Mol Genet Genomics
DOI 10.1007/s00438-006-0174-0

ORIGINAL PAPER

Molecular characterization and genomic distribution of *Isis*: a new retrotransposon of *Drosophila buzzatii*

M. P. García Guerreiro · A. Fontdevila

Received: 21 February 2006 / Accepted: 18 September 2006

Abstract A new transposable element, *Isis*, is identified as a LTR retrotransposon in *Drosophila buzzatii*. DNA sequence analysis shows that *Isis* contains three long ORFs similar to *gag, pol* and *env* genes of retroviruses. The ORF1 exhibits sequence homology to matrix, capsid and nucleocapsid *gag* proteins and ORF2 encodes a putative protease (PR), a reverse transcriptase (RT), an Rnase H (RH) and an integrase (IN) region. The analysis of a putative *env* product, encoded by the *env* ORF3, shows a degenerated protein containing several stop codons. The molecular study of the putative proteins coded by this new element shows striking similarities to both *Ulysses* and *Osvaldo* elements, two LTR retrotransposons, present in *D. virilis* and *D. buzzatii*, respectively. Comparisons of the predicted *Isis* RT to several known retrotransposons show strong phylogenetic relationships to gypsy-like elements, particulary to *Ulysses* retrotransposon. Studies of *Isis* chromosomal distribution show a strong hybridization signal in centromeric and pericentromeric regions, and a scattered distribution along all chromosomal arms. The existence of insertional polymorphisms between different strains and high molecular weight bands by Southern blot suggests the existence of full-sized copies that have been active recently. The presence of euchromatic insertion sites coincident between *Isis* and *Osvaldo* could indicate preferential insertion sites of *Osvaldo* element into *Isis* sequence or vice versa. Moreover, the presence of *Isis* in different species of the buzzatii complex indicates the ancient origin of this element.

Keywords Transposable elements · Retrotransposons · *Drosophila buzzatii* · *Isis* · Genomic distribution

Communicated by G. Reuter.

Sequence data from this article have been deposited in the EMBL/GenBank Data Librairies with the accession number DQ408427.

M. P. G. Guerreiro (✉) · A. Fontdevila
Departament de Genètica i Microbiologia,
Edifici C. Universitat Autònoma de Barcelona,
Bellaterra, Barcelona 08193, Spain
e-mail: mariapilar.garcia.guerreiro@uab.es

Introduction

Retrotransposons, or class I elements, are mobile elements that transpose through an RNA intermediate. As a consequence of their replicative transposition mechanism, they are generally present in higher number than transposable elements (TE) of other classes. These TE retrotransposons represent the largest proportion in *Drosophila* genome representing at least 15% (Kaminker et al. 2002). They also account for a large fraction of other eukaryotic genomes; for example, they represent more than 50% of the maize genome (SanMiguel et al. 1996, 1998) and nearly half (42%) of the human genome (IHGSC 2001).

The *Drosophila buzzatii* species belong to the *D. repleta* group (Wasserman 1982) and shows a nonsatellite repetitive DNA content (19–26%) higher than that found in *D. melanogaster* (Marin et al. 1992). Different TEs were isolated in this species including elements of class I (Francino et al. 1993; Labrador and Fontdevila 1994; Pantazidis et al 1999; García Guerreiro and Fontdevila 2001) and class II (Cáceres et al. 1999; Casals et al. 2005). The best-characterized retrotransposon is the *Osvaldo* element, isolated from a recent insertion site, that constitutes a new

family of insect retroviruses not closely related to the *gypsy* element (Pantazidis et al. 1999). In a previous study, we identified an *Osvaldo* element inserted in the 2F4a chromosomal site in different colonizer populations, flanked by a putative repetitive sequence. Here, we have sequenced and characterized this new repetitive and unannotated sequence that we name *Isis.* Structural and phylogenetic analyses show that *Isis* is a new retrotransposon that belongs to the gypsy-like family and is phylogenetically close to the *Ulysses* retrotransposon of *D. virilis*.

Materials and methods

Drosophila stocks

Ca107J9 is an inbred line of *D. buzzatii* originated from a population collected in Carboneras (Spain) in 1996 (García Guerreiro and Fontdevila 2001) and maintained by brother–sister matings.

63 42/7 F81 originated from *D. buzzatii* stock Bu42 (28/7) collected in Maimará (Argentina) and maintained by brother–sister matings during the initial 63 generations and thereafter by mass-culturing (Labrador and Fontdevila 1994; Labrador et al. 1998). This line is devoid of *Osvaldo* euchromatic insertions (Labrador et al. 1998; Labrador and Fontdevila 1994).

Carboneras, Puebla del Río, Cullera, Mazarrón and Sanlúcar are sets of *D. buzzatii* isofemale lines established from the laboratory stocks derived from natural populations collected at these localities in Spain. These isofemale lines were maintained by brother–sister matings and numbered as shown in Table 2.

Berna and Castelli are *D. buzzatii* laboratory stocks derived from natural populations collected in these localities in Argentina in November 1995 and 2006, respectively.

The other species of the *D. buzzatii* complex are as follows (the stock reference number is given in parentheses in cases where strains were obtained from the National *Drosophila* Services Resource Center): *D. koepferae* from S. Luis, Argentina; *D. serido* from Rio Paraguassa, Brazil (15081–1431.1); *D. venezolana* and *D. martensis* from Guaca, Venezuela; *D. uniseta* from Salamanca, Colombia; *D. stalkeri* from Saint Petersburg (15801–1321.0); *D. richardsoni* from La Parguera, Puerto Rico (15801–1421.0). We used also *D. mulleri* from Panuco, Mexico (15081–1371.1), a species of the mulleri complex.

Isolation, cloning and DNA sequencing of *Isis*

Total genomic DNA was extracted according to Piñol et al. (1988) from Ca107J9 line. DNA was partially digested with *Sau*3A and treated with RNAse and alkaline phosphatase according to Maniatis et al. (1982). A quantity of 0.3 μg genomic DNA was then ligated with 1 μg of lambda DASH II arms predigested with *BamH*I. Recombinant DNA was packaged using Gigapack III XL extracts (Stratagene). This *Ca107J9* genomic library was screened with an *Osvaldo* probe containing a 6.4 kb *EcoR*I *Osvaldo* fragment (Pantazidis et al. 1999) inserted in the Bluescript KS+ plasmid. A positive clone was selected that showed a repetitive in situ hybridization pattern on the 63 42/7 F81 stock. Because of the absence of euchromatic *Osvaldo* positions in this empty stock, the hybridization patterns observed suggested the presence of a repetitive element different but associated with an *Osvaldo* high insertion frequency site located at 2F4a cytological position in the *Ca107J9* line.

Different overlapping restriction fragments of the clone were cloned in pBSK plasmid and then sequenced. Sequencing was performed in one strand by Sanger's dideoxynucleotide chain-termination method (Sanger et al. 1977) using the ABI Prism Big-Dye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems) with an automatic ALF sequencer (Pharmacia Biotech, Piscataway, NJ, USA).

Sequences and phylogenetic analysis

Sequences were analyzed using BLAST programs to search sequence similarities to DNA (BLASTN) or proteins (BLASTP). Multiple sequences were aligned by Clustal W program (Higgins et al. 1996). Alignments were edited using the Genedoc program (http://www.psc.edu/biomed/genedoc). ORFs were identified in the sequences with the ORF Finder program (http://www.ncbi.nlm.gov/gorf/gorf.html). A phylogenetic tree was constructed with the MEGA v.2.1 program (Kumar et al. 2001) using the neighbor-joining method (Saitou and Nei 1987). Multiple substitution sites were taken into account in distance calculation using the Poisson correction.

In situ hybridization

Polytene chromosome squashes from salivary glands of third-instar larvae were prepared as described in Labrador et al. (1990) and were hybridized with digoxigenin labeled probes from *Osvaldo* and *Isis*. The *Osvaldo* probe consisted of a 6.4 kb *EcoR*I central fragment of the element. The *Isis* probe consisted of a 1.4 kb PCR fragment which included envelope and integrase regions. Prehybridization solutions and posthybridization washes were done following the protocol of Schmidt (1992) by Roche.

