



Isolation and analysis of a new *hopper hAT* transposon from the *Bactrocera dorsalis white eye* strain

Alfred M. Handler

Center for Medical, Agricultural, and Veterinary Entomology, Agricultural Research Service, US Department of Agriculture, Gainesville, FL, USA; Address for correspondence: USDA-ARS, 1700 S.W. 23rd Drive, Gainesville, FL 32608, USA (Phone: +1 (352) 374-5793; Fax: +1 (352) 374-5781; E-mail: handler@nersp.nerdc.ufl.edu)

Received 29 March 2002 Accepted 27 June 2002

Key words: gene vector, *hobo* Activator *Tam3* family, Oriental fruit fly, Tephritidae, transposable element, transposase

Abstract

A new *hopper* element belonging to the *hAT* transposon family was isolated from the *white eye* mutant strain of the Oriental fruit fly, *Bactrocera dorsalis*. Using the original *hopper* element sequence from the wild type Kahuku strain as a template, the new *hopper* was isolated by inverse and direct PCR. Nucleotide sequence analysis reveals a 3131 bp element with terminal and subterminal inverted repeat sequences, an 8 bp duplicated insertion site, and a conceptual translation yielding a single uninterrupted 650 amino acid open reading frame. The *white eye hopper* has structure more consistent with function than the Kahuku element, indicating that *hopper* is not an ancient relic. The *hopper* element remains distantly related to other known *hAT* elements including those from insects, and presently it is most similar to *Activator*-related elements discovered in the human genome. DNA hybridization studies indicate, however, that elements closely related to *hopper* exist in another bactrocerid species, the melonfly, *B. cucurbitae*.

Abbreviations: *Ac* – *Activator*; bp – base pair(s); *hAT*, *hobo*, *Ac*, *Tam3*; kb – kilo base pair(s); nt – nucleotide(s); ORF(s) – open reading-frames; PCR – polymerase chain reaction; ITR – inverted terminal repeat.

Introduction

One of the most widespread families of known transposable elements is the *hobo*, *Activator*, *Tam3* (*hAT*) family that exists in organisms ranging from fungi to humans (see Kempken & Windhofer, 2001; Rubin et al., 2001). They were first recognized by comparison of *hobo* from *Drosophila melanogaster* to *Activator* (*Ac*) from maize and *Tam3* from the snapdragon (Calvi et al., 1991; Feldmar & Kunze, 1991). Although distantly related, conservation of amino acid sequence motifs allowed further identification of related elements by gene amplification. The first of these included elements found in insects. A transpositionally active *Hermes* element from *Musca domestica* (Warren et al., 1994), and a complete *Homer* element from *Bactrocera tryoni*, (Pinkerton et al., 1999)

were derived in each species from independent PCR products. Partial sequences were isolated for *Hermit* from *Lucilia cuprina* (Coates et al., 1996), and discrete *hobo-Ac*-related sequences have been identified as well from *Heliothis virescens* (DeVault & Narang, 1994), *M. vetustissima* (Warren et al., 1995), and several tephritid fruit fly species (Handler & Gomez, 1996). Elements within the *hAT* family have since been found to be widespread among plants, fungi, worms and vertebrates (see Kempken & Windhofer, 2001; Rubin et al., 2001).

Using an internal PCR fragment of a *Bactrocera dorsalis hobo*-related element (Bd-HRE) as probe, a complete element, named *hopper*, and adjacent insertion site DNA were isolated from a *B. dorsalis* Kahuku wild strain genomic library (Handler & Gomez, 1997). This element was of particular interest since it

represented the most distantly related *hAT* element among insects, being equally distant from *hobo* and other insect *hAT* elements, and those found in plants. The Kahuku *hopper* (*hopper*^{Bd-Kah}) is 3120 bp in length with 19 bp inverted terminal repeat sequences having a single base pair mismatch. Several overlapping open reading frames were linked to create a single 1.9 kb consensus reading frame with similarities to those that exist in other *hAT* elements, and notably, this *hopper* does not have a canonical 8 bp duplicated insertion site found in all *hAT* elements. The *hopper*^{Bd-Kah} element is very likely non-functional, though its reading frames include discrete amino acid sequence motifs highly similar or identical to those found in other *hAT* elements from a wide range of organisms.

Most transposable elements exist as multiple genomic copies with many being defective and non-autonomous (see Robertson & Lampe, 1995), and previous hybridization studies showed that multiple *hopper* elements exist in several strains of *B. dorsalis* (Handler & Gomez, 1996). To better grasp the evolutionary history of *hopper* it is important to isolate these elements and compare their sequences to one another and other *hAT* elements. Defining the distribution of *hopper* is also important, and preliminary PCR studies in the melonfly, *B. cucurbitae*, suggested the existence of *hopper* in this species as well. Should functional elements exist in either of the species, they may also provide tools for transposon-mediated transformation and mutagenesis. To begin to isolate additional *hopper* elements, a survey for *hopper* in several *B. dorsalis* strains was initiated using inverse PCR. Here, we report the isolation of a new *hopper* element from the *white eye* strain having structural characteristics more consistent with a functional transposon.

