J Mol Evol (1998) 46:215–224

JOURNAL OF MOLECULAR EVOLUTION

Horizontal Escape of the Novel Tc1-Like Lepidopteran Transposon TCp3.2 into *Cydia pomonella* Granulovirus

Johannes A. Jehle,[1,2]* Antje Nickel,[1] Just M. Vlak,[2] Horst Backhaus[1]

[1] Institut für Biochemie und Pflanzenvirologie, Biologische Bundesanstalt für Land- und Forstwirtschaft, Messeweg 11–12, 38104 Braunschweig, Germany
[2] Department of Virology, Agricultural University Wageningen, P.O.B. 8045, 6700 EM Wageningen, The Netherlands

Received: 27 February 1997 / Accepted: 16 May 1997

Abstract. We characterized an insertion mutant of the baculovirus *Cydia pomonella* granulovirus (CpGV), which contained a transposable element of 3.2 kb. This transposon, termed TCp3.2, has unusually long inverted terminal repeats (ITRs) of 756 bp and encodes a defective gene for a putative transposase. Amino acid sequence comparison of the defective transposase gene revealed a distant relationship to a putative transposon in *Caenorhabditis elegans* which also shares some similarity of the ITRs. Maximum parsimony analysis of the predicted amino acid sequences of Tc1- and mariner-like transposases available from the GenBank data base grouped TCp3.2 within the superfamily of Tc1-like transposons. DNA hybridization indicated that TCp3.2 originated from the genome of *Cydia pomonella*, which is the natural host of CpGV, and is present in less than 10 copies in the *C. pomonella* genome. The transposon TCp3.2 most likely was inserted into the viral genome during infection of host larvae. TCp3.2 and the recently characterized Tc1-like transposon TC14.7 (Jehle et al. 1995), which was also found in a CpGV mutant, represent a new family of transposons found in baculovirus genomes. The occasional horizontal escape of different types of host transposons into baculovirus genomes evokes the question about the possible role of baculoviruses as an interspecies vector in the horizontal transmission of insect transposons.

Key words: *Cydia pomonella* granulovirus — Baculovirus — Insect transposons — Tc1

Introduction

Accumulating data on the phylogenetic distribution of transposable elements revealed interspecific discontinuity of many transposon families. From the presence of closely related transposons in taxonomically distantly related hosts it was hypothesized that transposons are not only vertically transmitted but that they also might be horizontally transferred across taxonomic borders. Some evidence for a horizontal transposon transfer has been provided for P-elements of *Drosophila spec.* (Daniels et al. 1990; Hagemann et al. 1996), mariner (Maruyama and Hartl 1991; Robertson 1993; Brunet et al. 1994), Tc1-like elements (Robertson 1995), jockey (Mizrokhi and Mazo 1990) or hobo elements (Calvi et al. 1991). A horizontal transposon transfer, however, necessitates at least an ecological relationship between the species concerned. Shared parasites and pathogens have been suggested as transposon vectors (Kidwell 1992; Houck et al. 1991), however, the molecular mechanisms involved in such a horizontal transfer have not yet been elucidated.

A possible candidate acting as a vector of insect transposons might be baculoviruses. Baculoviruses form a

* *Present address:* Staatliche Lehr- und Forschungsanstalt (SLFA), Breitenweg 71, 67435 Neustadt a. d. Weinstraße, Germany
Note: The nucleotide sequence reported here will appear in the EMBL, GenBank, and DDBJ Nucleotide Databases under Accession No. Z95154.
Correspondence to: J.A. Jehle; e-mail: slfa.jehle@t-online.de

large group of dsDNA viruses which primarily infect insect larvae (mainly Lepidoptera), but also occur on a few Crustacea. A unique characteristic of baculoviruses is that they spontaneously accommodate host transposable elements into their genome (for review, see Fraser 1986; Friesen 1993; Jehle and Vlak 1996). To date, the best characterized transposable element found in *Autographa californica* nucleopolyhedrovirus (AcMNPV) [Baculoviridae: Nucleopolyhedrovirus], is the class I retrotransposon TED, a 7.5-kb Ty3-like retroelement, which exhibits transcription and expression activities even after insertion into the viral genome (Miller and Miller 1982; Lerch and Friesen 1992). Beside this retrotransposon, a number of class II transposons, which originated from *Trichoplusia ni* or *Spodoptera frugiperda* cells, were detected in mutants of AcMNPV and *Galleria melonella* nucleopolyhedrovirus (GmMNPV). Most of these are characterized by short inverted repeat sequences of 13–15 bp, the duplication of a TTAA tetranucleotide target site and the lack of open reading frames, which could encode for a functional transposase gene (Carstens 1987; Cary et al. 1989; Wang et al. 1989; Beames and Summers 1990; Schetter et al. 1990). The only of these transposons, which was shown to be active is IFP2 (Fraser et al. 1995). Most of these transposons were detected when plaque morphology mutants of AcMNPV and GmMNPV, which derived from serial passages through cultured insect cells, were analysed. These plaque morphology mutants contain significantly less polyhedra per cell than do cells infected with wild-type virus (Fraser et al. 1983).

More recently, it was demonstrated that horizontal escape of a host transposon into a baculovirus genome is not restricted to cell culture systems but also occurs during infection of host larvae. Jehle et al. (1995) co-infected larvae of the false codling moth, *Cryptophlebia leucotreta*, with *C. leucotreta* granulovirus (CrleGV) and *Cydia pomonella* granulovirus (CpGV) [both Baculoviridae: genus Granulovirus]. Single CpGV and *CrleGV* genotypes were then isolated using an *in vivo* cloning procedure in the course of which larvae of *C. leucotreta* or *C. pomonella* were infected with very low doses of the progeny virus of the mixed infection. These experiments resulted in several CpGV mutants carrying insertions between one and several kb. One of these insertions, designated TC14.7, was characterized to be a 4726-bp-long transposable element containing 29-bp-long inverted terminal repeat sequences. It was integrated at a TA dinucleotide into a noncoding region of the CpGV genome (Jehle et al. 1995). TC14.7 contained a defective open reading frame with striking homologies to members of the Tc1-like transposons found in *Caenorhabditis* and *Drosophila* species, which belong to the recently established superfamily of Tc1/mariner-like transposons (Doak et al. 1994; Radice et al. 1994; Robertson 1995).