PCR amplification

PCR reactions were carried out in a final volume of 50 µl, including 1× activity buffer (Ecogen), 1.6 mM $MgCl_2$, 0.2 mM of each dNTP (Roche), 0.4 µM primer (Roche), template DNA (~10–20 ng), and 0.04 units of Taq polymerase (Ecotaq from Ecogen) per microliter. Amplifications were run in a MJ Research Inc. thermocycler programmed as follows: 5 min of preliminary denaturation at 94°C, 30 cycles of 45 s at 94°C (denaturation), 45 s at specific PCR annealing temperatures, 1.5 min at 72°C (extension) and a final extension for 10 min at 72°C. The amplified samples were then stored at 4°C. Afterwards the PCR product was gel purified with the Geneclean kit (Bio 101) and labeled by random primer with digoxigenin-11 dUTP

Southern blot analysis

Approximately 2.5 µg of genomic DNA was cut either by the enzyme *Sal*I, which has no restriction sites within the *Isis* sequence, or by the enzyme *Xba*I which cuts this sequence in four sites and gives 0.8, 1.2 and 1.5-kb-long fragments and then DNAs were transferred to nylon positively charged membranes (Roche). We used a dig-labeled DNA probe of 3.4 kb (Fig. 1) including the *pol–env* region. Hybridization was performed in a solution containing 5× SSC, 50% of formamide, 0.1% of *N*-laurylsarcosine, 0.02% of SDS and 5% of blocking reagent. Hybridization was allowed for 12–16 h at 42°C and filters were washed twice for 5 min in 2× SSC, 0.1% SSC and then twice in 0.1× SSC, 0.1% SDS for 15 min at 68°C. Interspecific hybridization, with the same probe, was carried out at 37°C and the filters were washed under medium astringency conditions (2× SSC, 0.1% SDS 10 min, room temperature and 0.5× SSC, 0.1% SDS 30 min, 50°C).

RNA isolation and RT-PCR amplification

Total RNA was isolated from five adult flies using RNeasy mini kit from Qiagen and treated with DNase I from DNA free kit (Ambion) to eliminate DNA contamination. First-strand cDNAs were synthesized with the Invitrogen Thermoscript RT-PCR system using a specific reverse primer for *Isis* (isis1: 5′ AAG GCA TTT GAT GTG AGG TC 3′) located at the end of the reverse transcriptase (RT) region of *Isis*. Negative reactions without retrotranscriptase were carried out to test for DNA contamination. For the PCR reactions we used the reverse primer used for cDNA synthesis plus the forward primer (solap: 5′ ATC AGT CCC TGG AGT TCA CC 3′) designed at the beginning of the RT coding region. These pairs of primers amplify the corresponding band of 1.2 kb. Parallel amplifications of a 0.4 kb fragment of the housekeeping gene *Gapdh* were done, as controls of cDNA synthesis and PCR amplification, using primers Gapdh1 (5′ ATG TCG AAG ATT GGT ATT AAT GG 3′) and Gapdh2 (5′ GTT CGA CAC GAC CTT CAT GT 3′)

Results

Structural characteristics of *Isis* retrotransposon

The sequenced *Isis* retrotransposon is 5,570-bp long and was isolated from a clone that did not contain the 3′ LTR sequence. Instead, an *Osvaldo* element is inserted in *Isis env*-like region as represented in Fig. 1. The *Isis* 5′ LTR sequence is 742-bp long, begins and ends with the nucleotides TGT and TACA, respectively, and contains different transcription signals: a putative TATA box (TATAA) located at position 48,

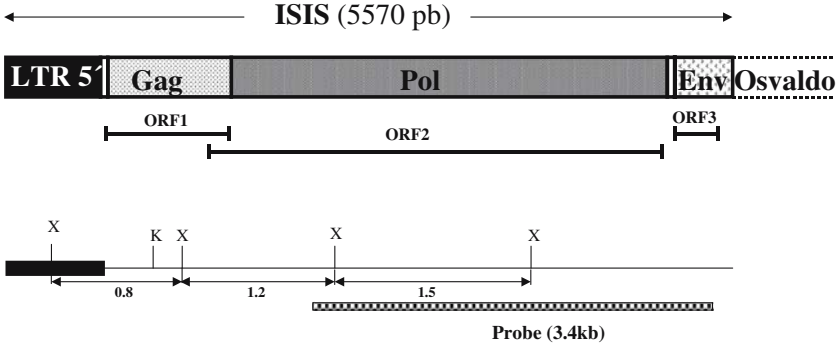

Fig. 1 Structural organization and restriction map of an *Isis* retrotransposon copy isolated from *D. buzzatii*. *Black box* represents 5′ LTR. *Gag* and *env*-like ORFs are depicted by different *dotted rectangles* and *pol* by a *gray rectangle*. *Dotted lines* represent 3′ LTR of an *Osvaldo* insertion. Below are depicted the restriction map of *Isis* and the Southern probe used. Abbreviations of restriction enzymes are as follows: *X Xba*I, *K Kpn*I

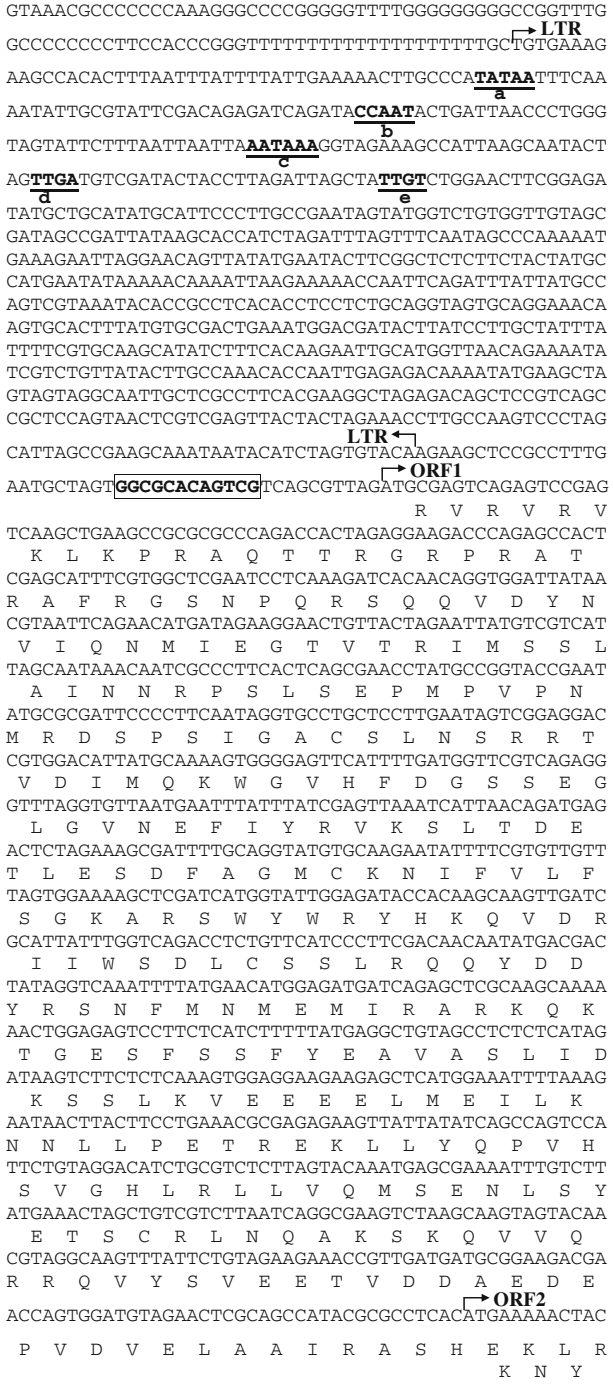

Fig. 2 *Isis* plus strand nucleotide sequence and predicted amino acid sequence. The nucleotide sequences corresponding to the TATA box, CAAT box, polyadenilation site and frameshift sequence are in *bold*, *underlined* and indicated as *a*, *b*, *c*, *d* and *e*, respectively. *Open box* indicates the putative minus DNA strand primer-binding site (PBS). *Asterisks* designate termination codons in ORF1 and ORF2 and stop codons in ORF3

Fig. 2 continued

a CCAAT box at position 88 followed 36-bp downstream by the AATAAA polyadenylation site (Fig. 2). Putative termination signals of RNA synthesis, similar to the TTGT consensus sequence (Temin 1981; Christy et al. 1985), are also found at positions 161 and 190. Downstream to the putative 5′ LTR end, we find a PBS-like sequence (GGCGCACAGTCG) in which 10 nucleotides are homologous to the *Drosophila melanogaster* $tRNA^{Ser}$ gene. The central part of *Isis* comprises three ORFs corresponding to the *gag*, *pol* and

