

# A new *hobo*, *Ac*, *Tam3* transposable element, *hopper*, from *Bactrocera dorsalis* is distantly related to *hobo* and *Ac*

Alfred M. Handler \*, Sheilachu P. Gomez

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

Received 24 June 1996; revised 22 August 1996; accepted 23 August 1996; Received by A. Benardi

## Abstract

A new transposable element from the *hobo*, *Ac*, *Tam3* transposon family was isolated as a genomic clone from the oriental fruit fly, *Bactrocera dorsalis*. It is approximately 3.1 kb in length with 19-bp inverted terminal repeat sequences having a single mismatch. Though sharing several amino acid sequence identities with other *hAT* elements, it is distantly related to both *hobo* and *Ac*. Among *hAT* elements thus far described in insects, it is apparently the most distantly related to *hobo*.

**Keywords:** Transposons; *hobo*; *Ac*; *Tam3* family; Transposase; Oriental fruit fly; Tephritidae; Gene vector

## 1. Introduction

Transposable elements appear to be ubiquitous genomic components of most organisms, with many sharing common ancestors (see Berg and Howe, 1989). Determining the organismal range of specific transposon families and elucidating the structural and functional relationships of their constituent elements is important to understanding their evolution, transmission, and regulation. This information is, in turn, important to developing their potential use as gene transfer vectors, among other types of genetic manipulation.

Recently, a structural relationship has been established among transposons found in both plant and insect species. The *hobo* transposon from *Drosophila melanogaster* was shown to have amino acid sequence similarity to the *Activator* (*Ac*) transposable element from maize, with a more distant relationship to the *Tam3* element from the snapdragon (Calvi et al., 1991). Utilizing some of the common amino acid sequences between *Ac* and *hobo*, gene amplification strategies have allowed the subsequent identification of a complete *hobo*-related element, *Hermes*, from *Musca domestica* (Warren et al., 1994) and *Hermit*, from *Lucilia cuprina* (Coates et al.,

1996). Discrete *hobo-Ac* related sequences have been identified as well from *Heliothis virescens* (DeVault and Narang, 1994), *M. vetustissima* (Warren et al., 1995), and several tephritid fruit fly species (Handler and Gomez, 1996). These elements are now considered to be members of the *hobo*, *Activator*, *Tam3* (*hAT*) transposon family. Of the known complete elements from insects, all exhibit at least a 50% amino acid sequence identity to *hobo* within their transposase coding regions, while identity to *Ac* is no greater than 25%. This suggests a clear divergence between these elements in plants and animals. However, translation of the *hAT* polymerase chain reaction (PCR) sequences from two tephritid fruit flies, *Anastrepha suspensa* and *Bactrocera dorsalis*, indicate an equally distant relationship to both *hobo* and *Ac* (Handler and Gomez, 1996). To better define the relationships and possible interactions among *hAT* elements, we have begun to isolate the complete elements from these tephritid species, and here we describe one from *B. dorsalis* which we designate as *hopper*.

## 2. Experimental and discussion

### 2.1. Cloning and sequencing

Utilizing a *hobo-Ac* PCR sequence fragment from *B. dorsalis* (Bd-HRE; Handler and Gomez, 1996) as a probe, we isolated several genomic clones from a  $\lambda$ GEM12 (Promega) *B. dorsalis* Kahuku strain genomic

\* Corresponding author. Tel. +1 352 3745793; Fax +1 352 3745781; e-mail: handler@nervm.nerdc.ufl.edu