In this study we present the detailed characterization and sequence analysis of the insertion of a further CpGV mutant, which appeared to harbor a lepidopteran transposable element with inverted terminal repeats of more than 750 bp. This element also contains a transposase-like open reading frame of the Tc1/mariner-type. Our findings demonstrate the horizontal escape of Tc1/mariner-like transposons from the genome of the host into the baculovirus genome and supports the hypothesis that baculoviruses might be involved in horizontal transfer of insect transposons.

Materials and Methods

Virus and Larval Stocks

Viral DNA used in this study derived from different *in vivo* cloned genotypes of *Cryptophlebia leucotreta* granulovirus (*CrleGV*-CV3) (described in Jehle et al. 1992) and from *Cydia pomonella* granulovirus (CpGV-M) (Tanada, 1964). The insertion mutant CpGV-MCp4 as well as the other mutants (CpGV-MCp5, CpGV-MCp10, CrleGV-MC13) were isolated from an infection experiment of *C. leucotreta* larvae with CpGV- and CrleGV-CV3 and a subsequent *in vivo* cloning procedure in *C. pomonella* larvae, which was previously described in Jehle et al. (1995).

Cloning and Sequencing of the Insertion

A 5.7-kb *Bam*HI/*Eco*RI fragment which encompassed the complete insertion, termed TCp3.2, in the mutant CpGV-MCp4 was cloned into vector pGem-Zf(+) using standard methods (Sambrook et al. 1989). For nucleotide sequence determination, nested deletion clones were prepared (Henikoff 1987) and sequenced as double-stranded plasmid templates using the dideoxy-chain termination method (Sanger et al. 1977). Both strands of the plasmid DNAs were sequenced. The sequence data were compiled and analyzed using the software packages of UWGCG (Devereux et al. 1984) and PAUP v3.1.1 (Swofford 1991).

Southern Hybridization

For Southern hybridization, a cloned fragment, which derived from one of the exonuclease clones was prepared and labelled with DIG-dUTP (Boehringer Mannheim) using the random primer labelling protocol (Feinberg und Vogelstein 1983). This fragment, which did not contain any CpGV sequences, was used as a TCp3.2-specific probe and hybridized to viral and larval genomic DNA, which was blotted to Nylon membrane (Southern 1975). Viral and larval DNA were prepared according to Jehle et al. (1992; 1995). Detection of DNA hybridization was accomplished by using chemiluminescence methods.

PCR Analysis

The 5′ region of the transposase genes of host copies of TCp3.2 was amplified by employing the polymerase chain reaction. The amplification was performed according to standard conditions using 1 μg of genomic DNA of *C. pomonella* as template, the primers 5′-GAAGCATTTGCGATCTTAGTA, 5′-GTAGTGAAATGCGTG-CATGTT, and SuperTaq polymerase (HT Biotechnology). After amplification (30 cycles: 1′ at 90°C, 1′ at 50°C, 1′ at 72°C), the PCR products were phenolized, ethanol precipitated, and cloned into *Esch-*

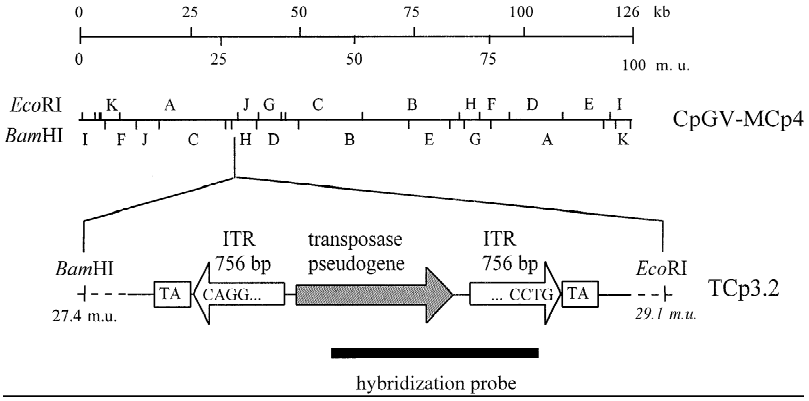

Fig. 1. Restriction map of mutant CpGV-MCp4 containing the TCp3.2 transposon. The inverted terminal repeats (ITR) and the transposase pseudogene are indicated by *arrows.* The sequence used as hybridization probe for analyzing the origin of TCp3.2 is indicated as a black bar (compare Fig. 3). kb = kilo base pairs, m.u. = map unit.

erichia coli using the pGemT vector system (Promega). Eight transformed colonies were randomly picked, and plasmid DNA was prepared and sequenced using standard techniques.

Results

Identification and Sequence Analysis of the Transposon TCp3.2

Recently, Jehle et al. (1995) reported on the *in vivo* isolation of *C. pomonella* granulovirus mutants which contained genomic insertions of nonvirus origin. We characterized one of these mutants (CpGV-MCp4) by restriction mapping and found that a fragment of 3.2 kb was inserted at about 28 map units of the CpGV restriction map (Fig. 1). For a detailed analysis of this insertion, a *Bam*HI/*Eco*RI fragment covering this region was cloned into *E. coli* and sequenced. In this region we identified an insertion of 3239 bp with characteristics of a transposable element: it carried inverted terminal repeats (ITRs) of 756 bp and contained a putative transposase gene (Figs. 1 and 2). This element was named TCp3.2.

The ITRs of TCp3.2 are 756 bp long and mismatched in only seven positions. The ITRs are flanked by a TA dinucleotide at both ends, as is typical for many members of the Tc1/mariner transposon superfamily (Doak et al. 1994). Their length is reminiscent to the transposons Tc3 (ITRS: 462 nt), Tc6 (765 nt), Minos (255 nt), Quatzal (236 nt), or Paris (242 nt), which all have rather atypical long ITRs among the Tc1/mariner-like transposons, whose ITRs generally count between 26 and 54 bp (Dreyfus and Emmons 1991; Collins et al. 1989; Franz and Savakis 1991; Radice et al. 1994).