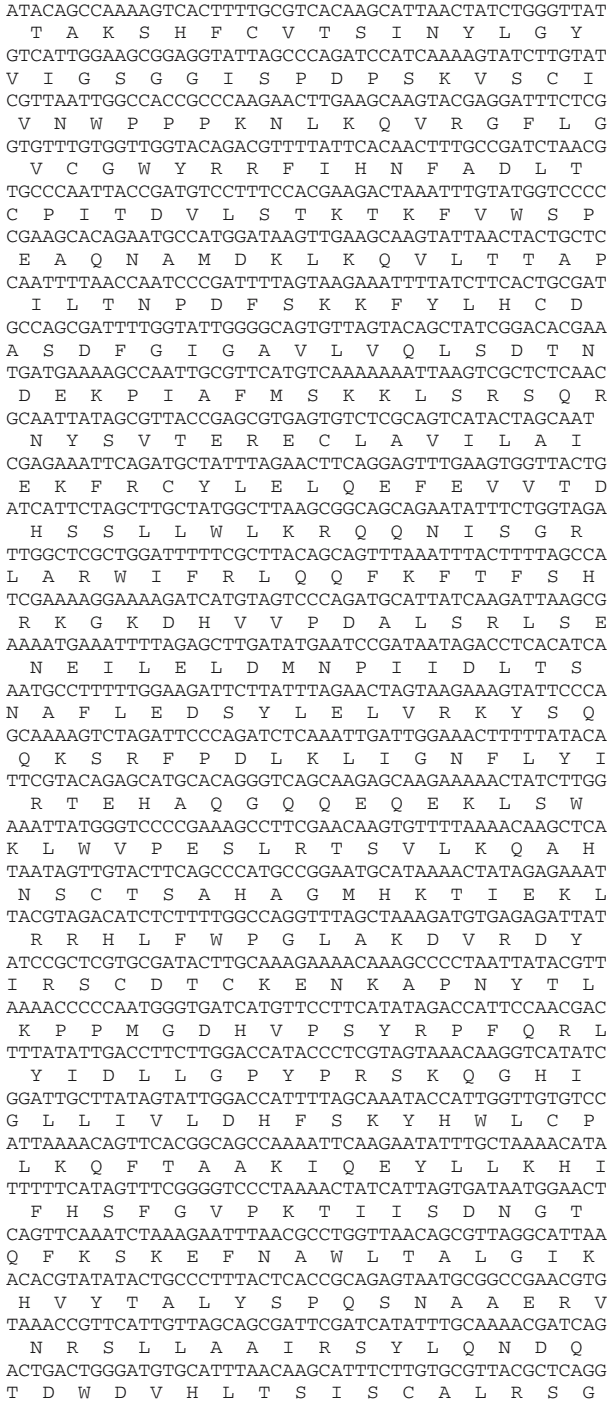

Fig. 2 continued

env retroviral genes, and encodes putative proteins whose amino acid sequences are depicted in Fig. 2. ORFs 1 and 2 are 999- and 3,436-bp long, respectively, and are out of phase by +1. Their overlapping region includes the sequence GAACT, involved in putative frame-shifting events (Yuki et al. 1986). The 355-bp ORF3 fragment is truncated by an inserted *Osvaldo*

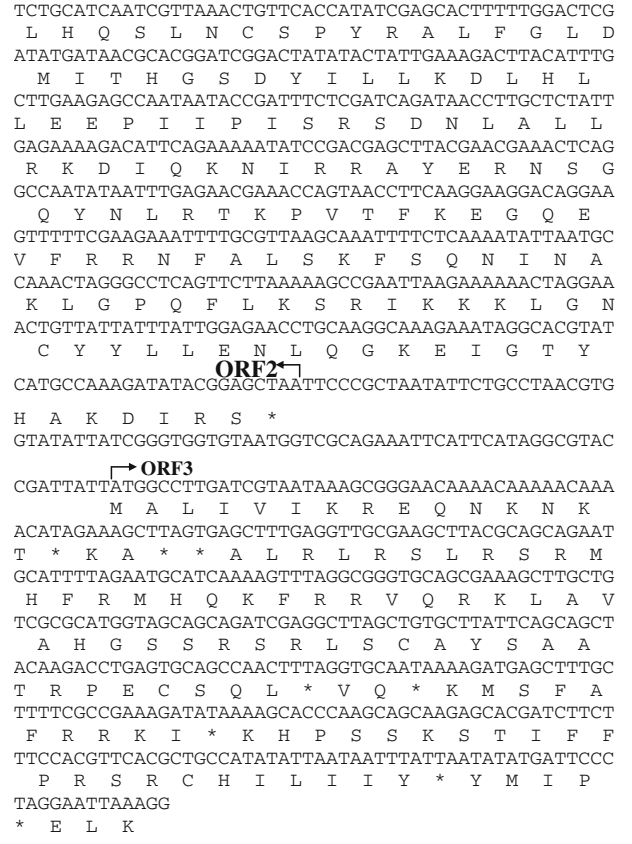

Fig. 2 continued

element. Moreover, *Isis* retrotransposon seems to be related to the gypsy-like elements because of the presence of an *env*-like gene and the order of the different protein domains is the same as in Metaviridae (PR-RT-RH-IN)

Coding region of *Isis*

Isis ORF1 is located in the same position as other retroviral genes and encodes a putative *gag* protein of 331 amino acids. This translational product shows sequence homology to matrix (MA), capsid (C) and nucleocapsid (NC) *gag* proteins of the retrotransposon *Osvaldo* from *D. buzzatii* and *Ulysses* from *D. virilis* (Fig. 3). However, the amino acid sequence differences make alignments difficult and the best alignments were obtained when sequences of the *gag* region were considered separately. Alignments of *Isis* to *Osvaldo* and *Ulysses* separately gave identities about 20% (data not shown). Moreover, a zinc-finger motif ($CX_2CX_4HX_4C$) was detected in the *Isis* NC region as observed in *Osvaldo* retrotransposon and some retroviruses.

Isis ORF2 is apparently complete and includes sequences encoding a putative protease (PR), a RT, an RNase H (RH) and an integrase (IN). The potential PR

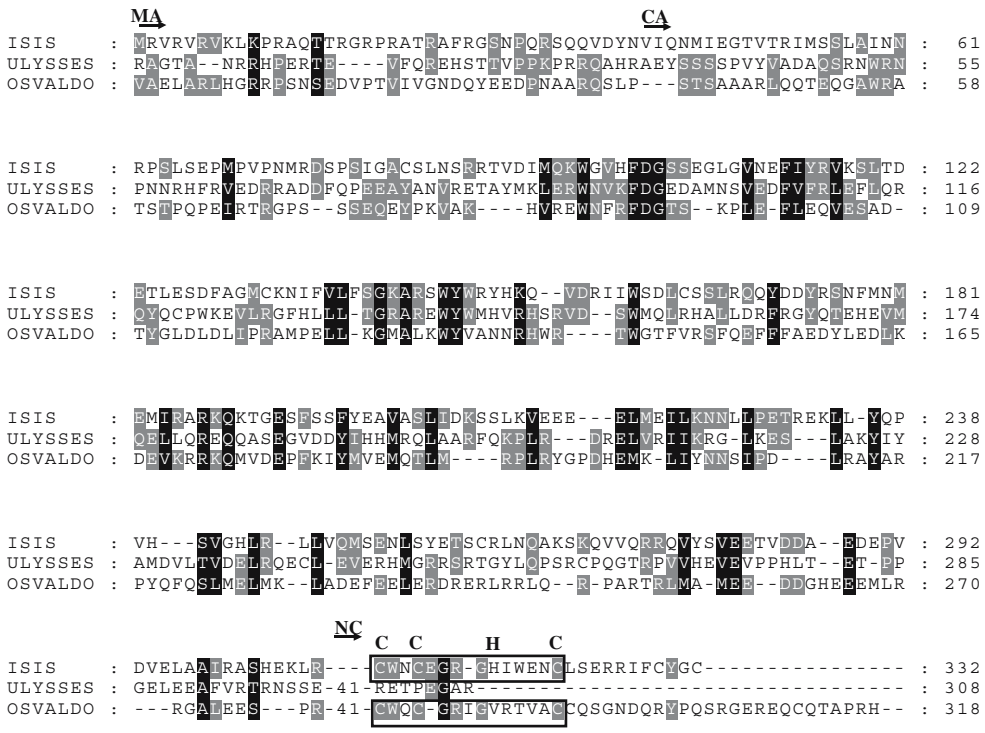

Fig. 3 Multiple sequence alignment of the putative *gag* products: matrix (MA), capsid (CA) and nucleocapsid (NC) of *Isis*, *Ulysses* (Evgen'ev et al. 1992) and *Osvaldo* (Pantazidis et al. 1999) retrotransposons. The positions at which *Isis* shares identical or chemically equivalent amino acids with one or both sequences are in *gray* and *black*, respectively. Just upstream of the CCHC box there are 41 amino acid of *Osvaldo* and *Ulysses* that cannot be aligned to *Isis*. *Open boxes* indicate the characteristic zinc-finger motif

