Journal of General Virology (2006), **87**, 3233–3250 DOI 10.1099/vir.0.82161-0

Genome of the most widely used viral biopesticide: *Anticarsia gemmatalis* multiple nucleopolyhedrovirus

Juliana Velasco de Castro Oliveira,[1] José Luiz Caldas Wolff,[2] Alejandra Garcia-Maruniak,[3] Bergmann Morais Ribeiro,[4] Maria Elita Batista de Castro,[5] Marlinda Lobo de Souza,[5] Flavio Moscardi,[6] James Edward Maruniak[3] and Paolo Marinho de Andrade Zanotto[1]

Correspondence
Paolo Marinho de Andrade Zanotto
pzanotto@usp.br

[1]Laboratório de Evolução Molecular e Bionformática, Departamento de Microbiologia, Instituto de Ciências Biomédicas, Universidade de São Paulo, SP, Brazil

[2]Laboratório de Virologia Molecular, Núcleo Integrado de Biotecnologia, Universidade de Mogi das Cruzes, Mogi das Cruzes, SP, Brazil

[3]Entomology and Nematology Department, PO Box 110620, University of Florida, Gainesville, FL 32611-0620, USA

[4]Departamento de Biologia Celular, Universidade de Brasília, Brasília, DF, Brazil

[5]Embrapa Recursos Genéticos e Biotecnologia-Núcleo Temático de Controle Biológico (NTCB), Brasília, DF, Brazil

[6]EMBRAPA-CNPSo, Londrina, PR, Brazil

Received 25 April 2006
Accepted 26 June 2006

The genome of *Anticarsia gemmatalis* multiple nucleopolyhedrovirus isolate 2D (AgMNPV-2D), which is the most extensively used virus pesticide in the world, was completely sequenced and shown to have 132 239 bp (G+C content 44·5 mol%) and to be capable of encoding 152 non-overlapping open reading frames (ORFs). Three ORFs were unique to AgMNPV-2D, one of which (ag31) had similarity to eukaryotic poly(ADP-ribose) polymerases. The lack of *chiA* and *v-cath* may explain some of the success and growth of the AgMNPV biological control programme, as it may explain the high recovery of polyhedra sequestered inside dead larvae in the field, which are collected and used for further application as biological pesticides in soybean fields. The genome organization was similar to that of the *Choristoneura fumiferana* defective MNPV (CfDefNPV). Most of the variation between the two genomes took place near highly repetitive regions, which were also closely associated with *bro*-coding regions. The separation of the NPVs into groups I and II was supported by: (i) a phenogram of the complete genomes of 28 baculovirus and *Heliothis zea* virus 1, (ii) the most parsimonious reconstruction of gene content along the phenograms and (iii) comparisons of genomic features. Moreover, these data also reinforced the notion that group I of the NPVs can be split further into the AgMNPV lineage (AgMNPV, CfDefNPV, *Epiphyas postvittana* NPV, *Orgyia pseudotsugata* MNPV and *C. fumiferana* MNPV), sharing eight defining genes, and the *Autographa californica* MNPV (AcMNPV) lineage (AcMNPV, *Rachiplusia ou* NPV and *Bombyx mori* NPV), sharing nine defining genes.

INTRODUCTION

Baculoviruses are arthropod-specific viruses (Tanada & Kaya, 1993) that have been used intensively as expression vectors (King & Possee, 1992; Kost *et al.*, 2005), as models of genetic regulatory systems (Lu & Miller, 1997; Miller, 1999), as biological-control agents against insect pests (Podgewaite, 1985; Moscardi, 1999) and as potential non-human viral DNA vectors for gene delivery (Tani *et al.*, 2003). Since the first baculovirus genome sequence from *Autographa californica* multiple nucleopolyhedrovirus (AcMNPV) was published (Ayres *et al.*, 1994), the number of complete genomes available has risen to 30. The use of the *Anticarsia gemmatalis* MNPV (AgMNPV) in the control of velvetbean caterpillar, *A. gemmatalis*, has been the most successful example of a virus used as a biological pesticide (Moscardi, 1999). In Brazil, it is currently applied over

The GenBank/EMBL/DDBJ accession number for the sequence determined in this work is DQ813662.

Downloaded from www.microbiologyresearch.org by
IP: 200.202.165.220
On: Tue, 22 May 2018 19:05:29

approximately 2 million hectares per year, providing effective control of larvae of the key crop defoliator of soybean fields. Since the initial isolation of AgMNPV from infected larvae in Brazil (Allen & Knell, 1977; Carner & Turnipseed, 1977), its production, commercialization and field application have increased constantly (Moscardi, 1999). Comparisons among AgMNPV temporal isolates have shown that the viral heterogeneity of the commercial preparation has increased and that changes are concentrated at 'hot spots' (Maruniak *et al.*, 1999). AgMNPV isolate 2D (AgMNPV-2D) was cloned by plaque purification and chosen as the prototype (Maruniak, 1989; Maruniak *et al.*, 1999) as it was the major genotype present among a wild-type population in 1972 (Allen & Knell, 1977). The UFL-AG-286 cell line (Sieburth & Maruniak, 1988a, b), established from *A. gemmatalis*, produces high titres of AgMNPV and is suitable for *in vitro* experiments with AgMNPV (Pombo *et al.*, 1998; Castro *et al.*, 2006). Moreover, genetic modification of AgMNPV has allowed studies of its pathology in *A. gemmatalis* larvae (Soares & Ribeiro, 2005), leading to the characterization of many *A. gemmatalis* larvae haemocytes, all of which have been shown to be susceptible to AgMNPV infection (da Silveira *et al.*, 2003, 2004).

Several individual genes of AgMNPV-2D have been sequenced and analysed: *polh* (polyhedrin) (Zanotto *et al.*, 1992), *gp41* (Liu & Maruniak, 1999), *egt* (ecdysteroid UDP-glucosyltransferase) (Rodrigues *et al.*, 2001), *p10* (Razuck *et al.*, 2002), *gp64* (Pilloff *et al.*, 2003; Slack *et al.*, 2004), *v-trex* (viral 3′ repair exonuclease) (Slack & Shapiro, 2004), *p143* (helicase) (Lima *et al.*, 2004), *dnapol* (DNA polymerase) (Dalmolin *et al.*, 2005), *iap-3* (Carpes *et al.*, 2005) and *p74* (Belaich *et al.*, 2006). In addition, a homologous region (*hr4*) of AgMNPV-2D has been characterized (Garcia-Maruniak *et al.*, 1996). Phylogenetic analysis of *polh* indicates that AgMNPV belongs in the group I NPVs (Zanotto *et al.*, 1993) and comparisons of some genes have shown that AgMNPV is closely related to *Choristoneura fumiferana* defective MNPV (CfDefNPV) (Lima *et al.*, 2004; Lauzon *et al.*, 2005). AgMNPV-2D has been modified genetically and is a viable expression system for heterologous genes (Arana *et al.*, 2001; Ribeiro *et al.*, 2001). The need to sequence the complete genome of AgMNPV was justified, given the increase in basic research on this virus and its increasing importance as a biological control agent. In addition, AgMNPV is probably one of the most important baculovirus systems under study today, as it is one of the few viruses that will allow integrating studies from large-scale field application to the genomic and post-genomic levels. In this work, the complete genome sequence of AgMNPV-2D and its genetic organization are presented.

METHODS

Virus and insects. AgMNPV-2D (Johnson & Maruniak, 1989) was grown in UFL-AG-286 cells. Viral DNA was purified from infected *A. gemmatalis* larvae injected with the viral isolate, as described previously (Garcia-Maruniak *et al.*, 2004).

AgMNPV-2D DNA library construction. Our sequencing strategy took advantage of the availability of the complete AgMNPV-2D *Hin*dIII genomic library (Johnson & Maruniak, 1989; Maruniak *et al.*, 1999). Subclones were obtained from *Hin*dIII clones with the four-cutter restriction enzymes *Hae*III, *Rsa*I and *Sau*3AI, cloned into pGEM-3Z (Promega). Clones were sequenced using primers binding to the T7 and SP6 promoter sequences. To create several sequencing start points, EZ::TN transposons (Epicentre) were inserted in the large *Hin*dIII A, C, D and E clones, as described by Garcia-Maruniak *et al.* (2004).

DNA sequencing. Internal sequences were obtained by: (i) primer walking using AgMNPV-2D-specific primers, (ii) further subcloning of smaller subfragments or (iii) by closing gaps or resolving ambiguities by PCR amplification of AgMNPV-2D DNA using specific primers. Sanger reactions were done by cycle sequencing using PTC-200 thermocycler machines (MJ Research) with an ABI PRISM Big Dye Terminator Sequencing Ready Reaction kit versions 2 and 3 (Applied Biosystems). Electrophoresis was carried out on ABI 3100 DNA sequencers.

DNA sequence assembly. Each nucleotide position was sequenced at least three times on each strand. Base calling (Q=26) was done using the PHRED program (Ewing & Green, 1998; Ewing *et al.*, 1998) and end clipping, cloning site, vector trimming and contig assembly were done using the ALIGNER program, version 1.3.4 (CodonCode Corp.).

Genome annotation. The complete genome was compared with other baculovirus genomes using the Artemis Comparative Tool release 1 (The Sanger Centre) (http://www.sanger.ac.uk/Software/ACT/) and then annotated with the Artemis release 5 program (http://www.sanger.ac.uk/Software/Artemis/). The Tandem Repeats Finder (http://tandem.bu.edu/trf/trf.html; Benson, 1999) and Dotter (Sonnhammer & Durbin, 1995) programs were used to locate homologous regions (*hr*s) and direct repeats (*dr*s). ORF Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) was used to find open reading frames (ORFs) of over 50 aa and their homology to sequences in GenBank was analysed using the link to BLASTP (Altschul *et al.*, 1997). Alignment of homologous amino acid sequences to highlight conserved regions was done using the conserved domain database (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml; Marchler-Bauer *et al.*, 2005). DNA and protein sequences were aligned using CLUSTAL X (Thompson *et al.*, 1997) with default settings.

Phylogenomics. In order to establish the relationship of AgMNPV-2D to 27 sequenced baculoviruses and *Heliothis zea* virus 1 (HzV-1), pairwise distance matrices were calculated with the BLASTPHEN pipeline. For each pair of genomes, bit scores, S' (Altschul *et al.*, 1990; Pertsemlidis & Fondon, 2001), were calculated for all local high-scoring pairs, which were collected to obtain a global genomic similarity either by central tendency statistics (median, mean and mode) or by comparing S' distributions. As S', being a measure of similarity, is obviously inversely proportional to evolutionary distance, we assumed $1/S'$ as a measure of evolutionary distance (k), and outputs were parsed with BLASTPHEN and the distributions of S' were obtained for each pair of genomes. These distributions were subjected to a range of data analyses prior to clustering and their efficacy in generating informative genome clusters was evaluated (P. M. A. Zanotto, M. A. C. Baccaro, R. N. Pereira & D. Krakauer, unpublished results). Dendrograms for complete baculovirus genomes were produced with ultrametric clustering algorithms (UPGMA, neighbour joining and their derivatives) (Li & Graur, 1991).

Gene content analysis. In order to study gene acquisition and loss during baculovirus evolution, a gene-content dataset expanded from Herniou *et al.* (2001, 2003) was used, including 665 genes

Downloaded from www.microbiologyresearch.org by
IP: 200.202.165.220
On: Tue, 22 May 2018 19:05:29

from all sequenced baculoviruses (except *Chrysodeixis chalcites* NPV, *Agrotis segetum* NPV and *Trichoplusia ni* SNPV). It also included a synthetic outgroup (i.e. a 'basal' root in which all genes are absent), in order to define gain and loss of genes in more detail among baculoviruses and other cellular organisms (both prokaryotes and eukaryotes) and DNA and RNA viruses.

RESULTS AND DISCUSSION

Nucleotide sequence analysis

The complete AgMNPV-2D genome comprised 132 239 bp, with a C+G content of 44·5 mol%. This falls within the mean C+G content of baculoviruses, which ranges from 32·4 mol% in *Cryptophlebia leucotreta* granulovirus (CrleGV) to 57·5 mol% in *Lymantria dispar* MNPV (LdMNPV) (Kuzio *et al.*, 1999; Lange & Jehle, 2003). Fig. 1 shows regions of the genome with higher than average G+C content, such as the largest non-coding region around 126 kbp, and regions that are distinctly AT-rich such as the *hr*s. AgMNPV-2D contained a total of 152 non-overlapping ORFs of ⩾50 aa, of which 81 (53 %) were oriented clockwise and 71 (47 %) anticlockwise (Fig. 1). The putative coding regions were numbered sequentially starting with the *polh* gene in the antisense direction as ORF1 (ag1). The 29 genes shared by all baculovirus genomes, including the dipteran and hymenopteran baculoviruses (Herniou *et al.*, 2003; Garcia-Maruniak *et al.*, 2004), were also found in the AgMNPV-2D genome. The phenogram shown in Fig. 2 summarizes the relationship of AgMNPV to 27 other complete genomes of baculoviruses and HzV-1. The topology of the genome phenogram was congruent with trees for individual genes (Zanotto *et al.*, 1993) or trees for concatemers of the 29 baculovirus conserved genes (Garcia-Maruniak *et al.*, 2004). Separation of the lepidopteran baculoviruses as NPVs of groups I and II and granuloviruses (GVs) was confirmed, as was the distant relationship that exists between the lepidopteran, hymenopteran and dipteran baculoviruses (Herniou *et al.*, 2001, 2003; Garcia-Maruniak *et al.*, 2004).