Materials and methods

Insect strains

B. dorsalis and *B. cucurbitae* came from laboratory wild type and mutant strains maintained at the Department of Plant and Environmental Protection Sciences, University of Hawaii, Honolulu (for descriptions see McCombs & Saul, 1992; McCombs et al., 1995). All mutant strains were isolated as spontaneous mutations from wild strains originally found on the islands of Hawaii.

Polymerase chain reaction

Inverse PCR was performed by initial digestion of 1–3 µg of *B. dorsalis white eye* genomic DNA with *MspI* which does not cut within the *hopper*^{Bd-Kah} transposon. After 4 h digestion, restriction fragments were purified by phenol–chloroform extraction and ethanol precipitation and circularized by ligation at 12°C for 16 h. PCR was performed on the circularized fragments using primer sequences proximal to the *hopper*^{Bd-Kah} termini and in outward facing orientation. These included the reverse primer (556R) 5'-TGCCAAATACCGATGAC-3' proximal to the 5' terminus and the forward primer (2311F) 5'-GCCAAATCGACCTTCTTA-3' proximal to the 3' terminus. Cycling conditions consisted of initial denaturation at 94°C for 5 min, and 32 cycles of denaturation at 94°C for 40 s, annealing at 55°C for 40 s, and elongation at 72°C for 4 min, with a final elongation for 10 min. PCR products were analyzed by agarose gel electrophoresis and directly subcloned into the pCRII vector (Invitrogen). The vector termini and adjacent insertion sites were sequenced from the subclones using M13 forward and reverse primers.

To isolate the internal *hopper* sequence the forward primer (–136F), 5'-CCTGGCATTACTTAGTGTTCC-3', from chromosomal DNA 136 bp upstream from the 5' terminus was used with the reverse primer (2707R), 5'-CGGAGGGAAGTGAAGTAGCAG-3', from internal *hopper* sequence.

pBSshopper^{Bd-we} construction and sequence analysis

To create a complete *hopper* element from PCR products, the 3' end of *hopper* was isolated as a 725 bp *SacII/XbaI* fragment from the inverse PCR product and ligated into the *SacII/XbaI* of pBSKS (Stratagene) to create pBSBdhop3'. This fragment includes 665 bp of 3' *hopper* terminal DNA and 60 bp of genomic sequence adjacent to the 3' insertion site. The 5'-end of *hopper* was amplified from *B. dorsalis white eye* DNA using primers that amplified 136 bp of 5' chromosomal insertion site DNA and 2707 bp of *hopper* starting at the 5' terminus. The subcloned 2.8 kb PCR product was digested with *NcoI* and blunted, and *SacII* to give a 2622 bp fragment that was ligated into pBSBdhop3' at a *SacI*-blunt/*SacII* site. This linked the 5' and 3' *hopper* fragments at their unique internal *SacII* site in pBSKS to create pBSshopper^{Bd-we}. The *hopper* element and adjacent insertion site DNA within pBSshopper^{Bd-we} was then re-sequenced. Nucleotide and amino acid sequence analysis and

comparisons were performed using GeneWorks 2.5 (Oxford Molecular Group) and MegAlign (DNASTAR, Inc.) software, and BLASTP (Altschul et al., 1997), ClustalV (Higgins et al., 1992), and ClustalW (Thompson et al., 1994) analysis. Parsimony analysis of ClustalW multiple sequence alignments was performed with PAUP v.3.1.1 (Swofford, 1993).

Southern hybridization

Approximately 5 µg of genomic DNA was digested with *Nsi*I restriction enzyme and separated on 0.8% agarose gels, blotted to nylon filters and immobilized by ultraviolet irradiation. The hybridization probe was a 2.0 kb *Bst*Z17 I *hopper* fragment digested from pBShopper^{Bd-we} and isolated by agarose gel-elution. Probe DNA was radiolabeled with [³²P]-dCTP by random priming (Gibco BRL) according to the manufacturer's specifications. Hybridizations were performed 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 −70°C.

Results

Isolation from PCR products and sequence analysis

A new genomic *hopper* element was isolated from the *B. dorsalis white eye* strain by initially performing inverse PCR using primers to the original *hopper* from the wild Kahuku strain (GB acc #U88976; Handler & Gomez, 1997). By cutting genomic DNA with a restriction enzyme lacking a recognition site within *hopper*, terminal sequences of the same element could be isolated from circularized DNA after amplification. Using primers to proximal genomic DNA and internal *hopper* sequence, a complete element was constructed from PCR products.