region was detected by the presence of the DTA amino acid sequence found in other retrotransposons and retroviruses (Abe et al. 2000). Multiple sequence alignments of the RT amino acid sequences of *Isis* with *Osvaldo* and other *gypsy*-related elements plus *copia* and *Ty1* are shown in Fig. 4. The seven domains identified in RT sequences (Xiong and Eickbush 1990) are present in *Isis* as detected using CLUSTAL W program alignments modified manually with the help of the GENEDOC program. Sequence identities and similarities are given in Table 1, showing the maximal identity (46%) and similarity (61%) with *Ulysses*. The third putative protein coded by the *pol* gene shows the conserved DAS and GAVL motifs described in RH sequences of retrotransposable elements (Abe et al. 1998). The integrase amino acid sequence, shown in Fig. 5, presents the HHCC and DD35E motifs, characteristic of some retrotransposons and retroviruses (Khan et al. 1990; Capy et al. 1996).

The *Isis* ORF3 looks like a degenerated *env* gene, which includes the glycosylation site (NKT) conforming to the consensus sequence Asn-Xaa-Ser/Thr, typical of *env* proteins. Moreover, the putative degenerated translated product is homologous to a short sequence of the HIV 1 envelope protein (data not shown) and shows homologies to trans-membrane polypeptides of *Osvaldo env* domain (Fig. 6).

Phylogenetic relationships to other retrotransposons

To analyze the phylogenetic relationships between *Isis* and other retroelements, we performed a multiple alignment of RT sequences from different LTR retrotransposons as shown in Fig. 4. The neighbor-joining method (Saitou and Nei 1987) implemented in the MEGA v.2.1 program (Kumar et al. 2001) was used to construct the phylogenetic tree. The tree in Fig. 7 is rooted using a non-LTR retrotransposon (*jockey*) as outgroup to compare different LTR elements. The phylogenetic tree is strongly supported by the branching that separates *Ty1-copia* from *Ty3-gypsy* group. Another well-supported branching separates *Kabuki* from the remaining *Ty3-gypsy*-like retrotransposons in which *Isis* is included. *Isis* clearly clusters to *Ulysses* retrotransposon and its location is stable regardless of whether RT or IN sequences are considered (not shown). In all cases, *Isis* appears more closely related to *micropia* and *Osvaldo* than to other *gypsy*-related retrotransposons, but nodes are not robust.

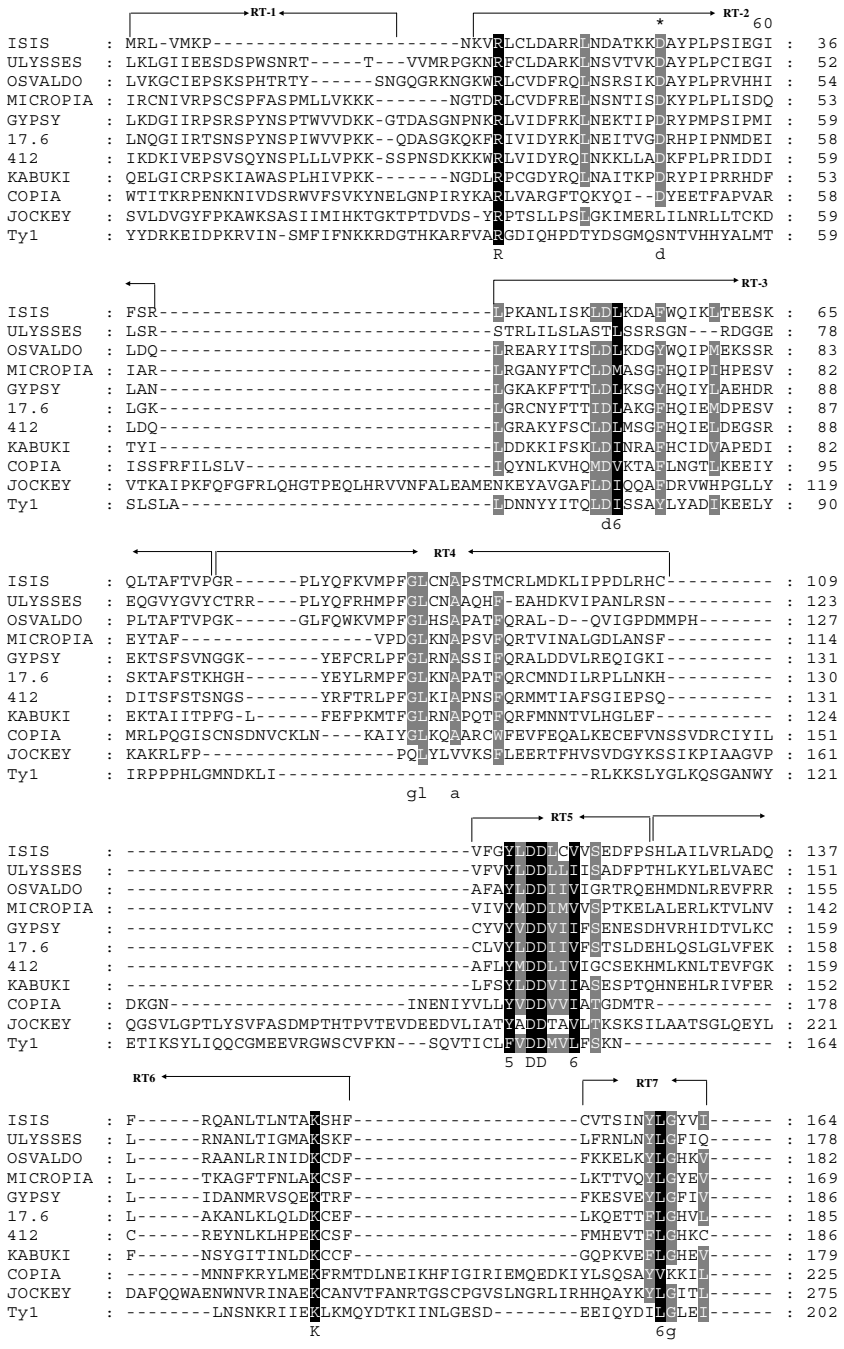

Fig. 4 Multiple sequence alignment of reverse transcriptases (RTs). *Arrows* delimit the seven conserved domains (from RT-1 to RT-7) identified after Xiong and Eickbush 1990. *Numbers* below sequences are biochemically similar amino acids: *5* F, Y, W; *6* L, I, V, M. Elements used in alignment are: *Ulysses* from *D. virilis* (Evgen'ev et al. 1992); *Osvaldo* from *D. buzzatii* (Pantazidis et al. 1999); *micropia* from *D. melanogaster* (Lankenau et al. 1988), *gypsy* from *D. melanogaster* (Marlor et al. 1986); *17.6* from *D. melanogaster* (Saigo et al. 1984), *412* from *D. melanogaster* (Yuki et al. 1986), *Kabuki* from *B. mori* (Abe et al. 2000), *copia* from *D. melanogaster* (Mount and Rubin 1985); *jockey* from *D. melanogaster* (Priimaegi et al. 1988) and *Ty1* from *S. cerevisiae* (Boeke et al. 1988)

Distribution of *Isis* in the genome of *D. buzzatii* and closely related species, and transcriptional activity

To estimate the mean insertion site number of *Isis* along chromosomes, we hybridized different isofemale lines of *D. buzzatii* by in situ hybridization using an *Isis* probe. *Isis* shows a strong hybridization signal in centromeric and pericentromeric regions of all chromosomes, and some euchromatic insertions in chromosomal arms, as shown in Table 2. The mean euchromatic insertion site number per genome (3.7) is higher than that observed for *Osvaldo* (2.1) in the same lines and some chromosomal insertion sites are coincident for both elements. Insertions of the *Isis* element are also present in the line devoid of *Osvaldo* euchromatic copies (63 42/7), suggesting that an *Osvaldo* sequence is not always required for *Isis* insertion.