Abbreviations: *Ac*, *Activator*; *B.*, *Bactrocera*; *C.*, *Caenorhabditis*; bp, base pair(s); *hAT*, *hobo*, *Ac*, *Tam3*; kb, kilobase(s) or 1000 bp; *M.*, *Musca*; ORF(s), open reading-frame(s); PCR, polymerase chain reaction; IS, insertion site; ITR, inverted terminal repeat

library. The Kahuku strain is a wild strain collected originally in 1986 in Oahu, Hawaii, and is maintained at the University of Hawaii. *EcoRI* and *BamHI* restriction fragments of the genomic clones Bd3-2 and Bd6-1 were sequenced after subcloning into pUC19. Sequencing primers were initially derived from the Bd-HRE sequence, with subsequent primer sequencing accomplished in both directions and on both strands. Sequences from the two genomic clones were identical, and the same element was also present within another clone based on adjacent chromosomal DNA sequence. Inverted terminal repeat (ITR) sequences were identified by alignment of complementary distal sequences, which yielded 19-bp inverted repeat sequences (below) containing a single mismatch nucleotide (lower case, bold).

```

5'-IS          ITR
cttcgtac/TAGTGGTGGGgACTATCGA-----/
/-----TCGATAGTCCCAACTA/gtaagtag
          3'-ITR          IS

```

Interestingly, the chromosomal insertion site (IS) DNA adjacent to the termini of this element does not consist of an 8-bp direct repeat (in lower case), which

was verified in another genomic clone only having its termini and insertion sites sequenced. A 50% identity does exist between the 8-bp sequences present.

2.2. Transposase coding region analysis

As shown in Fig. 1 the terminal sequences encompass a 3120-bp element which encodes two large open reading frames (ORF) having a 10-bp overlap (starting at the first ATG, extending from base positions 856 to 1572 and 1563 to 2354) as well as additional smaller ORFs. Comparison of all ORFs to GenBank sequences using BLAST (Altschul et al., 1990) indicated highest homology to other *hAT* elements in the first two ORFs as well as three small C-terminal ORFs, two of which overlap. A consensus reading frame containing these ORFs is realized after introducing two frameshifts between the first and second, and third and fourth ORFs yielding a 1.9-kb coding sequence with three internal stop signals. The specific positions of the frameshifts and stop codons are shown in Fig. 1, as well as the amino acid sequence motifs having highest homology to those in other *hAT* elements. Utilizing the conceptual translation of this consensus ORF for molecular comparisons, GAP analy-

