Phylogenetic analysis of conserved genes within the ecdysteroid UDP-glucosyltransferase gene region of the slow-killing Adoxophyes orana granulovirus

S. L. Wormleaton and D. Winstanley

Horticulture Research International (HRI), Wellesbourne, Warwickshire CV35 9EF, UK

A physical map of the genome of Adoxophyes orana granulovirus (AoGV) was constructed for the restriction enzymes BamHI, BglII, EcoRI, PstI and SacI using restriction endonuclease analysis and DNA hybridization techniques. This enabled the size of the AoGV genome to be estimated at 100–9 kbp. A plasmid library covering 99.9% of the AoGV genome was constructed using five restriction enzymes. The ecdysteroid UDP-glucosyltransferase gene (egt) was located by hybridization with the egt gene of Cydia pomonella granulovirus. The sequence of 6000 bp of the egt region is presented and compared to the equivalent area in other GVs. Database searches showed that this region contained eight open reading frames (ORFs) similar to the baculovirus genes egt, granulin, pk-1, me53 and four ORFs of Xestia c-nigrum granulovirus (ORF 178, ORF 2, ORF 7 and ORF 8). The egt gene was shown to encode an active EGT using an EGT assay. Phylogenetic trees of the granulovirus genes egt, granulin, pk-1 and me53 were constructed using maximum parsimony and distance analyses. These analyses indicated that AoGV genes may be more closely related to other tortricid-infecting GVs than to GVs that infect other lepidopteran families.

Introduction

Baculoviruses belong to a group of insect viruses that are used in many countries for the control of pests in agriculture and forestry. The granuloviruses (GVs) form one of two major groups in the family Baculoviridae (Murphy et al., 1995). Three types of GV are recognized based on their tissue tropism and have been categorized by Federici (1997). Type 1 infects the fat body usually resulting in a slow speed of kill; type 2 infects most tissues resulting in a faster speed of kill; type 3 infects only the midgut and at present contains only one member, Harrisina brillians GV (HbGV).

The entire sequences of Xestia c-nigrum GV (XcGV) and Plutella xylostella GV (PxGV) have been published recently (Hayakawa et al., 1999; Hashimoto et al., 2000) and the entire sequence of Cydia pomonella GV (CpGV) will soon be available (T. Luque, R. Finch, N. E. Crook, D. R. O’Reilly & D. Winstanley, unpublished results). DNA sequence analysis is only now beginning to distinguish different groups of GVs and nucleopolyhedroviruses (NPVs) and may provide a more appropriate method of differentiating the GVs taxonomically.

The summer fruit tortrix moth, Adoxophyes orana L. (Lepidoptera; Tortricidae) is a major fruit pest in Europe and Asia, costing the apple and pear industry millions of pounds in fruit losses (Yamada & Ono, 1973; Cross et al., 1999). It was first recorded in the United Kingdom in 1950 and has now spread through most of the south-east of England (Cross, 1994). With the increasing withdrawal of chemical pesticides, alternative methods of control are in demand.

A granulovirus from A. orana, A. orana granulovirus (AoGV), is already used as a biocontrol agent in Europe and Japan, with the potential to be used in the UK and elsewhere for the control of summer fruit tortrix.

Another tortricid species closely related to A. orana is Adoxophyes honnai or the smaller tea tortrix, which is pest of tea trees and is particularly prevalent in Japan (Nishi & Nonaka, 1996). There are only slight morphological differences between the two species, and the adults of the two species can reproduce but form sterile offspring (Honma, 1970). In Japan, a GV isolated from A. honnai is used as a biocontrol agent of smaller tea tortrix (Nishi & Nonaka, 1996).

