Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Comparative analysis of selected genes from *Diachasmimorpha longicaudata* entomopoxvirus and other poxviruses

Y. Hashimoto, P.O. Lawrence*

Department of Entomology and Nematology, University of Florida, Gainesville, FL 32611-0620, USA

Received 14 October 2004; accepted 22 October 2004

**Abstract**

The *Diachasmimorpha longicaudata* entomopoxvirus (DlEPV) is the first symbiotic EPV described from a parasitic wasp. The DlEPV is introduced into the tephritid fruit fly larval host along with the wasp egg at oviposition. We sequenced a shotgun genomic library of the DlEPV DNA and analyzed and compared the predicted protein sequences of eight ORFs with those of selected poxviruses and other organisms. BlastP searches showed that five of these are homologous to poxvirus putative proteins such as metalloprotease, a putative membrane protein, late transcription factor-3, virion surface protein, and poly (A) polymerase (PAP) regulatory small subunit. Three of these are similar to those of other organisms such as the gamma-glutamyltransferase (GGT) of *Arabidopsis thaliana*, eukaryotic initiation factor 4A (*Caenorhabditis briggsae*) and lambda phage integrase (λ-Int) of *Enterococcus faecium*. Transcription motifs for early (TGA,A/T,XXXXA) or late (TAAATG, TAAT, or TAAAT) gene expression conserved in poxviruses were identified with those ORFs. Phylogenetic analysis of multiple alignments of five ORFs and 20 poxvirus homologous sequences and of a concatenate of multiple alignments suggested that DlEPV probably diverged from the ancestral node between the fowlpoxvirus and the genus B, lepidopteran and orthopteran EPVs, to which *Amsacta moorei* and *Melanoplus sanguinipes* EPV, respectively, belong. The DlEPV putative GGT, eIF4A, and λ-Int contained many conserved domains that typified these proteins. These homologues may be involved in either viral pathogenicity or enhancing parasitism via the gamma-glutamyl cycle and compensation of eIF4A levels in the parasitized fly, or via the integration of a portion of the viral genome into the wasp and/or parasitized fly.

© 2004 Elsevier Ltd. All rights reserved.

**Keywords:** Anastrepha suspensa; Parasitism; Phylogenetic analysis; Symbiotic EPV; DlEPV; Hymenoptera; Parasitic wasp

1. Introduction

The *Poxviridae* has an extremely wide host range in both vertebrates and invertebrates and possesses a relatively large dsDNA genome (130–300 kb) (Moss, 2001). The family consists of two subfamilies, *Chordopoxvirinae* (ChPV) or vertebrate poxviruses and *Entomopoxvirinae* (EPV) or insect poxviruses. Eight genera of ChPV and three genera (EPV A, B, and C, hereafter group A, B, and C) of EPV, have been established based on virus structure and host range (Esposito et al., 2000).

Group A [type species, *Melolontha melolontha* EPV (MmEPV)], has an ovoid virion and infects coleopteran insects. Group B [type species, *Amsacta moorei* EPV (AmEPV)], also has an ovoid virion but infects lepidopteran and orthopteran insects. Group C [type species, *Chironomus luridus* EPV (CIEPV)], has a brick-shaped virion and infects dipteran insects (Esposito et al., 2000). Both the MsEPV and AmEPV genomes have been completely sequenced (Afonso et al., 1999; Bawden et al., 2000) and the analysis of their genome organization and gene composition demonstrated a lack of significant co-linearity between them, suggesting an inadequate classification of group B entomopoxviruses (Afonso et al., 1999; Bawden et al., 2000). The sequence and phylogenetic relationships among single genes of...
several EPVs in groups A and B have been reported. These include *spheroïdin*, a major component of the large crystalline occlusion body that is characteristic of EPVs and absent from ChPVs (Hall and Moyer, 1991; Sanz et al., 1994; Li et al., 1997; Sriskantha et al., 1997; Mitsuhashi et al., 1998; Hernandez–Crespo et al., 2000); *fusolin*, a spindle protein found primarily in lepidopteran EPVs (Dall et al., 1993; Hayakawa et al., 1996; Mitsuhashi et al., 1997); DNA polymerase (Mustafa and Yuen, 1991; Gauthier et al., 1995); poly(A) polymerase (PAP) regulatory small subunit, 30k virion protein and a 17k ORF (Sriskantha et al., 1997; Crnov and Dall, 1999); thymidine kinase (Lytvyn et al., 1992); nucleoside triphosphate phosphohydrolase I (NPH I) (Yuen et al., 1991; Sriskantha et al., 1997; Li et al., 1998); and the rifampicin resistance protein (Osborne et al., 1996). No genes have been sequenced from group C EPVs except for those reported previously by our laboratory for *Diachasmimorpha longicaudata* EPV (DlEPV). DlEPV homologues of poxvirus genes that encode putative proteins include DNA helicase, a *Drosophila* related oligoribonuclease, transcription factor, and other DlEPV-specific sequences (Mwaengo and Lawrence, 2003), DNA dependent RNA polymerase (Lawrence, 2002), and rifampicin-resistance protein (AF159588).

DlEPV was originally found in the poison gland apparatus of the female braconid wasp, *Diachasmimorpha longicaudata*, which parasitizes larvae of the Caribbean fruit fly, *Anastrepha suspensa* and other tephritid fruit flies (Lawrence and Akin, 1990). DlEPV is injected into larval hosts along with wasp eggs. The virus then invades and replicates in the hemocyte cytoplasm from which progeny bud into the hemolymph to infect other hemocytes (Lawrence, 2002). Infected hemocytes lose their adhesive property, and undergo apoptosis, and are unable to encapsulate the wasp eggs (Lawrence, 2005). These pathologies do not occur within the wasp where the virus is confined to the cuticle-lined poison gland (Lawrence and Akin, 1990; Lawrence, 2002). Analogous to other endoparasitoid viruses, DlEPV is considered to be a symbiotic agent of the parasitic wasp, because it disrupts normal hemocyte function in the host that results in host immunosuppression and parasite survival (Lawrence, 2005). DlEPV is unique in EPVs in that it does not produce an occlusion body in infected host hemocytes (Lawrence, 2002). Therefore we consider DlEPV to belong to group C of the subfamily *Entomopoxvirinae* or to a new group that reflects its unique biological property as a symbiotic EPV. Studies are in progress to determine how DlEPV progeny is acquired by neonate wasps or is maintained throughout wasp development.

Many endoparasitic wasps studied to date, particularly those attacking lepidopterans, possess virus(es) that cause immunosuppression in their hosts (Lavine and Beckage, 1996). The best-characterized viral symbionts of parasitic wasps are the polydnaviruses (PDVs) that possess segmented DNA genomes and are transmitted vertically to wasp offspring through proviral integration in the wasp genome (Fleming and Summers, 1991; see also a review by Kroemer and Webb, 2004). Unlike DIEPV, PDVs do not replicate in their hosts (Webb, 1998; Whittfield and Asgari, 2003). However, some PDV messages are expressed in various host tissues (Le et al., 2003). Several putative genes have been characterized and proposed as part of the PDV immunosuppression mechanism. These include *cys-motif, rep, vinexin, vankyrin* gene in the *Ichnovirinae* subfamily and *egf, ptp* in the *Brachovirinae* subfamily (Kroemer and Webb, 2004).

An increasing number of symbiotic non-poly DNA viruses of parasitic wasps have been described to date (see other papers in this issue). These include the DNA viruses *Diadromus pulchellus* ascovirus-4 (DpAV4) (Bigot et al., 1997) and a putative baculovirus from wasps of coleopterans (Krell and Stoltz, 1979; Stoltz, 1981). RNA viruses including reo viruses (Bigot et al., 1995; Stoltz and Makkay, 2000), a cypovirus of *D. pulchellus* (RV-2) (Renault et al., 2003), and a picorna-like virus have also been discovered from wasps of lepidopterans (see papers in this issue). DIEPV and DIRhV, a rhabdovirus from *D. longicaudata* (Lawrence and Akin, 1990), coronavirus-like and reovirus-like particles from *Psytalia* (*Opium*) *concolor* (Jacas et al., 1997), and virus-like particles (VLPs) from several *Leptopilina* sp. parasites of *Drosophila* sp. (Rizki and Rizki, 1990; Dupas et al., 1996; Morales et al., 2005) have been described in dipterans. Unlike PDVs, these latter viruses apparently replicate in their dipteran hosts but very few of their genes have been sequenced and their methods of transgenerational transmission are unknown.

The goal of this report is to analyze shotgun sequence data of DIEPV DNA in order to identify DIEPV homologues of poxvirus genes and to determine their phylogenetic relationship to those of EPVs and other organisms. We found five poxvirus conserved genes including *metalloprotease* (homologue of vaccinia virus [VV] G1L), a putative membrane protein (VV F9L homologue), *late transcription factor-3* (VV A2L homologue), *viroin surface protein* (VV A28L homologue), *poly (A)* polymerase (PAP) regulatory small subunit (VV J3R homologue), and three putative genes not found in poxviruses; *gamma-glutamyltransferase* (GGT), *putative eukaryotic initiation factor 4A* (eIF4A) and *lambda phage integrase* (%-Int).