Gene content comparisons in group I NPVs

In order to analyse change in gene content in group I NPVs, we used the topology shown in Fig. 2 to infer the most parsimonious reconstructions of the presence or absence of each gene (represented as discrete unordered characters). This was done to investigate patterns of independent gain and loss of orthologous genes (defined as sharing significant BLAST search scores) along lineages of baculovirus and several viral and cellular outgroups. The pattern of gene acquisition in group I NPVs is shown in Fig. 3. Group I NPVs had 26 synapomorphies (i.e. defining shared ORFs). Arrows pointing up in Fig. 3 represent synapomorphies that underwent secondary loss in later descendants [e.g. ep98 lost in *Orgyia pseudotsugata* MNPV (OpMNPV)]. Genes shared elsewhere are represented as arrows pointing down (e.g. *gp64* is only present in group I NPVs and in Thogoto virus (*Orthomyxoviridae*) (Morse *et al.*, 1992). Seven defining genes represented as boxes were unique to group I NPVs (e.g. *odv-e26*). Moreover, group I NPVs had two lineages: (i) the AcMNPV lineage (defined by nine synapomorphies), including AcMNPV, *Rachiplusia ou* NPV (RoMNPV) and *Bombyx mori* NPV (BmNPV), and (ii) the AgMNPV lineage (defined by eight synapomorphies) in which the CfDefNPV belongs as a sister group of AgMNPV and which included OpMNPV, *Choristoneura fumiferana* MNPV (CfMNPV) and *Epiphyas postvittana* NPV (EppoNPV). The division of group I NPVs into two different lineages has been delineated in recent studies (Jehle *et al.*, 2006; Landais *et al.*, 2006). ORFs for genes having orthologues involved in known functions were found and are discussed below.

Genomic organization

Syntenic maps of AgMNPV-2D, CfDefNPV, OpMNPV and AcMNPV were generated with the Artemis Comparative Tool by plotting the high similarity scoring pairs along pairs of genomes that were obtained with TBLASTX (using default settings) using the encoded proteins in all six reading frames and indicated that their genomes were extensively collinear (Fig. 4). Nevertheless, inversions of large portions of the genomes were observed in a region between the *polh* (ag1) and *p47* (ag44) genes, shared by all four genomes. Two inversions differentiated AgMNPV-2D, CfDefNPV and OpMNPV from AcMNPV. The first inversion of 7·8 kbp was located between the first repeat region (comprising five copies of a 30 bp imperfect palindromic sequence that included residue 1 of the linearized genome) and the next two copies of a 30 bp imperfect palindromic sequence between nt 7747 and 7864 of the AcMNPV genome. The second, of 11·5 kbp, started near the *pkip* gene and ended at the beginning of the *p47* gene. This inversion did not appear to be associated with repetitive elements, but contained the highly diverse *sod* locus flanked by *hr*s. Insertion of the AgMNPV-2D unique ag31 (PARP; see below) took place in this region of low synteny. Another significant inversion, located between the previous two mentioned above, differentiated AgMNPV-2D and CfDefNPV from all of the other group I NPVs. It comprised a region of 9·6 kbp flanked by *hr*s similar to *hr1a* in AcMNPV. Most of the other differences in organization among the four genomes could be explained by insertions and deletions (Fig. 4). The majority of the regions lacking synteny (70 %) in comparison with the other group I NPVs were near *hr*s or *dr*s, supporting previous findings that DNA insertions and deletions occur in these regions, contributing to the plasticity of the viral population (Garcia-Maruniak *et al.*, 1996; Li *et al.*, 2002b).

Homologous regions

*Hr*s are made of redundant direct repeat sequences with imperfect palindromes. They may function as origins of DNA replication and as transcriptional enhancers (reviewed by Possee & Rohrmann, 1997). Nine *hr*s were found in the AgMNPV-2D genome, but four (*hr1*–*hr4*) adjacent to baculovirus repeat ORFs (*bro*) have already been mapped

Downloaded from www.microbiologyresearch.org by
IP: 200.202.165.220
On: Tue, 22 May 2018 19:05:29

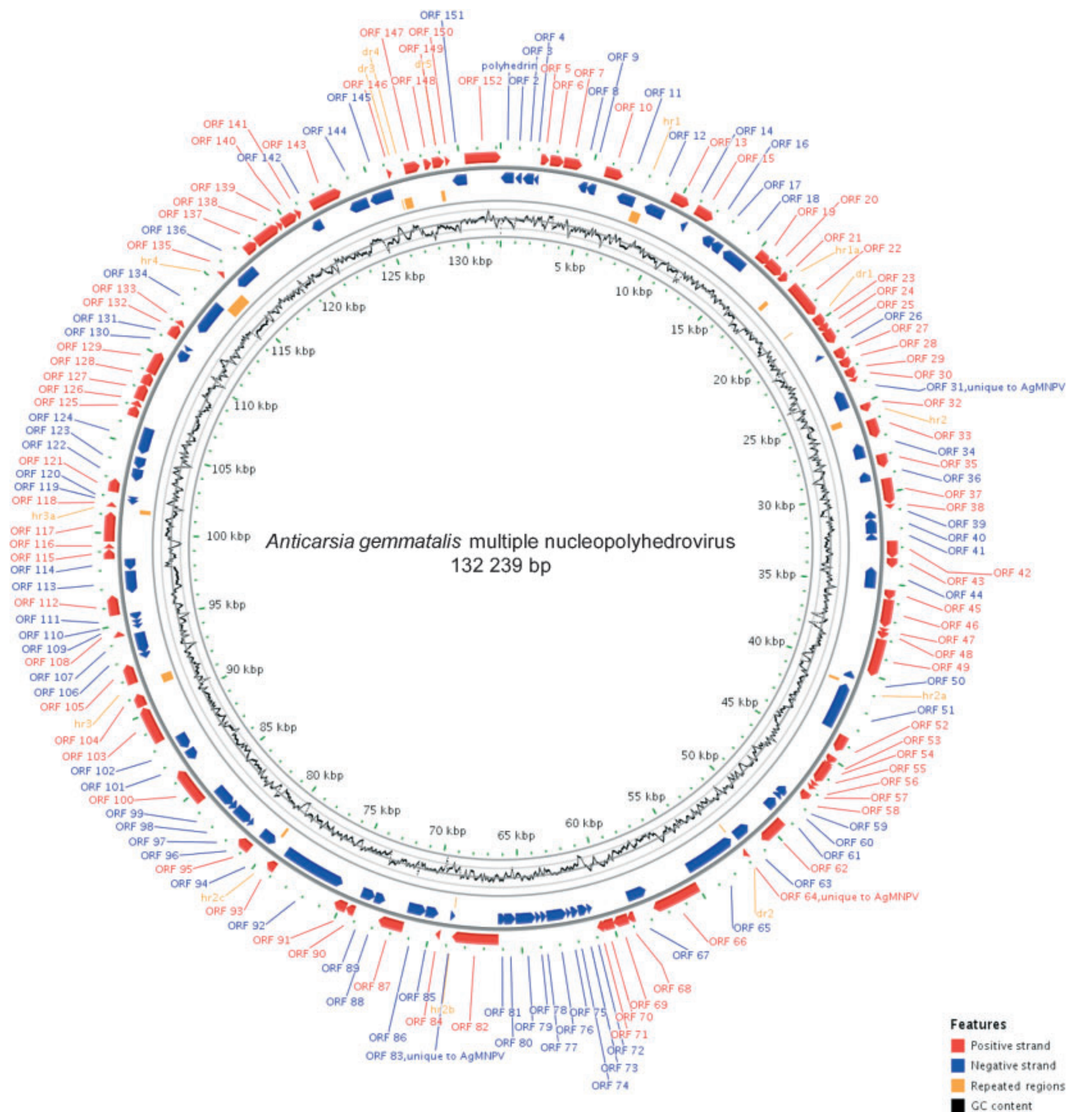


Fig. 1. Circular map of AgMNPV-2D. Arrows represent ORF positions and coding direction. Genomic features are shown concentrically. Repeat sequences (*dr*s and *hr*s) are indicated in the next circle as rectangles. The G+C content (mol%) plot of the genome is shown in the inner layer before the genome position in kbp. Peaks above the centre line correspond to regions that have a G+C content of >50 mol%. The figure was generated using the CGVIEW program (Stothard & Wishart, 2005).

(Garcia-Maruniak *et al.*, 1996). The consensus sequence had 92 nt for *hr1* (inverted in relation to the three other *hr*s) and *hr2*, 84 nt for *hr3* and 127 nt for *hr4* (Fig. 5a). The majority of the repeats were tandem repeats, except for *hr3*. A comparison of the 84 nt shared by the four consensus sequences (Fig. 5a) resulted in 82 % sequence identity, increasing to 90 % for the imperfect palindrome region comparison. Five other *hr*s (*hr1a*, *-2a*, *-2b*, *-2c* and *-3a*) were found that had imperfect palindrome sequences, but were not located in the core of tandemly repeated sequences, and this was similar to the structure of the *hr*s of CfDefNPV (Lauzon *et al.*, 2005). Three imperfect palindromes were present in *hr1a*, *-2a*, *-2c* and *-3a*, whilst only one copy was found in *hr2b* (Fig. 5b). The relative position of the AgMNPV-2D *hr*s matched the location of nine of the 13 *hr*s found in CfDefNPV. AgMNPV *hr2*, *hr2b*, *hr3a* and *hr4* were located downstream from the *sod*, *p95*, *pif* and *p74* genes, respectively, in conserved locations relative to other group I NPVs (Lauzon *et al.*, 2005).

The number of *hr*s found in baculoviruses ranges from three in CrleGV (Lange & Jehle, 2003) to 17 in *Spodoptera litura* NPV (SpltNPV) (Pang *et al.*, 2001). All group I NPVs have *hr*s, but six baculovirus genomes do not have typical *hr*s: *Cydia pomonella* GV (CpGV) (Luque *et al.*, 2001), *Adoxophyes orana* granulovirus (AdorGV) (Wormleaton *et al.*, 2003), *Agrotis segetum* NPV (AgseGV), *Chrysodeixis chalcites* NPV (ChchNPV) (van Oers *et al.*, 2005), *T. ni* single NPV

Downloaded from www.microbiologyresearch.org by
IP: 200.202.165.220
On: Tue, 22 May 2018 19:05:29

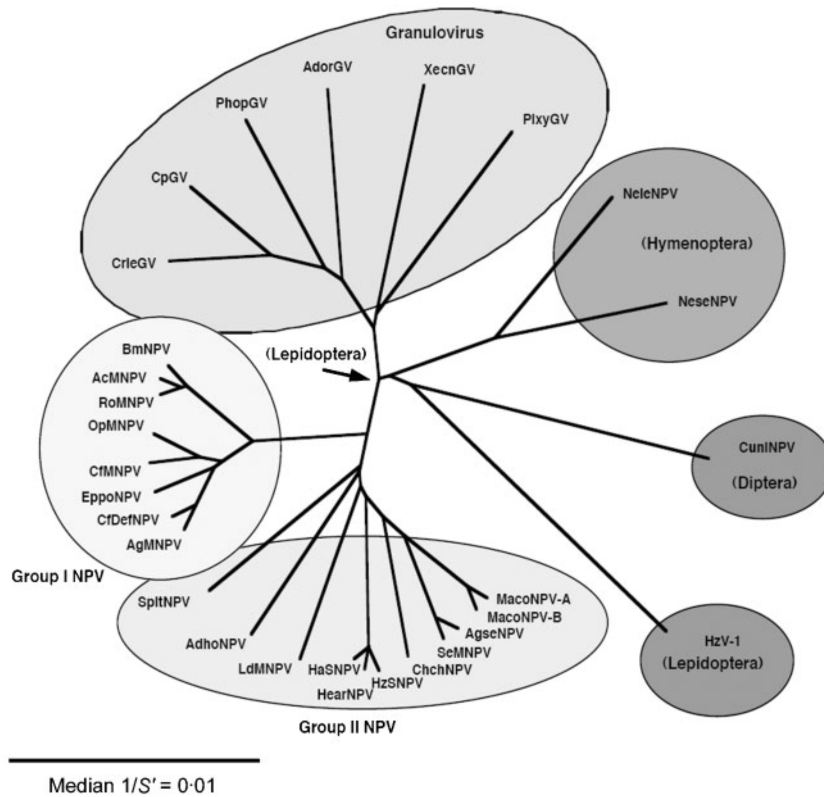


Fig. 2. Phenogram showing the relationship of AgMNPV-2D to 27 other complete baculovirus genomes and HzV-1. Distances in the phenogram were derived from matrices containing the reciprocal value of the median of all bit scores (S') from all high-scoring pairs for the pairwise comparison with TBLASTX of each genome calculated with BLASTPHEN. All major groups of baculovirus were recovered, including groups I and II of the NPVs. HaSNPV, *Helicoverpa armigera* SNPV; HearNPV, *H. armigera* NPV; HzSNPV, *Helicoverpa zea* SNPV; MacoNPV-A, *Mamestra configurata* NPV A; MacoNPV-B, *M. configurata* NPV B; NeseNPV, *Neodiprion sertifer* NPV; PlxyGV, *Plutella xylostella* GV; SeMNPV, *Spodoptera exigua* MNPV. Please refer to the text for all other abbreviations.

(Willis *et al.*, 2005) and *Neodiprion lecontei* NPV (NeleNPV) (Lauzon *et al.*, 2004). Moreover, five *dr*s were also found in the AgMNPV genome. The locations of the *hr*s and *dr*s are specified in Table 1.