The egt genes of baculoviruses encode ecdysteroid UDP-glucosyltransferase (EGT) enzymes, which belong to the UDP-...
glycosyltransferase superfamily. These transferases conjugate small lipophilic compounds with various sugars (O’Reilly & Miller, 1989). EGT is secreted into the insect’s haemolymph and it is here that it catalyses the conjugation of ecdysteroids, the insect moulting hormones, with the sugar moiety donated from UDP-glucose or UDP-galactose (O’Reilly et al., 1992). This inactivates the moulting hormone and thereby prevents moulting and pupation (O’Reilly & Miller, 1989, 1991). Therefore, larvae usually die in the instar in which they were infected. Uninfected larvae cease feeding just before and during a moult. In contrast, larvae infected with a wild-type virus expressing EGT rarely moult and therefore continue to feed, getting larger and subsequently increasing the yield of virus. Infection of larvae by a recombinant Autographa californica MNPV lacking the gene results in reduced feeding and an increased speed of kill (O’Reilly & Miller, 1991; Flipsen et al., 1995).

The egt gene has been identified in many baculoviruses. The only completely sequenced genome that does not contain an egt gene is that of XcGV (Hayakawa et al., 1999). XcGV is a slow-killing virus and does not kill within the infected instar so the absence of an egt gene was not unexpected. AoGV is also a slow-killing virus; infected larvae always die late in the final instar, irrespective of the instar infected (S. L. Wormleaton & D. Winstanley, unpublished results). Therefore it was anticipated that AoGV, like XcGV, would lack a functional egt gene. However, an egt homologue has been found in the same relative location as CpGV egt and the presence of a functional egt gene in the AoGV genome is also reported.

In this study, the AoGV genome has been further characterized and compared to other GVs. A restriction endonuclease map has been produced and the 6 kbp region containing the egt gene sequenced and analysed. The gene order in the egt/gramalin region has been compared to sequenced regions in other GVs. A phylogenetic analysis of conserved genes contained within this region has also been performed to determine the relationship between different GVs at the genetic level compared to their biological properties.

Methods

**Virus.** The AoGV-E isolate was recovered from overwintering A. orana larvae found in Kent, England in 1993 and was propagated in laboratory stocks of A. orana larvae maintained on artificial diet (Guennelon et al., 1981). A cloned genotype, AoGV-E1, was obtained by three successive rounds of in vivo cloning, using the limiting dilution method described by Smith & Crook (1988a). The virus occlusion bodies were purified using sucrose and glycerol gradients as described previously (Crook & Payne, 1980) and the DNA was extracted and purified as described previously (Smith & Crook, 1988b).

**Restriction enzyme analysis and hybridization studies.** DNA was digested with restriction enzymes and electrophoresed in 0.7% (w/v) agarose gels using standard procedures (Sambrook et al., 1989). Restriction endonuclease (RE) fragment sizes were calculated using the program given by Schaffer & Sederoff (1981). For Southern blotting, DNA was capillary transferred onto nylon membranes (Roche Diagnostics) (Southern, 1975). Probes were labelled, by the random-primed labelling method, with digoxigenin-11-dUTP (DIG) using a DIG labelling kit (Roche Diagnostics). Hybridizations and washes were carried out under high stringency conditions at 65 °C according to the manufacturer’s protocol.

**Construction of genomic DNA libraries.** Libraries of AoGV DNA fragments were constructed by ligating restricted viral DNA fragments into the pHBluescript II SK(+) plasmid vector (pBSK(+)) (Stratagene) using T4 DNA ligase (Life Technologies). Recombinant plasmids were cloned and propagated in Escherichia coli DH5α and purified by alkaline lysis.

**Sequencing.** The nucleotide sequence of double-stranded DNA fragments of the AoGV 6 kbp region was determined utilizing the dideoxynucleotide chain termination method (Sanger et al., 1977). Sequencing reactions using ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer Applied Biosystems) were set up according to the manufacturer’s recommendations. The sequencing reactions were carried out using the GeneAmp PCR Systems 9600 and analysed on an ABI 377 automated DNA Sequencer (Applied Biosystems). Universal primer forward and reverse primers and custom primers were used for sequencing.

