Cloning and sequence analysis of the *Antheraea pernyi* nucleopolyhedrovirus *gp64* gene

WENBING WANG, SHANYING ZHU, LIQUN WANG, FENG YU and WEIDE SHEN*,†

Institute of Life Sciences, Jiangsu University, Xuefu Road 301, Zhenjiang 212013, People's Republic of China

**School of Agricultural Science and Technology, Suzhou University, Shizhi Road 1, Suzhou, People's Republic of China*

†*Corresponding author (Fax, 86-512-65112235; Email, shenwd@suda.edu.cn)*

Frequent outbreaks of the purulence disease of Chinese oak silkworm are reported in Middle and Northeast China. The disease is produced by the pathogen *Antheraea pernyi* nucleopolyhedrovirus (AnpeNPV). To obtain molecular information of the virus, the polyhedra of AnpeNPV were purified and characterized. The genomic DNA of AnpeNPV was extracted and digested with *Hind*III. The genome size of AnpeNPV is estimated at 128 kb. Based on the analysis of DNA fragments digested with *Hind*III, 23 fragments were bigger than 564 bp. A genomic library was generated using *Hind*III and the positive clones were sequenced and analysed. The *gp64* gene, encoding the baculovirus envelope protein GP64, was found in an insert. The nucleotide sequence analysis indicated that the AnpeNPV *gp64* gene consists of a 1530 nucleotide open reading frame (ORF), encoding a protein of 509 amino acids. Of the eight *gp64* homologues, the AnpeNPV *gp64* ORF shared the most sequence similarity with the *gp64* gene of *Anticarsia gemmatalis* NPV, but not *Bombyx mori* NPV. The upstream region of the AnpeNPV *gp64* ORF encoded the conserved transcriptional elements for early and late stage of the viral infection cycle. These results indicated that AnpeNPV belongs to group I NPV and was far removed in molecular phylogeny from the BmNPV.

[Wang W, Zhu S, Wang L, Yu F and Shen W 2005 Cloning and sequence analysis of the *Antheraea pernyi* nucleopolyhedrovirus *gp64* gene; *J. Biosci.* **30** 605–610]

1. Introduction

The Baculoviridae family virusus are mostly pathogenic to insects, containing double-stranded circular DNA molecules ranging in size from 80 to 180 kb (Ayres *et al* 1994; Ahrens *et al* 1997). Baculovirus infection progresses in three phases – early, late and very late. The infectious virion is characterized by two phenotypes – budded virus (BV) and occlusion-derived virus (ODV) (Federici and Hice 1997; Lu 1998). During the early stage of infection BVs are predominantly responsible for the systemic infection of insect cells and tissues *in vivo*. The BVs are characterized by nucleocapsids surrounded by a membrane, the main protein of which is named GP64 or GP67 (Ayres *et al* 1994; Gomi *et al* 1999; Pilloff *et al* 2003). The ODVs are occluded in polyhedra and divided into

Keywords. *Antheraea pernyi* nucleopolyhedrovirus; baculovirus; *gp64* gene; Lepidoptera

Abbreviations used: AnpeNPV, *Antherea pernyi* nucleopolyhedrovirus; BmNPV, *Bombyx mori* nucleopolyhedrovirus; BV, budded virus; NPV, nucleopolyhedrovirus; ORF, open reading frame; PCR, polymerase chain reaction.

two classes: nucleopolyhedrovirus (NPV) and granulovirus (GV). The lepidopteran NPVs can be further subdivided into groups I and II, according to the size, shape and similarity of the genes (Hayakawa *et al* 2000; Herniou *et al* 2001). The BVs in group I possess the membrane protein GP64. The GP64 protein is expressed both in early and late stage of infection and is transported to and incorporated into the cell membrane. Group II is characterized by the fusion protein instead of GP64 protein functioning in membrane fusion (Bulach *et al* 1999; Kingsley *et al* 1999; Monsma and Blissard 1999).