Southern blot analyses were performed to find out whether full-sized *Isis* copies are present in the genome. When *Xba*I enzyme was used (Fig. 8a) and the filter hybridized with the 3.4-kb probe, the 1.5-kb expected band was observed in all stocks. The other two bands

Table 1 Comparison of the amino acid sequence of the RT domain of *Isis* with other retrotransposons

Retrotransposon	Identity[a] (%)	Similarity[b] (%)
Ulysses	46	61
Osvaldo	39	55
Micropia	34	48
Gypsy	32	54
17.6	32	54
412	30	50
Kabuki	31	46
Copia	4	9

[a] Percent of identical amino acids

[b] Percent of biochemically equivalent amino acids

were not detected because they present no homology (0.8-kb band) or only 300-bp (1.2-kb band) homology with the used probe (Fig. 1). The 1.2-kb band, even if it has a little homology, is not detected perhaps due to the severe astringency washes used after the hybridization process. Digestions with *Sal*I (Fig. 8b) gave high molecular weight bands representing probably full-length copies in all strains excepted in Castelli and Berna, where an extra 4.8-kb band corresponding probably to noncanonical elements was observed.

Southern blot of different species of the buzzatii complex (Fig. 9) showed the presence of *Isis* in all species analyzed. Moreover, the two expected bands (1.5 and 1.2 kb) revealed after digestion with *XbaI* were present in all species excepted in *D. mulleri* where a smear was observed and *D. richardsoni* (1.2-kb band). It is necessary to note that in this case the 1.2-kb band is observed in general because of the medium severity washes used after the hybridization process. Intensity differences could depend on the copy number of *Isis* elements with a canonical structure and the smear could be a result of the existence of elements, highly heterogeneous, located at different chromosomal sites.

To test whether *Isis* elements were transcribed in *D. buzzatii*, we performed RT-PCR with primers of the RT region (see Materials and methods). A product of the expected size (1.2 kb) was observed in the three RNA samples (Fig. 10) as well as in the genomic DNA used as positive control. Negative control (without RT) yielded no product, excluding the existence of DNA contamination in RNA samples.

Discussion

In this paper we describe a new *Isis* retrotransposon isolated from the 2F4a site of high insertion frequency of *Osvaldo* retrotransposon in a population of *D. buzzatii*.

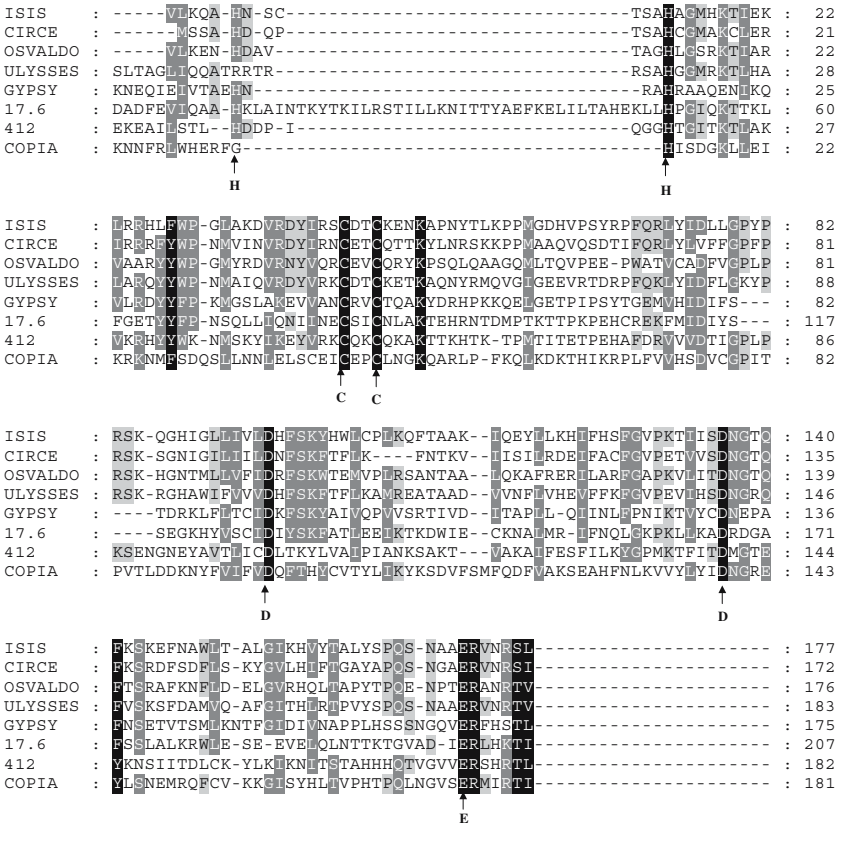

Fig. 5 Multiple sequence alignment of integrase. References of the elements used in alignment are as in Fig. 4 plus *circe* element from *D. melanogaster* (Losada et al. 1997). Common or chemically equivalent amino acids in all sequences are represented in *black*. *Gray* positions show common or chemically equivalent amino acids in most of the analyzed sequences. The characteristic conserved regions HHCC and DD35E are indicated by *arrows*

```
          *       540          *       560          *
Isis  : -----ALIVIKREQNKNKT*KA**ALRLRSLRSRMHFRMHQKFR :  36
Osv   : PSGPGASI---RHQLGLEEWKLTLKIRKRKLKKNIIFTYQQ--- : 566
              A I   R Q       K    6R R L4  6 F   Q

                 580          *       600          *
Isis  : RVQRKLAVAHGSSRSRLSCAYSAATRPECSQL*VQ*KMSFAFRR :  78
Osv   : RTE--LPAANG--KQLVSSAALLTER------------------ : 588
          R 2   L  A G  4  6S A     R
```

Fig. 6 Sequence alignment of *Isis* and *Osvaldo env* products. *Black shaded* amino acid represent common or chemically equivalent positions. *Numbers* below sequences are biochemically similar amino acids: *2* E, Q; *4* K, R; *6* L, I, V, M. *Open box* indicates a putative glycosilation site of *env* gene. *Asterisks* represent stop codons

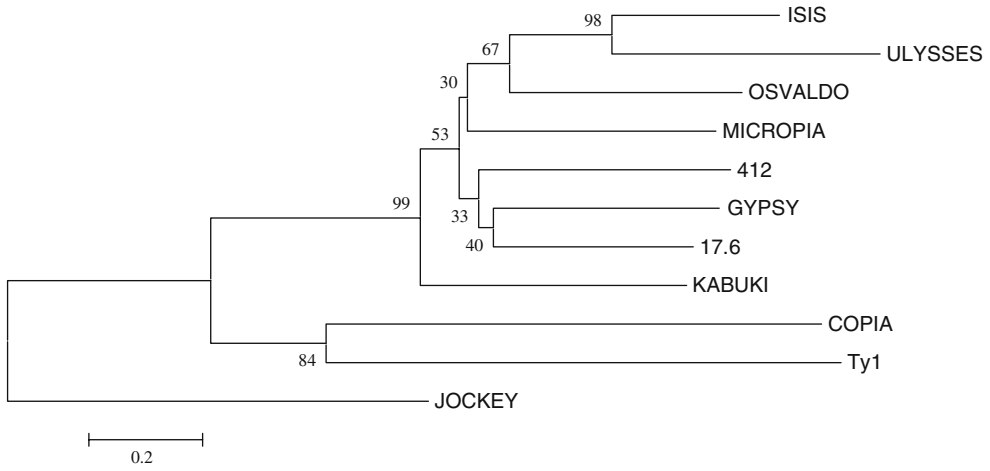

Fig. 7 Neighbor-joining tree based in reverse transcriptase sequence similarity, bootstrap values for 1,000 replicates are indicated at each *node*

Isis has structural characteristics of gypsy-like elements, with an ORF3 corresponding to a degenerated *env*-gene showing eight stop codons. The fact that the putative protein encoded by this ORF presents a glycosylation site, several homologies to *Osvaldo env* protein and the IVIKLREQYKNKT *env* sequence of HIV 1 suggests that it corresponds to an ancient, degenerated *env* gene. No evident homologies were found between this *Isis* and *env* regions of other retrotransposons; most probably because the amino acid sequence conservation of its *env* region is too low (Doolittle et al. 1989; McClure 1991, 1992). It is indeed known that this gene region is the target of high selection pressures that promote a high variability (Hirsch and Curran 1990).