Double-stranded DNA sequences were assembled using the SeqMan II sequence analysis package (Lasergene software version 4.03; DNAStar). The coding regions were predicted using the package GeneQuest II (DNAStar) by locating translation start and stop codons of open reading frames (ORFs) of 50 or more amino acids (aa). Database searches using the program BLASTP were used to identify proteins sharing similarity. Percentage pairwise identities were calculated using the GAP program of Wisconsin Package version 10.0, Genetics Computer Group (GCC), Madison, WI, USA (Devereux et al., 1984) with default settings. Multiple sequence alignments of protein ORFs were prepared using the Clustal W program version 1.81 (Thompson et al., 1994). There were no poorly aligned regions that needed re-alignment. Protein alignments were analysed by using the PHYLYP (Phylogeny Inference Package) version 3.5c suite of programs (Felsenstein, 1989). PROTDIST was used to produce a distance matrix of each sequence using Kimura’s distance formula, which is based on the PAM matrices of Margaret Dayhoff for aa substitutions. This information was used to generate a distance tree using the program NEIGHBOR. Parsimony analysis (PROTPARS) was used to produce a phylogenetic tree. The multiple data sets were re-sampled 100 times (SEQBOOT) and the random input order of sequences was jumbled 10 times (PROTPARS). The consensus tree was obtained (CONSENSE), and bootstrap values presented from this analysis. Trees were drawn using the package TreeView (Page, 1996).

**EGT enzyme assay.** The EGT assay was done according to O’Reilly & Miller (1989). Haemolymph was collected from A. orana larvae that had been inoculated with an LD 50 dose of AoGV as fourth instar. The collection time-points were 5, 10 and 15 days post-inoculation (p.i.). The haemolymph was transferred immediately into an Eppendorf tube, containing a crystal of phenylthiocarbamide, on ice. Haemolymph from uninfected fourth instar larvae collected at 5 days p.i. was used as a negative control. At all the haemolymph collection time-points the larvae had reached fifth instar. A lysed sample of cells from AcMNPV-infected Sf9 cells was used as a positive control. These samples were incubated in a 10 mM Tris–malate buffer with a mixture of UDP-glucose, UDP-galactose and (4H)ecdysone for 1 h. Ecdysone is the preferred ecdysteroid for AcMNPV EGT (Evans & O’Reilly, 1990) and was therefore used in
this assay. The reaction products were separated by thin-layer chromatography and the silica gel plates were exposed to $^{3}$H-sensitive phosphoimager screens and read with a Fujifilm BAS-1500 phosphoimager. The computing program used to visualize the products was TINA 2.0.

Results and Discussion

Cloned strains of AoGV-E1

The original field isolate of AoGV-E appeared homogeneous, having no submolar bands in REN fragment profiles using 12 different restriction enzymes. Even so, the virus was cloned three times in vivo to select a single genotype, AoGV-E1. No other genotypes were observed throughout the cloning and the original profiles for AoGV-E were the same as for AoGV-E1.

Physical map of the AoGV genome

Digests of the AoGV genome with the restriction enzymes BamHI, BglII, EcoRI, PstI and SacI resulted in a total of 51 fragments (Fig. 1). The fragments were designated alphabetically starting with A for the largest fragment for each REN digest, as proposed by Vlak & Smith (1982). The locations of the restriction sites were determined using REN double digests and hybridization data. In a few cases, where ambiguity remained, partial digests were performed on cloned fragments of the AoGV genome to provide additional data. The zero point for the map, using the convention of Vlak & Smith (1982), was the smallest fragment containing the start codon of the granulin gene, which was the BamHI-G fragment of 362 bp (Fig. 2). This was confirmed by hybridization to the CpGV granulin gene and sequencing of the AoGV granulin region. The size of the AoGV genome was calculated to be 100–9 kbp, based on the physical mapping; one of the smallest baculovirus genomes found to date.