The Chinese oak silkworm (tussah), *Antheraea pernyi*, is commercially cultivated mainly in Middle and Northeastern China (Lu 1982). The purulence disease was reported in cultures of both *Bombyx mori* and *Antheraea pernyi* (Herniou *et al* 2003). The *B. mori* NPV (BmNPV) did not infect the larva of *A. pernyi* and vice verse (Lu 1998). The BmNPV genome was sequenced and analysed in 1999 (Gomi *et al* 1999), while that of *A. pernyi* NPV (AnpeNPV) is still not elucidated.

In this report, we describe the *A. pernyi* polyhedra purification, genomic DNA digestion, cloning and sequencing of the AnpeNPV *gp64*. The complete sequence of the *gp64* and coded product were comparatively analysed with the corresponding sequences of the other NPVs by bioinformatics tools.

2. Materials and methods

2.1 *Purification of the ApNPV polyhedra and preparation of polyhedrovirus DNA*

The larvae of *A. pernyi* infected with AnpeNPV were ground into small pieces, and diluted with sterilized water. The liquid was filtrated with three layers of flossy paper to eliminate grainy impurities. The procedures for both polyhedra and viral DNA preparation were as described by O'Reilly *et al* (1994).

2.2 *Digestion of AnpeNPV DNA*

Two μg DNA was digested with an appropriate restriction endonuclease *Hind*III overnight. The restriction fragments were resolved by 1% agarose gel electrophoresis, stained with ethidium bromide and visualized under UV. The genomic library was generated by cloning the *Hind*III digested fragments from the isolated AnpeNPV into a pUC18 plasmid. Genomic and cloned DNA inserts were confirmed by digestion by *Hind*III and were sequenced using the Taq DyeDeoxy Terminator Cycle Sequencing Kit and analysed on a DNA sequencer.

2.3 *Cloning of the full length fragment of* gp64 *gene*

According to the sequences of the genomic library, two specific primers were designed to amplify the full length fragment of *gp64* by using a PCR (primer forward: GTG CGC TGC GCT GCC GTC G; primer reverse: AAC TAT TAA TTT TTT TTT ATG TAT). The fragment was cloned into T-vector and was sequenced.

2.4 *Gene sequences and computational analysis*

The nucleotide sequences obtained from the inserts were compared with the GenBank database using the BLASTN, BLASTP and BLASTX algorithms at the NCBI web site. Multiple sequence alignments, at nucleotide and amino acid level, were carried out using the DNAstar program. The group I NPVs *gp64* were used for alignment available in the GenBank: AcNPV (*Autographa californica* NPV, L22858), AgNPV (*Anticarsia gemmatalis* NPV, AY123150), AnfaNPV (*Anagrapha falcifera* NPV, AFU64897), BmNPV (*Bombyx mori* NPV, NC_001962), CfMNPV (*Choristoneura fumiferana* MNPV, NC_004778), EppoNPV (*Epiphyas postvittana* NPV, NC_003083), HycuNPV (*Hyphantria cunea* NPV, AF190124), and OpNPV (*Orgyia pseudotsugata* NPV, NC_001875). Putative post-translational modification sites were searched using the *PROSITE* database at the *Expasy* Proteomics web server.

3. Results

3.1 *AnpeNPV genomic DNA isolation and estimation of the genome size of the virus*

The purified AnpeNPV particles were triangular or quadrangular in shape and varied in size from 2 to 5 μm (figure 1). The genomic DNA of AnpeNPV was extracted and was digested with *Hind*III. According to the analysis of the fragments based on gel migration, there are twenty three of the DNA fragments digested by *Hind*III bigger than 564 bp (figure 2). The genome size of AnpeNPV is estimated at 128 kb according to the calculation by the AlphaEase™ software (Alpha Innotech Corparation San Leandro, CA, USA).

3.2 *Cloning and sequencing of the AnpeNPV gp64 locus*

The *Hind*III digested fragments from the isolated AnpeNPV DNA were inserted into a pUC18 plasmid. A *Hind*III genomic library was generated, and the positive clones were sequenced and analyzed. An insert sequence

was identified to be homologous with baculovirus *gp64* gene, which contains 5′ flank and the most parts of the open reading frame (ORF) of *gp64* (accession number AY854039). However, the 3′ terminal sequence was missing. A sequence containing 3′ terminal of AnpeNPV *gp64* was found in GenBank (accession number: AB072731). The primers were designed to amplify a putative *gp64* ORF specific fragment (about 1500 bp long). After purifying, cloning and sequencing of the amplified fragment, we confirmed that the polymerase chain reaction (PCR) product corresponds to the AnpeNPV *gp64* ORF.