Using the inverted terminal sequences (ITR) of the original 3120 bp *hopper*^{Bd-Kah} as landmarks, the *white eye hopper* (*hopper*^{Bd-we}) is 3131 bp in length, and the sequences of the two elements are approximately 97% identical and share the same imperfect 19 bp ITR and 16 bp subterminal inverted repeat sequences (Figure 1). In *hopper*^{Bd-we} the subterminal repeats are 92 bp from the 5' ITR (nt 111–126) and 50 bp from the 3' ITR (nt 3047–3062). While the two elements share the same inverted repeat sequences, nucleotide changes exist in between the terminal and subterminal

sequences. Several A/TCG sequences, implicated in transposase binding (Becker & Kunze, 1997), exist in the terminal regions and they are conserved within the subterminal and terminal sequences of the two elements. The *hopper* ITR sequence shows greatest similarity to ITRs in *hAT* elements found in plants (see Rubin et al., 2001), sharing the TAGxGxTG terminal sequence consensus. This is in contrast to the CAGAGAxTG terminal repeat consensus shared by all the other known insect *hAT*s.

A single conceptual open reading frame of 650 amino acids is encoded by nt 859–2811, that corresponds to a frameshifted consensus reading frame in *hopper*^{Bd-Kah}. Consistent with a transcriptional unit, a putative CAAT site exists at nt 706–709, a TATA site at nt 835–838, and a polyadenylation signal site at nt 2964–2969. Notably, *hopper*^{Bd-we} is inserted into a perfectly duplicated 8 bp genomic site (CTAAAAAA) which is consistent with all other *hAT* elements, but which differs from the non-duplicated insertion site of *hopper*^{Bd-Kah}.

Transposase coding region

The 650 amino acid open reading frame from *hopper*^{Bd-we} is generally consistent with the consensus *hAT* reading frame constructed for *hopper*^{Bd-Kah} from five reading frames (Handler & Gomez, 1997). Thus, the original comparisons between the *hopper* and other *hAT* transposase coding regions remain valid, though more extensive and accurate comparisons are now possible. The *hopper*^{Bd-Kah} element was found to have conserved amino acid sequence motifs with those found in most *hAT* elements, as well as to those found solely in *hAT*s from insects, plants, or *C. elegans* (see Figures 1 and 2). These comparisons and overall sequence similarities indicated that *hopper* was evolutionarily equidistant from *hAT*s in insects and plants, with GAP analysis indicating the strongest relatedness to the HFL1 *hobo* element from *Drosophila*. BLASTP analysis of the genomic database, as well as ClustalW and parsimony relationships among *hAT* coding regions is generally consistent with the earlier study, but now shows the highest identity of *hopper*^{Bd-we} to several *Ac*-related elements found in humans (Figure 3 and Table 1). While the other insect *hAT* elements share identities of 40–55% to one another, the similarity to *hopper* is in the range of 15–20%, which is similar to the relationship of *hopper* to *Ac* and the *C. elegans Ac*-related element (Table 1).