The Origin of Transposon TCp3.2

Since the TCp3.2-carrying mutant CpGV-MCp4 was obtained after a serial passage of CpGV and *CrleGV,* first in *C. leucotreta* larvae and then in *C. pomonella* larvae (for detailed description of these experiments see Jehle et al. 1995), the question arose as to what the origin of this element is. This was determined by a comparative Southern hybridization of the transposon TCp3.2 to viral DNA and genomic insect DNA, respectively. For this, a 1900-bp fragment, which derived from an exonuclease III treated sequencing subclone and did not contain any CpGV-specific border sequences, was prepared, DIG-labelled, and used as a hybridization probe. As shown in Figure 3, this fragment only hybridized to the digests of CpGV-MCp4 (as a positive control) as well as to the *Hin*dIII digest of *Cydia pomonella* DNA. It did not hybridize to the DNAs of CrleGV and CpGV (the parental virus genotypes used in the infection experiment, Jehle et al. 1995), nor to DNA of *C. leucotreta* (the other host used in the infection experiments), nor to DNA of the other virus mutants, which were isolated during the infection experiments (Jehle et al. 1995). This result provided strong evidence that TCp3.2 originated from the genome of *C. pomonella.* When the TCp3.2-specific probe was hybridized to *C. pomonella* DNA digested with *Eco*RI and *Nde*I, which both do not cut inside of TCp3.2, only four to six hybridization signals were observed. Each of these signals could be caused by single-copy genes, suggesting that TCp3.2 is present in only a few copies, presumably less than 10 in the genome of *C. pomonella.*

The Transposase Gene of TCp3.2 Is Defective

Within TCp3.2, we identified three open reading frames (Fig. 2, ORF1-3). Data base comparison using the programs BLAST and FASTA revealed significant similarities of the predicted amino acid sequence of ORF2 and ORF3 to the transposase genes of the IS630-Tc1/mariner transposon superfamily (Fig. 4). The observed similarities included the recently characterized ''D,D35E'' motif consisting of three highly conserved glutamic acid (D) and aspartic acid (E) residues (Doak et al. 1994). However, the distance between the second glutamic acid residue and the aspartic acid residue of the D,D35E motif counts 43 amino acid residues and, hence, it slightly exceeds the ''35-amino-acid-rule.'' Interestingly, both

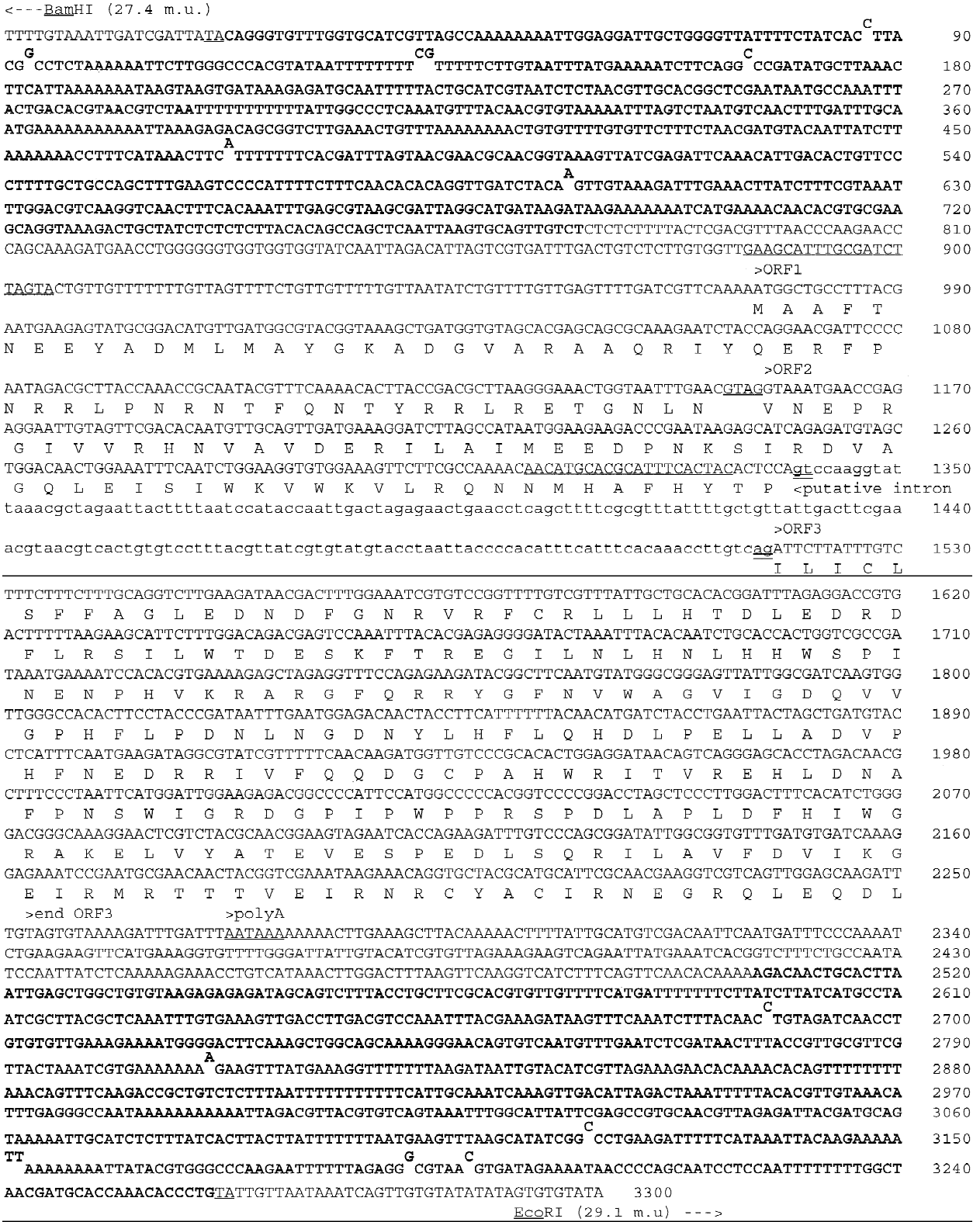

Fig. 2. Nucleotide and amino acid sequence of transposon TCp3.2. The suggested TA insertion sites are underlined. The inverted repeat sequences of TCp3.2 are boldface and the mismatching nucleotides are indicated by superior. Three open reading frames (ORF1–ORF3) are indicated and translated into the amino acid sequence. The putative 4-bp insertion at position 1153–1156 is underlined, and the invariant splicing borders of the suggested intron are double underlined. The location of the primers (position 885–995 and 1313–1333) used for the PCR of genomic DNA are underlined.