3.3 *General sequence analysis*

The PCR product of *gp64* was inserted into the plasmid pUC18. The sequencing result showed that the insert contains a 1530 nucleotide long ORF flanked by 187 and 15 bp of 5′ UTR and 3′ UTR regions, respectively (figure 3). The *gp64* ORF codes for the predicted 509 amino-acid-long polypeptide with an estimated molecular mass of 58·7 kDa. This sequence includes a putative signal peptide corresponding to the 18 most amino terminal residues.

A comparative analysis with the other available baculovirus sequences shows that the similarity of AnpeNPV *gp64* ORF to the other NPVs ranged from 69·0% to 80·1% (table 1). The AnpeNPV *gp64* ORF nucleotide sequence and the deduced amino acid sequence are more similar to those of AgNPV than to the homologous gene of the other NPVs (80·1% and 85·4%, respectively).

3.4 *Transcriptional regulatory regions*

The AnpeNPV *gp64* 5′ UTR contains an early INR (transcription initiator) element (CAGT) located at position – 29 and two late INR elements (TAAG) located at positions – 110 and – 117. In addition to the putative transcriptional regulatory elements found in other NPVs *gp64* locus (Kogan and Blissard 1994), the AnpeNPV *gp64* UTR contains two GATA motifs (positions – 72 and – 120), one CGT motif (position – 77) and one TATA box (position – 57).

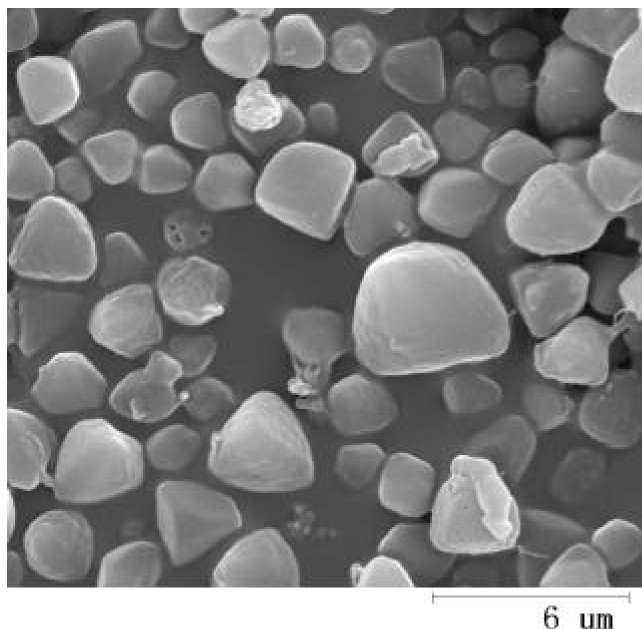

Figure 1. Observation of the AnpeNPV polyhedra by electron microscope. The particles appear in the shape of triangles and quadrangles varying in size from 2 to 5 μm.