1 TAGTGTGGGGACTATCGAATCAAAACAATCAAGTTATTTCGGTGGTTCAGCAATTTTCATTCATTTCGATAGTTTTCAAGAAATCACTCATTACTATTTTTTCGTTTCATTCGAATTTTTTC 120
121 GTTCATACCAATATTTTCGTTCAAACGAATATTTTGTTCGATACTAATTCCTCGTTTATTACTATTTTTGCATTCATAGTATTTTTTCGCTATTGAATGATTTTTTCGAATGAATGATTT 240
241 TATGCAAGTAGATATCTAGCAAAGATTTTTGTACGAATTCATTTATCCATTTCAAAGTAAATGGTATTATATTCAGCAAAATGTCAGAAATGTCGAAATTTGTTGATTTTTCAAATTA 360
361 TGTTGAATGATTTATGTGTGATAATGAAATTAATGAATGCATGTATGAGTGGTAGAAGTTAATAGATGCTTAAATAGATTTAAGTAAAAGTATTAATAATAATAATATAGACTT 480
481 ATAAAGCGATAAATGTAGGAGGCATTATAATTAATAAGATGGCTAAAGTAGATTATATGGTTCATCGGTAATTTGGCAAGATTAAATGACAGTTTCTAGTCGGTTTGAATAAATCCAATGGC 600
601 TTTGAAATTAAGCAGGTTGGCAACTTGGTTACGTATTGCGATTTCGAAATTTGATAGTATGTATACATTCACTTAACCGTTTTAACTGTCCGTTATTTTGCAGCATCAATAGTAAAACATT 720
721 TGAATGGGTCGCAACTTATATAAATATGGTTTTAGATTTTTAAATGCGAGTCTTCTGTGGTAATATAAAATGATTTATGATTTTTGTAAGAAAATGCTGAAAATTTTTTATACA 840
841 TAATTAGATAAAAACAAAATGCGCCCTGGATTAAGAGATCAGCCGTGTGGGAGTTTTTAACTAAGAATGGTAATAACGTCAACTGCAATTTTGTAAAAAAATTAAGTTTGCCTGGT 960
M R P G L K R S A V W E F F T K N G N N V N C N I C K K N L K F A G
961 AACACATCAAAATGAAATGATCACCTACGACGAAGGCATCCATCATCACATAGTGTGGGGGAGACTTCTTCGGCGATAGTCGAACGAGCACCTGCTGTAAGTGACTTAGATATGCCAAGC 1080
N T S N M N D H L R R R H P S S H S V G E T S S A I V E R A P A V S D L D M P S
1081 TGTTCTTCAACGCAAAGAGATGTGAACCTTAACAGTGGAACTATTGAAAGTGGATAAGTAGTGTCTTGGTACCTCCTGTGAACGTAGAAACTACAAGTTCTTCAACAATPCCAGT 1200
C S S T Q R D V N L T V E T I E S E I S S A S L V P P V N V E T T S S S T I P S
1201 AGAAATATTTTCAGTGTGGACCATTAAAAAGAAAGGGCAATGCAACAAAATTTATTTGTACAAGCAGCGGATCTGAGCTTTCGGAAACTGAAAAGAGAAGTATAGATGAGTCTCTTATT 1320
R N I S G G G P L K R R A M Q T K L F V T S T R S E L S E T E K R S I D E S L I
1321 AAAATGGTTACAAGAAATGCAACCACTCTCTATTGTTGAAAAAGAGGATTTTCGAGAATATACAAAAGCTTCAGCCTTTGTATTCAAATCCAAACAGAAAACTCTTTCCAAACACA 1440
K M V T R D M Q P L S I V E N E G F R E Y T K K L Q P L Y S I P N R K L L S N T
1441 ATGTTGCCTTCAAAATATAACGAGACACGAAAAAGCTTCATGCCATTTTGCAAAATATTTACACCTTTCTATAACGACAGATATGTGGACTACTGCAGCCAGAACTTTTTTAACT 1560
M L P S K Y N E T R K K L H A I T L L Q N I S H L S I T T T D M W T T D S Q K S F L T
1561 GTAACAGTCAATTTTCATTTGGGAAAGCAAATGAACCTGAGCTGCAAGTAAAGTAGTTCGGCTCCACTGCTCAAAATATAGCCACAGAATTAAGAACTTTAAAGCATTTTTGACGAG 1680
V T S H F I W E S K M N S A V L A T K V V F G S H T A Q N I A T E L K S I F D E
1681 TGGTCAATTTTTAAACAATAGTGACGATAGTGAAGTAAACGGTGCATAATTAATAAAGCGATAAGGGATATACTCCAAAACACCACCACCATGTGTGCTCACACCCTAAATTTA 1800
W S I F N K I V T I V S D N G A N I K K A I R D I L Q K A H H H P C V A H T L N L
1801 TCGTTGTGGATGCGATAAAGACTGTTCTCAGATTTTAGAGTTATAAAGCAATGTAGAGCTATAGTAACTATTTCCATCAGTCTCAAGCAGCAGAAAACTAAAAAATATGCAA 1920
C V V D A I K T V P Q I L E L I T K C R A I V T Y F H H S S Q A A E K L K N M Q
1921 AAGCAAATGGGAGTAGCTGAACCTTAAGATGAAGCAAGACGTTGCTACTAGATGGAACCTGCGCTTATAATGATGGAACGTATATGCTTGATAAAGAACCCTCTCTGCTGTAATACT 2040
K Q M G V A E L K M K Q D V A T R W N S G L I M M E R I C L I K E P L S A V L T
2041 TCCTTACCTAGTGACCCGAACCTTTCTGAATGCATCAGAATGGGAAAGGTTACGTGATTTCCACTGTATTGAAGCCGATAGAGCATATGACAATAGAACTTTGAGCACAACCACTACCCA 2160
S L P S A P N F L N A S E W E R L R D S I T V L K P I E H M T I E L S S A Q K Y P
2161 ACGATGTCAGTGGTGCCTATAGTCAGAGGACTGCAATATGCAATAAGATCCCAGCAAATGAAAACACGGAGGGAGAATGCTTGAAAAGTAGTTTGTGGAGTAATTTTCGAGACGA 2280
T M S L V V P I V R G L Q Y A I R S Q Q M K T T E G E C L K S S L L E V I S R R
2281 CTTGGCCAACCAGAGTCTGACAAGATGTGTGCCAAATCGACCTTCTTAGACCCCGGTTTAAAAAGATTGCCTTTGGAAATGAAAGTAACTCTAGTAATGCTCAAAAATGGTTGGGAGAA 2400
L G Q P E S D K M C A K S T F L D P R F K K I A F G N E S N S S N A Q K W L G E
2401 GAGGTATCAGCATTCATAGAGCGGAACCAAAGGACTGCTACAGCCCCAGTCATAGAATTACCCGCGGATAAAAGTAAATATCTCTTTGGACTCTACTTGACCAAAAAGTAGCGGAAGCA 2520
E V S A F I E R N Q R T A T A P V I E L P A D K S K L S L W T L L D Q K V A E A
2521 AAAACTATTTGTCAACGCCCTAGTGTAAATGCACATATTTTCGTTGGAACAATATCTTAGACAAGATTTTCGTTGAGAGACATCAAAACCCGTTAACTATTTGGGACAGCAAAAAGGCA 2640
K T I C H N A P S V N A H I S L E Q Y L R A Q D F V E R H Q N P L N Y W D S K K A
2641 ACTTTTCCAGAACTCTACGAGCTTTCCAACAAAATATTTATGTATACCTGCTACTTCAGTTCCTCCGAAAGGGTTTTTCTAAAGCTGGGCAAATAATAAATGATAGAAGAAATAGACTT 2760
T F P E L Y E L S N K Y L C I P A T S V P S E R V F S K A G Q I I N D R R N R L
2761 AAAGGTGAAAAGCTAGATCAAAATAATGTTTTTAAATAGCAATTTTAAATATAAATTAATATTTTTTTCGCTTCTCTGATATTTATATTTATATTTATATTTATATGATATTGAA 2880
K G E K L D Q I M F L N S F N I *
2881 TATCTTTTCTGTGCTTTACTATTTTTATGTTTTATTAATAAATATTGAATTAATAAATTAACCTTTGTATTGGTTAATTAATAAATGAATAAATGCCAGTACTTTTTTGTGTTTGAAGAGCA 3000
3001 CGCCTTGTCTTAGTTATTTGAATTTGATTAATAAATGAATGAATGAATGAACGAAAAATTCATTTCGATAGTCAAAATCATTTCAGTCATTAACAATCGATAGTTGAAATCACTCGATAGT 3120
3121 TCCCAACTA