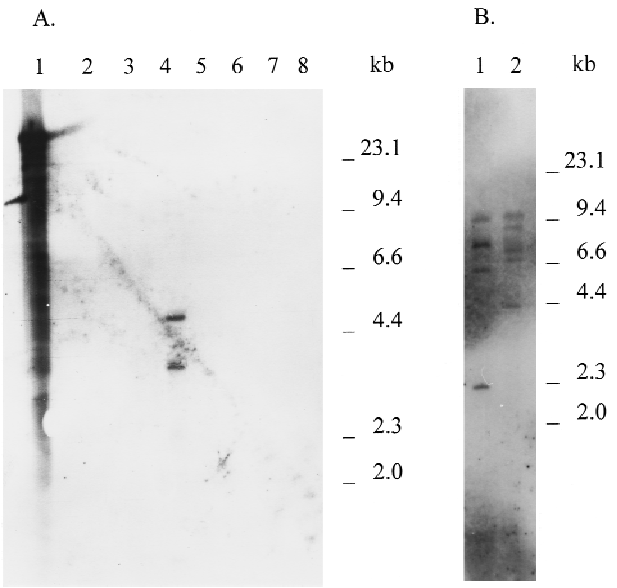

Fig. 3. Southern hybridization of the biotinylated TCp3.2-specific probe to **(A)** *Eco*RI digests of prepared DNA of MCp4 (lane 1, positive control), CpGV (lane 2), CrleGV (lane 3), MCp5 (lane 6), MC13 (lane 7), MCp10 (lane 8) (each 1 μg DNA), and *Hin*dIII digests of genomic DNA of *C. pomonella* (lane 4), and *C. leucotreta* (lane 5) (each 10 μg DNA). **(B)** Southern hybridization of the biotinylated TCp3.2-specific probe to *Eco*RI (lane 1) and *Nde*I (lane 2) digests of genomic DNA of *C. pomonella.*

the TCp3.2 transposase and the related Tc108 transposase (see below) did not contain a glutamic acid residue (E) but a further aspartic acid residue (D) in the D,D35E motif as it is typical for many mariner elements (Fig. 4).

The 3′ end of ORF1 (which covers the N-terminus of the predicted transposase gene) overlaps with the 5′ end of ORF2 in a stretch of 5 nt (Fig. 2). In this region, a frame shift mutation between ORF1 and ORF2 may have occurred suggesting that the putative transposase gene of TCp3.2 is defective. ORF2 and ORF3 are separated by an intron of 178 nucleotides with the invariant splicing borders (GT . . . AG). Apparently, the transposase of TCp3.2 is encoded by two exons, exon-1 consisting of ORF1 and ORF2, and exon-2 consisting of ORF3.

Location of the Frame-Shift Mutation in TCp3.2

The overlap between ORF1 and ORF2 and the very limited homology of the N-termini of the Tc1/mariner-like transposases did not allow a precise prediction of the site of the supposed frame shift mutation within exon-1. Therefore, we analyzed the sequences of further TCp3.2 transposon copies from the host *C. pomonella* by employing the polymerase chain reaction (PCR). DNA fragments of about 450 nt (corresponding to nucleotide position 885–1333 in Fig. 2) were amplified from genomic DNA of *C. pomonella* using two primers that covered the 5′ end of exon-1. The fragments were cloned in *E. coli* using the pGemT vector, and the inserts were sequenced. The sequence alignment and the resulting consensus of eight randomly picked and sequenced clones is shown in Figure 5. Only the sequence of clone TCp02 was identical to TCp3.2. The other clones showed heterogeneity at various nucleotide positions, which might be caused by genetic heterogeneity among the different transposon copies in the *C. pomonella* genome. More notably, all other clones missed 3–4 nt, which are located within the overlap of ORF1 and ORF2 (Figs. 2 and 5). The consensus of the sequenced PCR clones suggest a 4-bp insertion within exon-1 of the TCp3.2 transposase gene. A conceptual deletion of these 4 bp from the sequence of TCp3.2 results in a continuous exon1, whereby ORF1 and ORF2 are in frame.

Phylogenetic Inference of Transposon TCp3.2

The phylogenetic relationship of transposon TCp3.2 to other Tc1/mariner-like transposons was analyzed by maximum parsimony analysis of the aligned sequences of Fig. 4 using PAUP v3.1.1 (Swofford 1991). We refrained from including the N-termini of the transposases into our analyses, since the N-terminal part of Tc1/mariner transposases are uncertain to align due to their great divergence and, hence, they are of doubtful value for phylogenetic inference (Robertson 1995). A single most parsimonious tree was obtained by performing the heuristic search algorithm of PAUP, employing random addition of sequences and TBR branch swapping. The sequence of transposon ‘impala’ of *Fusarium oxysporum* was included for outgroup rooting and 100 bootstrap replications were done in order to analyze the support of the data on the tree topology.

The most parsimonious tree shows a clear distinction into two subtrees each containing the Tc1 and mariner transposon families, respectively (Fig. 6). TCp3.2 splits off as a branch within the Tc1-like transposons. However, the topology of this most parsimonious tree is only well supported by bootstrap analysis for the branching between the Tc1 and the mariner families (99% bootstrap value), as well as for the branching among the mariner families itself. On the other hand, bootstrap analysis using the 50% majority-rule (Swofford 1991) was not able to resolve the tree branching within the Tc1-subtree indicating a low robustness of the branching topology among these transposons. Only the sister grouping of TCp3.2 and Tc108, as well as the S element of *Drosophila melanogaster* and transposon Paris of *D. virilis* are reasonably supported by bootstrap values of 38% and 100%, respectively. Apparently, TCp3.2 is most closely related to a Tc1-like element of the free-living nematode *C. elegans*, which we termed Tc108 (indicating the GenBank ID*108*6846 of the amino acid sequence). This element was sequenced upon the nematode sequencing project and is located on *C. elegans* cosmid T14G12.

