HaPLE1. Tpase	HvPLE 1.1. Tpase	yabusame-1. Tpase	yabusame-W. Tpase
IDT. Tpase	100/100	91/95	91/94	26/46	26/47	27/47	24/44	24/43
TN5B. Tpase		100/100	99/99	25/45	25/46	27/47	23/44	23/43
SF21. Tpase			100/100	25/44	25/46	27/47	23/45	23/44
piggyBac. Tpase				100/100	63/78	33/51	36/51	35/50
HaPLE1. Tpase					100/100	32/51	33/50	33/50
HvPLE1.1. Tpase						100/100	40/59	40/58
yabusame-1. Tpase							100/100	99/99
yabusame-W. Tpase								100/100

analysing the protein sequence at the PredictNLS server (prediction and analysis of nuclear localization signals) at the web site: http://cubic.bioc.columbia.edu/predictnls/. The MCC tree obtained by maximum likelihood gave a high support ($P = 1.0$) for the association of the UFL-AG-286 Tpase to the IDT Tpase (Fig. 5). These IDT-related Tpases were a sister group of the Tpase from *Nasonia vitripennis* ($P = 0.89$) and had a distant association to the Tpases of mammals ($P = 0.57$) (see Table 2 and Fig. 5). Crucially, our tree indicated that several highly divergent groups of *piggyBac*-related Tpases do infect animals, since there was not clear association of deep lineages to a specific group of animals.

Discussion

The AgMNPV mutant virus known as vApAg was isolated in 1999 and shown to cause apoptosis in UFL-AG-286 insect cells (Silveira *et al*., 1999) and we have previously demonstrated that the baculovirus AgMNPV has an active inhibitor of apoptosis gene known as *iap*-3 (Carpes *et al*., 2005). PCR amplification of the *iap* genes of the wild-type and vApAg virus, revealed a TE of 2531 bp inserted in the middle of the *iap*-3 gene of the mutant virus, disrupting the *iap*-3 ORF (Figs 1, 2).

The IAP-3 proteins from AgMNPV (Carpes *et al*., 2005), *Orgya pseudotsugata multiple nucleopolyhedrovirus* (OpMNPV) (Birnbaum *et al*., 1994), *Leucania separata multiple nuclear polyhedrovirus* (LsMNPV) (Kim *et al*., 2007), *Hyphantria cunea nucleopolyhedrovirus* (HycuNPV) (Ikeda *et al*., 2004), *Cydia pomonella granulovirus* (CpGV) (Crook *et al*., 1993) have been shown to have antiapoptotic activity. Furthermore, the silencing of the OpMNPV *iap*-3 gene results in apoptosis during infection of *Lymantria dispar* insect cells (Ld652Y) (Means *et al*., 2003). However, removal of the *iap-1* and *iap-2* from AcMNPV had no effect on the ability of the virus to block apoptosis and on virus replication in permissive and semi-permissive cell lines (Griffiths *et al*., 1999; McLachlin *et al*., 2001).

The sequence analysis of the IDT TE showed that it has two ITRs (inverted terminal repeats) of 16 bp, and two internal inverted repeats of 25 bp (Figs 2, 3). By comparing the IDT TE with the *piggyBac* TE from *T. ni*, we found that they have a very similar structure, but are very different at the sequence level. The *piggyBac* TE has ITRs of 13 bp and internal repeated sequences of 19 bp (Fraser *et al*., 1996). In most *piggyBac*-related TEs, the ITR regions begin with a CCC/GGG trinucleotide which is the case for IDT, but some TEs can begin with the CAC/GTG trinucleotide (Li *et al*., 2005). These inverted terminal palindromic regions are essential to the transposition event (Gueguen *et al*., 2005). The total size of the two elements is almost the same, around 2500 bp, but the putative Tpase protein of IDT is 22 amino acids shorter (572 amino acids) than the *piggyBac* Tpase (594 amino acids), and has only 26% amino acid identity by BLASTP analysis (Table 1). Despite the low amino acid identity, the IDT TE has the same basic features of the *piggyBac* TE and may represent a new member of this family. Since the isolation of *piggyBac*, *piggyBac*-related sequences to this TE were found in a wide array of organisms, including mammals (Sarkar *et al*., 2003).