DNA sequence of the granulin gene region

The cloned restriction fragments EcoRI-D, EcoRI-T and BamHI-F, which spanned the egt/granulin region, were sequenced in both directions by primer walking. A PCR product was sequenced to confirm the junction of EcoRI-D and EcoRI-T. The ORFs found within this region were named using the convention of Crook et al. (1997). Each ORF was named according to the map position of its start codon and its direction of transcription, assuming that the zero point on the genome is the start codon of the granulin gene (Table 1).

Twelve putative ORFs of 50 aa or greater with start and stop codons were predicted; four of these were contained within larger ORFs. In these cases, the larger ORFs were assumed to be the most likely to be transcribed and translated. All of the proteins have homologues in other GV genomes, which supports the computer predictions. The ORFs were

---

Fig. 1. Restriction enzyme profiles of AoGV-E1 DNA for BamHI, BglII, EcoRI, PstI and SacI. λ DNA digested with HindIII is included for molecular size standards.
aligned contiguously with small intergenic regions (up to 72 bp) or small overlaps (up to 17 bp). The potential promoters for each ORF are indicated in Table 1. Potential promoters were identified in the 120 nucleotides upstream of the start codon for each ORF. These were identified as the late promoters using the consensus (A/T/G)TAAG and early promoters using the consensus structure of a TATA box [TATA(A/T)(A/A)] and an mRNA start site, CA(G/T), 25–35 nucleotides downstream.

ORF 99L as an EGT homologue

One ORF of particular interest is the homologue to the ecdysteroid UDP-glucosyltransferase (egt) gene of AcMNPV. The AoGV egt gene we have sequenced has two potential late promoters; one at position −18 (ATAAG) and one at position −56 (TTAAG) relative to the A of the start codon. It also has a potential early promoter, a TATA box at position −45 and a possible transcription start site 35 nucleotides downstream at position −10 (CATA). In infected Lycanobia oleracea larvae, the LoGV egt is transcribed predominantly as a late transcript (Smith & Goodale, 1998). This is in contrast to NPV egt genes studied to date, which have been shown to be transcribed from early promoters. AcMNPV EGT is secreted and has an N-terminal signal sequence that is cleaved from the mature enzyme (O’Reilly et al., 1992). All baculovirus EGT proteins have been found to possess hydrophobic sequences resembling signal peptides at their N termini. Likewise, AoGV EGT has 9 hydrophobic residues in its 15 aa N-terminal signal sequence. Like other baculovirus EGT sequences, AoGV also lacks the polar sequence found at the C terminus of mammalian UDP-glucuronosyltransferases, which acts as a membrane anchor.

Ten conserved regions (I–X) have been identified among EGT proteins (Hu et al., 1997). AoGV EGT lacks domain X as do the EGT proteins of LoGV, PxGV and CpGV, the other GV EGT proteins sequenced to date (Smith & Goodale, 1998; Hashimoto et al., 2000; D. Winstanley & N. E. Crook, unpublished results). There are seven conserved aa for all UDP-glucosyltransferases (Hu et al., 1997). These were all present in AoGV EGT. The conserved region II which is present in other baculoviruses has 9 aa absent in AoGV EGT and 7 aa absent in PxGV (Hashimoto et al., 2000). The significance of this is not yet clear.

A phylogenetic analysis has been performed on the EGT protein of AoGV and is shown in Fig. 3(a–b). In the analysis, the GVs and NPVs group separately as expected. Interestingly, AoGV and PxGV, and CpGV and LoGV group together in the parsimony analysis (Fig. 3a). The AoGV protein sequence is most similar (42% aa identity) to the EGT protein of CpGV, which is another tortricid-infecting virus. However, from the distance analysis it can be seen that the AoGV EGT sequence has similar aa identities to all of the GV EGT sequences (38–42% aa identity). More GV EGT sequences are required before firm conclusions regarding the relationship between GV EGT proteins can be drawn.