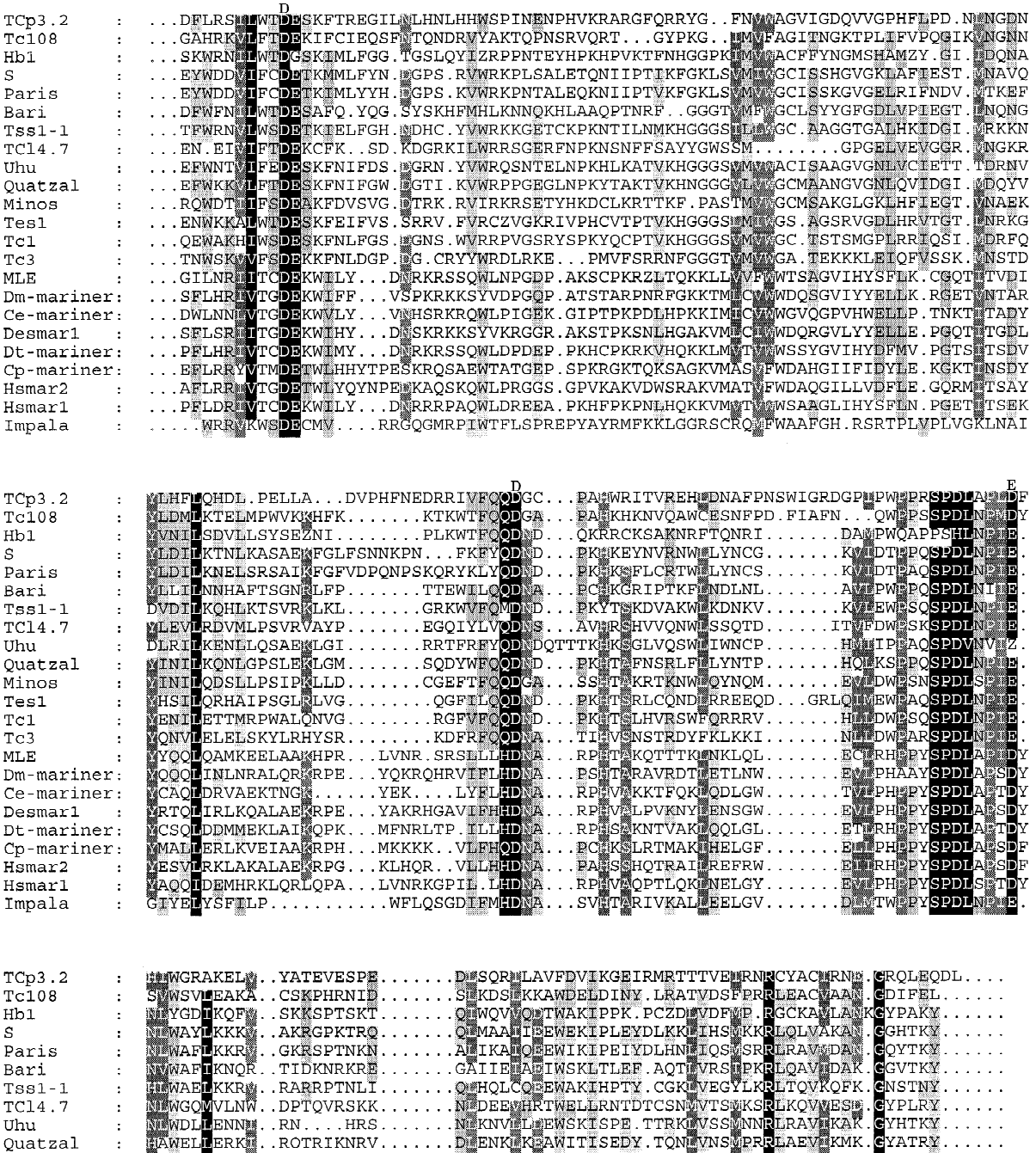

Fig. 4. Alignment of the C-terminal part of the predicted amino acid sequences of different Tc1- and mariner-like transposases: Tc108 (*Caenorhabditis elegans*, U41268), Hb1 (*Drosophila melanogaster*, X01748), S element (*D. melanogaster*, U33469), Paris, (*Drosophila virilis*, Z49253), Bari (*D. melanogaster*, X67681), Tss1-1 (*Salmo salar*, L12206), TC14.7 (CpGV, X79733), Uhu (*Drosophila heteroneura*, X63028), Quatzal (*Anopheles albimanus*, L76231), Minos (*Drosophila hydei*, Z29098), Tes1 (*Eptatretus stouti*, B46189), Tc1 (*C. elegans*, Z70783), Tc3 (*C. elegans*, U41105), MLE (*Hyalophora cecropia*, L10444) Dm-mariner (*Drosophila mauritiana*, M14653), Ce-mariner (*Caenorhabditis elegans*, Z50874), Desmar1 (*Mayetiola destructor*, U24436), Dt-mariner (*Dugesia tigrina*, X71979), Cp-mariner (*Crysoperla plorabunda*, U11650), Hsmar2 (*Homo sapiens*, U49974), Hsmar1 (*H. sapiens*, U52077), Impala (Fusarium oxysporum, S75106). The alignment starts at amino acid 155 (nucleotide position 1620 in Fig. 2) in respect to the TCp3.2 transposase. The ''D,D35E'' motif is indicated above. The alignment was generated using PILEUP and was manually adjusted. The shading was performed by GENEDOC software at shading levels (95%, 70%, and 50% amino acid similarity). Stop codons or frame shift mutations were denoted by ''Z.''

```
            1138                                  1167
TCp3.2    ...ACTGGTAATTTGAAC GTAG GTAAATGAACC...
TCp02u    ...ACTGGTAATTTGAAC GTAG GTAAATGAACC...
TCp01u    ...ACTGGTAATTTGAAC ---- GTAAATGAACC...
TCp03u    ...ACTGGTAATTTGAAC ---- GTAAATGAACC...
TCp04u    ...ACTGGTAATTTGAAC ---- GTAAATGAACC...
TCp05u    ...ACTGGTAATTTGAAC ---- GTAAATGAACC...
TCp06u    ...ACTGGTAATTTGAAC ---- GTAAATGAACC...
TCp07u    ...ACTGGTAATTTGAAC ---C GTAAATAAACC...
TCp08u    ...ACTGGTAATTTGGAA ---C GTAAATGAACC...
Consensus ...ACTGGTAATTTGAAC ---- GTAAATGAACC...
```

Fig. 5. Alignment of eight randomly cloned and sequenced PCR fragments (TCp01u–TCp08u) of genomic DNA of *C. pomonella* with the sequence of the viral TCp3.2 copy. The nucleotide positions in respect to the TCp3.2 sequence are given above, the consensus sequence is given below. Differing nucleotide positions are in boldface.