NPVs have many bZIP (basic region/leucine zipper)-binding motifs at *hr*s, as has been shown in AcMNPV *hr5* (Landais *et al.*, 2006). They appear specifically to bind *cis*-regulatory host factors. AgMNPV-2D *hr*s had similarity to the 12-*O*-tetradecanoylphorbol-3-acetate (TPA) response elements (TREs), which belong to the family of bZIP motifs. The TRE motif TGA(C/G)TCA was present 18 times in AgMNPV-2D, nine (50 %) of which were 10–30 bases away from the palindromes in *hr1*, *hr2*, *hr2a*, *hr2c* and *hr3* (Fig. 5). A second type of bZIP motif, the cyclic AMP response element (Landais *et al.*, 2006), was not present in AgMNPV-2D *hr*s. According to the sole presence of TRE or the combination of both motifs, Landais *et al.* (2006) found two phylogenetic clusters within the group I NPVs, resembling the AcMNPV and AgMNPV lineages presented here (Figs 2 and 3).

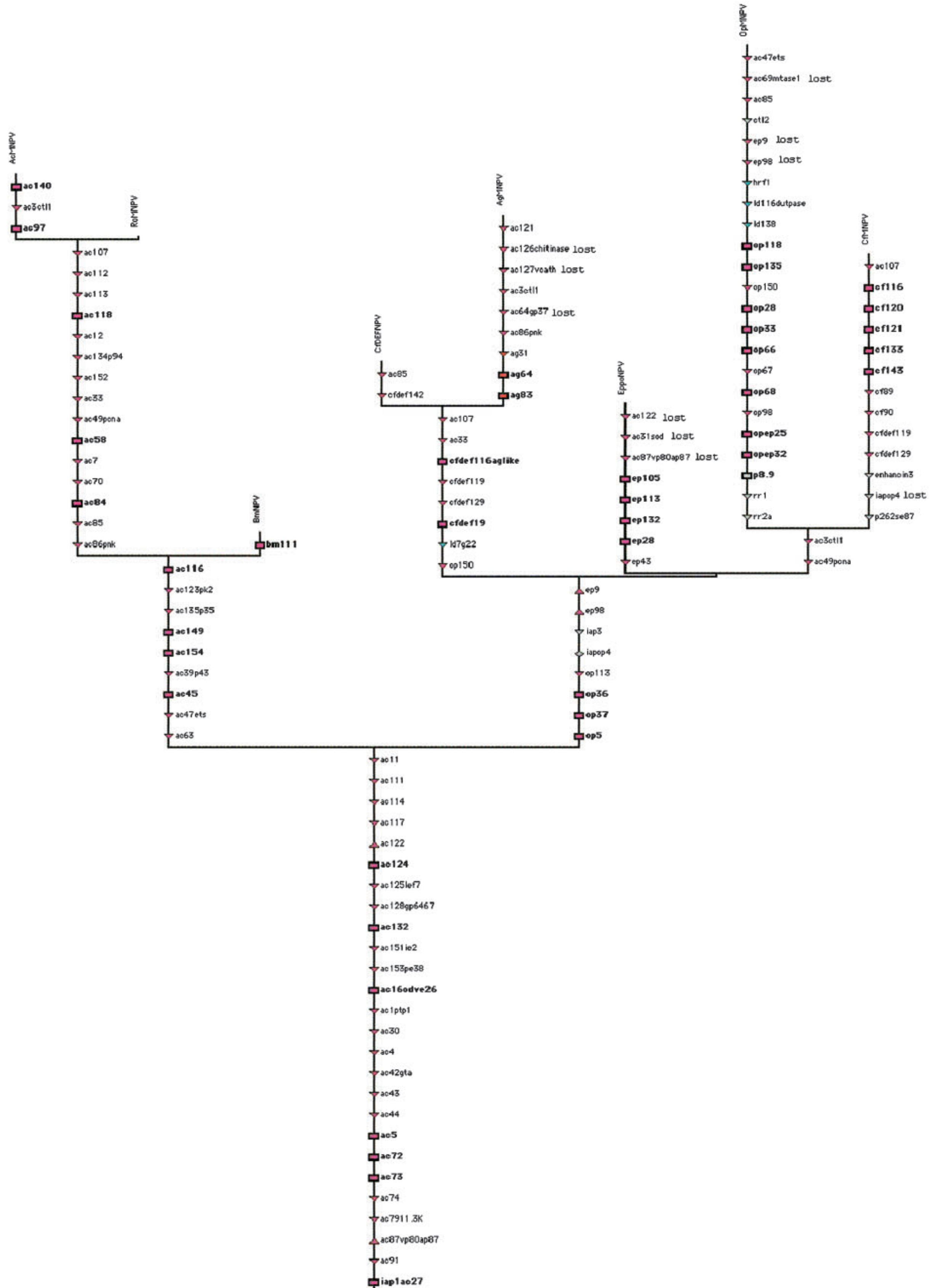
AgMNPV genes encoding structural proteins

Twenty-four of the 152 ORFs identified encoded known viral structural proteins. These included the nine conserved structural genes found in all baculoviruses and the six additional structural genes conserved in lepidopteran baculoviruses (Herniou *et al.*, 2003). These 24 genes are also present in CfDefNPV (Lauzon *et al.*, 2005). Sequence similarity among the structural proteins of AgMNPV-2D and CfDefNPV was generally above 92 % (Table 1), suggesting that they may be structurally similar. Interestingly, the putative major DNA binding protein, p6.9 (ag96), of both viruses was identical, being more conserved than the polyhedrin protein. The p6.9 and polyhedrin proteins were highly conserved among all of the group I NPVs (about 90 % similarity). In contrast, the AgMNPV-2D 1629 capsid (ag152) protein showed the lowest levels of sequence conservation among the group I NPV structural proteins (Table 1). In AcMNPV, this structural protein, which is present in budded viruses and occlusion-derived viruses, is essential for virus replication (Possee *et al.*, 1991). The putative p15 capsid protein (ag84), present only in group I NPVs, was 40–50 % smaller in size in both AgMNPV-2D and CfDefNPV.

Genes involved in replication and transcription

The AgMNPV-2D genome encoded the essential genes *lef-1* (ag19), *lef-2* (ag3), *lef-3* (ag67), *dnapol* (ag65), *helicase* (ag92) and *ie-1* (ag143) (Lu *et al.*, 1997). It also had the DNA binding protein (*dbp*) (ag42) (Mikhailov *et al.*, 1998), which unwinds DNA during replication, and *lef-11* (ag24), which is associated with the expression of late genes but is also essential for DNA replication (Todd *et al.*, 1995; Lin & Blissard, 2002). Moreover, both *pe-38* (ag148 and ag149) and *ie-2* (ag145) genes, which increase viral DNA replication (Kool *et al.*, 1994; Ahrens & Rohrmann, 1995), were present (Table 1). Interestingly, *pe-38* was split into two ORFs

Downloaded from www.microbiologyresearch.org by
IP: 200.202.165.220
On: Tue, 22 May 2018 19:05:29



Downloaded from www.microbiologyresearch.org by
IP: 200.202.165.220
On: Tue, 22 May 2018 19:05:29

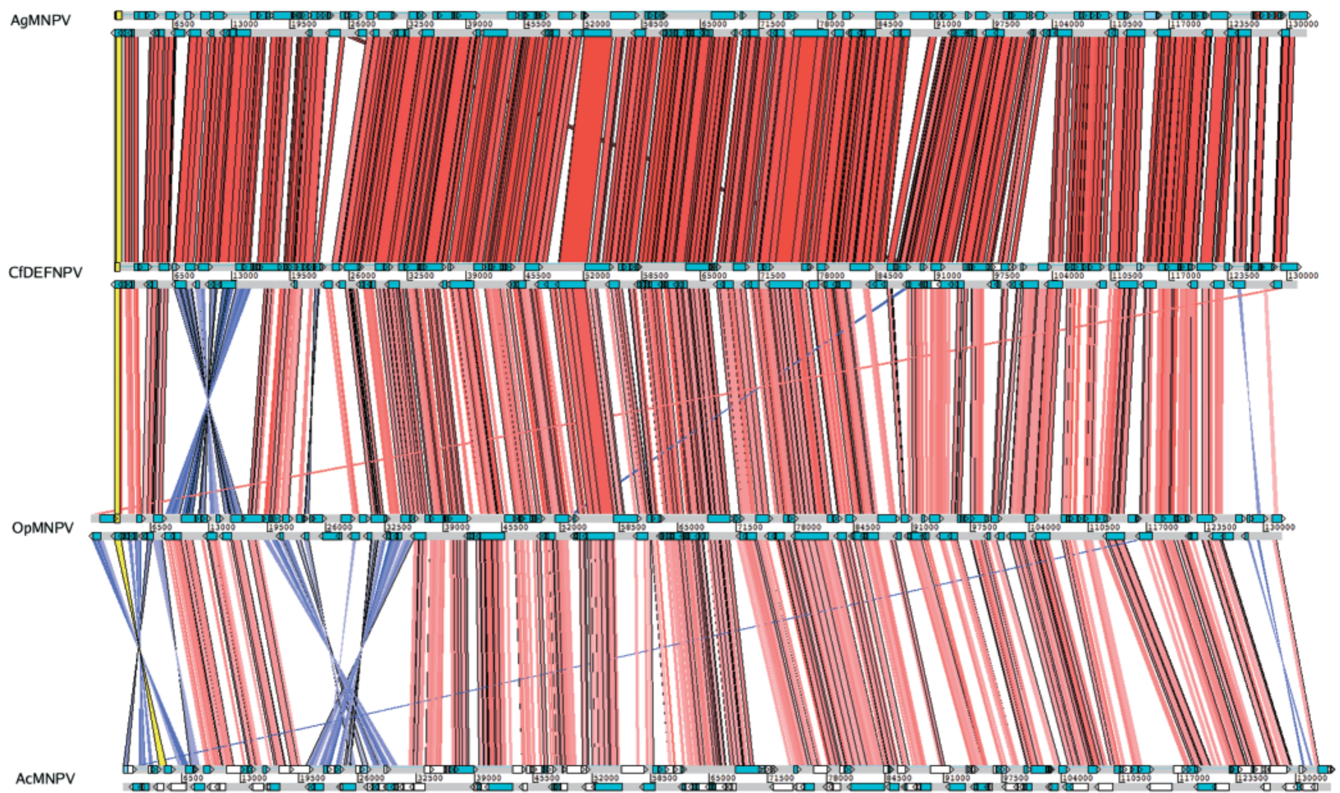


Fig. 4. Syntenic maps of four group I NPV representatives obtained with the Artemis Comparative Tool (version 7.0). Genomes are shown in a linear fashion, as in their GenBank files. Syntenic regions, indicated by stripes connecting genome overlays, indicate areas of extensive sequence conservation, irrespective of the actual coding regions, as found with the tBLASTX program. Main inversions are indicated by twisted sets of connecting lines. Hits running along diagonals indicate differences in the position of linearization (e.g. CfDefNPV vs OpMNPV) or BLAST hits among *bro* genes (e.g. AgMNPV vs CfDefNPV). Purple lines indicate inversions, red lines indicate collinear sections and the yellow line connects the polyhedrin gene of all four genomes.

(Table 1). The *pe-38* gene was found in group I NPVs and in an apical cluster of GVs that included CpGV, *Phthorimaea operculella* GV (PhopGV) and CrleGV (Fig. 2 and data not shown). The *ie-2* gene was found only in the group I NPVs and therefore seems to have a more restricted distribution among the baculoviruses compared with *pe-38* and *ie-1*. The late expression factor genes [*lef-1* to *lef-12* (ag45)], *39K (pp31)* (ag25) and *p47* (ag44) were also found (Table 1). The *lef-4* (ag87), *lef-8* (ag51), *lef-9* (ag62) and *p47* genes encode the DNA-dependent RNA polymerase and are found in all sequenced baculoviruses (Lu & Miller, 1997; Garcia-Maruniak *et al.*, 2004). However, only *lef-8*, which also is present in HzV-1, was not unique to the baculoviruses, having homologues in both eukaryotes and prokaryotes, as indicated by our gene-content analysis (data not shown).

The *lef-1*, *-2*, *-3* and *-11*, *helicase* and *ie-1* genes were present in all baculoviruses but absent in HzV-1. Moreover, the typical baculovirus type B DNA polymerase was quite distinct from that found in HzV-1 (Garcia-Maruniak *et al.*, 2004). HzV-1 shares sequence similarity with 10 baculovirus genes (Cheng *et al.*, 2002), but has a distinct set of core functions involved in replication. Possibly, the shared functions allow it to infect insects but do not necessarily imply common ancestry with the *Baculoviridae*.

Fig. 3. Most parsimonious reconstructions showing patterns of independent gene gain and loss events in group I NPVs. Boxes represent gene functions not found to be shared elsewhere (i.e. mapped as synapomorphies, which are derived character states shared by members of a group). Arrows pointing down represent genes shared elsewhere in the tree and arrows pointing up indicate genes shared by group I NPVs that will be lost above (i.e. mapped as homoplasies, which are character states that evolve more than once in different branches of the tree). Diamonds indicate that genes are homoplastic above and below their branch. Gene names are shown along branches. Only genes that were lost along a lineage are labelled.

