

Identification, Characterization and Phylogenic Analysis of Conserved Genes within the p74 Gene Region of *Choristoneura fumiferana* Granulovirus Genome

Kianoush Khajeh Rashidan^{†,*}, Nasha Nassoury[‡], Paresa N. Giannopoulos[‡], Yves Mauffette[§] and Claude Guertin[†]

[†]Institut national de la recherche scientifique-Institut Armand-Frappier 531 Blvd des Prairies, H7V 1B7, Laval, Canada

[‡]Université de Montréal, 4101 Sherbrook east, H1X 2B2 Montreal, Canada

[§]Université du Québec à Montréal, Succursale Centre Ville CP8888, H3C 3P8 Montreal, Canada

Received 4 May 2004, Accepted 21 June 2004

The genes located within the p74 gene region of the *Choristoneura fumiferana* granulovirus (ChfuGV) were identified by sequencing an 8.9 kb *Bam*HI restriction fragment on the ChfuGV genome. The global guanine-cytosine (GC) content of this region of the genome was 33.02%. This paper presents the ORFs within the p74 gene region along with their transcriptional orientations. This region contains a total of 15 open reading frames (ORFs). Among those, 8 ORFs were found to be homologues to the baculoviral ORFs: Cf-i-p, Cf-vi, Cf-vii, Cf-viii (*ubiquitin*), Cf-xi (*pp31*), Cf-xii (*lef-11*), Cf-xiii (*sod*) and Cf-xv-p (p74). To date, no specific function has been assigned to the ORFs: Cf-i, Cf-ii, Cf-iii, Cf-iv, Cf-v, Cf-vi, Cf-vii, Cf-ix and Cf-x. The most noticeable ORFs located in this region of the ChfuGV genome were *ubiquitin*, *lef-11*, *sod*, *fibrillin* and p74. The phylogenetic trees (constructed using conceptual products of major conserved ORFs) and gene arrangement in this region were used to further examine the classification of the members of the granulovirus genus. Comparative studies demonstrated that ChfuGV along with the *Cydia pomonella* granulovirus (CpGV), *Phthorimaea operculella* granulovirus (PhopGV), *Adoxophyes orana* granulovirus (AoGV) and *Cryptophlebia leucotreta* granulovirus (CIGV) share a high degree of amino acids sequence and gene arrangement preservation within the

studied region. These results support a previous report, which classified a granuloviruses into 2 distinct groups: Group I: ChfuGV, CpGV, PhopGV and AoGV and Group II: *Xestia c-nigrum* granulovirus (XcGV) and *Plutella xylostella* granulovirus (PxGV). The phylogenetic and gene arrangement studies also placed CIGV as a novel member of the Group I granuloviruses.

Keywords: Baculovirus, *Choristoneura fumiferana* granulovirus, Gene arrangement, Phylogeny

Introduction

Granuloviruses (GV) are members of the family *Baculoviridae*, which also includes nucleopolyhedroviruses (NPV) (Miller, 2001). The GV genome is a double-stranded super-coiled DNA (usually > 100 kb) that is packaged within a single nucleocapsid (Hayakawa *et al.*, 1999; Hashimoto *et al.*, 2000; Luque *et al.*, 2001; Lange and Jehle, 2003; Wormleaton *et al.*, 2003).

Generally, there is only one nucleocapsid per envelope, and together they are occluded within a paracrystalline protein matrix that is called granulin. Like the polyhedrin of the NPVs, granulin protects the virions against environmental effects and is involved in the horizontal transmission of the virus (Miller, 2001).

Baculoviral gene expression generally occurs in three stages, namely the early, late and very late stages, which is controlled in a regulatory cascade manner (Blissard, 1996; Hess and Falcon, 1997). All the genes are located randomly throughout the genome, even though their products have distinct regulatory or structural functions (Ayres *et al.*, 1994; Ahrens *et al.*, 1997; Gomi *et al.*, 1999; Hayakawa *et al.*, 1999;

The nucleotide sequence data reported in this paper has been submitted to the GenBank nucleotide sequence database and has been assigned the accession numbers ANN77187, ANN77188, ANN77189, ANN77190, ANN77191, ANN77192, ANN77193, ANN77194, ANN77195, ANN77196, ANN77197, ANN77198, ANN77199 and ANN77200.

*To whom correspondence should be addressed.

Tel: 1-514-987-3000; Fax: 1-514-987-4647

E-mail: kiarashidan@juno.com

Ijkel *et al.*, 1999; Kuzio *et al.*, 1999; Hashimoto *et al.*, 2000; Luque *et al.*, 2001; Pang *et al.*, 2001; Chen *et al.*, 2002; Li *et al.*, 2002; Lange and Jehle, 2003; Wormleaton *et al.*, 2003).

The *Choristoneura fumiferana* granulovirus (ChfuGV) is pathogenic towards the spruce budworm (*Choristoneura fumiferana*, (Clem.), (Lepidoptera: Tortricidae), which is responsible for the defoliation of coniferous trees in eastern North American forests. ChfuGV is believed to be an alternative means to control spruce budworm populations to chemical insecticides.