GGT has been found in many organisms other than viruses and functions in amino acid transport and the metabolism of glutathione, an antioxidant and detoxifying agent (Ikeda et al., 1995a). eIF4A is a component of eIF4F, which mediates the binding of mRNA to ribosomes (Svitkin et al., 2001). No eIF4A homologue
has been found in viruses. λ-Int induces the insertion of bacteriophage DNA into host bacterial chromosomal DNA through site-specific recombination (Campbell, 1992; Nunes-Duby et al., 1998).

To examine the gene structure and relative phylogenetic position of DlEPV within the Poxviridae, we performed multiple alignments of the protein sequences and constructed phylogenetic trees of each of five DlEPV ORFs with their homologues from 20 selected poxviruses whose genomes have been completely sequenced. Conserved domain (CD) analysis was performed with three DlEPV ORFs not found in poxviruses, to evaluate a possible role of their putative gene products within the context of the DlEPV-wasp-dipteran host interaction. The data indicate that DlEPV is undoubtedly a member of the Poxviridae but this virus also contains non-poxvirus genes, suggesting its possible acquisition of genes from its dipteran host and/or host-related microflora, that probably reflect an earlier evolutionary relationship with dipterans. The phylogenetic relationships identified between DlEPV and other EPVs suggest that DlEPV is different from the group B EPVs and should be classified as a group C EPV since it is pathogenic to its dipteran host and not to its hymenopteran vector.

2. Materials and methods

2.1. Insects and source of female accessory gland

A laboratory colony of D. longicaudata was maintained as previously described (Lawrence et al., 1976; Lawrence, 1988). The accessory glands (gland reservoir and filaments) from 5 to 10-day-old female wasps were collected into TE buffer (10mM Tris–HCl, 1mM EDTA, pH8.0), then frozen at −80°C until used as previously described (Lawrence, 2002).

2.2. Shotgun sequencing of DlEPV DNA

DlEPV genomic DNA (250–300 kb) was isolated from wasp glands by pulse field gel electrophoresis as previously described (Lawrence, 2002) and gel plugs containing DNA were processed for construction of a shotgun genomic library at the Interdisciplinary Center for Biotechnology Research (ICBR) at the University of Florida (UF) according to their established protocol. Briefly, the gel plugs were melted and about 2.5–5 µg of the released DNA was sheared (sonicated) and the resulting fragments were size fractionated by agarose gel electrophoresis and gel filtration to yield fragments of 1–3 kb. Fragments were blunt ended, then cloned into pUC19 and the ligated plasmid DNA was used to transform DH10B competent cells. The transformants were screened by LB agar plate containing 25 µg/ml ampicillin, oxacillin (25 µg/ml), and 800 mg/plate each IPTG and X-gal (2%). White colonies were isolated and grown overnight in Terrific broth medium containing ampicillin. Two hundred to 500 µg of plasmid DNA were prepared from the culture by the BioRobot 9600 (QIAgen) and used as template for labeling with ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction kit [Perkin-Elmer Applied Biosystems (ABI), Foster City, CA]. Cycle sequencing was carried out using a PTC-200 DNA Engine (MJ Research, Watertown, MA) (25 cycles; 1°C/s to 96°C; 96°C for 10 s; 1°C/s to 60°C; 60°C for 4 min). Dye terminator was removed by Sephadex G-50 superfine column (Sigma, St. Louis, MO) in water. The reactions were electrohoresed on an ABI 377 sequencer and the chromatograms were edited with Analysis version 1.2.1 (ABI) and assembled into consensus contiguous DNA sequences (contigs) using a package of Phred/Phrap/Consed software (Ewing and Green, 1998; Ewing et al., 1998; Gordon et al., 1998).

2.3. Analysis of DlEPV DNA sequences

Contigs larger than 1 kb were analyzed for ORFs using a web-based software (ExPASy Molecular Biology Server, http://us.expasy.org/). The DNA sequences were subjected to BlastX analysis and homologues of the resulting translated protein sequences were identified in the NCBI (http://www.ncbi.nlm.nih.gov/) protein and poxvirus genome database (http://www.poxvirus.org/) with a threshold E-value < 0.01.

Protein sequences were analyzed by PROSITE (Hofmann et al., 1999), CD-search (RPSBLAST; Marchler-Bauer et al., 2003), PSSM (Kelly et al., 1999), and Swiss Model (Guex and Peitsch, 1997). Poxvirus promoters around the ORFs were also identified.

The alignments of multiple protein sequences were performed by ClustalX (Thompson et al., 1997) using the default program parameters. Concatenated sequences of several genes have been used to more accurately describe the phylogenetic relationship among proteins from Chordopoxviruses (Gubser et al., 2004). Therefore, to generate concatenated multiple alignments, we retrieved individual sequences with their respective gaps from a FASTA file of the pre-aligned proteins, then sequentially combined them to create a new FASTA file. Phylogenetic analysis of the alignment was performed by the Bootstrap neighbour joining (NJ) Tree in the ClustalX package using random number generator seeds at 111 and 1000 bootstrap trials. Drawing and manipulation of phylogenetic trees were done with TreeView 1.5 (Page, 1996) and Adobe Illustrator 8.0.1. (Adobe Systems Inc., San Jose, CA).

The GenBank accession numbers of the poxvirus DNA genome sequences used for the multi-alignments
were: AmEPV AF250284; Camelopox virus isolate M-96 (CMPV-M96), AF438165; Cowpox virus strain Brighton Red (CPXV-BR), AF482758;ECTomelia virus Moscow strain (ECTV-MOS), AF012825; Fowlpox virus (FPV-FCV), AF198100; Lumpy skin disease virus strain Meethling isolate 2490 (LSD-NEE), AF325528; Molluscum contagiosum virus subtype 1 (MCU-SB1), U60315; Monkeypox virus strain Zaïre-96-I-16 (MPXV-ZRE), AF380138; MsEPV, AF063866; Myxoma virus strain Lausanne (MYX-LAU), AF170726; Rabbit fibroma virus (RFB-KAS), AF170722; Sheeppox virus TU-V02127 (SPPV-TU), AY077832; Swinepox virus isolate 17077-99 (SWPV-NEB), AF410153; Vaccinia virus strain Ankara (VAC-ANK), U94848; Vaccinia virus strain Copenhagen (VAC-COP), M35027; Vaccinia virus strain Tian Tan (VAC-TAN), AF095689; Variola major virus strain Bangladesh-1975 (VAR-BSH), L22579; Variola major virus strain India-1967 (VAR-IND), X69198; Variola minor virus strain Garcia-1966 (VMN-GAR), Y16780; and Yaba-like disease virus (YAB-YLD), AJ293568.

2.4. DNA sequences

GenBank accession numbers of the DlEPV DNA sequences in this study are: YAB-YLD, AJ293568.

3. Results

3.1. Relationship of five DlEPV genes to poxviruses

BlastX searches identified homologues to eight DlEPV putative proteins with E-values ≤ 2.00E−3 (Table 1). The homologues encoding metalloprotease, a putative membrane protein, late transcription factor-3, virion surface protein, and PAP regulatory small subunit were present in all poxvirus sequences. These homologues had ORFs of similar size to those of DlEPV and E-values of 2.00E−03 to 1.00E−127 that demonstrated high homology with the DlEPV ORFs. However, homologues of putative GGT, eIF4A, and λ-Int were absent from the poxvirus protein database. The evidence that the GGT and eIF4A ORFs belong to DlEPV is supported by their presence on the same sequence contig and as those ORFs that had a hit to poxvirus proteins in the Blast search (Table 1). The contig (ID#137) containing λ-Int ORF had no poxvirus blast hit (Table 1). However, at least three other ORFs within the same contig showed homology to Plasmodium falciparum putative proteins (data not shown). This eliminates the possibility of bacterial contamination.

Either TAAATG, TAAT, or TAAAT, poxvirus late transcription motifs (Rosel et al., 1986), was present in five DlEPV ORFs (Table 1). These DlEPV putative late genes contained a region of multiple consecutive Ts separated by eight to 24 nt from the first A of the late transcription motif (data not shown), suggesting the presence of strong late promoters similar to those found in vaccinia (Davison and Moss, 1989). The early poxvirus promoter, TGAANNNNA, that occurs upstream of several EPV early genes (Afonso et al., 1999; Bawden et al., 2000; Gruidl et al., 1992; Lytvyn et al., 1992; Mustafa and Yuen, 1991; Li et al., 1997) was also present upstream of the ORFs for the DlEPV eIF4 and PAP regulatory small subunit homologues (data not shown). No poxvirus transcription motifs were found within 100 nt upstream of the translation initiation codon of the ORF of the putative GGT.