Figure 1. Nucleotide sequence and conceptual translation of the putative transposase reading frame of the *B. dorsalis hopper*^{Bd-we} transposable element. In the nucleotide sequence the inverted terminal repeat sequences are double-underlined, and the sub-terminal inverted repeat sequences single underlined. Putative CAAT, TATA, and AATAAA polyadenylation signal site sequences are underlined and in bold. In the amino acid sequence, shown below the nucleotide sequence, the putative BED zinc finger DNA binding domain is underlined with conserved aromatic acid (VWF) and cysteine/histidine (CH) residues in bold. Sequences have been deposited in GenBank accession #AF486809.

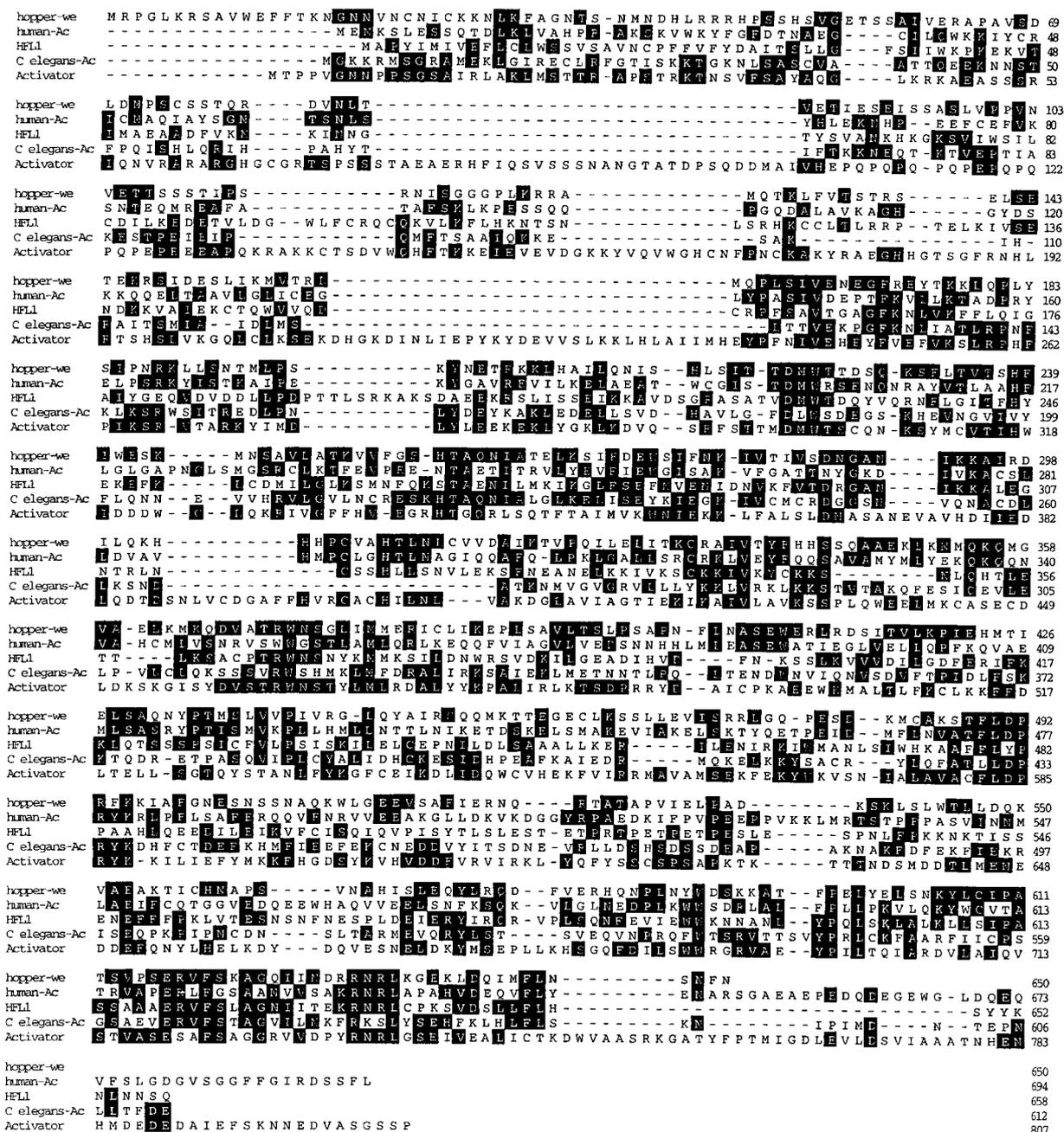


Figure 2. ClustalV multiple sequence alignment with PAM250 residue weight table of the transposase and putative transposase reading frames of *hopper*^{Bd-we}, Human-Ac (GB acc #NM_004729), HFL1 (*hobo*) (GB acc #A39652), *C. elegans*-Ac (GB acc #NP_509255), and *Activator* (GB acc #P08770). Introduced gaps are shown by hyphens and amino acids boxed in black represent majority consensus residues. The ClustalV method is used for improved alignment of N-terminal sequences.

Conserved regions within the *hAT* coding region have generally been divided into three domains, with the highest conservation existing in the C-terminal domain, designated as region 3 (Calvi et al., 1991; see Kempken & Windhofer, 2001). Some of the strongest

regions of *hopper* conservation exist, however, outside these domains. These include the HTAQNIA motif (aa 259–265) that is identical in *C. elegans*-*hobo* (GB acc #NP_509255) and GANIKKA (aa 290–296) that is identical in HFL1 *hobo* in *D. melano-*

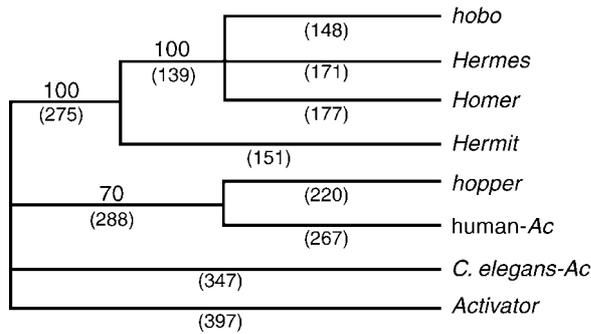


Figure 3. Phylogenetic relationships of indicated *hAT* amino acid sequences by parsimony analysis of a ClustalW multiple sequence alignment using PAUP v. 3.1.1. The *Activator* sequence is the defined outgroup with the percent support given for branches after bootstrap analysis with 1000 replications. The number of amino acid changes are indicated as branch lengths given in parentheses.