Downloaded from www.microbiologyresearch.org by
IP: 200.202.165.220
On: Tue, 22 May 2018 19:05:29

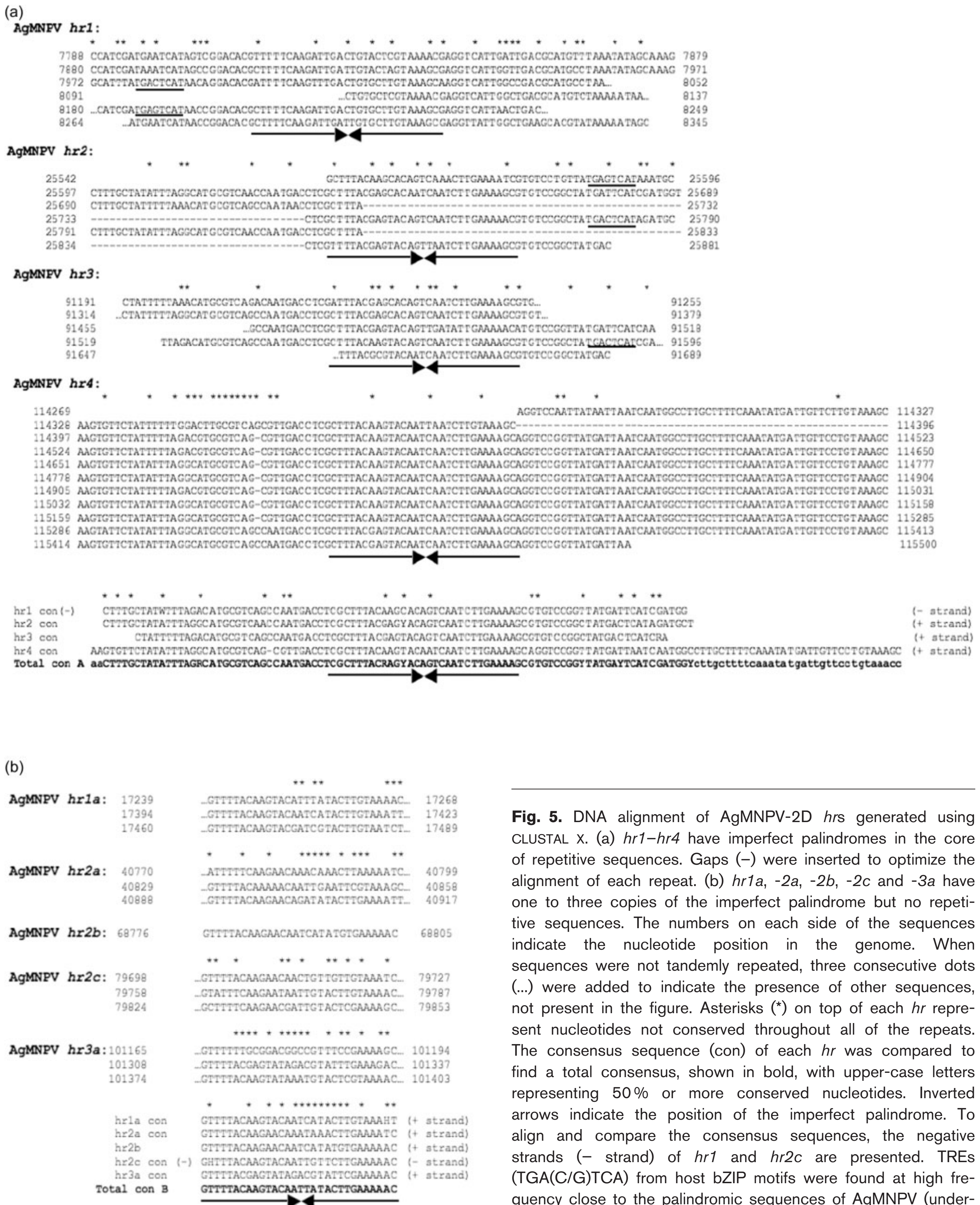


Fig. 5. DNA alignment of AgMNPV-2D *hr*s generated using CLUSTAL X. (a) *hr1*–*hr4* have imperfect palindromes in the core of repetitive sequences. Gaps (–) were inserted to optimize the alignment of each repeat. (b) *hr1a*, *-2a*, *-2b*, *-2c* and *-3a* have one to three copies of the imperfect palindrome but no repetitive sequences. The numbers on each side of the sequences indicate the nucleotide position in the genome. When sequences were not tandemly repeated, three consecutive dots (...) were added to indicate the presence of other sequences, not present in the figure. Asterisks (*) on top of each *hr* represent nucleotides not conserved throughout all of the repeats. The consensus sequence (con) of each *hr* was compared to find a total consensus, shown in bold, with upper-case letters representing 50 % or more conserved nucleotides. Inverted arrows indicate the position of the imperfect palindrome. To align and compare the consensus sequences, the negative strands (– strand) of *hr1* and *hr2c* are presented. TREs (TGA(C/G)TCA) from host bZIP motifs were found at high frequency close to the palindromic sequences of AgMNPV (underlined). W = A or T; Y = T or C; R = G or A; H = A or T or C.

Downloaded from www.microbiologyresearch.org by
IP: 200.202.165.220
On: Tue, 22 May 2018 19:05:29

bro genes

Eight *bro* genes [*bro-a* (ag6), *bro-b* (ag7), *bro-c* (ag11), *bro-d* (ag33), *bro-e* (ag63), *bro-f* (ag105), *bro-g* (ag115) and *bro-h* (ag135)] were found in AgMNPV-2D (Table 1). It should be noted that *bro-h* is very small and seems to be the remnant of a *bro* gene. This gene was located near *hr4* at a highly variable region (Garcia-Maruniak *et al.*, 1996; Maruniak *et al.*, 1999). Baculovirus *bro* genes are assumed to constitute a multigene family, composed of redundant ORFs dispersed along the genome. The number of *bro* genes is quite variable, ranging from none in GVs (Hashimoto *et al.*, 2000) to 16 in LdMNPV (Kuzio *et al.*, 1999). They may influence baculovirus genome diversity and be involved in recombination between baculovirus genomes (Li *et al.*, 2002a, b, 2005) and may cause strain heterogeneity (Zhang *et al.*, 2005). However, contradictory results have been reported on the role of *bro* genes in viral infection and replication, both *in vivo* and *in vitro* (Kang *et al.*, 1999; Zemskov *et al.*, 2000; Bideshi *et al.*, 2003). Thus, their role is still not clear and not all *bro* genes appeared to be homologues upon close scrutiny during our analyses.

Anti-apoptotic genes

Three inhibitor of apoptosis protein genes [*iap-1* (ag40), *iap-2* (ag70) and *iap-3* (ag34)] were found in AgMNPV-2D (Table 1). The *iap-3* gene of AgMNPV-2D encodes a functional anti-apoptotic protein (IAP) (Carpes *et al.*, 2005). IAPs were first discovered in baculoviruses (Crook *et al.*, 1993) and later found in various animal species, including insects and humans (Clem, 1997; Uren *et al.*, 1998). IAPs have ring fingers (zinc-binding motif) at the C terminus, which appear to be involved in their own ubiquitination (Yang & Li, 2000; Bergmann *et al.*, 2003). Homologues of *iap* genes have been found in most sequenced baculovirus genomes to date with the exception of *Culex nigripalpus* NPV (CuniNPV) (Afonso *et al.*, 2001). Analysis of the putative AgMNPV-2D-encoded IAPs showed that both IAP-1 and IAP-3 had two baculoviral *iap* repeat (BIR) motifs in the N-terminal region and a RING finger motif at its C-terminal end. However, IAP-2 had only one BIR motif at the N terminus and a RING finger motif at the C terminus.

In contrast, no homologues for the anti-apoptotic p35 protein (Clem & Miller, 1994) were present in AgMNPV-2D. So far, the *p35* gene has been found in AcMNPV, BmNPV and RoMNPV, three closely related group I NPVs, and has homologues in SpltNPV (group II NPV) and in *Xestia c-nigrum* GV (XecnGV), as well as in the entomopoxviruses (Clem *et al.*, 1991; Kamita *et al.*, 1993; Du *et al.*, 1999; Pang *et al.*, 2001). p35 inhibits caspase activity, blocking apoptosis (Bump *et al.*, 1995).

Auxiliary genes

Auxiliary genes are not essential for virus replication, but may provide a selective advantage (O'Reilly, 1997). Two auxiliary genes, *cathepsin* (Hill *et al.*, 1995) and *chitinase* (Hawtin *et al.*, 1995), found in most lepidopteran NPVs [with the exception of *chitinase* in *Adoxophyes honmai* NPV (AdhoNPV)] were not present in the AgMNPV-2D genome. The genomic region that harbours these two genes is highly conserved in group I NPVs. Previous studies by Slack *et al.* (2004) could not locate the genes around the AgMNPV *gp64* locus and demonstrated a lack of enzymic activity in AgMNPV-infected cell cultures. The absence of both the *chitinase* and *cathepsin* genes may be responsible for the lack of liquefaction of *A. gemmatalis* larvae killed by AgMNPV. Nevertheless, other genes, such as the *25K* gene, or host apoptotic responses may also be involved and may explain the liquefaction of *T. ni* by AgMNPV (Katsuma *et al.*, 1999; Slack *et al.*, 2004). When used as a biological pesticide, AgMNPV usually kills *A. gemmatalis* larvae within 7–10 days of virus application (Moscardi & Sosa-Gómez, 1992). The bodies of infected larvae usually remain intact for at least 2 days after death. After that period and under humid conditions, the larval body darkens and lyses due to putrefaction. However, under dry conditions, the bodies of the larvae dry up and shrink, keeping polyhedral inclusion bodies packed within them. Interestingly, the lack of liquefaction is of the utmost importance for the biological control programme using AgMNPV in Brazil (Moscardi, 1999). This is because it is necessary to collect dead and dying larvae for the preparation of future viral stocks, after applying viruses in the field (Moscardi & Sosa-Gómez, 1992; Moscardi, 1999). This has the benefit of reducing production costs to the farmers and may help in keeping fit viruses selected under natural conditions. The infectivity of AgMNPV and the lack of infected larvae liquefaction were probably key factors in allowing the expansion and current success of the biocontrol programme.

The AgMNPV-2D genome had several auxiliary genes that are also found in other baculoviruses (Table 1). Two ORFs encoding protein tyrosine phosphatases, *ptp-1* (ag9) and *ptp-2* (ag8), were found. PTP specifically removes phosphates from tyrosine residues and regulates tyrosine phosphorylation in concert with protein tyrosine kinases. They are associated with fibrillar structures in infected cells and can also be detected in budded viruses and occlusion-derived viruses, suggesting that they are a component of the viral capsid (Li & Miller, 1995). It has been shown that larvae of *B. mori* infected with BmNPV display enhanced locomotory activity due to a baculovirus-encoded PTP (Kamita *et al.*, 2005). Whilst the *ptp-1* gene was found in all other members of the group I NPVs, the *ptp-2* gene was found only in CfDefNPV, CfMNPV and OpMNPV (Table 1).

The *pnk*/*pnl* gene (ag103) had 66·7 % identity with *pnk*/*pnl* of both AcMNPV (ac86) and RoMNPV (ro83), which may be an RNA ligase possibly involved with RNA repair (Martins & Shuman, 2004). In phages, *pnk*/*pnl* activity is encoded by two different genes. All three group I NPV *pnk*/*pnl* genes had a high level of sequence conservation (79·6 %) suggestive of structural and functional equivalence. Interestingly, this gene was found to be present only in these three genomes. Therefore, it was either acquired independently by

Downloaded from www.microbiologyresearch.org by
IP: 200.202.165.220
On: Tue, 22 May 2018 19:05:29

Table 1. Putative ORFs identified in the AgMNPV-2D genome and comparison with the ORFs of group I NPVs

<, ORFs in negative strand; >, ORFs in positive strand; –, ORFs or *hr* sequences not found.