Assay of EGT activity to determine whether AoGV produces a functional EGT protein

Owing to the slow speed of kill of AoGV and the ability of infected larvae to undergo larval moults, it was not known if
AoGV produced a functional EGT protein. Therefore it was necessary to undertake an EGT assay to address this. The phosphoimage of the EGT assay is shown in Fig. 4. The results confirmed that AoGV was producing an active EGT in infected fifth instar larvae. No conjugation of glucose to [3H]ecdysone was observed in the negative control (haemolymph from mock infected A. orana larvae). Conjugation was also observed in the case of the positive control (AcMNPV-infected Sf9 cell extract) and in haemolymph from A. orana larvae infected with AoGV. Previously, larvae infected with a virus expressing EGT have been found to die within the instar that they were infected, or within the following instar. Therefore, the discovery of an egt gene in AoGV and the fact that AoGV expressed an active EGT were unexpected. This virus–host interaction in relation to EGT seems to be novel. The ability of larvae infected with AoGV to continue moulting post-infection may be because EGT does not accumulate to the threshold required to prevent a moult. This may be due to the narrow tissue tropism of infection and hence a lower output of EGT. It has been reported that a critical amount of ecdysone needs to be glucosylated to prevent a moult (O’Reilly et al., 1998). Very few external symptoms were observed in the infected larvae until the fifth instar, suggesting that the amount of virus replication and, as a result, the EGT production, is low. Once the larvae had reached fifth instar, the infection appeared to proceed rapidly and external symptoms were observed. The infected larvae failed to pupate when the control larvae pupated. This suggests that the EGT levels had risen and were able prevent the moult. The larvae remained in a prolonged fifth instar for several days, during which time many discharged large amounts of virus from their posterior end. This discharge may be a mechanism for disseminating virus to infect further larvae before they die.

**ORF OR as a granulin homologue**

ORF OR showed high aa identity (82–96%) to other granulin genes. It consisted of a 248 aa protein. A late promoter motif (ATAAG) was located within the DNA at position −27 relative to the A of the start codon. Database searches showed that AoGV ORF OR had the highest aa sequence identity to the granulin gene of CpGV (96–4%). A phylogenetic tree was constructed using the aa sequences of granulins of several GVs (Fig. 3c–d). The phylogenetic tree indicated that AoGV was most closely related to CpGV and the other tortricid-infecting GVs, Cryptophlebia leucotreta GV (CIGV), Choristoneura fumiferana GV (CIGV) and Epinotia aaporema GV (EpapGV), as well as GVs that infect other families, such as HbGV (which infects larvae of Zygaenidae) and Phthorimaea operculella GV (PhopGV) (which infects larvae of Gelechiidae). These constituted a clade that is clearly different from the noctuid-infecting GVs, supported by parsimony analysis. The relationship within the AoGV-containing clade is less well defined at the parsimony level. The
GVs from noctuids grouped together into two clades, namely *Spodoptera littoralis* GV (SpliGV), XcGV and *Trichoplusia ni* GV (TnGV) as one clade including PxGV (which infects larvae of Yponomeutidae), and LoGV and *Agrotis segetum* GV (AsGV) as a separate clade within the noctuids.

**ORF 99R as an ME53 homologue**

ORF 99R of AoGV showed relatively high similarity (48% identity at the aa level) to ORF 143 of CpGV. The function of this gene is unknown, but similar genes are present in other...
baculoviruses. Sequence analysis of the AoGV protein suggested two zinc-finger motifs, one at the N terminus (C-X$_{3}$-C-X$_{3}$-C; position 36) and one at the C terminus (C-X$_{2}$-C-X$_{14}$-C$_{2}$-C; position 229). These zinc-fingers are also conserved in other granulovirus homologues. The presence of zinc-fingers consisting only of clustered cysteine residues, as
promoter elements (Crook et al., 1994) and CpGV contains early but not late homologue also contains early and late promoter elements.