Five DlEPV ORFs contained CDs homologous to those of EPVs and ChPVs (Table 1) (Marchler-Bauer et al., 2003). The DlEPV ORF encoding the PAP regulatory small subunit possesses the PARP regulatory domain (NCBI, CCD entry # pfam01358; 294 aa) (E-value = 1e−47, 93.9% alignment to the consensus sequence of the domain, hereafter referenced as “aligned”) (Fig. 1a). Interestingly, the DlEPV putative metalloprotease possesses the peptidase M44 domain (NCBI, CCD entry # pfam03410; 590 aa) (E-value = 9e−5, 28.6% aligned) (Fig. 1b), which is lacking in the AmEPV and MsEPV metalloproteases (AmEPV ORF256, NP065038 and MsEPV ORF056 AAC97620, respectively). The ORF encoding a putative membrane protein possesses the L1L F9 C19 domain (NCBI, CCD entry # pfam03410; 590 aa) (E-value = 9e−5, 28.6% aligned) (Fig. 1c). The ORF encoding late transcription factor-3 homologue possesses the poxvirus late transcription factor domain (NCBI, CCD entry # pfam09497; 218 aa) (E-value = 1e−8, 91.7% aligned) (Fig. 1d) and the DlEPV virion surface protein homologue possesses the poxvirus A28 family domain (E-value = 4e−22, 97.1% aligned) (Fig. 1e). These results reveal that the five DlEPV ORFs contain the conserved motifs common in most homologous genes of poxviruses.

3.2. Phylogenetic relationship of five DlEPV genes to those of poxviruses

Phylogenetic trees were reconstructed from multiple alignments of five DlEPV putative proteins and 20–21 poxvirus homologues consisting of 2–3 EPVs and 18 ChPVs (Figs. 2a–c). The unrooted NJ trees of the 21 aligned poxvirus sequences for virion surface protein, metalloprotease, putative membrane protein and late transcription factor-3 were similar (data not shown). Therefore the tree for virion surface protein (Fig. 2a) is used as a representative to demonstrate this relationship. The phylogenetic trees of the virion surface protein homologue (Fig. 2a), PAP regulatory small subunit (Fig. 2b), and of the concatenated multiple alignments
Table 1
Selected complete ORFs and/or contigs from shotgun sequences of DlEPV DNA that are homologous with poxvirus proteins

| Contig ID (nts) | Selected ORF size (aa) | GenBank accession number | Poxvirus transcription motif | E-value | Highest BlastP hit (Gene ID, size in aa) | Highest Blast hit by partial or complete ORF(s) present in the DlEPV sequence contig (excluding the selected ORFs) |
|----------------|------------------------|--------------------------|----------------------------|---------|-----------------------------------------|----------------------------------------------------------------------------------------------------------------------------------|
| 151 (8558)     | 620                    | AY598433                 | E, L                       | 7.00E−04 | Canarypox virus, metalloprotease homolog (NP955131, CNPV108, 532) | AmEPV ORF230 (1.00E−06), VARPIND K4L (1.00E−68) |
| 150 (8175)     | 226                    | AY598434                 | L                          | 1.00E−17 | Melanoplus sanguinipes entomopoxvirus, putative membrane protein (NP048165, MSV094, 241) | AmEPV ORF181 (1.00E−17), MsEPV ORF089 (1.00E−68) |
| 150 (8175)     | 639                    | AY598437                 | —                          | 4.00E−17 | Arabidopsis thaliana, putative gamma-glutamyltransferase (NP195675, 574) | AmEPV ORF181 (1.00E−17), MsEPV ORF089 (1.00E−68) |
| 149 (7508)     | 379                    | AY598438                 | E, L                       | 1.00E−127 | Caenorhabditis briggsae, eukaryotic initiation factor 4a (CAE70046, 402) | AmEPV ORF091 (4.00E−11), AmEPV ORF166 (5.00E−07) |
| 145 (6044)     | 213                    | AY598435                 | E, L                       | 2.00E−05 | Melanoplus sanguinipes entomopoxvirus, putative late transcription factor 3 homolog (NP048136, MSV065, 218) | MsEPV ORF069 (1.00E−19), AmEPV ORF047 (5.00E−07) |
| 144 (5437)     | 155                    | AY598431                 | E, L                       | 2.00E−16 | Molluscum contagiosum virus, virion surface protein homolog (NP044085, MC134L, 141) | MsEPV ORF233 (1.00E−04), MsEPV ORF194 (2.00E−07) |
| 137 (4726)     | 302                    | AY598436                 | E, L                       | 2.00E−03 | Enterococcus faecium, site-specific recombinase XerD (ZP00036128, 490) | None |
| 130 (3780)     | 293                    | AY598432                 | E, L                       | 1.00E−57 | Ansacta moorei entomopoxivirus, poly(A) polymerase (PAP) regulatory small subunit (NP048442, AMV060, 295) | AmEPV ORF120 (9.00E−04), MsEPV ORF036 (3.00E−08) |

aa = length in amino acids; E, L = ORF containing early and late transcription motifs, respectively; nts = length in nucleotides; dashed line (−) = no transcription start site found.
of all five DIEPV poxvirus-related sequences (Fig. 2c), all corroborate established reports that fowlpox virus occurs in a clade that is the most distant of the non-orthopoxviruses (Fig. 2a). With the exception of the PAP sequences (Fig. 2b), the poxvirus homologues, suggest that DIEPV arose between fowlpox virus and the divergence of the group B EPVs (Figs. 2a and c). In contrast, the phylogenetic tree for the PAP regulatory small subunit showed a slight difference in the relationship of DIEPV to the group B EPVs (Fig. 2b). AmEPV possesses two PAP regulatory small subunit homologues, AmEPV ORF115 and ORF060 (NP 064897 and NP 064842, respectively). The former is closely related to MsEPV PAP regulatory small subunit (MsEPV ORF041, NP 048112) and the latter is closely related to Heliothis armigera EPV (HaEPV) PAP regulatory small subunit (Crnov and Dall, 1999). However, the low bootstrap value (49) for the MsEPV ORF041 and AmEPV ORF115 compared to the value (100) for the AmEPV ORF060 and HaEPV (Fig. 2b) along with the value (67) for DIEPV, may indicate that DIEPV PAP regulatory small subunit is more closely related to the AmEPV ORF060/HaEPV group than to the AmEPV ORF060/MsEPV ORF041 group (Fig. 2b). In any event, even the tree derived from the concatenated alignments of the five DIEPV ORFs (including the PAP sequences), confirms the result from the four individual phylogenetic trees (represented in Fig. 2a), that DIEPV is positioned between fowlpox virus and the group B EPVs (Fig. 2c).

3.3. Analysis of conserved domains of DIEPV homologues of GGT, eIF4A and λ-Int

The DIEPV ORF encoding a putative GGT homologue, had 83.9% (E-value = 7e–42) amino acid (aa) alignment to a gamma-glutamyltranspeptidase domain (NCBI, CCD entry # pfam01019, 504aa) (Fig. 1f). However, the DIEPV GGT possesses only five of the 15 conserved amino acids in the signature sequence of the GGT (Fig. 1g). Amino acid substitution analysis of human GGT revealed that Asp422 and Asp423 (Aspartic acid at positions #422 and 423, respectively) and Ser451 or Ser452 (serine at positions 451 or 452, respectively) are crucial for enzyme activity (Ikeda et al., 1995a, b). In a multiple alignment of the DIEPV GGT homologue and 12 GGT sequences from a variety of organisms such as Homo sapiens, D. melanogaster, and Vibrio parahaemolyticus, only the Ser residues were found in corresponding positions, but Asp422 and -423 in human GGT were instead replaced by Asn and Ser, respectively, in the DIEPV GGT homologue (Fig. 1g).