gaster (GB acc #A39652), both of which exist in between regions 1 and 2 (see Figures 1 and 2). An LDPR core motif (aa 490–493) in between regions 2 and 3, that was previously observed to be conserved only between *hopper*^{Bd-Kah} and noninsect *hAT*s (Handler & Gomez, 1997), can be extended to include proximal residues having identity and similarity to a sequence in the human-*Ac* element (Esposito et al., 1999). There is no known function for these consensus motifs, though one of the most highly conserved domains in region 3 has been shown, in *Activator*, to be involved in transposase dimerization (Essers et al., 2000). The highly conserved signature sequence motif WWxxxxxxxPxLxxxAxxxL is only partially conserved in *hopper*^{Bd-we} (xWxxxxxxxPxLxxxxxxxL; aa 588–607). Another conserved domain having an implied function within *hAT* elements is the BED zinc finger domain at the N-terminus (Aravind, 2000), which corresponds to amino acids 6–48 in *hopper*^{Bd-we} (Figure 1). BED fingers generally include 50–60 residues that share two highly conserved aromatic amino acids and a common pattern of cysteines and histidines to form a predicted zinc finger. DNA binding is suggested for the BED finger that is consistent with transposase interactions with the terminal repeat sequences.

Southern DNA hybridization

Genomic DNA from several strains of *B. dorsalis*, and the melonfly, *B. cucurbitae*, were hybridized to a 2.0 kb internal fragment of *hopper*^{Bd-we}. Although another *hobo/Ac*-related element was discovered in *B. cucurbitae* from the same PCR experiments that

revealed *hopper* in *B. dorsalis*, (Handler & Gomez, 1996), it was suspected that *hopper* might also exist in *B. cucurbitae* after sequence analysis of additional PCR products (A.M.H., unpublished). Genomic DNA was digested with *NsiI* which should yield an internal 1.7 kb fragment, and additional fragments for the 5' and 3' arms of the transposon. The probe used overlaps the 3' *NsiI* site and thus should hybridize to the internal *NsiI* fragment and the 3' arm of each genomic *hopper* element. All four strains of *B. dorsalis* showed identical hybridization patterns that included the 1.7 kb internal fragment and an additional six bands each representing an individual genomic element (Figure 4). This number of elements is generally consistent with previous hybridizations using the 436 bp internal Bd-HRE PCR product as probe to *BamHI* and *HindIII* digested DNA, though the previous study suggested that some strains might have additional elements or elements with varying genomic loci (Handler & Gomez, 1996). In contrast to *B. dorsalis*, the four *B. cucurbitae* strains all exhibited the internal 1.7 kb fragment with additional fragments consistent with two *hopper*-related elements in the genome of this species.