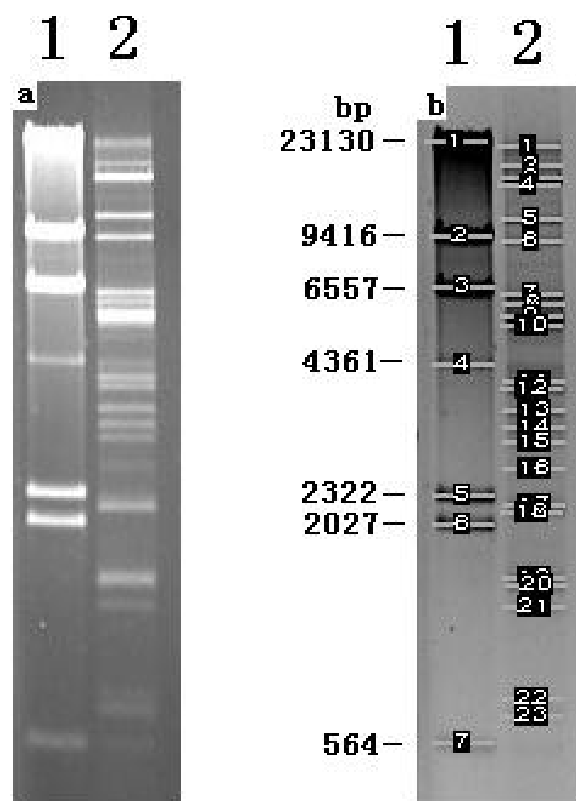

Figure 2. Restriction map of AnpeNPV *Hind*III genomic DNA fragments. Two μg DNA was digested with an appropriate restriction endonuclease *Hind*III overnight. Lane 1, Molecular weight marker (1DNA digested with *Hind*III); lane 2, AnpeNPV DNAs digested with restriction endonuclease *Hind*III. (**a**) The AnpeNPV DNA fragments resolved by 1% agarose gel electrophoresis; (**b**) the fragment size calculated by the AlphaEase^{TM} software (Alpha Innotech Corparation, USA). The genome size of AnpeNPV is estimated at 128 Kb pairs.