E. coli. To measure EGT activity, the radioactive EGT product was separated by TLC and quantified by Phosphoimage analysis.

Fig. 4. EGT activity in AoGV-infected A. orana haemolymph. Phosphoimage showing the separation of EGT reaction products on a TLC plate. E, free [3H]ecdysone; CE, [3H]ecdysone–sugar conjugates; P, a sample from AcMNPV-infected SF9 cells; M, mock-infected A. orana larvae; 5, 10 and 15 denote days p.i. of tested samples.

opposed to cysteine and histidine residues, has been observed in the steroid hormone nuclear receptor family of proteins; they are thought to be essential for sequence-specific recognition of DNA (Freedman et al., 1988). AoGV ORF 99R showed low aa identity to AcMNPV ME53 (22%). However, the presence of two cysteine zinc-fingers in AcMNPV ME53 indicates a possible functional similarity. Whether the putative zinc-fingers in ME53 play any functional role has yet to be determined. The AcMNPV me53 gene has been identified as an early gene in a transient expression assay (Knebel-Mörsdorf et al., 1993). However, it also contained a late promoter element. Other GV me53 homologues show little similarity in the upstream promoter region of this gene. CIGV ORF 891 has been shown to be transcribed from an early promoter but also contains a late promoter (Bah et al., 1999). The CIGV homologue also contains early and late promoter elements (Jehle & Backhaus, 1994) and CpGV contains early but not late promoter elements (Crook et al., 1997). AoGV has a potential early transcription start site that differs from the consensus. A cap site (CACT) is present at position −40 relative to the A of the start codon, and a TAATAT sequence 29 bp upstream at position −69 may act as the promoter rather than the normal TATA. There is another possible transcription start site (CATA) at position −135 and a TATA box 23 bp upstream at position −158, although these are unusually distant from the start codon. In addition, there is also a late promoter motif (ATAAG) which is also distant from the start codon at position −184. A phylogenetic analysis of ORF 99R is shown in Fig. 3(e–f). These analyses show a clear division between the GV and NPV proteins. There are two strongly supported clades within the GVs. One clade comprised ME53 homologues from noctuid and yponomeutid-infecting GVs and the other comprised GVs that infect tortricids, including AoGV. Parsimony analysis supports this division.

**AoGV ORF 1R as a protein kinase homologue**

A further gene found in the sequenced region of AoGV and other baculoviruses encodes a putative protein kinase. All protein kinase catalytic domains have been found to contain 11 conserved subdomains (Hanks et al., 1988). These subdomains include motifs for serine/threonine protein kinases and for ATP-binding (Hanks et al., 1988). AoGV contains 10 of the subdomains with 9 matching the consensus exactly. Consensus motif I is known to be involved in ATP-binding (Hanks et al., 1988). This region differed from the consensus in AoGV. However, inconsistencies are shown in this region in PxGV and XcGV (Hayakawa et al., 1999) and in protein kinases from other organisms (Russell & Nurse, 1987; Baylis et al., 1993). The consensus motif X is absent in AoGV but is also poorly conserved in other baculovirus proteins. It was found to be the least conserved of the 11 subdomains in other protein kinases (Hanks et al., 1988).

Phylogenetic analysis of protein kinase genes is shown in Fig. 3(g–h). In the analysis, the GVs and the NPVs group separately as expected. Within the GVs, the tortricid-infecting GVs, including AoGV, form a strongly supported clade as do the noctuid and yponomeutid-infecting GVs. There are differences in the temporal control of NPV protein kinases studied to date. *Lymantria dispar* (Ld)MNPV protein kinase is transcribed early and late in infection although a TATA box is not present (Bischoff & Slavicek, 1994). The AcMNPV protein kinase is expressed late in infection only (Reilly & Guarino, 1994), as is the *Bombyx mori* (Bm)NPV protein kinase (Zhang et al., 1998). The AoGV protein kinase contains just one late promoter motif at position −72 relative to the A of the start codon of the gene.