Since the IDT element was found inserted into the *iap-3* of the mutant vApAg virus and the tetranucleotide TTAA was found duplicated at the site of insertion, we believe that it has moved from the insect genome into the virus genome with the help of its own Tpase or another active insect Tpase. We amplified the IDT TE from total DNA of the UFL-AG-286 insect cells, which is the natural host cell for the AgMNPV wild type virus. We also amplified other IDT-like related Tpases in *T. ni* (BTI-Tn5B1-4)

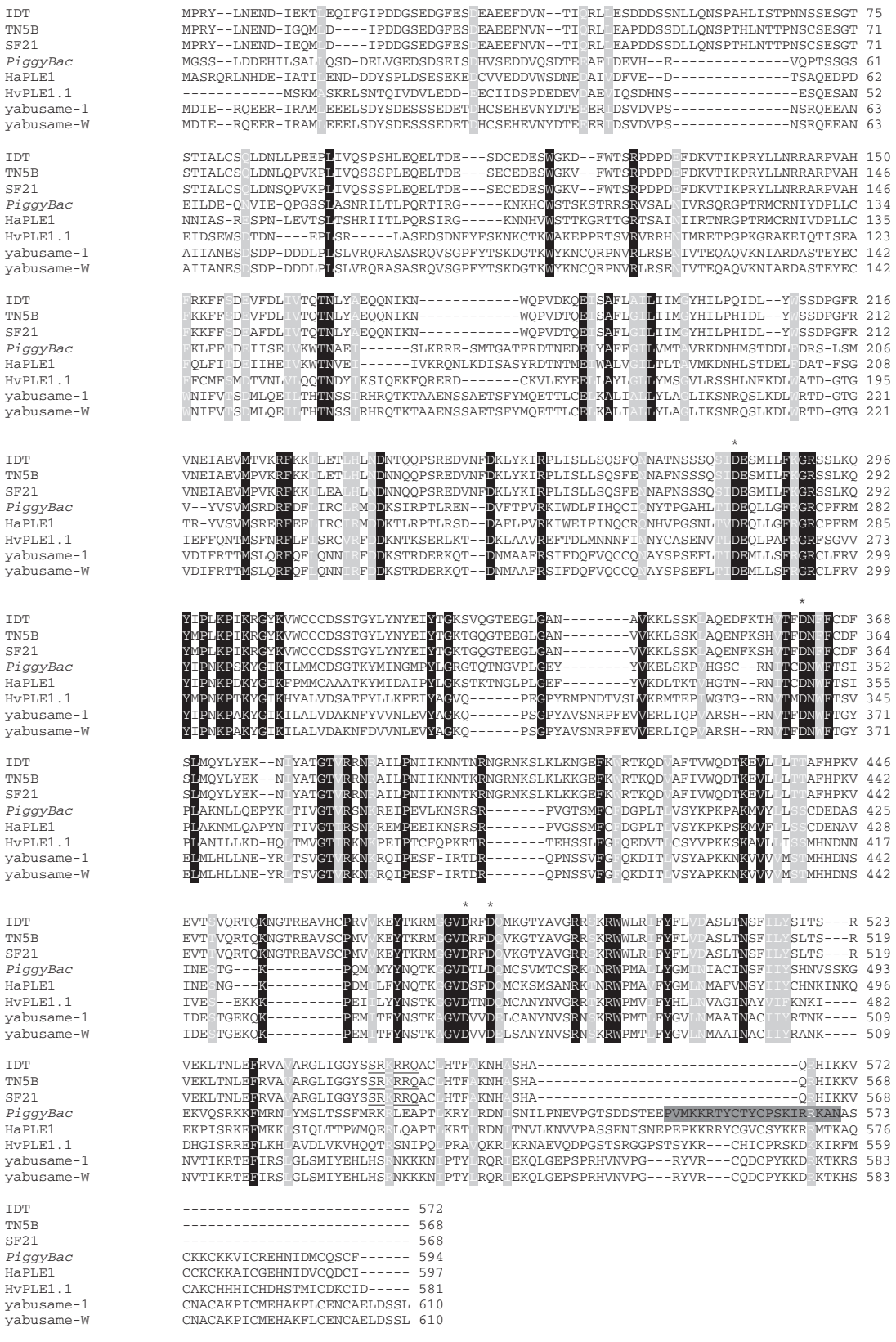

Figure 4. Multiple sequence alignment of Tpases from *piggyBac*-related TEs found in lepidopterans. Identical amino acids are shown in a black background and amino acids of the same chemical family are shown in a light grey background. The N-terminal and C-terminal regions are more variable with the central region (amino acids 150–520 related to the IDT Tpase) having more conserved amino acids. The conserved aspartic acid residues thought to be part of a putative DDD domain of Tpases from the *piggyBac* superfamily (Sarkar *et al*., 2003) are marked with an asterisk (*) above the sequences. The functional nuclear localization signal (NLS) at the C-terminal portion of the *piggyBac* Tpase between amino acids 551 and 571 (dark grey background) was identified by Keith *et al.* (2008b) and was not found in the IDT Tpase amino acid sequence. A predicted NLS (SRKRRQ) of the IDT, TN5B and SF21 Tpases are shown underlined.