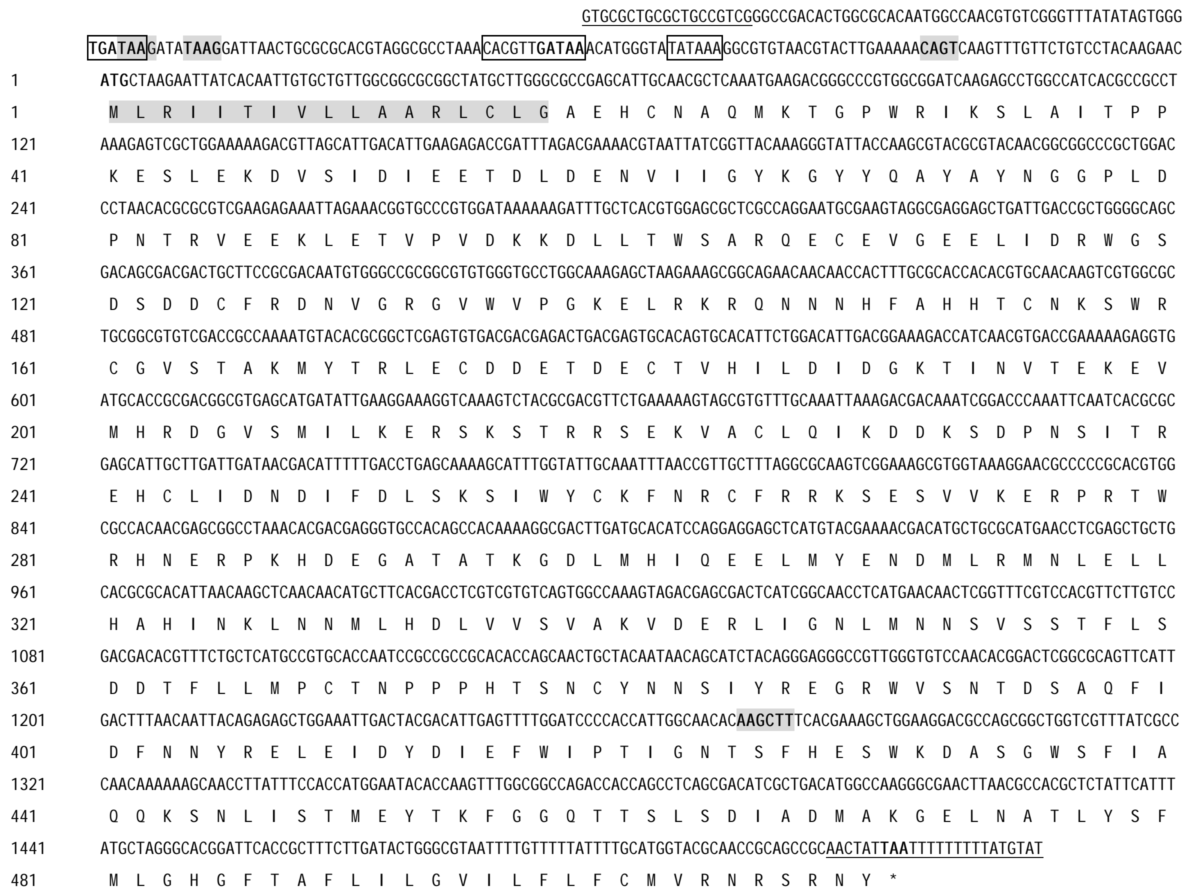

Figure 3. Nucleotide sequences of the gp64 ORF and untranslated flanking regions. The *gp64* ORF is included in the 1732 nucleotides region shown here; the deduced amino acid is presented below the coding region (the shaded N-terminal sequence indicates the putative signal peptide). The initiation and termination codons are in bold. The primer sequences are underlined. The first nucleotide of the translation initiation codon is designated as position 1. Putative early promoter elements (i.e. GATA motif: TGATAA, TATA box: TATAAA, CGT motif: CACGTT) are boxed. The sequence motifs associated with the initiation of transcription are black shaded (early INR: CAGT and late INR: TAAG), and the site of *Hind*III is black shaded as well.

4. Discussion

To date, over 500 NPVs have been described (Lu 1998). In the nineteen seventies, the NPVs were classified based on their shape visualized under microscope and the host larva infected. The nomenclature is easy to be confused because one NPV could infect some insects. It is also hard to predict the host range of the NPV. By now only about 30 NPVs' molecular information can be found in GenBank (Ayres *et al* 1994; Ahrens *et al* 1997; Gomi *et al* 1999; Ijkel *et al* 1999; Kuzio *et al* 1999; Taha *et al* 2000; Afonso *et al* 2001; Harrison and Bonning 2003). The knowledge of viral genes is very important for studying the host range and infection of the NPVs (Cheng *et al* 2005; Jakubowska *et al* 2005). Restriction enzyme mapping is useful to detecting the type of NPV. There are twenty three of the AnpeNPV DNA fragments digested by *Hind*III bigger than 564 base pairs. Upon comparing the physical maps of NPVs, we found that the AnpeNPV possesses a new type of NPV. We also report a new AnpeNPV *gp64* gene. It contains 509 amino acids. The AnpeNPV *gp64* ORF sequence is the most similar to that of AgNPV within the Lepidopteran group I NPVs. Interestingly, The GP64 from AgNPV is the smallest baculoviral envelope protein found to date, which contains 499 amino acids (Pilloff *et al* 2003). To this end, including AnpeNPV, nine of NPV GP64 proteins were analyzed. The length of *gp64* ORF varies from 499 to 530 amino acids. The results of homologue analysis showed that the similarity of GP64 within the NPVs is not related to the length of the *gp64* ORF. The cloning of *gp64* gene indicated that AnpeNPV belongs to group I NPV.

Table 1. Relative similarity of the group I NPV GP64s.

	AcNPV	AgNPV	AnfaNPV	BmNPV	CfMNPV	EppoNPV	HycuNPV	OpNPV
(A) Nucleic acid identity								
AnpeNPV	70·3	80·1	70·5	69·0	74·5	69·5	72·4	76·3
AcNPV		67·7	96·4	95·1	71·1	66·3	65·0	70·5
AgNPV			67·4	67·1	71·5	68·1	69·3	72·1
AnfaNPV				93·2	70·3	65·9	62·4	69·6
BmNPV					70·1	66·1	63·7	69·0
CfMNPV						69·7	73·3	80·7
EppoNPV							67·1	69·3
HycuNPV								74·6
(B) Amino acid identity								
AnpeNPV	75·6	85·4	75·6	73·9	78·8	75·6	76·6	80·4
AcNPV		73·7	98·2	94·7	79·8	74·1	72·9	78·6
AgNPV			73·5	73·7	75·4	74·3	73·3	76·8
AnfaNPV				94·3	79·4	74·1	73·3	78·2
BmNPV					78·2	73·7	71·5	78·2
CfMNPV						78·8	80·6	88·4
EppoNPV							74·3	78·4
HycuNPV								82·1

AcNPV, Autographa californica NPV; AgNPV, Anticarsia gemmatalis NPV; AnfaNPV, Anagrapha falcifera NPV; BmNPV, Bombyx mori NPV; CfMNPV, Choristoneura fumiferana MNPV; EppoNPV, Epiphyas postvittana NPV; HycuNPV, Hyphantria cunea NPV; and OpNPV, Orgyia pseudotsugata NPV.