AgMNPV ORFs	Name	Position (nt)	No. aa	ORFs of group I NPVs (% aa identity)				
				CfDefNPV	EppoNPV	OpMNPV	CfMNPV	AcMNPV
1	*polh*	1<738	245	1 (97)	1 (97)	3 (95)	1 (94)	8 (88)
2		829<1179	116	2 (92)	3 (67)	5 (65)	2 (59)	–
3	*lef-2*	1247<1861	204	3 (91)	4 (70)	6 (70)	3 (69)	6 (56)
4		1864<2130	88	4 (87)	5 (57)	7 (59)	4 (57)	5 (67)
5		2178>2615	145	5 (93)	6 (60)	8 (77)	5 (71)	4 (61)
6	*bro-a*	2674>3405	243	6 (58)	–	–	–	2 (50)
7	*bro-b*	3417>4406	329	6 (80)	103 (32)	–	–	2 (73)
8	*ptp-2*	4435<4917	160	7 (92)	–	9 (67)	8 (71)	–
9	*ptp-1*	4904<5428	174	8 (94)	7 (73)	10 (68)	9 (71)	1 (62)
10		5686>6669	327	9 (90)	8 (48)	11 (59)	10 (54)	11 (47)
11	*bro-c*	6704<7735	343	101 (76)	–	–	–	2 (21)
	hr1	7788–8345	–	–	–	–	–	–
12	*pif-2*	8381<9529	382	10 (96)	18 (85)	20 (88)	20 (88)	22 (81)
13	*arif-1*	9556>10536	326	11 (90)	17 (50)	19 (64)	19 (61)	21 (33)
14		10657<10971	104	12 (94)	16 (59)	18 (61)	18 (72)	19 (41)
15		10989>12032	347	13 (94)	15 (70)	17 (60)	17 (63)	18 (49)
16		12107<12742	211	14 (95)	14 (68)	16 (68)	16 (69)	17 (56)
17	*odv-e26*	12708<13355	215	15 (94)	13 (62)	15 (57)	15 (52)	16 (32)
18	*egt*	13532<15013	493	16 (95)	12 (75)	14 (73)	14 (73)	15 (64)
19	*lef-1*	15097>15858	253	17 (93)	11 (63)	13 (68)	13 (67)	14 (55)
20		15789>16763	324	18 (95)	10 (59)	12 (64)	12 (69)	13 (43)
21		16787>17230	147	19 (93)		–		–
				20 (44)	9 (32)		11 (39)	
	hr1a	17239–17489	–	–	–	–	–	–
22	*F protein*	17612>19570	652	21 (92)	19 (53)	21 (55)	21 (55)	23 (41)
	dr1 (2×19 bp)	19648–19685	–	–	–	–	–	–
23		19697>20344	215	22 (94)	20 (94)	22 (84)	22 (88)	38 (75)
24	*lef-11*	20274>20690	138	23 (93)	21 (71)	23 (69)	23 (75)	37 (52)
25	*39K/pp31*	20630>21427	265	24 (95)	22 (80)	23 (74)	24 (72)	36 (53)
26	*ubi*	21555<21791	78	25 (96)	23 (88)	25 (90)	25 (91)	35 (85)
27		21808>22419	203	26 (93)	24 (75)	26 (76)	26 (72)	34 (53)
28	*hisP*	22422>22967	181	27 (93)	–	–	–	33 (58)
29	*fgf*	23022>23576	184	28 (90)	25 (60)	27 (63)	27 (63)	32 (44)
30	*ctl-2*	23604>23798	64	–	–	30 (63)	–	–
31	*parp-like**	23857<24966	369	–	–	–	–	–
32	*sod*	25085>25537	150	29 (98)	–	29 (83)	28 (85)	31 (76)
	hr2	25542–25881	–	–	–	–	–	–
33	*bro-d*	25993>27015	340	101 (74)	–	67 (63)	–	–
				59 (43)				
34	*iap-3*	27052<27915	287	30 (78)	26 (56)	35 (55)	30 (57)	–
35		27942>28670	242	31 (81)	27 (51)	36 (55)	31 (39)	–
36		28708<29271	187	32 (85)	29 (48)	37 (47)	32 (48)	–
37		29270>30652	460	33 (94)	30 (65)	38 (68)	33 (66)	30 (51)
38		30697>30906	69	34 (100)	31 (64)	39 (64)	34 (66)	29 (53)
39	*lef-6*	30939<31370	143	35 (92)	32 (48)	40 (60)	35 (55)	28 (33)
40	*iap-1*	31370<32218	282	36 (95)	33 (74)	41 (73)	36 (70)	27 (57)
41		32215<32598	127	37 (96)	34 (75)	42 (70)	37 (69)	26 (61)
42	*dbp*	32654>33577	307	38 (98)	35 (78)	43 (71)	38 (70)	25 (44)
43	*pkip-1*	33591>34106	171	39 (92)	36 (56)	44 (60)	39 (64)	24 (45)
44	*p47*	34122<35306	394	40 (97)	37 (85)	45 (79)	40 (80)	40 (69)
45	*lef-12*	35305>35820	171	41 (94)	38 (67)	46 (67)	41 (74)	41 (51)

Downloaded from www.microbiologyresearch.org by
IP: 200.202.165.220
On: Tue, 22 May 2018 19:05:29

Table 1. cont.

AgMNPV ORFs	Name	Position (nt)	No. aa	ORFs of group I NPVs (% aa identity)				
				CfDefNPV	EppoNPV	OpMNPV	CfMNPV	AcMNPV
46	*gta*	35829 > 37325	498	42 (92)	39 (73)	47 (78)	42 (80)	42 (60)
47		37405 > 37608	67	43 (94)	40 (73)	48 (76)	43 (82)	43 (60)
48		37577 > 37948	123	44 (91)	41 (69)	49 (66)	44 (70)	44 (48)
49	*odv-e66*	37999 > 40047	682	45 (96)	42 (79)	50 (79)	45 (79)	46 (71)
50	*etm*	40154 < 40522	122	46 (92)	44 (58)	52 (74)	46 (69)	48 (69)
	hr2a	40770–40917						
51	*lef-8*	40929 < 43556	875	47 (98)	45 (86)	54 (80)	48 (80)	50 (70)
52		43586 > 44515	309	48 (86)	46 (59)	55 (43)	49 (45)	51 (31)
53		44844 > 45284	146	49 (97)	47 (84)	56 (85)	50 (82)	53 (64)
54	*lef-10*	45253 > 45486	77†	50 (96)	48 (75)	57 (53)	51 (71)	53a (57)
55	*vp1054*	45347 > 46495	382	51 (95)	49 (73)	58 (73)	52 (72)	54 (55)
56		46559 > 46768	69	52 (97)	50 (82)	59 (70)	53 (65)	55 (63)
57		46770 > 47018	82	53 (93)	51 (70)	60 (68)	54 (63)	56 (45)
58		47211 > 47699	162	54 (93)	52 (70)	61 (63)	55 (61)	57 (53)
59		47714 < 48214	166	55 (96)	53 (59)	62 (53)	56 (58)	59 (65)
60		48162 < 48473	103	56 (91)	54 (78)	63 (71)	57 (68)	60 (70)
61	*25k fp*	48592 < 49287	231	57 (93)	55 (82)	64 (77)	58 (75)	61 (71)
62	*lef-9*	49286 > 50758	490	58 (95)	56 (88)	65 (85)	59 (87)	62 (76)
63	*bro-e*	50797 < 51759	320	59 (84)	–	–	–	–
64	*	51807 > 52067	86	–	–	–	–	–
	dr2 (2 × 28 bp)	52009–52064	–	–	–	–	–	–
65	*dnapol*	52039 < 55014	991	61 (95)	58 (73)	70 (76)	61 (75)	65 (61)
66		55024 > 57684	886	62 (77)	59 (42)	71 (42)	62 (41)	66 (31)
67	*lef-3*	57677 < 58801	374	63 (95)	60 (71)	72 (63)	63 (61)	67 (44)
68		58803 > 59201	132	64 (96)	61 (74)	73 (81)	64 (83)	68 (63)
69	*met*	59158 > 59982	274	65 (92)	62 (73)	–	65 (72)	69 (62)
70	*iap-2*	59963 > 60688	241	66 (95)	63 (71)	74 (69)	66 (71)	71 (62)
71		60541 > 60894	117†	67 (79)	64 (62)	75 (57)	67 (60)	72 (53)
72		60891 < 61166	91	68 (91)	65 (30)	76 (32)	68 (32)	73 (31)
73		61163 < 61696	177	69 (94)	66 (65)	77 (63)	69 (55)	74 (50)
74		61715 < 62107	130	70 (99)	67 (76)	78 (55)	70 (56)	75 (51)
75		62115 < 62369	84	71 (100)	68 (95)	79 (91)	71 (94)	76 (85)
76	*vlf-1*	62380 < 63510	376	72 (98)	69 (89)	80 (89)	72 (90)	77 (79)
77		63516 < 63845	109	73 (92)	70 (73)	81 (69)	73 (68)	78 (63)
78		63842 < 64141	99	74 (95)	71 (88)	82 (79)	74 (83)	79 (68)
79	*gp41*	64138 < 65277	379	75 (97)	72 (85)	83 (86)	75 (82)	80 (68)
80		65270 < 65917	215	76 (97)	73 (84)	84 (84)	76 (84)	81 (75)
81	*tlp20*	65808 < 66269	153	77 (92)	74 (57)	85 (61)	77 (57)	82 (29)
82	*vp91/p95*	66238 > 68745	835†	78 (95)	75 (77)	86 (75)	78 (75)	83 (65)
	hr2b	68776–68805						
83	*	68754 < 68918	54	–	–	–	–	–
84	*p15*	69405 > 69638	77	80 (93)	–	88 (42)	79 (35)	87 (36)
85	*cg30*	69670 < 70392	240	81 (94)	76 (60)	89 (62)	80 (59)	88 (43)
86	*vp39 capsid*	70398 < 71420	340	82 (89)	77 (70)	90 (77)	81 (77)	89 (61)
87	*lef-4*	71432 > 72799	455	83 (95)	78 (75)	91 (70)	82 (70)	90 (55)
88		72786 < 73394	202	84 (79)	79 (58)	92 (70)	83 (54)	91 (52)
89		73408 < 74184	258	85 (96)	80 (89)	93 (76)	84 (80)	92 (84)
90	*p18*	74183 > 74662	159	86 (93)	81 (84)	94 (86)	85 (88)	93 (71)
91	*odv-e25*	74665 > 75354	229	87 (96)	82 (91)	95 (81)	86 (86)	94 (62)
92	*helicase*	75384 < 79049	1221	88 (97)	83 (82)	96 (76)	87 (73)	95 (59)
93		79039 > 79563	174	89 (96)	84 (85)	97 (85)	88 (82)	96 (73)
	hr2c	79698–79853						

Downloaded from www.microbiologyresearch.org by
IP: 200.202.165.220
On: Tue, 22 May 2018 19:05:29

Table 1. cont.

AgMNPV ORFs	Name	Position (nt)	No. aa	ORFs of group I NPVs (% aa identity)				
				CfDefNPV	EppoNPV	OpMNPV	CfMNPV	AcMNPV
94		79867<80799	310	90 (97)	85 (74)	99 (75)	91 (78)	98 (56)
95	*lef-5*	80746>81540	264	91 (94)	86 (81)	100 (72)	92 (72)	99 (62)
96	*p6.9*	81537<81698	53	92 (100)	87 (92)	101 (90)	93 (86)	100 (74)
97	*p40*	81741<82820	359	93 (96)	88 (78)	102 (78)	94 (77)	101 (63)
98	*p12*	82830<83174	114	94 (97)	89 (80)	103 (78)	95 (78)	102 (46)
99	*p48*	83152<84378	408	95 (97)	90 (79)	104 (77)	96 (76)	103 (55)
100	*p87*	84405>86438	677	96 (89)	91 (45)	105 (40)	97 (51)	104 (33)
101	*he65*‡	86475<87128	217	97 (92)	–	–	–	105 (54)
102	*he65*‡	87234<88121	295	98 (93)	–	–	–	105 (56)
103	*pnk/pnl*	88308>90344	678	–	–	–	–	86 (66)
104		90426>91154	242	100 (95)	93 (75)	107 (75)	98 (78)	106 (73)
	hr3	91191–91689		–	–	–	–	–
105	*bro-f*	91744>92820	358	101 (79)	–	–	–	2 (21)
106		92870<93214	114	102 (91)	94 (79)	108 (70)	99 (77)	108 (56)
107		93186<94364	392	103 (97)	95 (85)	109 (84)	100 (85)	109 (66)
108		94377>94670	97	104 (80)	–	110 (52)	–	150 (26)
109		94643<94825	60	105 (96)	96 (71)	111 (82)	101 (83)	110 (71)
110		94873<95088	71	106 (100)	97 (72)	112 (66)	102 (72)	111 (63)
111		95119<95526	135	107 (92)	98 (50)	–	103 (45)	–
112		95560>96648	362	108 (92)	99 (69)	113 (70)	104 (68)	–
113		96641<97903	420	109 (95)	100 (60)	114 (50)	105 (52)	114 (39)
114		97975<98586	203	111 (94)	102 (74)	115 (74)	106 (69)	115 (67)
115	*bro-g*	98611>99144	177	112 (93)	103 (77)	116 (80)	107 (70)	2 (33)
116		99175>99435	86	113 (93)	104 (44)	117 (35)	109 (37)	117 (41)
117	*pif*	99551>101146	531	114 (93)	106 (80)	119 (79)	110 (83)	119 (76)
	hr3a	101165–101403						
118		101427>101678	83	115 (91)	107 (54)	120 (61)	111 (60)	120 (54)
119	§	101661<101945	94	116 (74)	–	–	–	–
120		101887<102084	65	117 (85)	–	121 (50)	112 (59)	122 (43)
121		102237>102965	242	118 (83)	108 (57)	122 (62)	113 (61)	124 (39)
122	*v-trex*	103022<103714	230	119 (91)	–	–	114 (68)	–
123	*lef-7*	103740<104387	215	120 (60)	109 (29)	123 (53)	115 (51)	125 (29)
124	*gp64*	104602<106101	499	123 (87)	112 (74)	126 (77)	119 (75)	128 (74)
125	*p24*	106439>107017	192	124 (83)	114 (74)	127 (78)	122 (77)	129 (63)
126	*gp16*	107032>107346	104	125 (95)	115 (84)	128 (86)	123 (85)	130 (78)
127	*pp34/pep*	107394>108275	293	126 (93)	116 (79)	129 (81)	124 (74)	131 (52)
128	*p25*	108278>108934	218	127 (85)	117 (44)	130 (41)	125 (41)	132 (31)
129	*alk-exo*	108940>110202	420	128 (92)	118 (70)	131 (70)	126 (68)	133 (56)
130		110248<110883	211	129 (83)	–	–	127 (62)	–
131		110982<111206	74	–	–	–	–	134 (51)
132	*p26*	111218>111928	236	130 (87)	119 (59)	132 (66)	128 (72)	136 (55)
133	*p10*	111984>112265	93	131 (92)	120 (81)	133 (82)	129 (49)	137 (39)
134	*p74*	112268<114202	644	132 (94)	121 (84)	134 (86)	130 (84)	138 (79)
	hr4	114269–115500	–	–	–	–	–	–
135	*bro-h*	115493>115675	–	–	–	–	–	–
136	*me53*	115721<117064	447	133 (93)	122 (61)	137 (59)	132 (58)	139 (39)
137	*ie-0*	117314>118042	242	134 (95)	123 (78)	138 (69)	134 (71)	141 (58)
138	*p49*	118058>119518	486	135 (98)	124 (85)	139 (89)	135 (89)	142 (73)
139	*odv-e18*	119515>119778	87	136 (95)	125 (88)	140 (88)	136 (82)	143 (74)
140	*odv-e27*	119793>120686	297	137 (97)	126 (87)	141 (75)	137 (86)	144 (70)
141		120696>120983	95	138 (98)	127 (87)	142 (84)	138 (84)	145 (71)
142		121038<121676	212	139 (92)	128 (70)	144 (67)	139 (70)	146 (52)

Downloaded from www.microbiologyresearch.org by
IP: 200.202.165.220
On: Tue, 22 May 2018 19:05:29

Table 1. cont.