and *S. frugiperda* (IPLB-SF21-AE) insect cells. The putative Tpases of BTI-Tn5B1-4 and IPLB-SF21-AE cell lines characterized in this work have 91% amino acid identity to the IDT Tpase and may represent variants of the same protein. The alignment of Tpases from selected lepidopteran-*piggyBac*-related TEs showed regions of conserved amino acids, including the putative 'DDE/DDD' motif at amino acid positions D268, D346, and D447 of the *piggyBac* Tpase (Sarkar *et al*., 2003) and the putative nearly universally conserved 10 amino acid motif (GTVRxNKRxIP) among *piggyBac*-related Tpases, located at amino acid positions 369–739. Keith *et al.* (2008a) have recently shown that mutation of the conserved three aspartic acids and another highly conserved aspartic acid residue (D450) close to the third proposed aspartic acid (D447) are necessary for *piggyBac* Tpase activity. Keith *et al.* (2008b) have identified a functional nuclear localization signal at the C-terminal portion of the *piggyBac* Tpase between amino acids 551 and 571 which was not found in the IDT Tpase amino acid sequence. A possible zinc finger motif at the C-terminal of the *piggyBac* Tpase rich in cysteines was not found in the IDT Tpase either.

The *piggyBac* TE has been used as DNA transfer vector for germ-line or strain transformation of *Plasmodium falsiparum* (Balu *et al*., 2005), *Schistosoma mansoni* (Morales *et al*., 2007) and different insects (Handler *et al*., 1998; Berghammer *et al*., 1999; Handler & Harrell, 1999, 2001; Handler & McCombs, 2000; Peloquin *et al*., 2000; Tamura *et al*., 2000; Grossman *et al*., 2001; Hediger *et al*., 2001; Kokoza *et al*., 2001; Heinrich *et al*., 2002; Nolan *et al*., 2002; Perera *et al*., 2002; Sumitani *et al*., 2003; Allen *et al*., 2004). This TE is the only known active TE found in lepidopterans (Handler, 2002; Sun *et al*., 2008).

The UFL-AG-286 Tpase is identical to the IDT TE, which supports the hypothesis that this TE moved from the genome of this cell to the baculovirus AgMNPV genome during infection, and probably, by disrupting the *iap*-3 gene, the mutant vApAg became incapable of blocking apoptosis during infection of UFL-AG-286 cells. However, we cannot rule out the possibility of other mutations in the vApAg genome that may also be responsible for the altered phenotype. In order to confirm that, we plan to construct a recombinant AgMNPV with the *iap-3* deleted, or use RNAi technology to silence the gene.

Table 2. Twenty-three sequences used for phylogenetic reconstruction of *piggyBac*-related transposase (Tpase) proteins

Taxon on tree	Source	Genbank accession #
piggyBac	*piggyBac* transposase (*piggyBac* helper plasmid pBlu-uTp)	AAO43224.1
HaPLE1	*Helicoverpa armigera piggyBac*-like transposable element HaPLE1	EF593176
yabusame1	*Bombyx mori* putative transposase yabusame-1	AB162707.1
yabusameW	*B. mori* gene for putative transposase yabusame-W	AB159601
BmoriAB	*B. mori* gene for yabusame-2	AB162708
HvPLE1_1	*Heliothis virescens piggyBac*-like transposable element HvPLE1.1	DQ407726
IDT_TN5B	*Trichoplusia ni* IDT-like putative transposase	This work
IDT_SF21	*Spodoptera frugiperda* IDT-like putative transposase	This work
IDT_vApAg	Mutant AgMNPV virus IDT putative transposase	This work
Nvitripen	*Nasonia vitripennis* XP_001599370	XP_001599370
Mfasc	*Macaca fascicularis* testis cDNA clone QtsA-11460	BAE02063
HspNM53	*Homo sapiens piggyBac* transposable element derived 3 (PGBD3)	NM_170753
HspNM25	*H. sapiens piggyBac* transposable element derived 2 (PGBD2)	NM_170725
HspNM0325	*H. sapiens piggyBac* transposable element derived 1 (PGBD1)	NM_032507
rtNM17182	*Mus musculus piggyBac* derived element	NM_171824
HspNM0245	*H. sapiens piggyBac* transposable element derived 5 (PGBD5)	NM_024554
DmeAY1191	*Drosophila melanogaster*	AY119121.1
DmeAE0135	*D. melanogaster*	AE013599.4
DpulAY11	*Daphnia pulicaria* transposon Pokey 6.6kb element putative transposase gene	AY115589
HspNM1525	*H. sapiens piggyBac* transposable element derived 1 (PGBD1)	NM_152595
AgamXM312	*A. gambiae* str. PEST *piggyBac*-derived 1	XM_312615, XM_312615.2
AgamXM320	*A. gambiae* str. PEST *piggyBac*-derived 2	XM_320414.2
DmeAY0696	*D. melanogaster*	AY069639.1