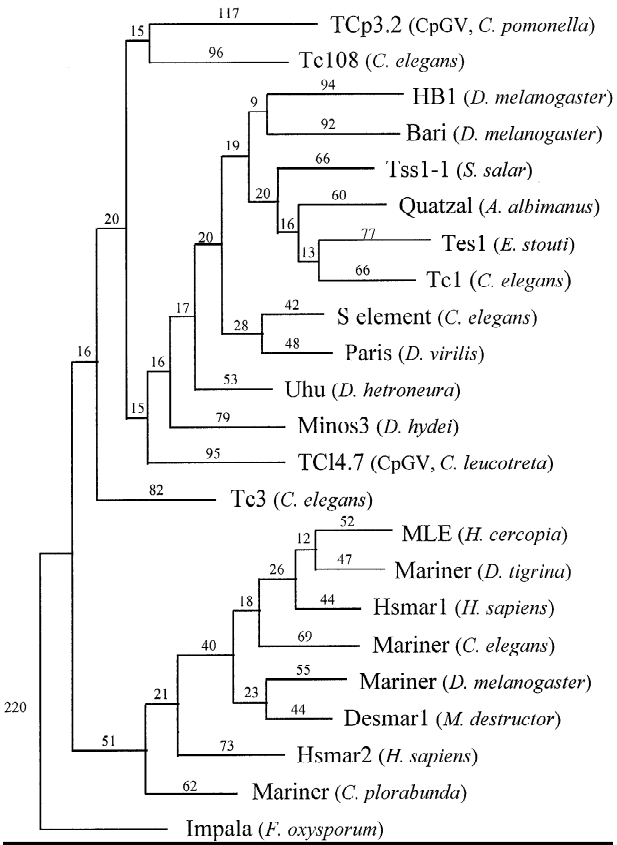

Fig. 6. Most parsimonious tree of the aligned transposase amino acid sequences obtained from the alignment of Figure 5 by heuristic search with random sequence addition (10 replicates) using PAUP v3.1.1 (Swofford 1991). The tree was constructed using ''Impala'' as an outgroup. Numbers above the lines are the numbers of changes between the node and the species.

Tc108 is clearly distinct from other *C. elegans* Tc1 or Tc3 elements because of its internal transposase sequence, inverted terminal repeats, and represents a new family of Tc1-like elements in *C. elegans.* Interestingly, the lepidopteran transposon TC14.7, which was also isolated from a CpGV mutant but originated from *C. leucotreta* appears to be less related to TCp3.2 than Tc108.

When the sequences of the ITRs of the various transposons are compared, it is striking that irrespective of the different lengths of the ITRs of Tc108 and TCp3.2, the most terminal 7 bp are identical and are more similar to each other than the terminal sequences of any of the other transposons (Fig. 7). The homology between the transposase genes of TCp3.2 and Tc108 is apparently also reflected in the conservation of the most terminal sequences of the ITRs.

Discussion

We identified a novel transposable element, termed TCp3.2, which is present in the genome of CpGV. TCp3.2 has ITR structures of 756 bp and contains a Tc1/mariner-like transposase pseudogene. The most terminal sequence of the ITRs consists of sequence 5′-TA CAG . . . -3′. This can be interpreted as a possible duplication of a transposition target site TA and the transposon borders CAG, as it is typical for many Tc1- and mariner-like transposons (Fig. 2, Fig. 7) (Radice et al. 1994; Robertson 1995). The putative transposase gene of TCp3.2 is defective because of a frame shift mutation near the 5′ end. This frame shift was confirmed by PCR analysis which showed that this mutated form is present in at least one of the copies in the genome of the CpGV host *C. pomonella.* Apparently, this mutated copy is already present in the host genome, suggesting that this frame shift mutation was not acquired upon the transposition into the viral genome. Hence, it is likely that the transposition of the defective TCp3.2 copy into the CpGV genome was mediated by an active transposase.

Functional studies of the transposase genes of Tc1 and Tc3 showed that the N-terminus of their transposases Tc1A and Tc3A, respectively, is involved in the specific recognition of ITRs of the transposons, whereas the more conserved C-terminal DDE motif of Tc1A is involved in endonuclease and phosphoryl transfer activity of this enzyme (Vos et al. 1993; Vos and Plasterk 1994; van Luenen et al. 1993; Colloms et al. 1994). The heterogeneity of the N-termini of these transposases might be explained by its specific binding function to the corresponding inverted repeats, which beside the highly conserved terminal 5-TA CAG . . . TG . . . CAA-3 motif (Fig. 7) do not share further homologies among the Tc1/mariner-like transposons. This is supported by our finding that transposon Tc108 of *C. elegans*, to which the transposase gene is most closely related to that of TCp3.2, also has the most similar terminal ITR sequence with seven identical nucleotides.

In Southern hybridization, TCp3.2 hybridized to preparations of genomic DNA of the host *C. pomonella* but not to DNA preparations of wild-type viruses of CpGV or CrleGV. This finding clearly indicates that TCp3.2 originated from the host genome, where it is present at a low copy number as deduced from the very few hybridization signals. This finding is remarkable

```
TCp3.2          ...ta-CAGGGTGTTTGGTGCATCGTTAGCCAAAAAAAATTGGAGGATTGCTG.
Tc108           ...ta-CAGGGTG--------------AGTCAAAATTATGGTAAGTCG......
Hb1             ...ta-CAGC-TG-------------TGTTCAGAAAAATAG-CAGTGC......
Bari            ...ta-CAGTCAT--------------GGTCAAAATTATTT-TCACAA......
Tss1-2          ...ta-CAGTTGA--------------AGTCGGAAGTTTAC-ATACAC......
Quatzal         ...ta-CACTTCT--------------CCACAAAAG--TGA-ATACACAGCAA.
Tes1            ...ta-CTCTACC--------------GGTCGAAAGTTTTA-GAACACCCCCA.
Tc1             ...ta-CAGTGCT--------------GGCCAAAAAGATAT-CCACTTTTGGT.
S element       ...ta-CAGTTTG----------------TCAAGAAACTGT-TTACACACCGC.
Paris           ...ta-CAGTTTG----------------TCAAGTAATTGT-TTGCAAAGTGA.
Uhu             ...ta-CAGTGTC--------------TCACAGCTCAACTG-GAACAGTGCCT.
Minos3          ...ta-CGAGCCC-------------CAACCACTATTAATTCGAACAGCAT...
TC14.7          ...ta-CAGTCAG--------------CTTCAGAGATATGT-GACCCCCC....
Tc3             ...ta-CAGTGCG---------------CCCAACTTCTATA-GCGCCCCCCTA.
Ce-mariner      ...taTCAGGTTG---------------TCCCATAAGTTTTTGTACTATTTT..
Desmar1         ...taTTGGG-TG---------------TACAACTTAAAAACCGGAATT.....
Dm-mariner      ...taCCAGG-TG---------------TACAAGTAGGGAATGTCGGTT.....
Cp-mariner      ...taTTAGGTTG----------------GCTGATAAGTCCCCGGTCT......
DL-mariner      ...taTTAGGTTG---------------TTCGATATGAAACGGGTCAAACTT..
Hsmar1          ...taTTAGGTTG--------------GTGCAAAAGTAATTGCGGTTTT.....
Hsmar2          ...taCAAGGGGT--------------CTTCAAAAAGTTCATGGAAAATG....
Impala          ...ta-CAGTGGG--------------GTACAATAAGTTTGAATACATCTGC..
Consensus       ...ta-CAG..TG.................CAA.A...T......CA.......
```