AgMNPV ORFs	Name	Position (nt)	No. aa	ORFs of group I NPVs (% aa identity)				
				CfDefNPV	EppoNPV	OpMNPV	CfMNPV	AcMNPV
143	*ie-1*	121615 > 123381	588	140 (95)	129 (72)	145 (66)	140 (68)	147 (50)
144	*odv-e56*	123437 < 124549	370	141 (95)	130 (84)	146 (82)	141 (85)	148 (69)
145	*ie-2/ie-n*	124690 < 126006	438	143 (83)	131 (36)	151 (35)	142 (41)	151 (33)
146		126124 > 126261	45	144 (82)	–	150 (76)	–	–
	dr3 (3 × 17 bp)¶	126402–126447	–	–	–	–	–	–
	dr4 (32 × 14 bp)¶	126568–127008	–	–	–	–	–	–
147		127014 > 127865	283	145 (84)	–	–	–	–
148	*pe-38-like*‖	128096 > 128491	131	146 (87)	133 (48)	152 (56)	144 (53)	153 (38)
149	*pe-38-like*‖	128550 > 129179	209	146 (49)	–	–	–	–
	dr5 (10 × 21 bp)¶	128754–128960	–	–	–	–	–	–
150		129263 > 129478	71	147 (97)	134 (35)	–	–	–
151	*pk-1*	129498 < 130310	270	148 (93)	135 (76)	1 (71)	145 (74)	10 (65)
152	*1629 capsid*	130312 > 5	643	149 (71)	136 (35)	2 (39)	146 (40)	9 (31)

*ORFs that are unique to AgMNPV-2D.
†ORFs containing more than one potential translation start site.
‡Ag101 and ag102 (as their homologues cfde97 and cfde98) had similarity to the 3′ and 5′ end of ac105, respectively, indicating a possible split in the ancestral lineage of both AgMNPV-2D and CfDefNPV.
§Ag119 is truncated but corresponds to cfdef116.
¶The difference between nucleotide positions and the length of the direct repeat is due to the presence of gaps.
‖Ag148 and ag149 had similarity to the 3′ and 5′ ends of cfdef146, respectively.

the AgMNPV lineage and the ancestral lineage of both AcMNPV and RoMNPV, or several successive independent losses have taken place along the NPV radiation.

The superoxide dismutase (*sod*) gene was found in AgMNPV-2D (ag32). The AcMNPV *sod* gene represents part of the enzymic defence against oxygen toxicity and is present in almost all aerotolerant organisms (Tomalski *et al.*, 1991). Homologues of the *sod* gene are found in several baculoviruses and in entomopoxviruses and insects. Although the function of the *sod* gene in baculoviruses remains unknown, it may protect occluded viruses in the environment from superoxide radicals generated by exposure to sunlight (Ignoffo & Garcia, 1994).

A putative *ctl* gene for a conotoxin-like peptide found in predatory marine snails was found (ag30), which had similarity mainly to the *ctl-2* gene of OpMNPV (op30). For that reason, we named this gene *ctl-2* (Table 1). OpMNPV also harbours the *ctl-1* gene, which has high levels of similarity to the *ctl* gene of AcMNPV (Eldridge *et al.*, 1992). The *ctl* gene is also present in the genome of other NPVs, GVs and entomopoxviruses. However, no *ctl* homologues are present in the CfDefNPV and EppoNPV genomes, indicative of independent secondary losses. The biological function of this protein is unknown, but it has been suggested that it may participate in the regulation of calcium levels during infection and induce some form of paralysis in infected insects (O'Reilly, 1997).

Ag122 encoded a putative 3′ repair exonuclease (*v-trex*) (Slack & Shapiro, 2004), which showed homology to eukaryotic 3′ exonuclease and was found in only two other closely related baculoviruses, CfDefNPV and CfMNPV.

A putative gene for ubiquitin (*ubi*) (ag26) was also found, as in all other lepidopteran baculoviruses sequenced to date. Ubiquitins are abundant in eukaryotic cells and mediate protein breakdown (Hershko & Ciechanover, 1992) and other cellular processes. However, little is known about their function in baculoviruses and they appear not to be essential for replication.

Genes with no homology to other baculovirus genes

Three AgMNPV-2D coding regions, ag31, ag64 and ag83, were unique and had no significant hits in the GenBank database. Ag64 and ag83 can encode 86 and 54 aa, respectively. Ag31 encoded a 369 aa polypeptide with significant similarity to the poly(ADP-ribose) polymerase (PARP) from several organisms, with a probability of being similar by chance to other known PARPs of $<10^{-23}$. The C terminus is the most highly conserved region of PARP proteins. An alignment of the C terminus of this protein of AgMNPV-2D and five other organisms comprising representatives of insects, plants, birds and nematodes is shown in Fig. 6. The AgMNPV-2D PARP-like protein was smaller than the homologous insect PARPs, but was similar in size to that of

Downloaded from www.microbiologyresearch.org by
IP: 200.202.165.220
On: Tue, 22 May 2018 19:05:29

vertebrates [e.g. chicken (*Gallus gallus*) with 358 aa]. The best match obtained by BLASTP was to the plant *Oryza sativa* with 37 % identity with a region of 212 aa from the total 655 aa protein. PARP-1 is a nuclear enzyme activated by nicked DNA molecules, possibly involved in DNA repair (D'Amours *et al.*, 1999; Smith, 2001). PARP binds to damaged DNA and catalyses the formation of ADP-ribose polymers that attach to its own glutamic acid residues, as well as to other proteins (Althaus & Richter, 1987; Smulson *et al.*, 1994). It is also thought to be a part of the base excision repair pathway (Oei & Ziegler, 2000), to promote transcription (Vispé *et al.*, 2000) and to be involved in apoptosis (Smith, 2001). PARP-1 is required for retroviral integration in the host genome (Ha *et al.*, 2001). In adenoviruses, inhibition of PARP-1 reduces viral infectivity (Dery *et al.*, 1986). It has been shown that PARP binds to simian virus 40 capsid proteins VP1 and VP3 and that the latter protein stimulates PARP activity, possibly leading the infected cell towards a necrotic pathway (Gordon-Shaag *et al.*, 2003). The function of the PARP-like protein in AgMNPV has not been studied and, as AgMNPV is the only baculovirus to date with this putative gene, further analysis is under way (T. S. Rizzi, B. M. Ribeiro, C. M. Romano, F. L. Melo, J. V. C. Oliveira & P. M. A. Zanotto, unpublished results).

Comparison of the AgMNPV and CfDefNPV genomes

We confirmed that the genomes of AgMNPV-2D and CfDefNPV were very similar, as indicated previously (Lima *et al.*, 2004; Lauzon *et al.*, 2005). Besides the unique genes found in AgMNPV-2D, there were other noteworthy differences between their genomes. AgMNPV-2D had additional *bro* genes (*bro-h*), a *ctl-2* gene and the *pnk/pnl* gene. On the other hand, CfDefNPV harboured the *chitinase* and *cathepsin* genes. The position of these genes in the CfDENPV genome suggests that they were lost from the ancestor baculovirus lineage that resulted in the AgMNPV-2D genome. CfDefNPV also harboured additional putative genes, located at positions equivalent to ag21 and between ag113 and 114 (Table 1). Several of the differences observed between the two genomes coincided with highly variable regions previously detected among AgMNPV variants (Maruniak *et al.*, 1999). For instance, the AgMNPV-2D *ctl-2* and *parp-1* genes were located in a variable region at 1·78 and 1·80 map units (m.u.), respectively. Likewise, the location of the expected *chitinase* and *cathepsin* genes (conserved in other baculoviruses) would correspond to another variable region of the AgMNPV-2D genome around 79·1 m.u. Moreover, ag21, near another variable region mapped at 13·32 m.u., seemed to be the combination of cfde19 and -20. Taken together, these data suggest that AgMNPV might indeed have hot spots that may favour evolutionary novelty.

In conclusion, as AgMNPV is the most widely used biopesticide in the world, it was essential that its DNA be sequenced so that we can scrutinize its gene content and function, unravel its genetic regulatory network and monitor genetic changes and evolution of the genome. Our

```
AgMNPV ORF 31    185 LHTTINALEPESP      RFALIRRYVDNSRPAMLK          QIFEIARPADDVRY        AAEL       GNRQLLWHGT 243
G. gallus        144 LRTDIKVVDKDSE      EAKIIKQYVKNTHAATHN.[  7] .EIFRIEREGESQRY.[ 2] .FKQL       HNRQLLWHGS 211
C. elegans       304 LPCHLEPVSEEIA.[ 4] .DCIAMRGPTHCYKLSLID.[ 16] .EVQEVPKKRGRKST.[ 2] .AAPT.[ 4] .TTKRLLWHGT 388
A. thaliana      421 LNCGLTPVGNDSE      EFSMVANYMENTHAKTHS.[  7] .QLFRASRAVEADRF.[ 2] .FSSS       KNRMLLWHGS 488
D. melanogaster  779 IKTQLVALDKNSE      EFSILSQYVKNTHASTHK.[  7] .DVFKVSRQGEARRF.[ 2] .FKKL       HNRKLLWHGS 846
S. peregrina     782 LKTKLEPLDKNSE      EYILLQKYVKNTHAETHK.[  7] .DIFKVARQGEARRY.[ 2] .FKKL       HNRRLLWHGS 849

AgMNPV ORF 31    244 SIGNVSRILHEGFKT.[ 2] .AKA      SG      MFGAGVYFSNAARKS       KSTLLFLCEVALGE         PLRC 298
G. gallus        212 RTTNFAGILSQGLRI.[ 2] .PEA.[ 2] .TG.[ 1] .MFGKGIYFADMVSKS.[ 10] .PIGLILLGEVALGN         MYEL 279
C. elegans       389 RVTNVPSILMNGLQF      PVG.[ 2] .CG.[ 1] .MFGNGVYFANVPTKS.[  9] .KRVFMLLCEVETAN.[ 3] .LYES 456
A. thaliana      489 RLTNWAGILSQGLRI.[ 2] .PEA.[ 2] .TG.[ 1] .MFGKGVYFADMFSKS.[ 10] .NDGVLLLCEVALGD         MNEL 556
D. melanogaster  847 RLTNFVGILSHGLRI.[ 2] .PEA.[ 2] .TG.[ 1] .MFGKGIYFADMVSKS.[ 10] .STGLMLLSEVALGD         MMEC 914
S. peregrina     850 RLTNFAGILSHGLKI.[ 2] .PEA.[ 2] .TG.[ 1] .MFGKGIYFADMVSKS.[ 10] .STGLMLLSEVALGD         MMEC 917

AgMNPV ORF 31    299          HAAR          SGALPDGKHSVHGVGALTP       AT.[ 2] .IID           DNVLVPSGK.[ 1] .VPANQAVTA 347
G. gallus        280 .[ 1] .NASH          ITKLPKGKHSVKGLGKTAP.[ 1] .PT.[ 2] .TTL           DGVEVPLGN.[ 1] .ISTGINDTC 330
C. elegans       457 .[ 1] .IDAD          EKMEKAKKTSVYAAGKHTP.[ 1] .DT          VEI           NGIPAFKSN          LETIEEETR 504
A. thaliana      557 .[ 1] .YSDY.[ 1] .ADNLPPGKLSTKGVGKTAP.[ 1] .PS.[ 2] .QTL.[ 1] .DGVVVPLGK.[ 1] .VERSCSKGM 609
D. melanogaster  915 .[ 1] .SAKY          INKLSNNKHSCFGRGRTMP.[ 1] .PT.[ 3] .IRS           DGVEIPYGE.[ 1] .ITDEHLKSS 966
S. peregrina     918 .[ 1] .AAKY          VTKLPNDKHSCFGRGRTMP.[ 1] .PS.[ 2] .IIR.[ 1] .DGVEIPLGK.[ 1] .ITNDSLKSS 969

AgMNPV ORF 31    348        FDEFVIYNTSQIKMRYLI.[ 4] .  369
G. gallus        331 .[ 2] .YNEYIVYDVAQVNLKYLL.[ 8] .  358
C. elegans       505 .[ 2] .YDEYVMFNKEHFKIKYVV.[ 8] .  532
A. thaliana      610 .[ 2] .YNEYIVYNVEQIKMRYVI.[ 8] .  637
D. melanogaster  967 .[ 2] .YNEYIVYDVAQVNIQYLE.[ 8] .  994
S. peregrina     970 .[ 2] .YNEFIIYDIAQVNIQYML.[ 7] .  996
```

Fig. 6. Protein alignments of ag31 and PARP-like proteins from other organisms generated with CLUSTAL X. Comparisons were made with *Gallus gallus* (chicken), *Caenorhabditis elegans* (nematode), *Arabidopsis thaliana* (thale cress), *Drosophila melanogaster* (fruit fly) and *Sarcophaga peregrina* (dipteran insect). Sequences in boxes are at least 50 % conserved. Amino acids underlined in the ag31 sequence are those conserved in all six organisms. Numbers at the beginning and end of each line represent amino acid positions. Numbers in square brackets represent amino acids inserted at each location.