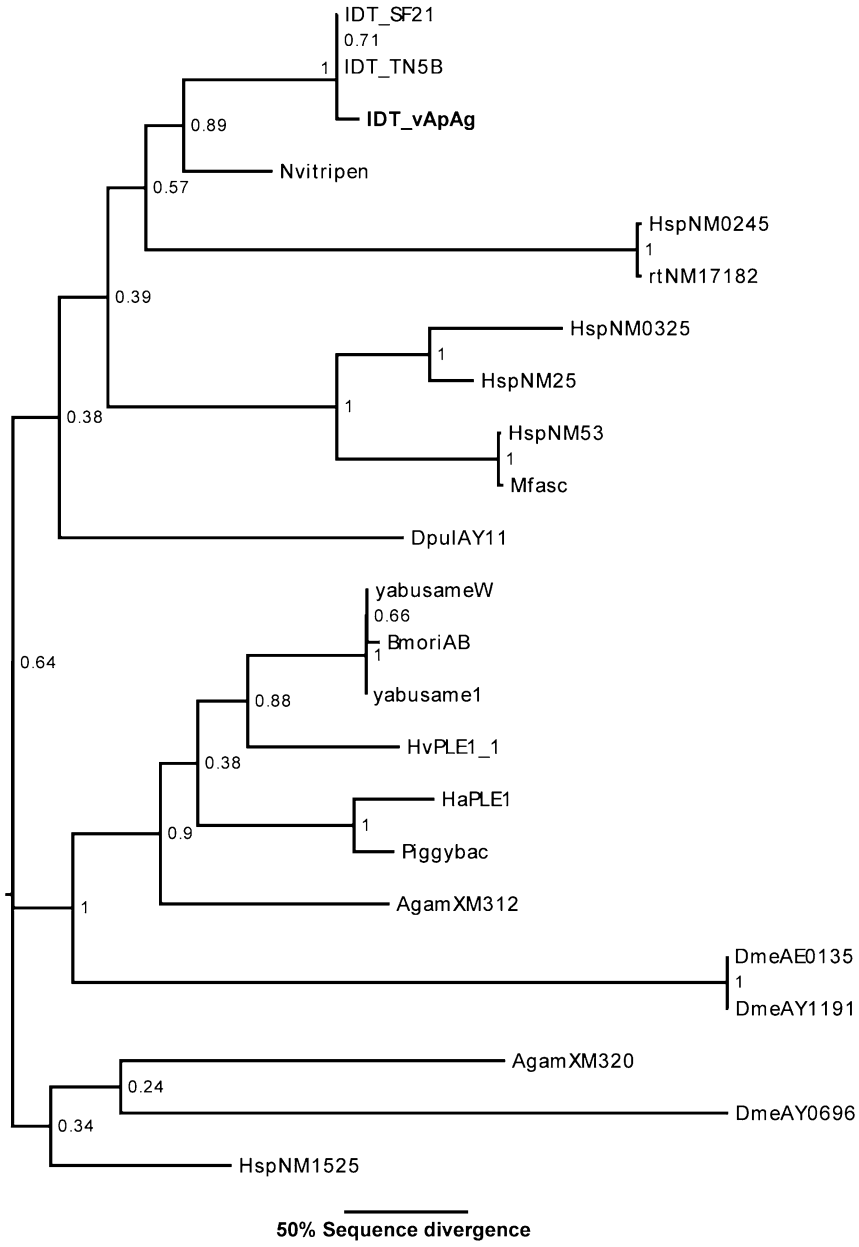

Figure 5. Maximum clade credibility (MCC) tree for 23 *piggyBac*-related animal transposases. The MCC tree was obtained after 100 non-parametric bootstrap iterations with the *PhyML* program (see methods). *PiggyBac*, Tpase from the *piggyBac* TE found in *Trichoplusia ni* (GenBank accession number J04364); HaPLE1, Tpase from the HaPLE1 TE found in *Helicoverpa armigera* (GenBank accession number EF593176); HvPLE1.1, Tpase from the HvPLE1.1 TE found in *Heliothis virescens* (GenBank accession number DQ407726); IDT, Tpase from the IDT TE found in the genome of the mutant baculovirus vApAg (this work), TN5B, Tpase from the IDT-like TE found in *T. ni* (this work); SF21, Tpase from the IDT-like TE found in *Spodoptera frugiperda* (this work); Yabusame-1, Tpase from the Yabusame-1 TE found in *Bombyx mori* (GenBank accession number BAD11135); Yabusame-W, Tpase from the Yabusame-W TE found in *Bombyx mori* (GenBank accession number AB159601).

Experimental procedures

Cells and virus

A. gemmatalis (UFL-AG-286) (Sieburth & Maruniak, 1988), *T. ni* (BTI-Tn5B1-4 or TN5B) (Granados *et al.*, 1994), and *S. frugiperda* (IPLB-SF21-AE or SF21) (Vaughn *et al.*, 1977) cells were grown in TC-100 medium (Gibco-BRL) supplemented with 10% fetal bovine serum at 27 °C. UFL-AG-286 and BTI-Tn5B1-4 served as hosts for the propagation of the baculovirus AgMNPV (Ag-2D) (Johnson & Maruniak, 1989) and the mutant vApAg (Silveira *et al.*, 1999).

Amplification and cloning of the three iap *genes of the mutant virus*

In order to amplify all three vApAg *iap* genes, the following specific oligonucleotides were designed: IAP-1F (5′-AAACGC

GCGTCAAGTTGGGCC- 3′) and IAP-1R (5′- GTACGCGT GTCCGTTGTACTGG- 3′) which anneal at nucleotides -82 to -61 and at +860 to +881 nucleotides of the *iap-1* ORF, respectively. IAP-2F (5′– TGGCGTTTGGTCAAACCGCC–3′) and the IAP-2R (5′– CCACTTGTTGTATCTTCAGG– 3′) which anneal at nucleotides -124 to -98 and at +850 to +869 nucleotides of the *iap-2* ORF, respectively. IAPAgFSalI (5′-CGTGTCGACACACACAATG-3′) and IAPAgRNCOI (5′-CCTCCATGGCTGAACG-3′), which anneal at -8 to +3 and at +884 to +899 nucleotides of the *iap-3* ORF, respectively. Oligos were used to a final concentration of 10 μM, and 50 ηg of template DNA were used in every reaction, according with the Taq DNA polymerase manufacturer's instructions (Invitrogen, Carlsbad, CA, USA). The PCR program used was: [94 °C/5 min; (94 °C/1 min, 50 °C/1.5 min, 72 °C/1 min) × 30; 72 °C/7 min.]. The *iap-3* gene fragment obtained from the mutant vApAg by PCR was cloned into the vector pGEM®-T Easy vector following the manufacturer's instructions (Promega, Madison, WI, USA) and used to transform DH5-α *Escherichia coli* competent cells (Life Technologies, Rockville, MD, USA) generating the plasmid pGEMvApiap-3. The DNA from the pGEMvApiap-3 plasmid was purified and sequenced (Mega BACE 1000, Amersham Biosciences, Little Chalfont, UK), following the manufacturer's instructions.

Amplification, cloning and sequencing of the piggyBac*-related TE from different insect cell lines*

UFL-AG-286, BTI-Tn5B1-4 and IPLB-SF21-AE genomic DNAs were purified according to Aljanabi & Martinez (1997) and used as templates in PCR reactions with oligonucleotides designed to anneal to the terminal inverted repeats of the IDT TE and to the Tpase ORF. The UFLAGTRANS oligonucleotide (5′-CCCTT TATAAGGCAGA-3′) anneals to the terminal inverted repeats of the TEs and the oligonucleotides TRASEF (5′-ATGCCGC GCTATTTGAAT-3′) and TRASER (5′-CTATAGCCACCAATC-3′) anneal at nucleotides +1 to +18 and at nucleotides +1619 to +1633 of the putative Tpase ORF. The PCR program used for the amplification of the TE was: [94 °C/5 min; (94 °C/40 s, 52 °C/ 1:30 min and 72 °C/1:30 min) × 30; 72 °C/7 min]. For the amplification of the Tpases the same program was used, except for the annealing temperature, which was 55 °C/1 min. The amplified DNA fragments were then cloned into the pGEM®-T Easy vector following the manufacturer's instructions (Promega) and sequenced as described above.