A CD analysis of the DIEPV λ-Int ORF indicated the presence of a phage integrase domain (pfam00589, 175aa) (E-value = 1e–9, 98.3% aligned) and a site-specific recombinase XerD (COG4974, 300aa) (E-value = 5e–9, 77.7% aligned) (Fig. 1h). The λ-Int is a prototype of a diverse family of tyrosine recombinases that rearranges DNA duplexes by means of conservative site-specific recombination reactions (Campbell, 1992). The λ-Int protein has four highly conserved residues, a tetrad, that includes tyrosine in an active site, which is directly involved in transesterification (Arg212, His308-XX-Arg311, Tyr343) (Argos et al., 1986; Abremiski and Hoess, 1992). λ-Int possesses three distinct structural domains: an N-terminal domain (aa residues 1–64) responsible for binding to distal arm sites on the phage DNA (Moitoso de Vargas et al., 1988), a central domain (aa residues 65–169) required for high-affinity binding to core-type sites, and the C-terminal catalytic domain (aa residues 170–356) (Tirumalai et al., 1998). The catalytic domain is known to be competent in specifically recognizing and cleaving its cognate core DNA substrates, either as full- or half sizes (Tirumalai et al., 1997). When a 3D structure of the DIEPV λ-Int homologue [predicted by 3D-psm (Fischer et al., 1999)] is manually aligned with the conserved tetrads in the catalytic domains of λ-Int (PDB entry #, 1AE9A)

![Fig. 1](http://www.ncbi.nlm.nih.gov/30/Structure/cdd/cdd.shtml). Red = identical; blue = similar; - = gap. (a) DIEPV poly (A) polymerase regulatory subunit (PAP) and the protein family (pfam), pfam01358 (Poly (A) polymerase regulatory subunit) domain. CD-Length = 294 residues, 93.9% aligned, E value = 1e–47. (b) DIEPV metalloprotease homologue (DIEPV metall) and pfam03410 (Peptidase M44 or Protein G1 that is a glycoprotein expressed by many Poxviridae) domain. CD-Length = 90 residues, 28.6% aligned, E value = 9e–45. (c) DIEPV membrane protein homologue (DIEPV memb) and pfam02442 (poxvirus unknown orf family) domain. CD-Length = 224 residues, 69.6% aligned, E value = 2e–17. (d) DIEPV late transcription factor homologue (DIEPV LTF) and pfam04947 (Poxvirus Late Transcription Factor VLT3 like) domain. CD-Length = 212 residues, 93.9% aligned, E value = 1e–9. (e) DIEPV metalloprotease homologue and pfam04947 (Poxvirus Late Transcription Factor VLT3 like) domain. CD-Length = 218 residues, 91.7% aligned, E value = 1e–08. (f) DIEPV virion surface protein homologue (DIEPV VSP) and pfam04584 (Poxvirus A28 family) domain. CD-Length = 140 residues, 97.1% aligned, E value = 4e–22. (g) DIEPV gamma-glutamyltransferase homologue (DIEPV GGT) and pfam01019 (Gamma-glutamyltransferase domain). CD-Length = 504 residues, 83.9% aligned, E value = 7e–42. (h) Upper alignment: DIEPV GGT and a signature pattern of CD of gamma-glutamyltransferase, BDCC00404. AA residues in [ ] represent possible alternative residues. Lower alignment: GGT protein sequences used for alignment are human GGT (AAH25927), Drosophila melanogaster GGT (CAA21412), and Vibrio parahaemolyticus GGT (NP799390). The regions used for alignment are indicated in parenthesis. (h) DIEPV λ-integrate homologue (DIEPV λ) and pfam00589 (Phage integrate) domain. CD-Length = 175 residues, 98.3% aligned, E value = 1e–09. (i) DIEPV λ-integrate homologue (DIEPV λ) and COG4974 (Site-specific recombinase XerD) domain. CD-Length = 300 residues, 77.7% aligned, E value = 5e–09. (i) DIEPV eIF4A homologue and cd00079 (Helicase superfamily c-terminal) domain. CD-Length = 131 residues, 99.2% aligned, E value = 1e–30. DIEPV eIF4A homologue and cd00268 (DEAD-box helicases) domain. CD-Length = 203 residues, 100.0% aligned, E value = 1e–59.
a) DlEPV PAP regulatory small subunit and pfam01358.

DlEPV PAP:  15 FYYVSDIQDRSVYNPNLAP--VKPFPKQGQLKLLLNEIRFFSESIDHDAKRKVYVTVL 72
pfam01358:  15 ILYFSEIKNELDYKKETENYIFFKFPYQGQLKLLGGELF--TVV 63

DlEPV PAP:  73 YIGSSPGLHIPFLIQLYSKYK--IKWWFYDPKPACQKLKDMAKKNKNIENLYTFRKDDI 130
pfam01358:  64 YIGSAPGYHIRYLMEHYSDLNVIIKWMLYDGRGHDDSLNGL----SNVSIVMRFVDYKLYLDGII--TVV 119

DlEPV PAP: 131 KTFKNTRDL--IFISDIRSSEDNSEPTTTNLLFDYKLQNEILLELFNYPKKLFRVPFPD 188
pfam01358: 120 ELYKNVKNLKILLISDIRSI-RGSEPSTKDLLEDYALQYNMLSLIPASLWKKRFYDPFVDD 178

DlEPV PAP: 246 NT-YQKK---ENDLVIAAHILNVYFRKEKGETYNLTKDTYMDLFKKIL 290
pfam01358: 236 NTKYRNKIVNNFDYPNQEYDYFMNKYLTVYLPFLKSF--DTNKNKVI 282

b) DlEPV metalloprotease and pfam03410.

DlEPV metal:   6 VLKNGITVFLQKNPEFPCSIHLKNHGSFRDIESNVSGFQHLAEHAIFVEGHILSRDTNAS 65
pfam03410:  3 VLENGVRVFINPGMGKDNYIGLSNFGFENDIPE-ILGIAHLLEH-ILISFDYTRFVANAS 60

DlEPV metal:  66 TDFAHMVLY------QDNVAILKSSLTNKHNQISLARNLSKDRIQGFIDELDNEFLYRQQ 119
pfam03410:  61 TARSYMSFWCRAIRAETYLDALRTLVSWFFARGALKDDFSRSRIRNHKELENNYTFH 120

DlEPV metal: 120 VNSANSPIYYLSTPFEVFYIGGNKLTLEKEREINQALNYQATSDVVFYLF 172
pfam03410:  121 VFHCMDILTFLGG--GDLYNGGRLSMLDQLDALELGRMRRIAGPNVVIFV 171

DlEPV memb:  13 LYEKIMNTFFTKLYISSAIASGTCNITIKRLVLKNSSTCKIDVKNVCMNSKSSTTTHLIE 72
pfam02442:  8 LYNTFNERYLRKLEQYAASASANCAIHIGNIIFGTNKNCNLTVLNMCSANAEASFDLVE 67

DlEPV memb:  73 SIQNEKFIKT-DOEKLRKLRPHIDKYQWNSNE---LRLTCEFASVETISTIDTFSIENC 128
pfam02442:  68 ALAEVSYLLEKPAATALEEGITDTLTTNDENFYVQCNKADAVNISQDQLTVVGEC 127

DlEPV memb: 129 YSSTDASIALNSLNSDAKCYSKYNIKVAKLASSOD 164
pfam02442: 128 SAPPQGMLFLQFVNGTANAKMCMAILDSLRTKRI 163

c) DlEPV membrane protein and pfam02442.

DlEPV LTF:   4 IKCSKCDSEDLIYSKKVLFCNRCGEHISNLNTFAETSK----NKILHNVNLKSMETVDN 59
pfam04947:  4 RHCKYCKSFGIIKEKDIYFCIDCQSVYTYSKRLKTSSFHVNNKLHNVLNLRKLSNQ1 54

DlEPV LTF:  60 FLLDLYKELKMKLKEYLHDDDDITDYHDYKLQEQMKNVTMFLHHNYWIDFYLCR 119
pfam04947:  64 SGDIIEFLEMAKMNULST---ITNIFYSEFARIGNYLRNLFELFEIINHKSANENFVLW 108

DlEPV LTF: 120 TKK-EXK1FRTFIDFDFTRYTVGAISHHLLTRYSTLYLIEHNLFSLNTRNHEER 177
pfam04947: 122 DOTTINENFDFFYVFCMYSYKSTYRSSFMDIPLDINNITSNLKPYVKKSHBNK 164

DlEPV LTF: 178 LIFDKFLLTNKRVTQGTLF 199
pfam04947: 182 IVPNIK---NDCIFLKSVE 139

d) DlEPV late transcription factor and pfam04947.

DlEPV VSP:   2 LVILVIFIIT--LCFSLFQFYEFYKNYDYILEFNDRYKKYEYVTTPGFNFIDKSIFDPND 59
pfam04584:  4 LSVFIIATVACCLLPIQLTYIUNYENQNINSKFLDTHNLSEYKSAGNLYLKDRAFDPD 54

DlEPV VSP:  60 EARTKEKRMCDQSKRWGSIGRSFV-NDGHHIAQWNNRRCRCLNTFIPSH-EILYNTPNC 117
pfam04584:  66 TYYDVQWMCRSQIYIVYIVYSVSMFQFKSSTSSNIRFTFTDLCIDYTDFTSTSHDIYNPC 108

DlEPV VSP: 118 KTDGKQSDNLKCNLKFNFKA 137
pfam04584: 124 IPVPNKS---NDCIFLKSVE 139

e) DlEPV virion surface protein and pfam04584.