Downloaded from www.microbiologyresearch.org by
IP: 200.202.165.220
On: Tue, 22 May 2018 19:05:29

previous research indicated areas of genome change over the years of virus application in soybean in Brazil that will now be studied. One of the reasons that the programme may have been so successful was identified during sequencing of the genome. Neither *chitinase* nor *cathepsin*, present in all other baculoviruses sequenced to date, nor other ORFs that would potentially fulfil this role were found in AgMNPV-2D. This allows field production and harvesting of relatively intact insects full of virus needed for subsequent applications.

ACKNOWLEDGEMENTS

This work was funded by the Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP) project 01/06051-9. The infrastructure was provided by FAPESP 00/004205-6. J. V. de C. O. has FAPESP Scholarship 04/12456-0. P. M. de A. Z. holds a CNPq PQ Research Scholarship. We would like to thank the two anonymous referees for their suggestions and comments.

REFERENCES

Afonso, C. L., Tulman, E. R., Lu, Z., Balinsky, C. A., Moser, B. A., Becnel, J. J., Rock, D. L. & Kutish, G. F. (2001). Genome sequence of a baculovirus pathogenic for *Culex nigripalpus*. *J Virol* **75**, 11157–11165.

Ahrens, C. H. & Rohrmann, G. F. (1995). Replication of *Orgyia pseudotsugata* baculovirus DNA: *lef-2* and *ie-1* are essential and *ie-2*, *p34*, and Op-*iap* are stimulatory genes. *Virology* **212**, 650–662.

Allen, G. E. & Knell, J. D. (1977). A nuclear polyhedrosis virus of *Anticarsia gemmatalis*: I. Ultrastructure, replication and pathogenicity. *Fla Entomol* **60**, 233–240.

Althaus, F. R. & Richter, C. (1987). ADP-ribosylation of proteins: enzymology and biological significance. *Mol Biol Biochem Biophys* **37**, 1–237.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* **215**, 403–410.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389–3402.

Arana, E. I., Albariño, C. G., O'Reilly, D., Ghiringhelli, P. D. & Romanowski, V. (2001). Generation of a recombinant *Anticarsia gemmatalis* multicapsid nucleopolyhedrovirus expressing a foreign gene under the control of a very late promoter. *Virus Genes* **22**, 363–372.

Ayres, M. D., Howard, S. C., Kuzio, J., Lopez-Ferber, M. & Possee, R. D. (1994). The complete DNA sequence of *Autographa californica* nuclear polyhedrosis virus. *Virology* **202**, 586–605.

Belaich, M. N., Rodríguez, V. A., Bilen, M. F., Pilloff, M. G., Romanowski, V., Sciocco-Cap, A. & Ghiringhelli, P. D. (2006). Sequencing and characterisation of *p74* gene in two isolates of *Anticarsia gemmatalis* MNPV. *Virus Genes* **32**, 59–70.

Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res* **27**, 573–580.

Bergmann, A., Yang, A. Y.-P. & Srivastava, M. (2003). Regulators of IAP function: coming to grips with the grim reaper. *Curr Opin Cell Biol* **15**, 717–724.

Bideshi, D. K., Renault, S., Stasiak, K., Federici, B. A. & Bigot, Y. (2003). Phylogenetic analysis and possible function of *bro*-like genes, a multigene family widespread among large double-stranded DNA viruses of invertebrates and bacteria. *J Gen Virol* **84**, 2531–2544.

Bump, N. J., Hackett, M., Hugunin, M. & 13 other authors (1995). Inhibition of ICE family proteases by baculovirus antiapoptotic protein p35. *Science* **269**, 1885–1888.

Carner, G. R. & Turnipseed, S. G. (1977). Potential of a nuclear polyhedrosis virus for control of the velvetbean caterpillar in soybean. *J Econ Entomol* **70**, 608–610.

Carpes, M. P., de Castro, M. E. B., Soares, E. F., Villela, A. G., Pinedo, F. J. R. & Ribeiro, B. M. (2005). The inhibitor of apoptosis gene (*iap-3*) of *Anticarsia gemmatalis* multicapsid nucleopolyhedrovirus (AgMNPV) encodes a functional IAP. *Arch Virol* **150**, 1549–1562.

Castro, M. E. B., Ribeiro, Z. M. A. & Souza, M. L. (2006). Infectivity of *Anticarsia gemmatalis* nucleopolyhedrovirus to different insect cell lines: morphology, viral production, and protein synthesis. *Biol Control* **36**, 299–304.

Cheng, C.-H., Liu, S.-M., Chow, T.-Y., Hsiao, Y.-Y., Wang, D.-P., Huang, J.-J. & Chen, H.-H. (2002). Analysis of the complete genome sequence of Hz-1 virus suggests that it is related to members of *Baculoviridae*. *J Virol* **76**, 9024–9034.

Clem, R. J. (1997). Regulation of programmed cell death by baculoviruses. In *The Baculoviruses*, pp. 237–266. Edited by L. K. Miller. New York: Plenum.

Clem, R. J. & Miller, L. K. (1994). Control of programmed cell death by the baculovirus genes *p35* and *iap*. *Mol Cell Biol* **14**, 5212–5222.

Clem, R. J., Fechheimer, M. & Miller, L. K. (1991). Prevention of apoptosis by a baculovirus gene during infection of insect cells. *Science* **254**, 1388–1390.

Crook, N. E., Clem, R. J. & Miller, L. K. (1993). An apoptosis-inhibiting baculovirus gene with a zinc finger-like motif. *J Virol* **67**, 2168–2174.

Dalmolin, C. C., da Silva, F. R., Mello, L. V., Rigden, D. J. & Castro, M. E. B. (2005). Nucleotide sequence and phylogenetic analyses of the DNA polymerase gene of *Anticarsia gemmatalis* nucleopolyhedrovirus. *Virus Res* **110**, 99–109.

D'Amours, D., Desnoyers, S., D'Silva, I. & Poirier, G. G. (1999). Poly(ADP-ribosyl)ation reactions in the regulation of nuclear functions. *Biochem J* **342**, 249–268.

da Silveira, E. B., Ribeiro, B. M. & Báo, S. N. (2003). Characterization of larval haemocytes from the velvetbean caterpillar *Anticarsia gemmatalis* (Hübner) (Lepidoptera: Noctuidae). *J Submicrosc Cytol Pathol* **35**, 129–139.

da Silveira, E. B., Cordeiro, B. A., Ribeiro, B. M. & Báo, S. N. (2004). Morphological characterization of *Anticarsia gemmatalis* M nucleopolyhedrovirus infection in haemocytes from its natural larval host, the velvet bean caterpillar *Anticarsia gemmatalis* (Hübner) (Lepidoptera: Noctuidae). *Tissue Cell* **36**, 171–180.

Dery, C. V., de Murcia, G., Lamarre, D., Morin, N., Poirier, G. G. & Weber, J. (1986). Possible role of ADP-ribosylation of adenovirus core proteins in virus infection. *Virus Res* **4**, 313–329.

Du, Q., Lehavi, D., Faktor, O., Qi, Y. & Chejanovsky, N. (1999). Isolation of an apoptosis suppressor gene of the *Spodoptera littoralis* nucleopolyhedrovirus. *J Virol* **73**, 1278–1285.

Eldridge, R., Li, Y. & Miller, L. K. (1992). Characterization of a baculovirus gene encoding a small conotoxinlike polypeptide. *J Virol* **66**, 6563–6571.

Ewing, B. & Green, P. (1998). Base-calling of automated sequencer traces using Phred. II. Error probabilities. *Genome Res* **8**, 186–194.

Ewing, B., Hillier, L., Wendl, M. C. & Green, P. (1998). Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* **8**, 175–185.

Downloaded from www.microbiologyresearch.org by
IP: 200.202.165.220
On: Tue, 22 May 2018 19:05:29

Garcia-Maruniak, A., Pavan, O. H. O. & Maruniak, J. E. (1996). A variable region of *Anticarsia gemmatalis* nuclear polyhedrosis virus contains tandemly repeated DNA sequences. *Virus Res* **41**, 123–132.

Garcia-Maruniak, A., Maruniak, J. E., Zanotto, P. M. A., Doumbouya, A. E., Liu, J.-C., Merritt, T. M. & Lanoie, J. S. (2004). Sequence analysis of the genome of *Neodiprion sertifer* nucleopolyhedrovirus. *J Virol* **78**, 7036–7051.

Gordon-Shaag, A., Yosef, Y., Abd El-Latif, M. & Oppenheim, A. (2003). The abundant nuclear enzyme PARP participates in the life cycle of simian virus 40 and is stimulated by minor capsid protein VP3. *J Virol* **77**, 4273–4282.

Ha, H. C., Juluri, K., Zhou, Y., Leung, S., Hermankova, M. & Snyder, S. H. (2001). Poly(ADP-ribose) polymerase-1 is required for efficient HIV-1 integration. *Proc Natl Acad Sci U S A* **98**, 3364–3368.

Hashimoto, Y., Hayakawa, T., Ueno, Y., Fujita, T., Sano, Y. & Matsumoto, T. (2000). Sequence analysis of the *Plutella xylostella* granulovirus genome. *Virology* **275**, 358–372.

Hawtin, R. E., Arnold, K., Ayres, M. D. & 7 other authors (1995). Identification and preliminary characterization of a chitinase gene in the *Autographa californica* nuclear polyhedrosis virus genome. *Virology* **212**, 673–685.

Herniou, E. A., Luque, T., Chen, X., Vlak, J. M., Winstanley, D., Cory, J. S. & O'Reilly, D. R. (2001). Use of whole genome sequence data to infer baculovirus phylogeny. *J Virol* **75**, 8117–8126.

Herniou, E. A., Olszewski, J. A., Cory, J. S. & O'Reilly, D. R. (2003). The genome sequence and evolution of baculoviruses. *Annu Rev Entomol* **48**, 211–234.

Hershko, A. & Ciechanover, A. (1992). The ubiquitin system for protein degradation. *Annu Rev Biochem* **61**, 761–807.

Hill, J. E., Kuzio, J. & Faulkner, P. (1995). Identification and characterization of the *v-cath* gene of the baculovirus, CfMNPV. *Biochim Biophys Acta* **1264**, 275–278.

Ignoffo, C. M. & Garcia, C. (1994). Antioxidant and oxidant enzyme effects of the inactivation of inclusion bodies of the *Heliothis* baculovirus by simulated sunlight-UV. *Environ Entomol* **23**, 1025–1029.

Jehle, J. A., Lange, M., Wang, H., Hu, Z., Wang, Y. & Hauschild, R. (2006). Molecular identification and phylogenetic analysis of baculoviruses from Lepidoptera. *Virology* **346**, 180–193.

Johnson, D. W. & Maruniak, J. E. (1989). Physical map of *Anticarsia gemmatalis* nuclear polyhedrosis virus (AgMNPV-2) DNA. *J Gen Virol* **70**, 1877–1883.

Kamita, S. G., Majima, K. & Maeda, S. (1993). Identification and characterization of the p35 gene of *Bombyx mori* nuclear polyhedrosis virus that prevents virus-induced apoptosis. *J Virol* **67**, 455–463.

Kamita, S. G., Nagasaka, K., Chua, J. W., Shimada, T., Mita, K., Kobayashi, M., Maeda, S. & Hammock, B. D. (2005). A baculovirus-encoded protein tyrosine phosphatase gene induces enhanced locomotory activity in a lepidopteran host. *Proc Natl Acad Sci U S A* **102**, 2584–2589.

Kang, W., Suzuki, M., Zemskov, E., Okano, K. & Maeda, S. (1999). Characterization of baculovirus repeated open reading frames (*bro*) in *Bombyx mori* nucleopolyhedrovirus. *J Virol* **73**, 10339–10345.

Katsuma, S., Noguchi, Y., Zhou, C. L. E., Kobayashi, M. & Maeda, S. (1999). Characterization of the 25K FP gene of the baculovirus *Bombyx mori* nucleopolyhedrovirus: implications for post-mortem host degradation. *J Gen Virol* **80**, 783–791.

King, L. A. & Possee, R. D. (1992). *The Baculovirus Expression Vector System: a Laboratory Guide.* London: Chapman & Hall.

Kool, M., Ahrens, C. H., Goldbach, R. W., Rohrmann, G. F. & Vlak, J. M. (1994). Identification of genes involved in DNA replication of the *Autographa californica* baculovirus. *Proc Natl Acad Sci U S A* **91**, 11212–11216.

Kost, T. A., Condreay, J. P. & Jarvis, D. L. (2005). Baculovirus as versatile vectors for protein expression in insect and mammalian cells. *Nat Biotechnol* **23**, 567–575.

Kuzio, J., Pearson, M. N., Harwood, S. H., Funk, C. J., Evans, J. T., Slavicek, J. M. & Rohrmann, G. F. (1999). Sequence and analysis of the genome of a baculovirus pathogenic for *Lymantria dispar*. *Virology* **253**, 17–34.

Landais, I., Vincent, R., Bouton, M., Devauchelle, G., Duonor-Cerutti, M. & Ogliastro, M. (2006). Functional analysis of evolutionary conserved clustering of bZIP binding sites in the baculovirus homologous regions (*hr*s) suggests a cooperativity between host and viral transcription factors. *Virology* **344**, 421–431.

Lange, M. & Jehle, J. A. (2003). The genome of the *Cryptophlebia leucotreta* granulovirus. *Virology* **317**, 220–236.

Lauzon, H. A. M., Lucarotti, C. J., Krell, P. J., Retnakaran, A. & Arif, B. M. (2004). Sequence and organization of the *Neodiprion lecontei* nucleopolyhedrovirus genome. *J Virol* **78**, 7023–7035.