Sequence comparison and phylogenetic analyses

The *iap-3* and putative Tpase ORFs were sequenced and analysed using the ORF Finder and BLAST programs (Altschul *et al*., 1990, 1997) at the NCBI home page (http://www.ncbi.nlm.nih.gov). Sequence similarity statistics among sequences closely-related to the IDT vApAg are shown in Table 1, and the sequence alignment, and phylogram were generated using the online multiple alignment program MAFFT version 6 at the website: http://align.bmr.kyushu-u.ac.jp/mafft/online/server/ and the Jalview Java Alignment Editor (Clamp *et al*., 2004). Moreover, the relationship of the IDT vApAg Tpase to 22 representative Tpases found with Blast (Table 2) was investigated by phylogenetic reconstruction. Tpase protein sequences were aligned using Clustal X 1.82 (Thompson *et al*., 1994) generating a dataset 959 amino-acids in length (including gaps). A maximum clade credibility (MCC) tree was inferred after 100 non-parametric maximum likelihood bootstrap iterations with the *PhyML* program using the JTT amino-acid transition model and gamma-distributed variable rates and proportion of invariable sites estimated from the data (Guidon & Gascuel, 2003). The MCC tree was obtained from the 100 non-parametric bootstrap trees using the *TreeAnnotator* v.1.4.8 program (Drummond & Rambaut, 2007).

Acknowledgements

This work was supported by the following Brazilian agencies: PRONEX, FAPDF, CNPq, FINATEC, FINEP and CAPES.

References

Aljanabi, S.M. and Martinez, I. (1997) Universal and rapid salt extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Res* **25**: 4692–4693.

Allen, M.L., Handler, A.M., Berkebile, D.R. and Skoda, S.R. (2004) *piggyBac* transformation of the New World screwworm, *Cochliomyia hominivorax*, produces multiple distinct mutant strains. *Med Vet Entomol* **18**: 1–9.

Altschul, S.F., Gish, W., Miller, W., Meyers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**: 586–605.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. *et al.* (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389–3402.

Balu, B., Shoue, D.A., Fraser, M.J. Jr. and Adams, J.H. (2005) High-efficiency transformation of *Plasmodium falciparum* by the lepidopteran transposable element *piggyBac*. *Proc Natl Acad Sci USA* **102**: 16391–19396.

Berghammer, A.J., Klinger, M. and Wimmer, E.A. (1999) A universal marker for transgenic insects. *Nature* **402**: 370–371.

Birnbaum, M.J., Clem, R.J. and Miller, L.K. (1994) An apoptosis-inhibiting gene from a nuclear polyhedrosis virus encoding a polypeptide with Cys/His sequence motifs. *J Virol* **68**: 2521–2528.

Carpes, M.P., Castro, M.E.B., Soares, E.F., Villela, A.G., Pinedo, F.J. and Ribeiro, B.M. (2005) The inhibitor of apoptosis gene (iap-3) of *Anticarsia gemmatalis multicapsid nucleopolyhedrovirus* (AgMNPV) encodes a functional IAP. *Arch Virol* **150**: 1549–1562.

Clamp, M., Cuff, J., Searle, S.M. and Barton, G.J. (2004) The Jalview Java Alignment Editor. *Bioinformatics* **20**: 426–427.

Clarke, T.E. and Clem, R.J. (2003a) Insect defenses against virus infection: the role of apoptosis. *Int Rev Immunol* **22**: 401–424.

Clarke, T.E. and Clem, R.J. (2003b) *In vivo* induction of apoptosis correlating with reduced infectivity during baculovirus infection. *J Virol* **77**: 2227–2232.

Clem, R.J. (2005) The role of apoptosis in defense against baculovirus infection in insects. *Curr Top Microbiol Immunol* **289**: 113–130.

Crook, N.E., Clem, R.J. and Miller, L.K. (1993) An apoptosis-inhibiting baculovirus gene with a zinc finger-like motif. *J Virol* **67**: 2168–2174.

Drummond, A.J. and Rambaut, A. (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* **7**: 214.

Feschotte, C. and Pritham, E.J. (2007) DNA transposons and the evolution of eukaryotic genomes. *Annu Rev Genet* **41**: 331–368.

Fraser, M.J., Smith, G.E. and 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., Elick, T. and 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.

Gao, H., Wang, Y., Li, N., Peng, W., Sun, Y., Tong, G. *et al.* (2007) Efficient gene delivery into mammalian cells mediated recombinant baculovirus containing a whispovirus ie-1 promoter, a novel shuttle promoter between insect cells and mammalian cells. *J Biotechnol* **131**: 138–143.