ORFs 1 and 2 encode the *gag* and *pol* proteins necessary to transposition and viral particle formation. Alignments of all conserved protein regions relate the *Isis* element to the gypsy-like group, which includes elements like *Ulysses* of *D. virilis*, *Osvaldo* of *D. buzzatii* and *gypsy* and *17.6* of *D. melanogaster*. Suprisingly, *Isis* shows *gag* sequence homology to *Ulysses* (Evgen'ev et al. 1992) and *Osvaldo* (Pantazidis et al. 1999) despite the low conservation of this region between close retroviruses and retrotransposons (Evgen'ev et al. 1992). In addition, the *gag* sequence has a perfect zinc-finger motif, or CCHC box (Covey 1986), characteristic of all retroviruses, except the spumaviruses (Vogt 1997), and some retrotransposons among them *Osvaldo* (Pantazidis et al. 1999). These motifs have a role in binding proteins to nucleic acids and interact between *gag* monomers (Vincent et al. 1994; Rein et al. 1998; Gorelick et al. 1999). However, in some gypsy-like retrotransposons as *gypsy, 297, ZAM*, *tom* and *17.6*, these motifs are absent and are in some cases substituted by basic amino acids (Leblanc et al. 1997).

Isis element encodes a putative RT that includes the classical seven conserved domains (Xiong and Eickbush 1990). Phylogenetic analysis of this region assigns this retrotransposon to the gypsy-like class, close to *Ulysses* of *D. virilis.* Since *Ulysses* lacks the third ORF found in *Isis*, this new element can be considered as a new retrovirus of *D. buzzatii* different from *Osvaldo*. *Isis* differs also from *Osvaldo* and *Ulysses* in that it has sequences homologous to $tRNA^{Ser}$, located 26 bp

Table 2 Chromosomal insertion profile of *Isis* and *Osvaldo* retrotransposons in different *D. buzzatii* lines

Population	Puebla del Río							
Line	2		3		6		10	
Probe	*Isis*	*Osvaldo*	*Isis*	*Osvaldo*	*Isis*	*Osvaldo*	*Isis*	*Osvaldo*
Chromosomes								
2	F4d	–	F4d	–	F1d	–	F4d	–
3	C1b	–	–	–	C4a	–	C1b	–
4	**C2g, F4b**	**C2g, F4b**	**C2g, F4b**	**C2g, F4b**	**C2g**	**C2g**	**C2g, F4b**	**C2g, F4b**
5	–	C2c	–	–	–	–	–	–

	Cullera				Sanlúcar			
	1		6		4		9	
	Isis	*Osvaldo*	*Isis*	*Osvaldo*	*Isis*	*Osvaldo*	*Isis*	*Osvaldo*
2	**F4a**	**F4a**	**F4a**	**F4a**	E6f, **F1d**	–	**F1d**, F4a	–
3	F4b	F1f	F4b	F1f	–	–	–	–
4	–	–	–	–	–	–	–	–
5	A4b, G1d	A4b	A4b, G1d	A4b	–	–	–	–

	Mazarrón		63 42/7		Ca107J9		Adeje	
	9						10	
	Isis	*Osvaldo*	*Isis*	*Osvaldo*	*Isis*	*Osvaldo*	*Isis*	*Osvaldo*
2	**E5a**, G2c	**E5A**	F1d	–	**F1d**, **F4a**, E4a	**F1d**, **F4a**, D5a, (E4a)	F1d, **F4a**	G2, **F4a**
3	E2f, F1f	–	–	–	–	F2b, B5f	–	–
4	–	–	F2b	–	**C2g**	**C2g**, E1b	–	–
5	–	–	–	–	B1a	A4b, (B1f), (C2a)	–	–

Coincident sites of *Osvaldo* and *Isis* are in bold. In parenthesis are represented the polymorphic sites

downstream of the 5′ LTR end instead of sequences homologous to the $tRNA^{Lys}$ that exist in *Osvaldo* and *Ulysses* and are implicated in the synthesis of the viral DNA minus strand. $tRNA^{Ser}$ is used too as a primer in other insect gypsy-like elements as *Tirant* (Cañizares et al. 2000), *ZAM* (Leblanc et al. 1997), *297* (Inouye et al. 1986), *tom* (Tanda et al. 1994), *17.6* (Saigo et al. 1984) and *TED* (Friesen and Nissen 1990).

The absence of stop codons in *gag* and *pol* regions of *Isis* suggests that the described *Isis* copy could have been active in a recent past. The structure of this copy is truncated by an *Osvaldo* insertion in the *env* region. This insertion probably inactivates the element because the PPT region, important for the synthesis of the plus strand of DNA, and the putative 3′ LTR are separated by the *Osvaldo* insertion. However, *Isis* presents insertion sites scattered over the chromosomes at a location that differs between laboratory lines from different geographical regions. This distribution suggests that *Isis* has been active recently and Southern blot analysis favor the idea of the presence of full-sized elements in the genome.

Isis seems to co-occur with *Osvaldo* at the same cytological site because 6 out of the 18 *Isis* different positions of Table 2 coincide with *Osvaldo* sites. By in situ hybridization, it is impossible to detect whether *Osvaldo* and *Isis* are inserted adjacent in the remaining five sites, not sequenced in this work, but it is not impossible. It is known that some retrotransposons insert inside others in maize (SanMiguel et al. 1996) and *B. mori* (Abe et al. 2000). We cannot discard the possibility that *Isis* sequence particularities could be attractive to *Osvaldo* insertion. Retrotransposon insertion preferential sites have indeed been reported for various elements. For example, *Penelope* tends to insert in sites where other *Penelope* copies are already present in *D. virilis* and *D. lummei* (Evgen'ev et al. 2000), *Zam* and *Idefix* within the white locus of the RevI strain of *D. melanogaster* (Conte et al. 2000), and *nomad* in the TANA host sequence (Whalen and Grigliatti 1998). On the other hand, analysis of the 5′ LTR flanking sequence of *Isis* shows microsatellite repetitive DNA as shown in retrotransposons of plants (Ramsay et al. 1999), humans (Nadir et al. 1996) and *Drosophila* (Kolesnikov et al. 1996). The mechanisms by which these elements select insertion sites with such fidelity are unknown and some authors argue that it is given by the integrase.

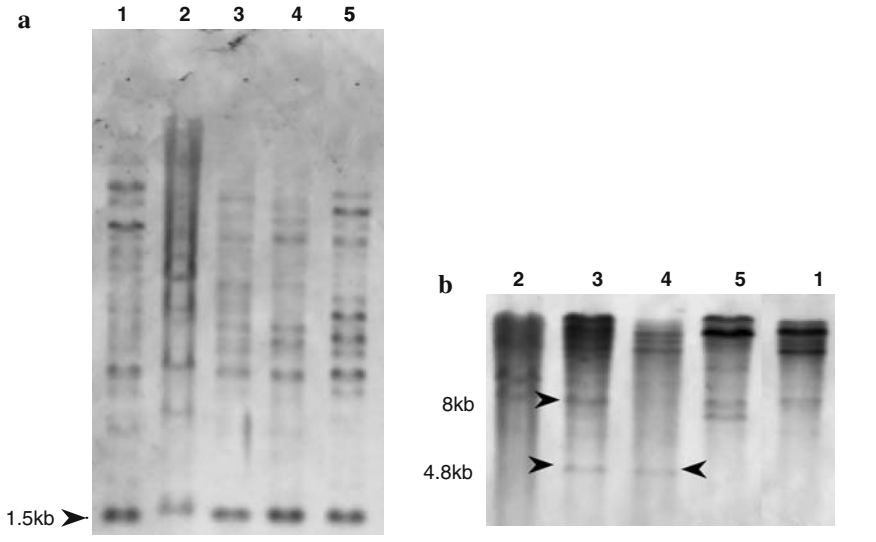

Fig. 8 Southern blot analysis of genomic DNA from different *D. buzzatii* strains hybridized with a 3.4-kb probe including *Isis* positions from 2,124 to 5,528 (Fig. 1). **a** Southern blot after digestion with *Xba*I. **b** Southern blot after digestion with *Sal*I. Lanes: *1* Adeje-10, *2* Ca107J9, *3* Castelli, *4* Berna, *5* 63 42/7