Lauzon, H. A. M., Jamieson, P. B., Krell, P. J. & Arif, B. M. (2005). Gene organization and sequencing of the *Choristoneura fumiferana* defective nucleopolyhedrovirus genome. *J Gen Virol* **86**, 945–961.

Li, W.-H. & Graur, D. (1991). Molecular phylogeny. In *Fundamentals of Molecular Evolution*, pp. 99–135. Sunderland, MA: Sinauer Associates.

Li, Y. & Miller, L. K. (1995). Expression and localization of a baculovirus protein phosphatase. *J Gen Virol* **76**, 2941–2948.

Li, L., Donly, C., Li, Q., Willis, L. G., Keddie, B. A., Erlanson, M. A. & Theilmann, D. A. (2002a). Identification and genomic analysis of a second species of nucleopolyhedrovirus isolated from *Mamestra configurata*. *Virology* **297**, 226–244.

Li, Q., Donly, C., Li, L., Willis, L. G., Theilmann, D. A. & Erlandson, M. (2002b). Sequence and organization of the *Mamestra configurata* nucleopolyhedrovirus genome. *Virology* **294**, 106–121.

Li, L., Li, Q., Willis, L. G., Erlandson, M., Theilmann, D. A. & Donly, C. (2005). Complete comparative genomic analysis of two field isolates of *Mamestra configurata* nucleopolyhedrovirus-A. *J Gen Virol* **86**, 91–105.

Lima, L., Pinedo, F. J., Ribeiro, B. M., Zanotto, P. M. A. & Wolff, J. L. (2004). Identification, expression and phylogenetic analysis of the *Anticarsia gemmatalis* multicapsid nucleopolyhedrovirus (AgMNPV) *helicase*. *Virus Genes* **29**, 345–352.

Lin, G. & Blissard, G. W. (2002). Analysis of an *Autographa californica* multicapsid nucleopolyhedrovirus *lef-6*-null virus: LEF-6 is not essential for viral replication but appears to accelerate late gene transcription. *J Virol* **76**, 5503–5514.

Liu, J.-C. & Maruniak, J. E. (1999). Molecular characterization of genes in the GP41 region of baculoviruses and phylogenetic analysis based upon GP41 and polyhedrin genes. *Virus Res* **64**, 187–196.

Lu, A. & Miller, L. K. (1997). Regulation of baculovirus late and very late gene expression. In *The Baculoviruses*, pp. 193–216. Edited by L. K. Miller. New York: Plenum.

Lu, A., Krell, P. J., Vlak, J. M. & Rohrmann, G. F. (1997). Baculovirus DNA replication. In *The Baculoviruses*, pp. 171–191. Edited by L. K. Miller. New York: Plenum.

Luque, T., Finch, R., Crook, N., O'Reilly, D. R. & Winstanley, D. (2001). The complete sequence of the *Cydia pomonella* granulovirus genome. *J Gen Virol* **82**, 2531–2547.

Marchler-Bauer, A., Anderson, J. B., Cherukuri, P. F. & 21 other authors (2005). CDD: a Conserved Domain Database for protein classification. *Nucleic Acids Res* **33**, D192–D196.

Downloaded from www.microbiologyresearch.org by
IP: 200.202.165.220
On: Tue, 22 May 2018 19:05:29

Martins, A. & Shuman, S. (2004). Characterization of a baculovirus enzyme with RNA ligase, polynucleotide 5′-kinase, and polynucleotide 3′-phosphatase activities. *J Biol Chem* **279**, 18220–18231.

Maruniak, J. E. (1989). Molecular biology of *Anticarsia gemmatalis* baculovirus. *Mem Inst Oswaldo Cruz* **84**, 107–111.

Maruniak, J. E., Garcia-Maruniak, A., Souza, M. L., Zanotto, P. M. A. & Moscardi, F. (1999). Physical maps and virulence of *Anticarsia gemmatalis* nucleopolyhedrovirus genomic variants. *Arch Virol* **144**, 1991–2006.

Mikhailov, V. S., Mikhailova, A. L., Iwanaga, M., Gomi, S. & Maeda, S. (1998). *Bombyx mori* nucleopolyhedrovirus encodes a DNA-binding protein capable of destabilizing duplex DNA. *J Virol* **72**, 3107–3116.

Miller, L. K. (1999). An exegesis of IAPs: salvation and surprises from BIR motifs. *Trends Cell Biol* **9**, 323–328.

Morse, M. A., Marriott, A. C. & Nuttall, P. A. (1992). The glycoprotein of Thogoto virus (a tick-borne orthomyxo-like virus) is related to the baculovirus glycoprotein GP64. *Virology* **186**, 640–646.

Moscardi, F. (1999). Assessment of the application of baculoviruses for control of lepidoptera. *Annu Rev Entomol* **44**, 257–289.

Moscardi, F. & Sosa-Gómez, D. R. (1992). Use of viruses against soybean caterpillars in Brazil. In *Pest Management in Soybean*, pp. 98–109. Edited by L. G. Copping, M. B. Green & R. T. Rees. London: Elsevier.

Oei, S. L. & Ziegler, M. (2000). ATP for the DNA ligation step in base excision repair is generated from poly(ADP-ribose). *J Biol Chem* **275**, 23234–23239.

O'Reilly, D. R. (1997). Auxiliary genes of baculoviruses. In *The Baculoviruses*, pp. 267–300. Edited by L. K. Miller. New York: Plenum.

Pang, Y., Yu, J., Wang, L. & 7 other authors (2001). Sequence analysis of the *Spodoptera litura* multicapsid nucleopolyhedrovirus genome. *Virology* **287**, 391–404.

Pertsemlidis, A. & Fondon, J. W., III (2001). Having a BLAST with bioinformatics (and avoiding BLASTphemy). *Genome Biol* **2**, reviews2002.1–2002.10.

Pilloff, M. G., Bilen, M. F., Belaich, M. N., Lozano, M. E. & Ghiringhelli, P. D. (2003). Molecular cloning and sequence analysis of the *Anticarsia gemmatalis* multicapsid nuclear polyhedrosis virus GP64 glycoprotein. *Virus Genes* **26**, 57–69.

Podgewaite, J. D. (1985). Strategies for field use of baculoviruses. In *Viral Insecticides for Biological Control*, pp. 775–797. Edited by K. Maramorosch & K. E. Sherman. Orlando, FL: Academic Press.

Pombo, V., Velloso, L. M., Ribeiro, B. M. & Báo, S. N. (1998). Structural and ultrastructural changes during the infection of UFL-AG-286 cells with the baculovirus AgMNPV. *J Invertebr Pathol* **72**, 239–245.

Possee, R. D. & Rohrmann, G. F. (1997). Baculovirus genome organization and evolution. In *The Baculoviruses*, pp. 109–140. Edited by L. K. Miller. New York: Plenum.

Possee, R. D., Sun, T.-P., Howard, S. C., Ayres, M. D., Hill-Perkins, M. & Gearing, K. L. (1991). Nucleotide sequence of the *Autographa californica* nuclear polyhedrosis 9·4 kbp *Eco*RI-I and -R (polyhedrin gene) region. *Virology* **185**, 229–241.

Razuck, F. B., Ribeiro, B., Vargas, J. H., Wolff, J. L. & Ribeiro, B. M. (2002). Characterization of the *p10* gene region of *Anticarsia gemmatalis* nucleopolyhedrovirus. *Virus Genes* **24**, 243–247.

Ribeiro, B. M., Gatti, C. D. C., Costa, M. H., Moscardi, F., Maruniak, J. E., Possee, R. D. & Zanotto, P. M. A. (2001). Construction of a recombinant *Anticarsia gemmatalis* nucleopolyhedrovirus (AgMNPV-2D) harbouring the *β*-galactosidase gene. *Arch Virol* **146**, 1355–1367.

Rodrigues, J. C. M., de Souza, M. L., O'Reilly, D., Velloso, L. M., Pinedo, F. J. R., Razuck, F. B., Ribeiro, B. & Ribeiro, B. M. (2001). Characterization of the ecdysteroid UDP-glucosyltransferase (*egt*) gene of *Anticarsia gemmatalis* nucleopolyhedrovirus. *Virus Genes* **22**, 103–112.

Sieburth, P. J. & Maruniak, J. E. (1988a). Growth characteristics of a continuous cell line from the velvetbean caterpillar, *Anticarsia gemmatalis* Hübner (Lepidoptera: Noctuidae). *In Vitro Cell Dev Biol* **24**, 195–198.

Sieburth, P. J. & Maruniak, J. E. (1988b). Susceptibility of an established cell line of *Anticarsia gemmatalis* (Lepidoptera: Noctuidae) to three nuclear polyhedrosis viruses. *J Invertebr Pathol* **52**, 453–458.

Slack, J. M. & Shapiro, M. (2004). *Anticarsia gemmatalis* multicapsid nucleopolyhedrovirus *v-trex* gene encodes a functional 3′ to 5′ exonuclease. *J Gen Virol* **85**, 2863–2871.

Slack, J. M., Ribeiro, B. M. & de Souza, M. L. (2004). The *gp64* locus of *Anticarsia gemmatalis* multicapsid nucleopolyhedrovirus contains a 3′ repair exonuclease homologue and lacks *v-cath* and *ChiA* genes. *J Gen Virol* **85**, 211–219.

Smith, S. (2001). The world according to PARP. *Trends Biochem Sci* **26**, 174–179.

Smulson, M., Istock, N., Ding, R. & Cherney, B. (1994). Deletion mutants of poly(ADP-ribose) polymerase support a model of cyclic association and dissociation of enzyme from DNA ends during DNA repair. *Biochemistry* **33**, 6186–6191.

Soares, J. S. & Ribeiro, B. M. (2005). Pathology of *Anticarsia gemmatalis* larvae infected by two recombinant *A. gemmatalis* multicapsid nucleopolyhedroviruses. *Res Microbiol* **156**, 263–269.

Sonnhammer, E. L. L. & Durbin, R. (1995). A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. *Gene* **167**, GC1–GC10.

Stothard, P. & Wishart, D. S. (2005). Circular genome visualization and exploration using CGView. *Bioinformatics* **21**, 537–539.

Tanada, Y. & Kaya, H. K. (1993). *Insect Pathology*. San Diego, CA: Academic Press.

Tani, H., Limn, C. K., Yap, C. C. & 9 other authors (2003). In vitro and in vivo gene delivery by recombinant baculoviruses. *J Virol* **77**, 9799–9808.

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

Todd, J. W., Passarelli, A. L. & Miller, L. K. (1995). Eighteen baculovirus genes, including *lef-11*, *p35*, *39K*, and *p47*, support late gene expression. *J Virol* **69**, 968–974.

Tomalski, M. D., Eldridge, R. & Miller, L. K. (1991). A baculovirus homolog of a Cu/Zn superoxide dismutase gene. *Virology* **184**, 149–161.

Uren, A. G., Coulson, E. J. & Vaux, D. L. (1998). Conservation of baculovirus inhibitor of apoptosis repeat proteins (BIRPs) in viruses, nematodes, vertebrates and yeasts. *Trends Biochem Sci* **23**, 159–162.

van Oers, M. M., Abma-Henkens, M. H. C., Herniou, E. A., de Groot, J. C. W., Peters, S. & Vlak, J. M. (2005). Genome sequence of *Chrysodeixis chalcites* nucleopolyhedrovirus, a baculovirus with two DNA photolyase genes. *J Gen Virol* **86**, 2069–2080.

Vispé, S., Yung, T. M. C., Ritchot, J., Serizawa, H. & Satoh, M. S. (2000). A cellular defense pathway regulating transcription through poly(ADP-ribosyl)ation in response to DNA damage. *Proc Natl Acad Sci U S A* **97**, 9886–9891.

Willis, L. G., Seipp, R., Stewart, T. M., Erlandson, M. A. & Theilmann, D. A. (2005). Sequence analysis of the complete genome of *Trichoplusia ni* single nucleopolyhedrovirus and the identification of a baculovirus photolyase gene. *Virology* **338**, 209–226.

Downloaded from www.microbiologyresearch.org by
IP: 200.202.165.220
On: Tue, 22 May 2018 19:05:29

Wormleaton, S., Kuzio, J. & Winstanley, D. (2003). The complete sequence of the *Adoxophyes orana* granulovirus genome. *Virology* **311**, 350–365.

Yang, Y. L. & Li, X. M. (2000). The IAP family: endogenous caspase inhibitors with multiple biological activities. *Cell Res* **10**, 169–177.

Zanotto, P. M. A., Sampaio, M. J. A., Johnson, D. W., Rocha, T. L. & Maruniak, J. E. (1992). The *Anticarsia gemmatalis* nuclear polyhedrosis virus polyhedrin gene region: sequence analysis, gene product and structural comparisons. *J Gen Virol* **73**, 1049–1056.

Zanotto, P. M. A., Kessing, B. D. & Maruniak, J. E. (1993). Phylogenetic interrelationships among baculoviruses: evolutionary rates and host associations. *J Invertebr Pathol* **62**, 147–164.

Zemskov, E. A., Kang, W. & Maeda, S. (2000). Evidence for nucleic acid binding ability and nucleosome association of *Bombyx mori* nucleopolyhedrovirus BRO proteins. *J Virol* **74**, 6784–6789.

Zhang, C.-X., Ma, X.-C. & Guo, Z.-J. (2005). Comparison of the complete genome sequence between C1 and G4 isolates of the *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus. *Virology* **333**, 190–199.

Downloaded from www.microbiologyresearch.org by
IP: 200.202.165.220
On: Tue, 22 May 2018 19:05:29