Granados, R.R., Guoxum, L., Dersksen, A.C.G. and McKenna, K.A. (1994) A new insect cell line from *Trichoplusia ni* (BTI-Tn-5B1-4) susceptible to *Trichoplusia ni single nuclear polyhedrosis* virus. *J Invertebr Pathol* **64**: 260–266.

Griffiths, C.M., Barnett, A.L., Ayres, M.D., Windass, J., King, L.A. and Possee, R.D. (1999) *In vitro* host range of *Autographa californica nucleopolyhedrovirus* recombinants lacking functional p35, iap1 or iap2. *J Gen Virol* **80**: 1055–1066.

Grossman, G.L., Rafferty, C.S., Clayton, J.R., Stevens, T.K., Mukabayire, O. and Benedict, M.Q. (2001) Germline transformation of the malaria vector *Anopheles gambiae*, with the *piggyBac* transposable element. *Insect Mol Biol* **10**: 597–604.

Gueguen, E., Rousseau, P., Duval-Valentine, G. and Chandler, M. (2005) The transpososome: control of transposition at the level of catalysis. *Trends Micobiol* **13**: 543–549.

Guidon, S. and Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* **52**: 696–704.

Handler, A.M. (2002) Use of the *piggyBac* transposon for germline transformation of insects. *Insect Biochem Mol Biol* **10**: 1211–1220.

Handler, A.M. and Harrell, R.A. (1999) Germline transformation of *Drosophila melanogaster* with the *piggyBac* transposon vector. *Insect Mol Biol* **8**: 449–458.

Handler, A.M. and Harrell, R.A. (2001) 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., Dübendorfer, 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.G. *et al.* (2002) Germ-line transformation of the Australian sheep blowfly, *Lucilia cuprina*. *Insect Mol Biol* **11**: 1–10.

Hu, Y.C. (2006) Baculovirus vectors for gene therapy. *Adv Virus Res* **68**: 287–320.

Ikeda, M., Yanagimoto, K. and Kobayashi, M. (2004) Identification and functional analysis of *Hyphantria cunea nucleopolyhedrovirus iap* genes. *Virology* **321**: 359–371.

Johnson, D.W. and Maruniak, J.E. (1989) Physical map of *Anticarsia gemmatalis nuclear polyhedrosis virus* (AgMNPV-2) DNA. *J Gen Virol* **70**: 1877–1883.

Keith, J.H., Cheryl, A., Schaeper, C.A., Fraser, T.S. and Fraser, M.J. (2008a) Mutational analysis of highly conserved aspartate residues essential to the catalytic core of the *piggyBac* transposase. *BMC Mol Biol* **9**: 73, doi:10.1186/1471-2199-9-73.

Keith, J.H., Fraser, T.S. and Fraser, M.J. (2008b) Analysis of the *piggyBac* Tpase reveals a functional nuclear targeting signal in the 94 c-terminal residues. *BMC Mol Biol* **9**: 72, doi:10.1186/1471-2199-9-72.

Kim, Y.S., Xiao, H.Z., Du, E.Q., Cai, G.S., Lu, S.Y. and Qi, Y.P. (2007) Identification and functional analysis of LsMNPV anti-apoptosis genes. *J Biochem Mol Biol* **40**: 571–576.

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.

Kost, T.A., Condreay, J.P. and Jarvis, D.L. (2005) Baculovirus as versatile vectors for protein expression in insect and mammalian cells. *Nat Biotechnol* **23**: 567–575.

Li, X., Harrell, R.A., Handler, A.M., Beam, T., Hennessy, K. and Fraser, M.J. (2005) *piggyBac* internal sequences are necessary for efficient transformation of target genomes. *Insect Mol Biol* **14**: 17–30.

McLachlin, J.R., Escobar, J.C., Harrelson, J.A., Clem, R.J. and Miller, L.K. (2001) Deletions in the Ac-iap1 gene of the baculovirus AcMNPV occur spontaneously during serial passage and confer a cell line-specific replication advantage. *Virus Res* **81**: 77–91.

Means, J.C., Muro, I. and Clem, R.J. (2003) Silencing of the baculovirus Op-iap3 gene by RNA interference reveals that it is required for prevention of apoptosis during *Orgyia pseudotsugata M nucleopolyhedrovirus* infection of Ld652Y cells. *J Virol* **77**: 4481–4488.

Means, J.C., Penabaz, T. and Clem, R.J. (2007) Identification and functional characterization of AMVp33, a novel homolog of the baculovirus caspase inhibitor p35 found in Amsacta moorei entomopoxvirus. *Virology* **358**: 436–447.