Conclusions

We describe here a new *hopper* element from the *Bactrocera dorsalis* white eye mutant strain, *hopper*^{Bd-we}, that shares the same terminal and subterminal inverted repeat sequences with *hopper*^{Bd-Kah}, but which has a single open reading frame and is found in a 8 bp duplicated insertion site consistent with other *hAT* elements. These features for *hopper*^{Bd-we} are consistent with transpositional activity, but this remains undetermined at present. While *hopper*^{Bd-we} is distantly related to all other known *hAT* elements, Southern hybridization showed that closely related elements exist in *B. cucurbitae*, suggesting co-evolution of *hopper* or a common ancestor in the two species. Another *hAT*-related element in *B. cucurbitae* discovered by PCR, Bc-HRE, also exists in *B. dorsalis* (Handler & Gomez, 1996), and partial sequencing of PCR products and a genomic clone for Bc-HRE indicates that it is much more closely related to the known insect *hAT* elements (Handler & Gomez, 1996; A.M.H., unpublished).

The lack of a duplicated insertion site in *hopper*^{Bd-Kah}, in addition to internal nucleotide differences with *hopper*^{Bd-we}, suggests that it has existed for a time sufficient to accumulate numerous muta-

Table 1. Percent identity of the indicated *hAT* amino acid sequences based on sequence pair distances from a ClustalW multiple sequence alignment^a

	1	2	3	4	5	6	7	8
1 <i>hopper-we</i>	–	23.7	18.6	18.1	20.7	15.0	19.3	20.0
2 Human-Ac		–	13.5	13.7	18.6	15.6	17.0	15.0
3 <i>hobo</i> (HFL1)			–	54.6	41.5	51.8	12.7	12.0
4 <i>Hermes</i>				–	39.8	51.0	13.7	12.9
5 <i>Hermit</i>					–	43.7	12.9	16.4
6 <i>Homer</i>						–	15.0	13.5
7 <i>C. elegans</i> -Ac							–	15.5
8 Activator								–

^a Transposase and putative transposase reading frames were obtained from GenBank submissions listed for Figure 2 in addition to *Hermes* (acc #U36211), *Hermit* (acc #AAA64851), and *Homer* (acc #AF110403).

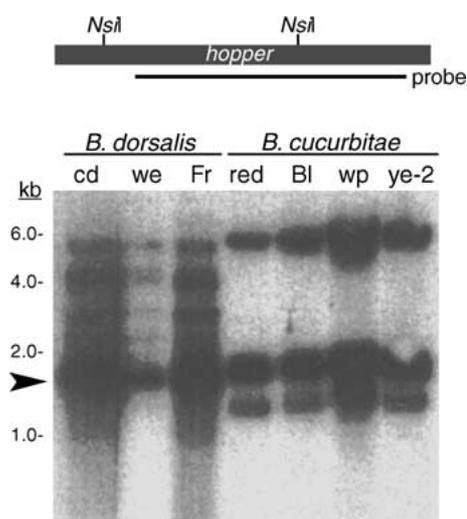


Figure 4. Southern blot DNA hybridization analysis of *hopper* in strains of *B. dorsalis* and *B. cucurbitae*. *B. dorsalis* strains are curled (cd), white eye (we), and furrow (Fr), and *B. cucurbitae* strains are reddish (red), blue (Bl), white puparium (wp), and yellow eye-2 (ye-2). On top is a schematic (not to scale) of the *hopper* element and relative locations of the *Nsi*I sites used to digest DNA, and the bar indicates the *Bst*Z17 I *hopper* fragment used as a hybridization probe. DNA size markers are shown to the left of the autoradiogram and 1.7 kb fragments are denoted by the arrow.

tions. The existence of *hopper*^{Bd-we}, having structure more consistent with functional elements indicates that *hopper* is not an ancient relic within the species, but its distance from all the other known insect *hAT* elements raises the possibility that it was introduced from a non-insect system. Clustal W alignment, parsimony analysis, and BLASTP comparisons show a weak relationship between *hopper* and *hAT* elements from other insects, as well as from other animals and plants. In these comparisons, *hopper* is most

closely related to an Ac-related element from humans (Esposito et al., 1999) and is most distantly related to *Homer* from another bactrocerid, *B. tryoni* (Pinkerton et al., 1999). The *hopper* ITR sequence also shares consensus identity with *hAT* ITRs from plants, and not from insects. These specific comparisons are consistent with more comprehensive comparisons made for 147 *hAT* elements from various plant and animals that place *hopper* in the Tramp cluster containing elements from both vertebrates and invertebrates (Rubin et al., 2001). Most of the *hAT* clusters from this analysis kept plant, fungal, and insect elements in independent groupings, and the authors concluded that no evidence exists for trans-kingdom horizontal transmission. However, the presence of *hopper*, as the only insect *hAT* in the Tramp cluster, with elements from humans and *C. elegans* suggests that the separation of *hopper* from the insect *hobo* cluster may have resulted from the introduction of *hopper*, or an ancestor, from a non-insect system. Alternatively, *hopper* may represent an early branch of *hAT* elements in insects that co-evolved with those that are more closely related to one another.