DlEPV GGT:   1 MLCEGVVNPQTSGIGGGFLGMLYMNKQ--ISMINAREKTPYLKYYEVITYTPGFHDFSKIDPDND 59
pfam01019:  18 ALCLGVVEPHSSGIGGGGFMLIYDAGTGKVLAIDARETPAAATPDMFDGLGDPKLSKTG 77

DlEPV GGT:  55 -RAAGIPGAIRGYHLLHQKWGNLGWHEIFEPILNLCKTPFRL-----RQSQANKELLREH 108
pfam01019:  78 GLAVGVPGEVAGLEEAHKRYGRLPWADLLAPAIKLARDGFPVSPYLAAALALEEPFAD 103

DlEPV GGT: 109 QWYNFSFIGYQYRLVQRIPFVYNPELCYVLGNISQHGPDYFYT-EVAHLISEEF---Q 164
pfam01019: 138 PGSAEIFLPTGRPPKAGDLL-KQPALAKTLELIAEEGPEAFYRGELAEALVGOAAG 150

DlEPV GGT: 178 LIFDKFLLTNKRVTQGTLF 199
pfam01019: 182 ELWNNYLIYLRHKNLKINNSYVF 203

f) DlEPV gamma-glutamyltransferase and pfam01019.

DlEPV GGT:   1 MLCEGVVNPQTSGIGGGFLGMLYMNKQ--ISMINAREKTPYLKYYEVITYTPGFHDFSKIDPDND 59
pfam01019:  18 ALCLGVVEPHSSGIGGGGFMLIYDAGTGKVLAIDARETPAAATPDMFDGLGDPKLSKTG 77

DlEPV GGT:  55 -RAAGIPGAIRGYHLLHQKWGNLGWHEIFEPILNLCKTPFRL-----RQSQANKELLREH 108
pfam01019:  78 GLAVGVPGEVAGLEEAHKRYGRLPWADLLAPAIKLARDGFPVSPYLAAALALEEPFAD 103

DlEPV GGT: 109 QWYNFSFIGYQYRLVQRIPFVYNPELCYVLGNISQHGPDYFYT-EVAHLISEEF---Q 164
pfam01019: 138 PGSAEIFLPTGRPPKAGDLL-KQPALAKTLELIAEEGPEAFYRGELAEALVGOAAG 150

DlEPV GGT: 178 LIFDKFLLTNKRVTQGTLF 199
pfam01019: 182 ELWNNYLIYLRHKNLKINNSYVF 203
g) DlEPV GGT and a signature pattern of PDOC00404.

DlEPV GGT: 259 R   E   S N   V     C    IKKN N   D  A I   C   F

DlEPV GGT: 279

DlEPV: 223 TILKTYNQASMQSFQDY-------------------YQHLTN--------QTIAQOFOSQNSY 258
pfam01019: 248 GSAEYVLHELAEKMLRALCRDRLYDPPDVPLVYNNLQEEAKNALNLKNAAFQSSS 307

DlEPV : 259 RESNV---------CIRKNNDAICTISSINREPGY-YVKSSLILNQWNLF--- 301
pfam01019: 308 YGGAAPAAPEDGTHSFVQDRGNVNASSTNQGLFGQVVPQGTVGNEMDFITP 367

DlEPV : 302 ------FPMHYRGQIDLQFPSSSGFLYIGKIPYVLAAATGKGAPFYSLHINLYFNYQYL 355
pfam01019: 368 GFGILDSHANAIFEPKRLPLMSFPTVLDGKFPVLVAGFGGRSITQFQLVQYVNLVDVG 427

DlEPV : 356 ACPQEAINAYRYY 368
pfam01019: 428 MEKLAQEVADFAPFRYY 440

Human GGT containing Asp-422 and 423 and Ser-451 and ß452 (underlined) and GGTs of
Drosophila melanogaster
Vibrio parahaemolyticus
and DlEPV.

Human (414-462)       GILFNNEMDD FSSPSITNE FGVPPSPANF IQPGKQPLSS

Drosophila (408-456)  GFILNDEMDD FSTPGKI-NA YGIPASPANF IKPGKRPMSS TCPSIVLD

Vibrio (438-482)      GFLLNNELTD FSFRTHQNGV P-----IANR LEPGKRPRSS MAPTIIMK

DlEPV (291-327)       GIILNNQLNS FPMHYRQ--- ---------- IQDLQPPSSS GITFLIYG

------------------------------------------------------------------------

h) DlEPV -integrase and pfam00589.

DlEPV:   122 LKNEELEKLYDELDSYDKIG------FLILITTGMRRGALIKCDRSKIDYSYKVPIIS-- 173
pfam00589:   1 LTEDQIEKLLAASELARPIGARDRAAVELLLLTGLRISELLSLRWSDIDFDKGTIFIPVR 60

DlEPV:   174 TIEKGNVLTKYTLNPEIIHLLKE-------NPTFNFYTAATL----------RTLINRCK 216
pfam00589:  61 TSGKGRKSRTVPLSDKAVEALKQYLEIYGRRDLLGGERSDALFPSAVGKPLSRRLLRRAG 172

DlEPV:   217 KILNKHIKTHIFRFTFARLVLNTKNNAQNVQVLHDLANSLTQTSTYIEKETYQD 269
pfam00589: 121 KDAGEELTPHDLRHTFATHLLEAGVDLRVIQKLLGHSSISMTQR-YTHVAAE 172

DlEPV -integrase and COG4974.

DlEPV:   51   TKIKTITDQDIFFFIEHRSKAYIFFLSYIAKRYLDIDLDPIKSIKKYRAHLDIISGN--- 107
COG4974:   47   TDLADATEADIREYLTEL-AEQGLSATSIARA-----LSALSPKLPKRLPKFLSEEEVEALLEAPDEDTPLGLRDRAMLELLYATGLRVSELVGLLS 100

DlEPV:   108  YLSYPVKEKKGTPYLKNEELEKLYDELDS------YDKIGFLILITTGMRRGALIKCDRS 161
COG4974:   101  LLDSPKLPKRLPKFLSEEEVEALLEAPDEDTPLGLRDRAMLELLYATGLRVSELVGLLS 160

DlEPV:   162  KIDYSYKVPISTIEKGN-------------VLTKYTLNPEIIHLLKENPTFNFY----- 203
COG4974:   161  DVDLRQGV--VRVKGKGNKERLVPFGEEAVEALEKY-LEEARPKLLKGKSSTDALFPNQR 217

DlEPV:   204  -TAATLRTLINRCKKI-----LNKHKITHIFRFTFARLVLNTKNNAQNVQVLHDLANSLTQTSTYIEKETYQD 257
pfam00589: 121 KDAGEELTPHDLRHTFATHLLEAGVDLRVIQKLLGHSSISMTQR-YTHVAAE 172

DlEPV χ-integrate and COG4974.

DlEPV χ:  51  TKITKTQDQFIFFYEIKRRLKYIYFLPSTIQQKAIFPCISGKDVIVQAQSGTGKTATYAISVLQQ 107
COG4974:  47  TDLDATADDRLSYELTEL-AEQGSLATSARA-----LSALSHQFYQLREGLEDSPTR 100

DlEPV χ: 108 YLSYPVKEKKGTPYLKNEELEKLYDELDS------YKIGFLILITTMGRLIKCDRS 161
COG4974: 101 LLDSPKLPKRLPKFLSEEEVEALLEAPDEDTPLGLRDRAMLELLYATGLRVSELVGLLS 160

DlEPV χ: 162 KIDYSYKVPISTIEKGN-------------VLTKYTLNPEIIHLLKENPTFNFY----- 203
COG4974: 161 DVDLRQGV--VRVKGKGNKERLVPFGEEAVEALEKY-LEEARPKLLKGKSSTDALFPNQR 217

DlEPV χ: 204 -TAATLRTLINRCKKI-----LNKHKITHIFRFTFARLVLNTKNNAQNVQVLHDLANSLTQTSTYIEKETYQD 257
pfam00589: 121 KDAGEELTPHDLRHTFATHLLEAGVDLRVIQKLLGHSSISMTQR-YTHVAAE 172

DlEPV χ and cd00079.

DlEPV eIF4A: 219 IRQYYINVKKNDFKAETLFDLYDHLSLT--QTLIFCNTRRQVDVLMECLTNRNFTASSIH 276
at00079:  2 IKQYVLPVE--DEKLEALLELLKEHLKKGGKVLIFCPSKKMLDELAELLRKPGIKVAALH 59

DlEPV eIF4A: 277 GDMSQQERDFIMKEFRDGKTRILLSTNLARGIDVQQISLVINYDLPHNRENYIHRIGRS 336
at00079:  60 GDGSQEEREEVLKDFREGEIVVLVATDVIARGIDLPNVSVVINYDLPHQSGQSPSYYIQIGRA 119

DlEPV eIF4A: 337 GRFGRKGIAINF 348
at00079: 120 GRAGQKGTAILL 131

DlEPV eIF4A and cd00268.