Morales, M.E., Mann, V.H., Kines, K.J., Gobert, G.N., Fraser, M.J. Jr, Kalinna, B.H. *et al.* (2007) *piggyBac* transposon mediated transgenesis of the human blood fluke, *Schistosoma mansoni*. *FASEB J* **21**: 3479–3489.

Moscardi, F. (1999) Assessment of the application of baculoviruses for control of Lepidoptera. *Annu Rev Entomol* **44**: 257–289.

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.

Oliveira, J.V., Wolff, J.L., Garcia-Maruniak, A., Ribeiro, B.M., Castro, M.E.B., Souza, M.L. *et al.* (2006) Genome of the most widely used viral biopesticide: *anticarsia gemmatalis multiple nucleopolyhedrovirus*. *J Gen Virol* **87**: 3233–3250.

Peloquin, J.J., Thibault, S.T., Staten, R. and Miller, T.A. (2000) Germline transformation of pink bollworm (Lepidoptera: Gelechiidae) mediated by the *piggyBac* transposable element. *Insect Mol Biol* **9**: 323–333.

Perera, O.P., Harrell, 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.

Sarkar, A., Sim, C., Hong, Y.S., Hogan, J.R., Fraser, M.J., Robertson, H.M. *et al.* (2003) Molecular evolutionary analysis of the widespread *piggyBac* transposon family and related 'domesticated' sequences. *Mol Genet Genomics* **270**: 173–180.

Sieburth, P.J. and Maruniak, J.E. (1988) Growth characteristic of a continuous cell line from the velvetbean caterpillar, *Anticarsia gemmatalis* Hübner (Lepidoptera: Noctuidae). *In Vitro Cell Dev Biol Plant* **24**: 195–198.

Silveira, E.B., Ribeiro, B.M. and Báo, S.N. (1999) Morphological studies of apoptosis in insect cells infected with vApAg, an *Anticarsia gemmatalis nucleopolyhedrovirus* mutant. *J Submicrosc Cytol Pathol* **31**: 1–8.

Silveira, E.B., Cordeiro, B.A., Ribeiro, B.M. and Báo, S.N. (2005) In vivo apoptosis induction and reduction of infectivity by an *Autographa californica* multiple nucleopolyhedrovirus *p35*– recombinant in hemocytes from the velvet bean caterpillar *Anticarsia gemmatalis* (Hubner) (Lepidoptera: Noctuidae). *Res Microbiol* **156**: 1014–1025.

Silveira, E.B., Cordeiro, B.A., Ribeiro, B.M., Castro, M.E.B., Soares, E.F. and Báo, S.N. (2007) An *Anticarsia gemmatalis multiple nucleopolyhedrovirus* mutant, vApAg, induces hemocytes apoptosis *in vivo* and displays reduced infectivity in larvae of *Anticarsia gemmatalis* (Hübner) (Lepidoptera: Noctuidae). *Virus Res* **130**: 182–192.

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.

Sun, Z.C., Wu, M., Miller, T.A. and Han, Z.J. (2008) *piggyBac*-like elements in cotton bollworm, *Helicoverpa armigera* (Hübner). *Insect Mol Biol* **17**: 9–18.

Tamura, T., Thibert, T., Royer, C., Kanda, T., Eappen, A., Kamba, M. *et al.* (2000) A *piggyBac* element-derived vector efficiently promotes germ-line transformation in the silkworm *Bombyx mori* L. *Nat Biotechnol* **18**: 81–84.

Thompson, J., Higgins, D. and Gibson, T. (1994) CUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**: 4673–4680.

Vaughn, J.L., Goodwin, R.H., Tompkins, G.J. and McCawley, P. (1977) The establishment of two cell lines from the insect *Spodoptera frugiperda* (Lepidoptera; Noctuidae). *In Vitro Cell Dev Biol Plant* **13**: 213–217.

Zhang, P., Yang, K., Dai, X., Pang, Y. and Su, D. (2002) Infection of wild-type *Autographa californica multicapsid nucleopolyhedrovirus* induces in vivo apoptosis of *Spodoptera litura* larvae. *J Gen Virol* **83**: 3003–3011.

Zimmermann, K.C., Bonzon, C. and Green, D.R. (2001) The machinery of programmed cell death. *Pharmacol Ther* **92**: 57–70.

Zimowska, G.J. and Handler, A.M. (2006) Highly conserved *piggyBac* elements in noctuid species of Lepidoptera. *Insect Biochem Mol Biol* **36**: 421–428.