Fig. 7. Alignment of the most 5′ sequences of the inverted terminal repeats of different Tc1/mariner-like transposons. The consensus is indicated below. Note the similarity of the terminal nucleotides of TCp3.2 and Tc108.

since another CpGV mutant (MCp5) isolated in the course of the same experiment also contained a Tc1-like transposon, TC14.7 (Jehle et al. 1995). TC14.7 and TCp3.2 are completely different in size, ITR, and transposase sequence. Even more notable, TC14.7 originated from the lepidopteran host, *C. leucotreta*, in which CpGV was also passaged during the infection experiment.

Our phylogenetic analyses of the transposase gene using most parsimony methods suggested that TCp3.2 is a member of the Tc1-subtree of the Tc1/mariner transposon superfamily (Fig. 6). It appeared to be most closely related to a transposon of the nematode *C. elegans*, which we termed Tc108. This relationship is reasonably supported by bootstrapping as well as by sequence similarity of the ITRs (Figs. 6, 7). Tc108 differs from other Tc1 elements by size, transposase, and ITR sequence and appears to be a further family of Tc1-like elements in the *C. elegans* genome. Among the numerous Tc1 sequences only this unique Tc108 sequence was found, so far, in the nematode suggesting that within its genome the copy number of Tc108 might be much lower than that of Tc1.

Bootstrap analysis of the Tc1/mariner superfamily resulted in a clear distinction of two subtrees, Tc1-, and mariner-like transposons. It also well supported the location of TCp3.2 within the Tc1 subtree. Whereas the topology of the mariner subtree is robust, the Tc1 subtree is only poorly resolved. Only the bifurcation of TCp3.2 and Tc108, on the one hand, and S element and Paris were reasonably supported. Radice et al. (1994) and Robertson (1995) already noticed the difficulties of a reliable resolution of the phylogeny of Tc1-like elements by most parsimony analysis. Possible reasons for this might be that Tc1-like elements are much more diverse than mariner transposons. This apparent heterogeneity is not only due to the use of inactive and possibly highly mutated transposase sequences of Tc1-like transposons in the phylogenetic analyses. It is also expressed by the variability of transposon lengths (1.6–4.7 kb), the variability of ITR lengths (30–several 100 bp), as well as by the presence or absence of introns within the transposase genes. Compared to this heterogeneous puzzle mariner transposons appear to be structurally much more homogenous and less prone to variability.

A number of different insect transposons have been found in baculovirus genomes so far, but they all were generated in cell culture systems. The physiological circumstances and the frequency of transposition have never been determined. The novel Tc1/mariner-like transposon TCp3.2 may have escaped from the host into the CpGV genome during the infection of host larvae. Together with TC14.7, it belongs to a new family of host transposons found in baculovirus genomes. Its capture during a single round of virus replication suggests that horizontal transmission of insect transposons into baculoviruses is an occasional event which is not restricted to a specific type of transposons. Although the question whether baculoviruses are able to shuttle host transposons still remains to be experimentally proven, there are some specific characteristics of the biology of baculoviruses that conform with this potential role. First, baculoviruses have a large dsDNA genome of 80–230 kb, which is able to accommodate additional sequences without interfering with viral functions. Its frequent use as an expression vector for foreign proteins underscores this role (King and Possee 1992). For example, the CpGV mutant harboring the transposon TCp3.2 did not show any differences in the median lethal dose compared to wild type CpGV (Jehle et al. 1995). Second, baculovirus DNA replication takes place in the host nucleus (Granados and Williams 1986), where there is a good chance to pick up transposable elements from the host genome or deliver transposons into the host genome.

Third, baculoviruses have been shown that during an abortive infection of a nonsusceptible host, transcription and expression of at least a few viral genes can occur (McClintock et al. 1986; Carbonell et al. 1985; Carbonell and Miller 1987). It is also conceivable that during such an abortive infection of a host by a transposon-carrying virus, the transposon is mobilized and transferred into the new host genome.

Acknowledgments. The authors are indebted to Dr. E. Fritsch (BBA Darmstadt) for providing viral DNA and larvae of *C. pomonella* and *C. leucotreta.* J.A.J. is a recipient of a grant of the Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (No. 0319306A) and a Human Capital and Mobility grant from the European Union (ERBCHBICT941143).

References

Beames B, Summers MD (1990) Sequence comparison of cellular and viral copies of host cell DNA insertions found in *Autographa californica* nuclear polyhedrosis virus. Virol 174:354–363

Brunet F, Godin F, David JR, Capy P (1994) The mariner transposable element in the Drosophilidae family. Heredity 73:377–385

Calvi B, Hong TJ, Findley SD, Gelbart WM (1991) Evidence for a common evolutionary origin in inverted repeat transposons in *Drosophila* and plants: hobo, Activator, and Tam3. Cell 66:465–471

Carbonell LF, Klowden MJ, Miller LK (1985) Baculovirus-mediated expression of bacterial genes in dipteran and mammalian cells. J Virol 56:153–160

Carbonell LF, Miller LK (1987) Baculovirus interaction with nontarget organisms: a virus-borne reporter gene is not expressed in two mammalian cell lines. Appl Environ Microbiol 53:1412–1417

Carstens EB (1987) Identification and nucleotide sequence of the regions of *Autographa californica* nuclear polyhedrosis virus genome carrying insertion elements derived from *Spodoptera frugiperda.* Virol 161:8–17

Cary LC, Goebel M, Corsaro BG, Wang H-G, Rosen E, Fraser MJ (1989) Transposon mutagenesis of baculoviruses: analysis of *Trichoplusia ni* transposon IFP2 insertion within the FP-locus of nuclear polyhedrosis virus. Virol 172:156–169