DlEPV eIF4A:   8 FDHMGLKKNILKGIYSCGFEKPSTIQQKAIFPCISGKDVIVQAQSGTGKTATYAISVLQQ 67
Cd00268:   1 FEELGfpsELKCCQGELVRGKLDLSKVKYLVLDEADRMLDMGFEDQIREILKLLPKDRQ 60

DlEPV eIF4A:  68 IDTSNSN--IQALILTPTRELALQAQRVLQTIGNYLYNFKCQVCIGGTSIKESQETLKK- 124
Cd00268: 61 LDPSPKKDGPQALILAPTRELALQAERARKLGKHT-NLKVVYYGGSIDQR1KLRKG 119

DlEPV eIF4A: 125 AQVLIGTPGRMIDLLTRKSIDTKAIKIVVIDEADEMLIDNFLDKIQDIFEFFESHVQVIL 184
Cd00268: 120 PHVATPPGRRDLIREGLKLGLSKVLYLCDERMLDMGFDQIREILKLLPKDRQ 179

DlEPV eIF4A: 185 LSATVPFSRVINTSQVMDPVPKIL 208
Cd00268: 180 FATMFPFVRDLKAFWAYPVRIL 203

Fig. 1. (Continued)
and Cre (PDB entry #, 1CRXA) tyrosine recombinases, the same tetrad is conserved in DlEPV (Fig. 3). The DlEPV λ-Int tetrads align closely with those of λ-Int and Cre recombinases, with the Arg (R)–Arg (R) spacing (78aa) and Arg (R)–Tyr (Y) spacing (32aa) of DlEPV integrase homologue falling within the range found in other tyrosine recombinases (Esposito and Scocca, 1997) (Fig. 3). Additionally, the secondary structure of the DlEPV λ-Int homologue consists of a similar number of helices and overall interhelical spacing as do the λ-Int and Cre recombinases (Fig. 3).

Several CDs were identified among the eIF4A homologues in the Blast search, including the helicase super family C-terminal domain (NCBI, CDD entry # cd00079, 322aa) (E-value = 1e−30, 99.2% aligned) and the DEAD-box helicase domain (NCBI, CCD entry # cd00268, 203aa) (E-value = 1e−59, 100.0% aligned) (Fig. 1i). eIF4A together with eIF4E and eIF4G consist of a eukaryotic initiation factor 4F that is involved in the recognition of the mRNA cap structure, ATP-dependent unwinding of the 5’-terminal secondary structures (helicases) and recruitment of the mRNA to the ribosome (Svitkin et al., 2001). The eIF4A is a prototype helicase of the DEAD box helicase family that consists of eight motifs (I, Ia, Ib, II, III, IV, V, and VI) (Tanner and Linder, 2001). The DlEPV eIF4A homologue has a single amino acid substitution in motif Ib, where the last amino acid is altered from Ile (I) to Met (M), and five amino acid substitutions in motif V (Fig. 4).

The distances between the N-terminus and motif I (49aa) of DlEPV and between motif VI and the C-terminus (39aa) along with the overall arrangement of the motifs, are very similar to those of DEAD box helicases (Nunes-Duby et al., 1998). Interestingly, the DlEPV sequences in all motifs are almost identical (except for six residues) with those of Drosophila melanogaster (Fig. 4), a
cyclophragnous dipteran like the A. suspensa host within which DIEPV is pathogenic (Lawrence, 2002).

When the sequence of DIEPV eIF4A is inferred with 3D-pssm (Kelley et al., 2000), there was a strong similarity in structural profile to that of yeast eIF4A (E-value, 0.00196). Furthermore, the Swiss Model for comparative protein modeling (Guex and Peitsch, 1997) indicated that a predicted 3D model of DIEPV eIF4A

---

Fig. 4. Alignments of the conserved motif of DEAD helicases with the eIF4A ORFs of DIEPV (379 aa) and Drosophila melanogaster (D.m.) (403 aa). In the DEAD motif, amino acids conserved at least 80% of the time are shown as capital letters while those conserved 50%-79% of the time are in lower case (Tanner and Linder, 2001). DEAD helicases have 5 motifs (boxed) that are separated into two domains by motif III. Non-conserved amino acid residues between DIEPV and D.m. are underlined. Numbers of amino acid residues between motifs and at both termini are shown in parenthesis.
consists of two globular domains that are connected by the motif III (data not shown), which is typical of 3D structures of DEAD box helicases.

4. Discussion

In this paper, we describe five DIEPV ORFs that encode putative proteins that are homologous to those conserved in poxviruses, including two homologues of the vaccinia virus (VV) putative membrane protein (VV F9L homologue) and virion surface protein (VV A28L homologue), two transcription/RNA modification proteins, late transcription factor 3 (VV A2L homologue), PAP regulatory small subunit (VV J3R homologue), and one viral enzyme, metalloprotease (VV G1L homologue). These sequences enrich previous DIEPV sequence data consisting of ORFs similar to DNA dependent RNA polymerase (VV J6R homologue) (Lawrence, 2002) and DNA helicase (Mwaengo and Lawrence, 2003) and an ORF similar to rifampicin resistant protein (AF159588, VV D13L homologue).

Phylogenetic analysis of single and concatenated multiple alignments of each of the 20–21 poxvirus homologues suggested that DIEPV diverged from an ancestor before AmEPV and MsEPV (Fig. 2). This pattern was consistent in four of five phylogenetic trees of single protein alignments and one from concatenated alignments (Fig. 2). Three DIEPV ORFs similar to GGT, λ-Int and eIF4A are not found in poxviruses including EPVs. However, these ORFs retain poxviral transcription motifs and CDs and/or secondary–tertiary structures specific to each of the proteins. These findings and the characteristics of the DIEPV genes presented here, along with earlier reports of DIEPV phylogenetic relationship with other EPVs (Lawrence, 2002; Mwaengo and Lawrence, 2003), strongly support our hypothesis that DIEPV is a member of the Entomopoxvirinae, but is distinct from the group B EPVs of AmEPV and MsEPV.

Interestingly, DIEPV showed a closer relationship to fowlpox virus than to other ChPVs in all phylogenetic trees (Fig. 2). Similar phylogenetic relationships were reported with other EPVs when homologues of NPH 1 and vaccinia virus VP8 were analyzed (Li et al., 1998; Crnogor and Dall, 1999). In other studies, phylogenetic analysis using completely sequenced poxvirus genomes revealed that AmEPV, MsEPV, and fowlpox virus have common genome characteristics in that they possess large size genomes (>230 kb) compared to other ChPVs (<230 kb) (Afonso et al., 2000). Additionally co-linearity of conserved genes on the genome is significantly reduced or lost in these viruses and it has been suggested that a number of genes are acquired (McLysaght et al., 2003; Gubser et al., 2004).

The presence of DIEPV genes that encode new poxvirus putative proteins, suggests that these proteins likely, play a role in the interactions of the virus with either its wasp vector, the dipteran host, or both. Since GGT, eIF4A, and λ-Int occur in bacteria, Drosophila, phage, and other organisms but not in poxviruses, it is likely that these genes were acquired from other organisms (probably the A. suspensa host and/or microflora) by DIEPV and used in some way, probably to facilitate aspects of viral pathogenicity and/or wasp survival. GGT has been found in many organisms except for viruses (PROSITE ID, PDOC00404). It catalyzes the transfer of the gamma-glutamyl moiety of glutathione to an acceptor that may be an amino acid, a peptide, or water (forming glutamate) (Tate and Meister, 1985). GGT plays a key role in the gamma-glutamyl cycle, a pathway for the synthesis and degradation of glutathione, which is an antioxidant and detoxifying agent (Tate and Meister, 1985). Thus, GGT could be employed to maintain a physiologically stable environment and enhance development of the wasp larva. Indeed, Nappi et al. (1995) have proposed that free radicals released from Drosophila hemocytes kill the Leptopilina sp. parasitic wasps in Drosophila larvae. Also, Li et al. (2003) reported that a major component of VLP, VLP1, from the endoparasitoid Venturia canesens, which has been reported as a non-enzymatic molecule of the hydroperoxide glutathione peroxidase, could function as an antioxidant that reduces peroxide phospholipids and cholesterol within membranes (Ursini et al., 1981; Ursini et al., 1986).

The DIEPV putative λ-Int presumably, like other integrases, catalyses site-specific DNA recombinations via the Holliday junction structure (Azano and Landy, 1997). The latter is recognized by poxvirus resolvases, that process concatameric viral DNA into unit-length genomes (Arazo and Landy, 1997; Garcia and Moss, 2001).

However, there are apparently no poxvirus integrases in the database and the closest homologue to DIEPV λ-Int was that of a bacteriophage (Fig. 3). It is unlikely that this ORF is a result of bacterial contamination of DIEPV genomic DNA because three other ORFs in the same contig are homologous to Plasmodium falciparum proteins (data not shown). This raises the possibility that DIEPV probably acquired its λ-Int from bacteria that occur in the host insect’s hemolymph and/or gut.

We can only speculate that DIEPV λ-Int may be used for DNA recombination in portions of the wasp’s genome as reported for PDVs that insert viral segments into the genome of their wasp vector (Belle et al., 2002). The PDV genes are then expressed in the host and disrupt the host’s immune response, thus protecting the developing wasp (Belle et al., 2002). However, unlike DIEPV, no PDV virion replication occurs in the host. It is not yet known whether portions of the DIEPV genome are integrated into the D. longicaudata wasp genome. However, this phenomenon has also been
suggested for other wasp symbiotic viruses (see Reineke and Asgari, 2005), and could be a shared mechanism underlying our term “accommodative symbiosis”. If such viral integration exists for viruses with non-segmented genomes, like DlEPV, then the integration could (1) prevent the wasp from recognizing the virus as non-self (hence “accommodative symbiosis”) or (2) facilitate the expression of viral genes early in the infection of the host (prior to virus morphogenesis and replication) to protect the wasp egg from encapsulation as suggested for PDVs (Webb, 1998).