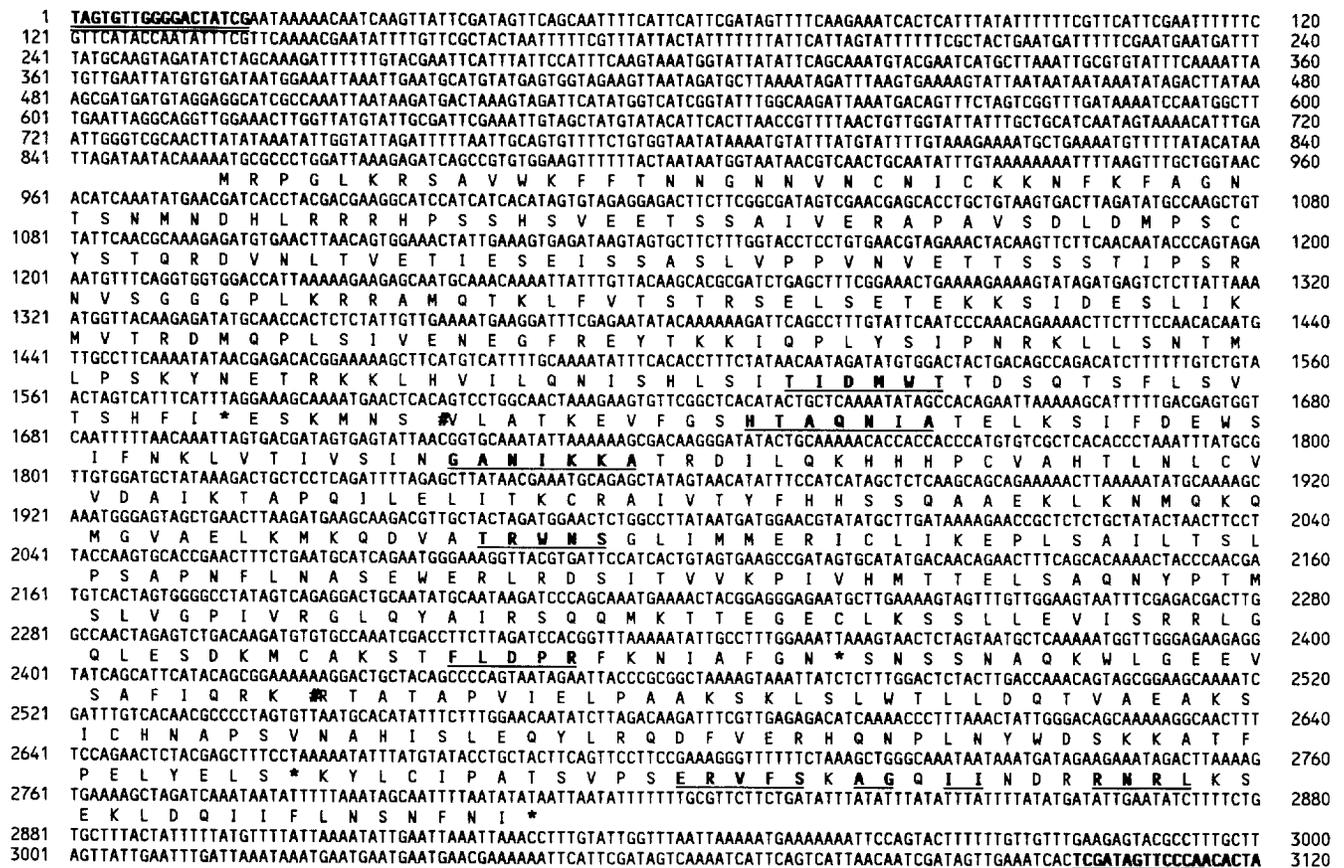


Fig. 1. The complete 3120-bp *hopper* genomic clone sequence from the *B. dorsalis* Kahuku strain with a consensus translation. The 19-bp inverted terminal repeat sequences are double underlined and in bold. Frameshifts introduced to maintain a *hAT* element consensus reading frame are designated by (#), and stop codons are designated by (\*). Amino acid sequence motifs having identity with other *hAT* element sequences are underlined and in bold, with further descriptions given in Section 2.2. GenBank accession number xx-xxxx.

Tam3	SSCLDPRVRL I
Ce-hobo	ATILDPRVVKV K
Ac	ACFLDPRVYK I
hopper	STFLDPRFKN I
Consensus	...LDPR.K. I

Fig. 2. Amino acid sequence alignment of the LPDR sequence and adjacent amino acids found in *hAT* elements distantly related to *hobo*.

sis (Devereux et al., 1984) yielded almost the same identity of 24% and similarity (identities plus positive substitutions) of 44% to both *hobo* and *Ac*. While this level of homology is not a strong indication of relatedness, the BLAST search of the entire GenBank (release 91) indicated highest amino acid sequence similarity to the *hobo* transposase ( $P_{\text{BLASTP}}=7.4e-26$ ), and interestingly, secondly to a cosmid sequence from the Nematode Sequencing Project ( $P_{\text{BLASTP}}=1.3e-25$ ). The same sequence was recently discovered and defined as a *hAT* element based on searches with *hobo* (Bigot et al., 1996). The third closest homology was to *Ac* ( $P_{\text{BLASTP}}=1.2e-24$ ). Succeeding high scoring homologies were solely to other *hAT* sequences including *Hermes* ( $P_{\text{BLASTP}}=7.7e-21$ ), *Hermit* from *Lucilia cuprina* ( $P_{\text{BLASTP}}=2.1e-18$ ), *Tam3* ( $P_{\text{BLASTP}}=1.1e-10$ ), a *Musca vetustissima* *hobo*-related sequence ( $P_{\text{BLASTP}}=0.0012$ ), and an *Ac*-related element in pearl millet ( $P_{\text{BLASTP}}=0.046$ ).