Collins J, Forbes E, Anderson P (1989) The Tc3 family of transposable genetic elements in *Caenorhabditis elegans.* Genetics 120: 621–627

Colloms SD, van Luenen HGAM, Plasterk RHA (1994) DNA binding activities of the *Caenorhabditis elegans* Tc3 transposase. NAR 22: 5548–5554

Daniels SB, Peterson KR, Strausbaugh LD, Kidwell MG, Chovnick A (1990) Evidence for horizontal transmission of the P transposable element between *Drosophila* species. Genetics 124:339–355

Devereux J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis for the VAX. NAR 12:387–395

Doak TG, Doerder FP, Jahn CL, Herrick G (1994) A proposed superfamily of transposase genes: transposon-like elements in ciliated protozoa and a common ''D35E'' motif. Proc Natl Acad Sci USA 91:942–946

Dreyfus DH, Emmons SW (1991) A transposon-related palindromic repetitive sequence from *C. elegans.* NAR 19:1871–1877

Feinberg AP, Vogelstein B (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132:6–13

Franz G, Savakis C (1991) Minos, a new transposable element from *Drosophila hydei,* is a member of the Tc1-like family of transposons. NAR 19:6646

Fraser MJ (1986) Transposon-mediated mutagenesis of baculoviruses: transposon shuttling and implications for speciation. Ann Entomol Soc Amer 79:773–783

Fraser MJ, Cary LC, Boonvisudhi K, Wang H-G (1995) Assay for movement of Lepidopteran transposon IFP2 in insect cells using a baculovirus genome as target DNA. Virol 211:397–407

Fraser MJ, Smith GE, Summers MD (1983) Aquisition of host cell DNA sequences by baculoviruses: relationship between host DNA insertions and FP mutants of *Autographa californica* and *Galleria mellonella* nuclear polyhedrosis viruses. J Virol 47:287–300

Friesen PD (1993) Invertebrate transposable elements in the baculovirus genome: characterization and significance. In: Beckage NE, Thompson SN, Federici BA (eds) Parasites and pathogens of insects, Vol. 2: Parasites. Academic Press, London, pp 147–178

Granados RR, Williams KA (1986) *In vivo* infection and replication of baculoviruses. In: Granados RR, Federici BA (eds) The biology of baculoviruses, Vol. I: biological properties and molecular biology. CRC Press, Boca Raton, pp 89–108

Hagemann S, Haring E, Pinsker W (1996) Repeated horizontal transfer of P transposons between *Scaptomyza pallida* and *Drosophila bifasciata.* Genetica 98:43–51

Henikoff S (1987) Unidirectional digestion with exonuclease III in DNA sequence analysis. Methods Enzymol 155:156–165

Houck MA, Clark JB, Peterson KR, Kidwell MG (1991) Possible horizontal transfer of *Drosophila* genes by the mite *Proctolaelaps regalis.* Science 253:1125–1129

Jehle JA, Backhaus H, Fritsch E, Huber J (1992) Physical map of the *Cryptophlebia leucotreta* granulosis virus genome and its relationship to the genome of *Cydia pomonella* granulosis virus. J Gen Virol 73:1621–1626

Jehle JA, Fritsch E, Nickel A, Huber J, Backhaus H (1995) TC14.7: a novel lepidopteran transposon found in *Cydia pomonella* granulosis virus. Virol 207:369–379

Jehle JA, Vlak JM (1996) Transposable elements in baculoviruses: their possible role in horizontal gene transfer. In: Schmidt ER, Hankeln T (eds) Transgenic organisms and biosafety. Springer-Verlag, Berlin, pp 31–41

Kidwell MG (1992) Horizontal transfer. Curr Opinion Genet Dev 2: 868–873

King LA, Possee RD (1992) The baculovirus expression system: a laboratory guide. Chapman and Hall, London

Lerch RA, Friesen PD (1992) The baculovirus-integrated retrotransposon TED encodes gag and pol proteins that assemble into viruslike particles with reverse transcriptase. J Virol 66:1590–1601

Maruyama K, Hartl DL (1991) Evidence for interspecific transfer of the transposable element mariner between *Drosophila* and *Zaprionus.* J Mol Evol 33:514–524

McClintock JT, Dougherty EM, Weiner RM (1986) Semipermissive replication of a nuclear polyhedrosis virus of *Autographa californica* in a gypsy moth cell line. J Virol 57:197–204

Miller DW, Miller LK (1982) A virus mutant with an insertion of a copia-like transposable element. Nature 299:562–564

Mizrokhi LJ, Mazo AM (1990) Evidence for horizontal transmission of the mobile element jockey between distant *Drosophila* species. Proc Natl Acad Sci USA 87:9216–9220

Radice AD, Bugaj B, Fitch DHA, Emmons SW (1994) Widespread occurrence of the *Tc1* transposon family: *Tc1*-like transposons from teleost fish. Mol Gen Genet 244:606–612

Robertson HM (1993) The mariner transposable element is widespread in insects. Nature 362:241–245

Robertson HM (1995) The Tc1-mariner superfamily of transposons in animals. J Insect Physiol 41:99–105

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

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

Schetter C, Oellig C, Doerfler W (1990) An insertion of insect cell DNA in the 81-map-unit segment of *Autographa californica* nuclear polyhedrosis virus. J Virol 64:1844–1850

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98:503–517

Swofford DL (1991) PAUP: Phylogenetic analysis using parsimony, Version 3.0. Computer program distributed by the Illinois Natural History Survey, Champaign, Illinois

Tanada Y (1964) A granulosis virus of the codling moth, *Carpocapsa pomonella* L. J Insect Pathol 6:378–380

van Luenen HGAM, Colloms SD, Plasterk RHA (1993) Mobilization of quiet, endogenous Tc3 transposons in *Caenorhabditis elegans* by forced expression of Tc3 transposase. EMBO J 12:2513–2520

Vos JC, Plasterk RHA (1994) Tc1 transposase of *Caenorhabditis elegans* is an endonuclease with a bipartite DNA binding domain. EMBO 7:1907–1911

Vos JC, van Luenen HGAM, Plasterk RHA (1993) Characterization of the *Caenorhabditis elegans* Tc1 transposase *in vivo* and *in vitro*. Genes Development 7:1244–1253

Wang HH, Fraser MJ, Csiszar Cary L (1989) Transposon mutagenesis of baculoviruses: analysis of TFP3 lepidopteran transposon insertions at the FP locus of nuclear polyhedrosis. Gene 81:97–108