The homologue encoding eIF4A is present only in DlEPV but not in other viruses. Although the DlEPV eIF4A ORF contained DEAD motifs that were homologous with those of DEAD helicases, the highly significant E-values of the eIF4A hit in the Blast search indicated that the DlEPV ORF encodes a putative protein that is more closely related to eIF4A proteins than to helicases. Furthermore, this close relationship is supported by the predicted secondary/tertiary structure similarities (Fig. 4). eIF4A has been identified as an essential single copy gene in D. melanogaster (Dorn et al., 1993). An alignment of DlEPV putative eIF4A protein sequence to DEAD helicase motifs and D. melanogaster eIF4A revealed a strong similarity between them (Fig. 4), and suggests a common ancestral origin of these proteins. Recent reports on the expression of eIF4A from D. melanogaster in different stages demonstrated the ubiquitous presence of two abundant eIF4A gene transcripts throughout development. However, the levels of the eIF4A protein decreased significantly (40%) only in third instar larvae (Dorn et al., 1993; Hernandez et al., 2004). The third instar is the last developmental stage of fruit fly larvae and is the stage normally parasitized by D. longicaudata females (Lawrence et al., 1976). This raises the question of whether a decrease of eIF4A may be replenished by DlEPV to strengthen the protein synthesis machinery of the host larvae, in order to support virus replication and wasp development. Further studies on the expression and function of these DlEPV genes are needed to clarify their role, if any, in these parasite-host interactions.

DlEPV propagates in the parasitized fruit fly host (Lawrence, 2002). This attribute clearly separates DlEPV from the extensively studied PDVs, whose virions do not normally replicate in the parasitized hosts. Therefore, immunosuppression, hemocyte blebbing, and apoptosis induced by DlEPV infection of the parasitized host are clear evidence of virus pathogenicity (Lawrence, 2005). Given the phylogenetic distinction of DlEPV from the group B EPVs (Fig. 2), the viral replication in the adult wasp without any evidence of viral-induced pathology (Lawrence and Akin, 1990), and pathogenicity in virus-infected host hemocytes (Lawrence, 2005) comitant with the successful development of the wasp larvae, we believe that DlEPV belongs to the group C (Diptera) EPVs or to a new EPV group that reflects its unique biology as the first symbiotic EPV that also does not express the typical EPV occlusion body (Lawrence, 2002).

Acknowledgements

Paid technical services were provided by the UF ICBR DNA Sequencing and Hybridoma Cores. Financial support from the National Science Foundation Grant IBN 9986076 to POL is gratefully acknowledged. This research is part of CRIS project ENY 03507, supported by the Florida Agricultural Experiment Station, and approved for publication as Journal Series No. R-10486.

References

Abremski, K.E., Hoess, R.H., 1992. Evidence for a second conserved arginine residue in the integrase family of recombination proteins. Protein Engineering 5, 87–91.
Afonso, C.L., Tulman, E.R., Lu, Z., Oma, E., Kutish, G.F., Rock, D.L., 1999. The genome of Melanoplus sanguinipes entomopoxvirus. Journal of Virology 73, 533–552.
Afonso, C.L., Tulman, E.R., Lu, Z., Zsak, L., Kutish, G.F., Rock, D.L., 2000. The genome of fowlpox virus. Journal of Virology 74, 3815–3831.
Argos, P., Landy, A., Abremski, K., Egan, J.B., Haggard-Ljungquist, E., Hoess, R.H., Kahn, M.L., Kalionis, B., Narayana, S.V., Pierson III, L.S., 1986. The integrase family of site-specific recombinases: regional similarities and global diversity. The EMBO Journal 5, 433–440.
Azaro, M.A., Landy, A., 1997. The isomeric preference of Holliday junctions influences resolution bias by lambda integrase. The EMBO Journal 16, 3744–3755.
Bawden, A.L., Glassberg, K.J., Diggans, J., Shaw, R., Farmerie, W., Moyer, R.W., 2000. Complete genomic sequence of the A. moorei entomopoxvirus: analysis and comparison with other poxviruses, Virology 274, 120–139.
Belle, E., Beckage, N.E., Rousselet, J., Poirie, M., Lemeunier, F., Drezen, J.M., 2002. Visualization of polydnavirus sequences in a parasitoid wasp chromosome. Journal of Virology 76, 5793–5796.
Bigot, Y., Drezen, J.M., Sizaret, P.Y., Rabouille, A., Hamelin, M.H., Periquet, G., 1995. The genome segments of DpRV, a commensal reovirus of the wasp Diadromus pulchellus (Hymenoptera). Virology 210, 109–119.
Bigot, Y., Rabouille, A., Sizaret, P.Y., Hamelin, M.H., Periquet, G., 1997. Particle and genomic characteristics of a new member of the Ascoviridae: Diadromus pulchellus ascovirus. Journal of General Virology 78, 1139–1147.
Campbell, A.M., 1992. Chromosomal insertion sites for phages and plasmids. Journal of Bacteriology 174, 7495–7499.
Crnov, R., Dall, D.J., 1999. Virions of Heliothis armigera entomopoxvirus contain a homologue of the vaccinia VP8 major core protein. Virus Genes 19, 23–31.
Dall, D., Srisanktha, A., Vera, A., Lai-Fook, J., Symonds, T., 1993. A gene encoding a highly expressed spindle body protein of Heliothis armigera entomopoxvirus. Journal General Virology 74, 1811–1818.
Davison, A.J., Moss, B., 1989. Structure of vaccinia virus late promoters. Journal of Molecular Biology 210, 771–784.

Dom, R., Morawietz, H., Reuter, G., Sauweber, H., 1993. Identification of an essential Drosophila gene that is homologous to the translation initiation factor eIF-4A of yeast and mouse. Molecular Genetics and Genomics 237, 233–240.

Dupas, S., Brehelin, M., Frey, F., Carton, Y., 1996. Immune suppressive virus-like particles in a Drosophila parasitoid; significance of their intraspecific morphological variations. Parasitology 113, 207–212.

Esposito, D., Scocca, J.J., 1997. The integrase family of tyrosine recombinases: evolution of a conserved active site domain. Nucleic Acids Research 25, 3605–3614.

Esposito, J.J., Baxby, D., Black D.N., Dales, S., Darai, G., Dumbell, F., Esposito, D., Scocca, J.J., 1997. The integrase family of tyrosine recombinases: evolution of a conserved active site domain. Nucleic Acids Research 25, 3605–3614.

Fleming, J.G., Summers, M.D., 1991. Polydnavirus DNA is integrated to the translation initiation factor eIF-4A of yeast and mouse. Journal of Molecular Biology 210, 771–784.

Gruidl, M.E., Hall, R.L., Moyer, R.W., Pickup, D.J., Robinson, A.J., Tripathy, D.N., 2000. The spheroidin of an entomopoxvirus isolated from the parasitoid wasp Apanteles melanoscelus. Isolation and preliminary characterization. Journal of Virology 29, 1118–1130.

Kroemer, J.A., Webb, B.A., 2004. Polydnavirus genes and genomes: emerging gene families and new insights into polydnavirus replication. Annual Review of Entomology 49, 431–456.

Lavine, M.D., Beckage, N.E., 1996. Polydnavirus: Potent mediators of host insect immune dysfunction. Parasitology Today 11, 362–368.

Lawrence, P.O., 1998. Ecdysteroid titers and integument changes in superparasitized puparia of Anastrepha suspensa (Diptera: Tephritidae). Journal of Insect Physiology 34, 603–608.

Lawrence, P.O., 2002. Purification and partial characterization of an entomopoxvirus (DIEPV) from a parasitic wasp of tephritid fruit flies. Journal of Insect Science 2 (10), 1–12.

Lawrence, P.O. 2005. Morphogenesis and cytopathic effects of the Diachasmimorpha longicuodata entomopoxvirus in host haemocytes. Journal of Insect Physiology, (accepted).

Lawrence, P.O., Akin, D.B., 1990. Virus-like particles from the poison glands of the parasitic wasp Biosteres longicaudatus (Hymenoptera: Braconidae). Canadian Journal of Zoology 68, 539–546.

Le, N.T., Asgari, S., Amaya, K., Tan, F.F., Beckage, N.E., 2003. Persistence and expression of Cotesia congregata polydnavirus in host larvae of the tobacco hornworm, Manduca sexta. Journal of Insect Physiology 49, 533–543.