**AoGV ORF 1L**

Only one baculovirus match for ORF 1L was found on the BLASTP database search: this was XcGV ORF 2, also located directly downstream of granulin. The first ORF immediately downstream from granulin in AoGV, CpGV, PxGV, LoGV, XcGV, SpiiGV and PhopGV varies in length from 106 aa in AoGV to 231 aa in XcGV. However, all show identity in the first 65 aa of the N-terminal region of the protein and all have a late promoter consensus motif. All of these ORFs are transcribed on the opposite strand to granulin and most either terminate very close to the end of the granulin sequence or overlap it slightly.

XcGV ORF 2 shows 22% identity to AcMNPV 1629 capsid protein (Hayakawa et al., 1999), which is also located between the polyhedrin gene and protein kinase gene (Possee et al., 1991). AcMNPV 1629 is expressed late in infection and is essential for AcMNPV viability (Possee et al., 1991). The protein is thought to be associated with the basal structure of the capsid (Russell et al., 1997). AoGV ORF 1L does not appear to show similarity with AcMNPV 1629 capsid or any other NPV 1629 capsid homologue although it does contain two late
Analysis of the AoGV egt region

Fig. 5. Comparison of the gene structure and orientation of the granulin-containing region of various GVs. The arrows represent the length and direction of the ORFs over 50 aa. The ORFs are evenly spaced for clarity. Ld151 (similarity to LdMNPV ORF 151), Xc2 and Xc20 (similarity to XcGV ORF 2 and ORF 20 respectively). Black arrows indicate ORFs with no baculovirus homologue.

promoter consensus motifs ATAAG at nucleotide positions −79 and −89 relative to the A of the initiation codon. The function of this ORF is unknown at present but may be conserved since all GVs in this study contained this ORF.

ORFs 96R, 2L and 2R

These ORFs all have homologues in CpGV, PxGV and XcGV. These appear to be GV-specific genes, as NPV homologues have not been identified to date. However, ORF 96R of AoGV showed 44% aa identity with 47 aa of the AcMNPV fibroblast growth factor (fgf). However, a second fgf homologue with higher identity to AcMNPV fgf (ORF 32) has been found elsewhere in the AoGV genome. This is more likely to be the true homologue (S. L. Wormleaton & D. Winstanley, unpublished results).

ORF 2L contains a typical early promoter and the homologues in CpGV and PxGV also have an early promoter. The XcGV homologue does not have a consensus early promoter, but could have a potential one, as it has the sequence TATA at position −74 and a CATT potential mRNA site 29 bp downstream at position −45. Therefore, this ORF could be an early gene that is unique to GVs.

Comparative analysis of baculovirus genomes

The phylogenetic analyses of the genes around the egt/granulin gene region were performed to determine where AoGV would lie in relation to other GVs. It was unclear whether it would be more related to GVs with small genomes, slow-killing GVs with similar tissue tropism or GVs that infect hosts of the same family. AoGV is slow to kill, infects a tortricid host and has a small genome. In contrast, other slow-killing GVs such as XcGV usually infect noctuid hosts and have genomes almost twice the size. Based on phylogenetic analyses and aa identity, the AoGV genome showed most similarity to CpGV, which is a fast-killing GV.