BLAST analysis alignments indicated strongest overall homologies in C-terminal sequences (or region 3; see Calvi et al., 1991) beginning with the ERVFS sequence, though discrete amino acid sequence identities are clearly apparent in regions 1 and 2, including those used for initial PCR screens (TIDMWT and TRWNS sequences; see Handler and Gomez, 1996). The HTAQNIA sequence is identical in *C. elegans-hobo*, and the GANIKKA sequence is identical in *hobo*. A specific sequence identity revealed in this analysis has LDPR as a core (Fig. 2), and is thus far observed only in those *hAT* elements most distantly related to *hobo*, including the *C. elegans hobo*, *Ac*, *Tam3*, and *hopper*.

### 3. Conclusions

We have described a new transposable element in *B. dorsalis*, which shares common features with other short inverted terminal repeat elements in the *hAT* transposon family. Its complete length is approximately 3.1 kb, it is bounded by short inverted terminal repeat sequences having the A2;G5 motif implicated in DNA splicing mechanisms (Warren et al., 1994), and it includes putative transposase-encoding ORFs sharing common sequences with those from *hobo*, *Ac*, as well as other *hAT* transposons. Previous studies with a *hopper* PCR subfragment showed that *hopper* exists as a repetitive element in the *B. dorsalis* genome (Handler and Gomez, 1996), and

we confirm here that, thus far, it is the most distantly related insect *hAT* element to *hobo*. Of interest is the finding that the element described does not contain an 8-bp insertion site direct repeat which is consistently present in all other *hAT* elements. Taken together with the sequence comparisons, *hopper* may be an ancient predecessor to other insect *hAT* elements, perhaps being a link or early branch from the non-insect elements. Sequence changes due to mutations or recombination may have collected in adjacent and internal sequences, and the transposase ORF discontinuity is not supportive of autonomous function. Previous hybridization studies indicated that five to 10 *hopper* elements exist in the genomes of several *B. dorsalis* strains. While it is curious that all of our genomic clones contained the same element, eventual isolation of the other elements will further elucidate the structure and function of *hopper*.

### Acknowledgement

Grateful appreciation is extended to Drs. Susan McCombs and Stephen Saul for providing pupae and to the US Department of Agriculture Competitive Grants Program for support.

### References

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignments search tool. *J. Mol. Biol.* 215, 403–410.
- Berg, D.E. and Howe, M.M. (1989) *Mobile DNA*. American Society for Microbiology, Washington, DC.
- Bigot, Y., Auge-Gouillou, C., Periquet, G. (1996) Computer analyses reveal a *hobo*-like element in the nematode *Caenorhabditis elegans*, which presents a conserved transposase domain common with the Tc1-mariner transposon family. *Gene* 174, 265–271.
- Calvi, B.R., Hong, T.J., Findley, S.D. and Gelbart, W.M. (1991) Evidence for a common evolutionary origin of inverted repeat transposons in *Drosophila* and plants: *hobo*, *Activator*, and *Tam3*. *Cell* 66, 465–471.
- Coates, C.J., Johnson, N.K., Perkins, H.D., Howells, A.J. and O'Brochta, D.A. (1996) The *hermit* transposable element of the Australian sheep blowfly, *Lucilia cuprina*, belongs to the *hAT* family of transposable elements. *Genetica* 23–31.
- Devereux, J.D., Haeblerli, P. and Smithies, O. (1984) A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12, 387–395.
- DeVault, J.D. and Narang, S.K. (1994) Transposable elements in Lepidoptera: *hobo*-like transposons in *Heliothis virescens* and *Helicoverpa zea*. *Biochem. Biophys. Res. Commun.* 203, 169–175.
- Handler, A.M. and Gomez, S.P. (1996) The *hobo* transposable element excises and has related elements in tephritid species. *Genetics* 143, 1339–1347.
- Warren, W.D., Atkinson, P.W. and O'Brochta, D.A. (1994) The *Hermes* transposable element from the house fly, *Musca domestica*, is a short inverted repeat-type element of the *hobo*, *Ac*, and *Tam3* (*hAT*) element family. *Genet. Res. Cambridge* 64, 87–97.
- Warren, W.D., Atkinson, P.W. and O'Brochta, D.A. (1995) The Australian bushfly *Musca vetustissima* contains a sequence related to transposons of the *hobo*, *Ac*, and *Tam3* family. *Gene* 154, 133–134.