Li, X., Barrett, J.W., Yuen, L., Arif, B.M., 1997. Cloning, sequencing and transcriptional analysis of the Choristoneura fumiferana entomopoxvirus spheroidin gene. Virus Research 47, 143–154.

Li, X., Wallis, J.L., Barrett, J.W., Krell, P.J., Arif, B.M., 1998. Characterization of the nucleoside triphosphate phosphohydrolase I gene from the Choristoneura fumiferana entomopoxvirus. Virus Research 56, 93–105.

Li, D., Blasievich, F., Theopold, U., Schmidt, O., 2003. Possible function of two insect phospholipid-hydroperoxide glutathione peroxidases. Journal of Insect Physiology 49, 1–9.

Lytvyn, V., Fortin, Y., Banville, M., Arif, B., Richardson, C., 1992. Comparison of the thymidine kinase genes from three entomopoxviruses. Journal of General Virology 73, 3235–3240.
Marchler-Bauer, A., Anderson, J.B., DeWeese-Scott, C., Fedorova, N.D., Geer, L.Y., He, S., Hurwitz, D.I., Jackson, J.D., Jacoby, A.R., Lanczycki, C.J., Liebert, C.A., Liu, C., Madej, T., Marchler, G.H., Mazumder, R., Nikolskaya, A.N., Panchenko, A.R., Rao, B.S., Shoemaker, B.A., Simonyan, V., Song, J.S., Thiessen, R.A., Vasudevan, S., Wang, Y., Yamashita, R.A., Yin, J.J., Bryant, S.H., 2003. CDD: a curated Entrez database of conserved domain alignments. Nucleic Acids Research 31, 383–387.

McLyshaght, A., Baldi, P.F., Gaut, B.S., 2003. Extensive gene gain associated with adaptive evolution of poxviruses. Proceedings of the National Academy of Sciences USA 100, 15655–15660.

Mitsuhashi, W., Saito, H., Sato, M., 1997. Complete nucleotide sequence of spheroidin gene of the cupreous chafer, Anomala cuprea Hope (Coleoptera: Scarabaeidae). Insect Biochemistry and Molecular Biology 27, 869–876.

Mitsuhashi, W., Saito, H., Sato, M., Nakashima, N., Noda, H., 1998. Complete nucleotide sequence of spheroidin gene of Anomala cuprea entomopoxvirus. Virus Research 55, 61–69.

Moitozo de Vargas, L., Pargellis, C.A., Hasen, N.M., Bushman, E.W., Landy, A., 1988. Autonomous DNA binding domains of lambda integrase recognize two different sequence families. Cell 54, 923–929.

Moraes, J., Hsiling, C., Oo, T., Plaza, R., Hoskins, S., Govind, S., 2005. Biogenesis, structure, and immune suppressive effects of virus like particles of a Drosophila parasitoid, Leptopilina victoriae. Journal of Insect Physiology, accepted.

Moss, B., 2001. Poxviridae: The viruses and their replication. In: Knipe, D.M., Howley, P.M. (Eds.), Fundamental Virology. Lippincott Williams & Wilkins, A Wolters Kluwer Company, pp. 1249–1283.

Mustafa, A., Yuen, L., 1991. Identification and sequencing of the Choristoneura biennis entomopoxvirus DNA polymerase gene. DNA Sequence 2, 39–45.

Mwaengo, D.M., Lawrence, P.O., 2003. A putative DNA helicase and novel oligoribonuclease in the Dicauhsimornphora longicuata entomopoxvirus (DIEPV). Archives of Virology 148, 1431–1444.

Nappi, A.J., Vass, E., Frey, F., Carton, Y., 1995. Superoxide anion generation in Drosophila during melanotic encapsulation of parasites. European Journal of Cell Biology 68, 450–456.

Nunes-Duby, S.E., Kwon, H.J., Tirumalai, R.S., Ellenberger, T., Landy, A., 1998. Similarities and differences among 105 members of the Int family of site-specific recombinases. Nucleic Acids Research 26, 391–406.

Oshlakae, R.J., Symonds, T.M., Sriskantha, A., Lai-Fook, J., Feron, C.A., Dall, D.J., 1996. An entomopoxvirus homologue of the vaccinia virus D13L-encoded ‘rifampicin resistance’ protein. Journal of General Virology 77, 839–846.

Page, R.D.M., 1996. TREEVIEW: an application to display phylogenetic trees on personal computers. Computer Applications in the Biosciences 12, 357–358.

Reineke, A., Asgari, S., 2005. Presence of a novel picorna-like virus in a laboratory culture of the endoparasitic wasp Venturia canescens (Hymenoptera: I chneumonidae). Journal of Insect Physiology, accepted.

Renault, S., Bigot, S., Lemesle, M., Sizaret, P.Y., Bigot, Y., 2003. The cytopivirus Diadromus pulchellus RV-2 is sporadically associated with the endoparasitoid wasp D. pulchellus and modulates the defense mechanisms of pupae of the parasitized leek-moth, Acrolepis assectella. Journal of General Virology 84, 1799–1807.

Rizki, R.M., Rizki, T.M., 1990. Parasitoid virus-like particles destroy Drosophila cellular immunity. Proceedings of the National Academy of Sciences USA 87, 8388–8392.

Risol, J.L., Earl, P.L., Weir, J.P., Moss, B., 1986. Conserved TAAATG sequence at the transcriptional and translational initiation sites of vaccinia virus late genes deduced by structural and functional analysis of the HindIII H genome fragment. Journal of Virology 60, 436–449.

Sanz, P., Veyrunes, J.C., Cousserans, F., Bergoin, M., 1994. Cloning and sequencing of the spherulin gene, the occlusion body major polypeptide of the Melolontha melolontha entomopoxvirus (MmEPV). Virology 202, 449–457.

Sriskantha, A., Osborne, R.J., Dall, D.J., 1997. Mapping of the Heliothis armigera entomopoxvirus (HaEPV) genome and analysis of genes encoding the HaEPV spheroidin and nucleoside triphosphate phosphohydrolase I proteins. Journal of General Virology 78, 3115–3123.

Stoltz, D.B., 1981. A putative baculovirus in the ichneumonid parasitoid, Mesoleius tentheadrini. Canadian Journal of Microbiology 27, 116–122.

Stoltz, D., Makkay, A., 2000. Co-replication of a reovirus and a polydnavirus in the ichneumonid parasitoid Hypsospy exiguia. Virology 278, 266–275.

Subramanian, K., Rutvisuttinunt, W., Scott, W., Myers, R.S., 2003. The enzymatic basis of processivity in lambda exonuclease. Nucleic Acids Research 31, 1585–1596.

Svitkin, Y.V., Pause, A., Haghhat, A., Pyronnet, S., Witherell, G., Belsham, G.J., Sonenberg, N., 2001. The requirement for eukaryotic initiation factor 4A (eIF4A) in translation is in direct proportion to the degree of mRNA 5’ secondary structure. RNA 7, 382–394.

Tanner, N.K., Linder, P., 2001. DExD/H box RNA helicases: from generic motors to specific dissociation functions. Molecular Cell 8, 251–262.

Tate, S.S., Meister, A., 1985. Gamma-Glutamyl transpeptidase from kidney. Methods in Enzymology 113, 400–419.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research 25, 4876–4882.

Tirumalai, R.S., Healey, E., Landy, A., 1997. The catalytic domain of lambda site-specific recombinase. Proceedings of the National Academy of Sciences USA 94, 6104–6109.

Tirumalai, R.S., Kwon, H.J., Cardente, E.H., Ellenberger, T., Landy, A., 1998. Recognition of core-type DNA sites by lambda integrase. Journal of Molecular Biology 279, 513–527.

Ursini, F., Maiorino, M., Ferri, L., Valente, M., Gregolin, C., 1981. Hydrogen peroxide and hematin in microsomal lipid peroxidation. Journal of Inorganic Biochemistry 15, 163–169.

Ursini, F., Maiorino, M., Gregolin, C., 1986. Phospholipid hydroperoxide glutathione peroxidase. International Journal of Tissue Reaction 8, 99–103.

Yuen, L., Noiex, M., Gomes, M., 1991. DNA sequence of the nucleoside triphosphate phosphohydrolase I (NPH I) of the Choristoneura biennis entomopoxvirus. Virology 182, 403–406.

Webb, B.A., 1998. Polydnavirus biology, genome structure, and evolution. In: Miller, L.K., Ball, A.L. (Eds.), The Insect Viruses. Plenum Pub. Coop., New York, pp. 105–139.

Whitfield, J.B., Asgari, S., 2003. Virus or not? Phylogenetics of polydnaviruses and their wasp carriers. Journal of Insect Physiology 49, 397–405.

Further reading

Lawrence, P.O., 2000. Parasitism specific proteins and a symbiotic entomopoxvirus from Dicauhsimornphora longicuata, a parasite of tephritid fruit flies. Abstracts of the XXI International Congress of Entomology and XVIII Brazilian Congress of Entomology, Iguassu Falls, Brazil. Abstract #2371, Book II, p. 600.