Acknowledgements

This work was supported by grants from the National High Technology and Development Program of China (No. 2004AA2Z1020), the Education Committee of Jiangsu Province (No. 02KJD18003) and the Nature science Foundation of Jiangsu University (04CX08).

References

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

Ahrens C H, Russell R L Q, Funk C J, Evans J T, Harwood S H and Rohrmann G F 1997 The sequence of the *Orgyia pseudotsugata* multinucleocapsid nuclear polyhedrosis virus genome; *Virology* **229** 381–399

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

Bulach D M, Kumar C A, Zaia A, Liang B and Tribe D E 1999 Group II nucleopolyhedrovirus subgroups revealed by phylogenetic analysis of polyhedrin and DNA polymerase gene sequences; *J. Invertebr. Pathol.* **73** 59–73

Cheng X, Carner G R, Lange M, Jehle J A and Arif B M 2005 Biological and molecular characterization of a multicapsid nucleopolyhedrovirus from *Thysanoplusia orichalcea* (L.) (Lepidoptera: Noctuidae); *J. Invertebr. Pathol.* **88** 126–135

Federici B A and Hice R H 1997 Organization and molecular characterization of genes in the polyhedrin region of the *Anagrapha falcifera* multinucleocapsid NPV; *Arch. Virol.* **142** 333–348

Gomi S, Majima K and Maeda S 1999 Sequence analysis of the genome of *Bombyx mori* nucleopolyhedrovirus; *Gen. Virol.* **80** 1323–1337

Harrison R L and Bonning B C 2003 Comparative analysis of the genomes of *Rachiplusia ou* and *Autographa californica* multiple nucleopolyhedroviruses; *J. Gen. Virol.* **84** 1827–1842

Hayakawa T, Rohrmann G F and Hashimoto Y 2000 Patterns of genome organization and content in lepidopteran baculoviruses; *Virology*. **278** 1–12

Herniou E A, Luque T, Chen X, Vlak J M, Winstanley D, Cory J S and 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 and O'Reilly D R 2003 The genome sequence and evolution of baculoviruses; *Annu. Rev. Entomol.* **48** 211–234

Ijkel W F, van Strien E A, Heldens J G, Broer R, Zuidema D, Goldbach R W and Vlak J M 1999 Sequence and organization of the *Spodoptera exigua* multicapsid nucleopolyhedrovirus genome; *J. Gen. Virol.* **80** 3289–3304

Jakubowska A, Oers M M, Cory J S, Ziemnick J and Vlak J M 2005 European *Leucoma salicis* NPV is closely related to North American *Orgyia pseudotsugata* MNPV; *J. Invertebr. Pathol.* **88** 100–107

Kingsley D H, Behbahani A, Rashtian A, Blissard G W and Zimmerberg J 1999 A discrete stage of baculovirus GP64-mediated membrane fusion; *Mol. Biol. Cell.* **10** 4191–4200

Kogan P H and Blissard G W 1994 A baculovirus gp64 early promoter is activated by host transcription factor binding to CACGTG and GATA elements; *J. Virol.* **68** 813–822

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

Lu H S 1982 Insect viruses and insect virus diseases (Beijing: Science Press)

Lu H S 1998 *Molecular biology of insect viruses* (Beijing: China Agricultural Scientech Press)

Monsma S A and Blissard G W 1999 Identification of a membrane fusion domain and an oligomerization domain in the baculovirus GP64 envelope fusion protein; *J. Virol*. **69** 2583–2595

O'Reilly D R, Miller L K and Luckow V A 1994 Baculovirus expression vectors: *A laboratory Manual* (New York: Oxford University Press)

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

Taha A, Nour-El-Din A, Croizier L, Ferber M L and Croizier G 2000 Comparative analysis of the granulin regions of the *Phthorimaea operculella* and *Spodoptera littoralis* granuloviruses; *Virus Genes* **21** 147–155

MS received 25 May 2005; accepted 9 September 2005

*e*Publication: 15 November 2005

Corresponding editor: SHAHID JAMEEL