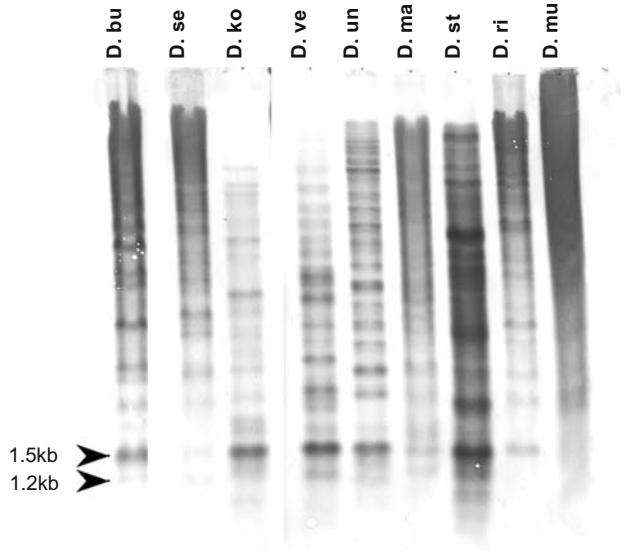

Fig. 9 Southern blot analysis of genomic DNA from species of the *D. buzzatii* complex digested with *Xba*I and probed with the 3.4-kb probe (described in Fig. 8). Lanes: *D. buzzatii* (*D. bu*), *D. serido* (*D. se*), *D. koepferae* (*D. ko*), *D. venezolana* (*D. ve*), *D. uniseta* (*D. un*), *D. martensis* (*D. ma*), *D. starmeri* (*D. st*), *D. richardsoni* (*D. ri*), *D. mulleri* (*D. mu*)

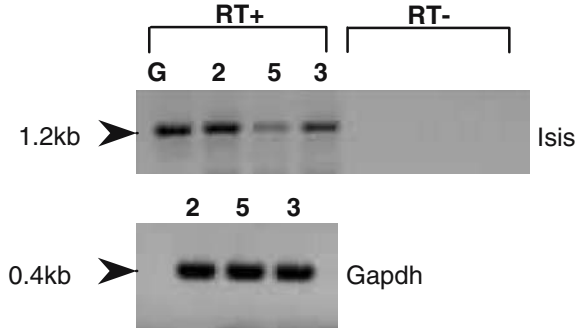

Fig. 10 RT-PCR carried out with total RNA from adult flies. Two sets of identical reactions differing only by the addition (RT+) or omission (RT−) of reverse transcriptase are presented. *Lane G* shows the results of amplification of genomic DNA and *lanes 2*, *5* and *3* show, respectively, amplification of RNA from *D. buzzatii* strains Ca107J9, 63 42/7 and Castelli. The housekeeping gene *Gapdh* was used as positive control of RNA amplification in the same *Drosophila* strains

In our case, where *Isis* and *Osvaldo* are inserted together at chromosomal site 2F4a, we have two possible hypotheses to explain the *Isis* insertion mechanism. First we consider that *Isis* is an old element because of the presence of stop codons in the *env* region. Even if this region has low sequence conservation, we cannot exclude, however, that some insertions or deletions could have been undetected in the other domains because we do not have a canonical *Isis* sequence to be compared. Moreover, it seems that the *Osvaldo* element could be inserted in the *env* region of *Isis*, even if the boundary sequences are not included in the analyzed clone because adjacently we found a 3′ *Osvaldo* LTR. Our previous analysis of this *Osvaldo* insertion showed that it is an old degenerated copy containing insertions, deletions and stop codons. If this hypothesis is correct, *Isis* insertion could be older than the *Osvaldo* one because its insertion occurred before. The second hypothesis considers that *Isis* insertion occurred independently, followed by an ectopic recombination with the *Osvaldo* element inserted in the genome. Ectopic recombination was proposed as a mechanism for tandem element and composite element formation (Ke and Voytas 1997; Mugnier et al. 2005), although the mechanisms by which tandem elements arise in *Drosophila* are not known.

Data presented here suggest that *Isis* could be an element that has been active recently in *D. buzzatii.* We have several lines of evidence: the existence of

polymorphic sites between different laboratory strains, the presence of high molecular weight bands by Southern blot when the DNA was digested with *Sal*I (6.2-kb bands are the minimum size expected for complete *Isis* elements) and finally the detection of *Isis* transcripts by RT-PCR.

Isis clearly hybridizes in the nine species tested in this work. This result could indicate that this element is an ancient component of the buzzatii complex and it has had a long history in this group of species. Moreover, the two expected bands revealed with the probe, after digestion with *Xba*I, were observed in all species of this complex excepted *D. mulleri* (mulleri complex). These results show that the element structure was quite conserved during the evolution. Moreover, the fact that other bands were shared between different species could suggest either the existence of old sites antedating speciation and retained in these species, or the association of *Isis* to heterochromatic regions in tandem. High degree of band conservation is not rare in retrotransposons and was reported for elements as *Osvaldo* (Labrador and Fontdevila 1994) and *Ulysses* (Zelentsova et al. 1999), two elements closely related with *Isis*.

In general, results of *Isis* element presented in this paper are encouraging; however, more work must be done to determine its transcriptional levels in different tissues, the distribution out of the buzzatii complex and why *Osvaldo* element seems to have *Isis* as an insertion target sequence.

Acknowledgments This research has been supported by grants BOS2000-0295-C02 and BOS2003-05904-C02 from Ministerio de Ciencia y Tecnología (MCYT) Spain and 2001SGR-00207 from the Generalitat de Catalunya to AF. We acknowledge C. Biémont and two anonymous reviewers for their helpful discussions and comments on the manuscript, and we thank L. Alarcón for providing us with DNA samples from different species of the buzzatii complex.

References

Abe H, Kanehara M, Terada T, Ohbayashi F, Shimada T, Kawai S, Suzuki M, T Sugasaki T, Oshiki T (1998) Identification of novel random amplified polymorphic DNAs (RAPDs) on the W chromosome of the domesticated silkworm, *Bombyx mori*, and the wild silkworm *B. mandarina*, and their retrotransposable-related nucleotide sequences. Genes Genet Sys 73:243–254

Abe H, Ohbayashi F, Shimada T, Sugasaki T, Kawai S, Mita K, Oshiki T (2000) Molecular structure of a novel gypsy- Ty3-like retrotransposon (Kabuki) and nested retrotransposable elements on the W chromosome of the silkwom *Bombyx mori*. Mol Gen Genet 263:916–924

Boeke JD, Eichinger D, Castrillon D, Fink GR (1988) The Saccharomyces cerevisiae genome contains functional and non functional copies of transposon *Ty1*. Mol Cell Biol 8:1432–1442

Cáceres M, Ranz JM, Barbadilla A, Long M, Ruiz A (1999) Generation of a widespread *Drosophila* inversion by a transposable element. Science 285:415–418

Cañizares J, Grau M, Paricio N, Moltó MD (2000) *Tirant* is a new member of the gypsy family of retrotransposons in *Drosophila melanogaster*. Genome 43:9–14

Capy P, Vitalis R, Langin T, Higuet D, Bazin C (1996) Relationships between transposable elements based upon the integrase-transposase domains: is there a common ancestor? J Mol Evol 42:359–368

Casals F, Caceres M, Manfrin MH, González J, Ruiz A (2005) Molecular characterization and chromosomal distribution of *Galileo*, *Kepler* and *Newton*, three *Foldback* transposable elements of the *Drosophila buzzatii* species complex. Genetics 169:2047–2059

Christy RJ, Brown AR, Gourlie BB, Huang RCC (1985) Nucleotide sequences of murine intracisternal A-particle gene LTRs have extensive variability within the R region. Nucleic Acids Res 13:289–302

Conte C, Calco V, Desset S, Leblanc P, Dastugue B, Vaury C (2000) Impact of multiple insertions of two retroelements, *Zam* and *Idefix* at an euchromatic locus. Genetica 109:53–59

Covey SN (1986) Amino acid homology in gag regions of reverse transcribing elements and the coat protein gene of cauliflower mosaic virus. Nucleic Acids Res 14:623–633

Doolittle RF, Feng DF, Johnson MS, McClure MA (1989) Origin and evolutionary relationships of retroviruses. Q Rev Biol 64:1–30

Evgen'ev MB, Corces VG, Lankenau DH (1992) *Ulysses* transposable element of *Drosophila* shows high structural similarities to functional domains of retroviruses. J Mol Biol 225:917–924

Evgen'ev MB, Zelentsova H, Poluectova H, Lyozin GT, Veleikodvorskaja V, Pyatkov KI, Zhivotovsky LA, Kidwell MG (2000) Mobile elements and chromosomal evolution in the *virilis* group of *Drosophila*. Proc Natl Acad Sci USA 10:11337–11342

Francino O, Cabré O, Fontdevila A (1993) Distribution of the *copia* transposable element in the *repleta* group of *Drosophila*. Genet Sel Evol 25:501–516

Friesen PD, Niessen M (1990) Gene organization and trascription of *TED*, a lepidopteran retrotransposon integrated within the baculovirus genome. Mol Cell Biol 10:3067–3077