Acknowledgements

Grateful appreciation is extended to Susan McCombs and Stephen Saul for providing *Bactrocera* pupae, Sheilachu Gomez for assistance with sequencing, Jamie Stephens for assistance with alignment analysis, and to Garnet Suck and Rod Nagoshi for comments on the manuscript. Support was provided by the USDA-NRI Competitive Grants Program and USDA-APHIS-PPQ.

References

- Altschul, S.F., T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller & D.J. Lipman, 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res.* 25: 3389–3402.
- Aravind, L., 2000. The BED finger, a novel DNA-binding domain in chromatin-boundary-element-binding proteins and transposases. *Trends Biochem. Sci.* 25: 421–423.
- Becker, H.A. & R. Kunze, 1997. Maize activator transposase has a bipartite DNA binding domain that recognizes subterminal sequences and the terminal inverted repeats. *Mol. Gen. Genet.* 254: 219–230.
- Calvi, B.R., T.J. Hong, S.D. Findley & W.M. Gelbart, 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., N.K. Johnson, H.D. Perkins, A.J. Howells & D.A. O'Brochta, 1996. The *hermit* transposable element of the Australian sheep blowfly, *Lucilia cuprina*, belongs to the *hAT* family of transposable elements. *Genetica* 97: 23–31.
- DeVault, J.D. & S.K. Narang, 1994. Transposable elements in Lepidoptera: hobo-like transposons in *Heliothis virescens* and *Helicoverpa zea*. *Biochem. Biophys. Res. Comm.* 203: 169–175.
- Esposito, T., F. Gianfrancesco, A. Ciccodicola, L. Montanini, S. Mumm, M. D'Urso & A. Forabosco, 1999. A novel pseudoautosomal human gene encodes a putative protein similar to Ac-like transposases. *Hum. Mol. Genet.* 8: 61–67.
- Essers, L., R.H. Adolphs & R. Kunze, 2000. A highly conserved domain of the maize activator transposase is involved in dimerization. *Plant Cell.* 12: 211–224.
- Feldmar, S. & R. Kunze, 1991. The ORF α protein, the putative transposase of maize transposable element Ac, has a basic DNA binding domain. *EMBO J.* 10: 4003–4010.
- Handler, A.M. & S.P. Gomez, 1996. The *hobo* transposable element excises and has related elements in tephritid species. *Genetics* 143: 1339–1347.
- Handler, A.M. & S.P. Gomez, 1997. A new *hobo*, *Activator*, *Tam3* transposable element, *hopper*, from *Bactrocera dorsalis* is distantly related to *hobo* and *Ac*. *Gene* 185: 133–135.
- Higgins, D.G., A.J. Bleasby & R. Fuchs, 1992. CLUSTAL V: improved software for multiple sequence alignment. *Comput. Appl. Biosci.* 8: 189–191.
- Kempken, F. & F. Windhofer, 2001. The *hAT* family: a versatile transposon group common to plants, fungi, animals, and man. *Chromosoma* 110: 1–9.
- McCombs, S.D. & S.H. Saul, 1992. Linkage analysis of five new genetic markers of the oriental fruit fly, *Bactrocera dorsalis* (Diptera: Tephritidae). *J. Heredity* 83: 199–203.
- McCombs, S.D., D.O. McInnis & S.H. Saul, 1995. Genetic studies of the melon fly, *Bactrocera cucurbitae* (Diptera: Tephritidae). In *Fruit Fly Pests A World Assessment of their Biology and Management*, edited by B.A. McPherson & G.S. Steck. St. Lucie Press, Clearwater, FL.
- Pinkerton, A.C., S. Whyard, H.A. Mende, C.J. Coates, D.A. O'Brochta & P.W. Atkinson, 1999. The Queensland fruit fly, *Bactrocera tryoni*, contains multiple members of the *hAT* family of transposable elements. *Insect Mol. Biol.* 8: 423–434.
- Robertson, H.M. & D.J. Lampe, 1995. Distribution of transposable elements in arthropods. *Annu. Rev. Entomol.* 40: 333–357.
- Rubin, E., G. Lithwick & A.A. Levy, 2001. Structure and evolution of the hAT transposon superfamily. *Genetics* 158: 949–957.
- Swofford, D.L., 1993. PAUP: Phylogenetic Analysis Using Parsimony, Version 3.1.1. Smithsonian Institute, Washington, DC.
- Thompson, J.D., D.G. Higgins & T.J. Gibson, 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl. Acids Res.* 22: 4673–4680.
- Warren, W.D., P.W. Atkinson & D.A. O'Brochta, 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. Camb.* 64: 87–97.
- Warren, W.D., P.W. Atkinson & D.A. O'Brochta, 1995. The Australian bushfly *Musca vetustissima* contains a sequence related to transposons of the *hobo*, *Ac*, and *Tam3* family. *Gene* 154: 133–134.