The most sequenced region of GVs is the granulin-containing region. The order of the genes may give an indication of the evolutionary differences between different GVs. An alignment of the sequenced granulin-containing regions in GVs is shown in Fig. 5. The gene organization in the granulin region of AoGV is similar to other tortricid-infecting viruses (CpGV, CIGV and CfGV) and PhopGV, which infects a pest of the family Gelechiidae (Taha et al., 2000). In contrast, the organization of this region in noctuid-infecting viruses (LoGV, SpilGV and XcGV) and in PxGV showed more diversity. Firstly, the noctuid-infecting viruses and PxGV do not have an me53 homologue directly upstream of the granulin gene. In XcGV, it was one ORF further upstream (Hashimoto et al., 2000). An me53 homologue was not identified within the granulin area of LoGV (I. R. L. Smith, personal communication), whereas in PxGV it was four ORFs further upstream (Hayakawa et al., 1999). Secondly, PxGV was found to have a possible p10 homologue within the granulin-containing region.
(Hashimoto et al., 2000). Thirdly, the egt gene, which is the second or third ORF upstream of the granulin gene in AoGV and other tortricid-infecting viruses, lies 8 kbp upstream of granulin in LoGV. The egt gene of PxGV is six ORFs upstream of its granulin. LoGV also contained an ORF with similarity to XcGV ORF 20 and an iap homologue within the granulin region (I. R. L. Smith, personal communication).

The only genes found in conserved positions, in relation to the granulin gene, in nearly all the GVs sequenced to date were ORF 1L and the protein kinase gene, located one and two ORFs downstream from the granulin, respectively. In LoGV there is an extra small ORF (Xc20) downstream of granulin which moves ORF 1L and protein kinase one ORF downstream (I. R. L. Smith, personal communication). The protein kinase genes in AcMNPV (Ayres et al., 1994), BmNPV (Gomi et al., 1999), LdMNPV (Kuzio et al., 1999), Orgyia pseudotsugata (Op)MNPV (Ahrens et al., 1997) and Spodoptera exigua (Se)MNPV (Ijkel et al., 1999) are also located two ORFs downstream of the polyhedrin (granulin-equivalent) gene and are transcribed in the same direction.

The gene arrangement around the granulin area of GVs shows that conservation of certain genes has occurred. Usually the same set of genes is present, with slight alterations or additions specific to each virus. At present, the gene arrangement in tortricid-infecting GVs appears to be more conserved than in the case of GVs that infect other families. The latter appear to have extra small ORFs that do not appear to have similarity to ORFs in other viruses.

The results from these GV gene comparisons imply that the original classification of GVs based on tissue tropism is invalid at the genetic level. The speed of kill and tissue tropism of GVs do not appear to be dependent on genome size. For example, the slow-killing GVs, AoGV and XcGV, have genome sizes of 100–9 kbp and 178–7 kbp, respectively. The speed of kill and tissue tropism are also not dependent on the presence of an egt gene. For example, AoGV contains a functional egt gene whereas XcGV does not have an egt gene. The speed of kill and tissue tropism are also not dependent on the host family of the virus. For example, AoGV and CpGV are both tortricid pests. However, AoGV is slow-killing with infection restricted to the fat body and CpGV is fast-killing with most tissues becoming infected. In the case of genes such as granulin, me35 and pk-1, the GVs seem more closely related based on their host family. AoGV tends to group with other tortricid-infecting GVs and XcGV tends to group with other nocticid-infecting GVs. It is therefore possible that AoGV may have evolved from an ancestral virus similar to the tortricid-infecting GVs.

It appears from the phylogenetic analysis of ORFs in the egt/granulin region and the gene arrangements that GVs are more closely related depending on the family of Lepidoptera they infect rather than their speed of kill or the tissues they infect within their target species. However, there is a need for comparison of further sequences before firm conclusions can be drawn.

The authors would like to thank the following: Dr I. R. L. Smith (Central Science Laboratory, York) for providing the unpublished sequence of the granulin gene region of LoGV; Dr M. López-Ferber (Laboratoire de Pathologie Comparée, France) for the unpublished granulin sequences of PhopGV and SpliGV; Dr V. Romanowski (IBBM, Argentina) for the unpublished granulin sequence of EpapGV and Dr D. R. O’Reilly and Dr T. Luque for their assistance with the EGT assay, which was performed at Imperial College, London. This work was supported by funding from BBSCC.

References


Received 15 February 2001; Accepted 1 June 2001