García Guerreiro MP, Fontdevila A (2001) Chromosomal distribution of the transposable elements *Osvaldo* and *blanco* in original and colonizer populations of *Drosophila buzzatii*. Genet Res 77:227–238

Gorelick RJ, Gagliardi TD, Bosche WJ, Wiltrout TA, Coren LV, Chabot DJ, Lifson JD, Henderson LE, Arthur LO (1999) Strict conservation of the retroviral nucleocapsid protein zinc finger is strongly influenced by its role in viral infection processes: characterization of HIV-1 particles containing mutant nucleocapsid zinc-coordinating sequences. Virology 256:92–104

Higgins DG, Thompson JD, Saigo K (1996) Using CLUSTAL for multiple sequence alignments. Methods Enzymol 266:383–402

Hirsch MS, Curran J (1990) Human immunodeficiency viruses. In: Fields BN, Knipe DM (eds) Virology. Raven Press, New York

IHGSC (International Human Genome Sequencing Consortium) (2001) Initial sequencing and analysis of the human genome. Nature 409:860–921

Inouye S, Yuki S, Saigo K (1986) Complete nucleotide sequence and genome organization of *Drosophila* transposable genetic element, *297*. Eur J Biochem 154:417–425

Kaminker JSC, Bergman CM, Kronmiller B, Carlson J, Svirskas R, Patel S, Frise E, Wheeler DA, Lewis SE, Rubin GM, Ashburner M, Celniker SE (2002) The transposable elements of the *Drosophila melanogaster* euchromatin: a genomics perspective. Genome Biol 3: Research0084

Ke N, Voytas DF (1997) High frequency cDNA recombination of the *Saccharomyces* retrotransposon Ty5: the LTR mediates formation of tandem elements. Genetics 147:545–556

Khan E, Mack JPG, Katz RK, Kulkosky J, Skalka AM (1990) Retroviral integrase domains: DNA binding and the recognition of LTR sequences. Nucleic Acids Res 4:851–860

Kolesnikov AV, Ponomarenko NA, Churikov NA (1996) Analysis of copies of the suffix-retrotransposon of *Drosophila*, produced using PCR on genomic DNA. Genetika 32:150–153

Kumar SK, Tamura IB, Jakobsen M, Nei M (2001) MEGA-2: Molecular Evolutionary Genetics Analysis software. Arizona, USA

Labrador M, Fontdevila A (1994) High transposition rates of *Osvaldo*, a new *D. buzzatii* retrotransposon. Mol Gen Genet 245:661–674

Labrador M, Naveira H, Fontdevila A (1990) Genetic mapping of the *Adh* locus in the *repleta* group of *Drosophila* by in situ hybridization. J Hered 81:83–86

Labrador M, Seleme MC, Fontdevila A (1998) The evolutionary history of *D. buzzatii*. XXXIV. The distribution of the retrotransposon *Osvaldo* in original an colonizer populations. Mol Biol Evol 15:1532–1547

Lankenau DH, Hauijser P, Jansen E, Miedema K, Henning W (1988) *Micropia*: a retrotransposon of *Drosophila* combining structural features of DNA viruses, retroviruses and non-viral transposable elements. J Mol Biol 204:233–246

Leblanc P, Desset S, Dastugue B, Vaury C (1997) Invertebrate retroviruses: *ZAM* a new candidate in *D. melanogaster*. EMBO J 16:7521–7531

Losada A, Abad JP, Villasante A (1997) Organization of DNA sequences near the centromere of the *Drosophila melanogaster* Y chromosome. Chromosoma 106:503–512

Maniatis TE, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor

Marin I, Labrador M, Fontdevila A (1992) The evolutionary history of *Drosophila buzzatii*. XXII. High content of nonsatellite repetitive DNA in *D. buzzatii* and its sibling *D. koepferae*. Genome 35:967–974

Marlor RL, Parkhust SM, Corces VG (1986) The *Drosophila melanogaster gypsy* transposable element encodes putative gene products homologous to retroviral proteins. Mol Cell Biol 6:1129–1134

McClure MA (1991) Evolution of retrotransposons by acquisition or deletion of retrovirus-like genes. Mol Biol Evol 8:835–856

McClure MA (1992) Sequence analysis of eukaryotic retroid proteins. Math Comput Model 16:121–136

Mount SM, Rubin GM (1985) Complete nucleotide sequence of the *Drosophila* transposable element *copia*: homology between *copia* and retroviral proteins. Mol Cell Biol 5:1630–1638

Mugnier N, Biémont C, Vieira C (2005) New regulatory regions of *Drosophila* 412 retrotransposable element generated by recombination. Mol Biol Evol 22:747–757

Nadir E, Margalit H, Gallil T, Ben-Sasson SA (1996) Microsatellite spreading in the human genome. Evolutionary mechanisms and structural implications. Proc Natl Acad Sci USA 93:6470–6475

Pantazidis A, Labrador M, Fontdevila A (1999) The retrotransposon *Osvaldo* from *D. buzzatii* displays all structural features of functional retrovirus. Mol Biol Evol 16:909–921

Piñol J, Francino O, Fontdevila A, Cabré O (1988) Rapid isolation of *Drosophila* high molecular weight DNA to obtain genomic libraries. Nucleic Acids Res 16:2736–2737

Priimaegi AF, Mizhrokhi LJ, Ilyin YV (1988) The *Drosophila* mobile element *jockey* belongs to LINEs and contains coding sequences homologous to some retroviral proteins. Gene 70:253–262

Ramsay L, Macaulay M, Cardle L, Morgante M, Ivanissevich S, Maestri E, Powell Waugh R (1999) Intimate association of microsatellite repeats with retrotransposons and other dispersed repetitive elements in barley. Plant J 17:415–425

Rein A, Henderson LE, Levin JG (1998) Nucleic-acid-cahaperone activity of retroviral nucleocapsid proteins: significance for viral replication. Trends Biochem Sci 23:297–301

Saigo K, Kugimiya W, Matsuo Y, Inouye S, Yoshioka K, Yuki S (1984) Identification of the coding sequence for a reverse transcriptase-like enzyme in a transposable genetic element in *Drosophila melanogaster*. Nature 312:659–661

Saitou N, Nei M (1987) The Neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467

Schmidt ER (1992) A simplified and efficient protocol for non-radiactive in situ hybridisation to polytene chromosomes with a DIG.-labeled DNA probe. In: Non-radioactive in situ hybridisation application manual. Roche, pp 36–38

SanMiguel P, Tikhonov A, Jin YK, Motchoulskaia N, Zakharov D, Melake-Berhan A, Springer PS, Edwards KJ, Lee M, Avramova Z, Bennetzen JL (1996) Nested retrotransposons in the intergenic regions of the maize genome. Science 274:765–768

SanMiguel P, Gaut BS, Thikonov A, Nakajima Y, Bennetzen JL (1998) The paleontology of intergene retrotransposons of maize. Nat Genet 20:43–45

Tanda S, Mullor JL, Corces VG (1994) The *Drosophila tom* retrotransposon encodes an envelope protein. Mol Cell Biol 14:5392–5401

Temin HM (1981) Structure variation and synthesis of retrovirus long terminal repeats. Cell 27:1–3

Vincent JP, Girdham CH, O´Farrell PH (1994) A cell-autonomous ubiquitous marker for the analysis of *Drosophila* genetic mosaics. Dev Biol 164:328–331

Vogt VM (1997) Retroviral virions and genomes. In: Coffin J, Hughes SH, Varmus HE (eds) Retroviruses. Cold Spring Harbor Laboratory press, New York

Wasserman M (1982) Evolution of the *repleta* group. In: Ashburner M, Carlson HL, Thompson JN (eds) The genetics and biology of *Drosophila*. Academic, London, pp 62–139

Whalen JH, Grigliatti TA (1998) Molecular characterization of a retrotransposon in *Drosophila melanogaster*, *nomad*, and its relationship to other retrovirus-like mobile elements. Mol Gen Genet 260:401–409

Xiong Y, Eickbush TH (1990) Origin and evolution of retroelements based upon their reverse transcriptase sequences. EMBO J 9:3353–3362

Yuki S, Inouye S, Ishimaru S, Saigo K (1986) Nucleotide sequence characterization of a *Drosophila* retrotransposon, *412*. Eur J Biochem 158:403–410

Zelentsova H, Poluectova H, Mnjoian L, Lyozin G, Veleikodvorskaja V, Zhivotovsky L, Kidwell MG, Evgen'ev MB (1999) Distribution and evolution of mobile elements in the virilis species group of *Drosophila*. Chromosoma 108:443–456