There is limited knowledge about the molecular characteristic of the ChfuGV. Therefore, as part of an ongoing investigation into the nature of the ChfuGV, this study concentrated on the molecular characterization of this virus. A previous communication reported the identification and characterization of the ORFs located within the *odvp-6e/odv-e56* gene region of the ChfuGV (Rashidan *et al.*, 2004). This paper reports the identification and characterization of the ORFs located within the p74 gene region of the ChfuGV. Gene p74 is a highly conserved baculoviral gene coding an envelope protein, and the complete sequence and characterization of this gene along with its product in the ChfuGV has been reported recently (Rashidan *et al.*, 2003). The ORFs located within the p74 gene region of the ChfuGV are positioned downstream to those within the *odvp-6e/odv-e56* gene region.

Materials and Methods

In vivo production and purification of virus and DNA extraction *Choristoneura fumiferana* fourth-instar larvae was infected by ChfuGV using a virus contaminated artificial diet (Forté *et al.*, 1999). The virus isolated from the infected larvae was carried out as previously described (Bah *et al.*, 1997).

DNA extraction, cloning and sequencing The ChfuGV genomic DNA was extracted from a purified enveloped nucleocapsids, as already described (Rashidan *et al.*, 2002). An 8.9 kb *Bam*HI restriction fragment was cloned into a pBlueScript-SK⁺ cloning vector (Stratagen, La Jolla, USA), transformed into *E. coli* XL-1 Blue (Stratagen) and then sequenced using the techniques reported elsewhere (Rashidan *et al.*, 2002).

Computer analysis The sequence data assembly and analysis were performed using Sequencer software version 4.0.5 (Gene Codes Corp.) and the MacVector program version 4.5.0 (Eastman Kodak). The nucleotide sequence and its predicted amino acid sequence were compared with the homologues in GenBank/EMBL and SWISSPROT using BLAST (Altschul *et al.*, 1990). CLUSTALX (1.81) (Thompson *et al.*, 1997) was used for the multiple amino acid sequence alignments. BOXSHADE version 3.21 was used for the similarity shading and scoring among the aligned sequences. The alignments were then used as the input to construct the phylogenetic tree using the Branch and Bound search settings of PAUP 4.0b4a employing Neighbor joining (NJ) method (Swofford, 2000). The reliability of the phylogenetic trees was evaluated by bootstrap analysis using 1,000 bootstrap replicates.

Assignment of putative ORFs The information obtained from the sequencing of the 8.9 kb genomic fragment of the ChfuGV was used to detect the homologue sequences in the Genbank database (whole database and restricted database containing only virus sequences) using the BLASTN (Altschul *et al.*, 1990) and Fasta3 (Pearson and Lipman, 1988) programs. On the other hand, the putative ORFs were first translated from the ChfuGV nucleotide sequences, and amino acid sequences were compared with the existing sequences in Genbank database using the BLASTP program (Altschul *et al.*, 1990). These results along with the result obtained from the gene order from the other granulovirus genomes were used to identify the ChfuGV ORFs.

Results and Discussion

The open reading frames (ORFs) located within the p74 gene region of the *Choristoneura fumiferana* granulovirus (ChfuGV) were identified by sequencing an 8.9 kb *Bam*HI restriction fragment on the ChfuGV genome (Fig. 1). The global GC content of this region of the genome was 33.02%. The region examined contained a total of 15 ORFs. Among those, 8 ORFs were found to be homologues to aculoviral ORFs such as Cf-i-p (*p* for partial), Cf-vi, Cf-vii, Cf-viii (*ubiquitin*), Cf-xi (*pp31*), Cf-xii (*lef11*), Cf-xiii (*sod*) and Cf-xv-p (p74). To date, no specific function has been assigned to the ORFs, Cf-i, Cf-ii, Cf-iii, Cf-iv, Cf-v, Cf-vi, Cf-vii, Cf-ix and Cf-x (Table 1).

One of the ORFs codes was identified as a conceptual protein similar to the ubiquitin protein (*ubique*, Latin for everywhere). Ubiquitin is a globular protein that is found in all eukaryotic cells and has a sequence that is extremely well conserved from simple organisms to more sophisticated organisms such as vertebrates. A putative ORF (Cf-viii) of 288 nt located 7 ORFs upstream to p74 potentially encodes a baculoviral ubiquitin-like protein homologue. This is a highly conserved gene that is found in all known baculoviral genomes (Ayres *et al.*, 1994; Ahrens *et al.*, 1997; Gomi *et al.*, 1999; Hayakawa *et al.*, 1999; Ijkel *et al.*, 1999; Kuzio *et al.*, 1999; Hashimoto *et al.*, 2000; Luque *et al.*, 2001; Pang *et al.*, 2001; Chen *et al.*, 2002; Li *et al.*, 2002; Lange and Jehle, 2003; Wormleaton *et al.*, 2003).

In eukaryotic cells, ubiquitin plays important roles in several cellular processes, such as the ATP-dependent selective degradation of cellular proteins, the maintenance of the chromatin structure, regulation of gene expression, programmed cell death, stress response, survival of protein-damaging stresses and ribosome biogenesis DNA repair. The key role of ubiquitin in these procedures is to provide a tag that earmarks proteins for degradation by the proteasome. The targets of the ubiquitin system include proteins that are malfunctioning due to a mutation or the exposure of the cell to physical or chemical stresses, or are naturally short-lived proteins such as cellular regulators. Beside the primary function of ubiquitin, which is to target the proteins for degradation by the proteasome, other noncanonical functions

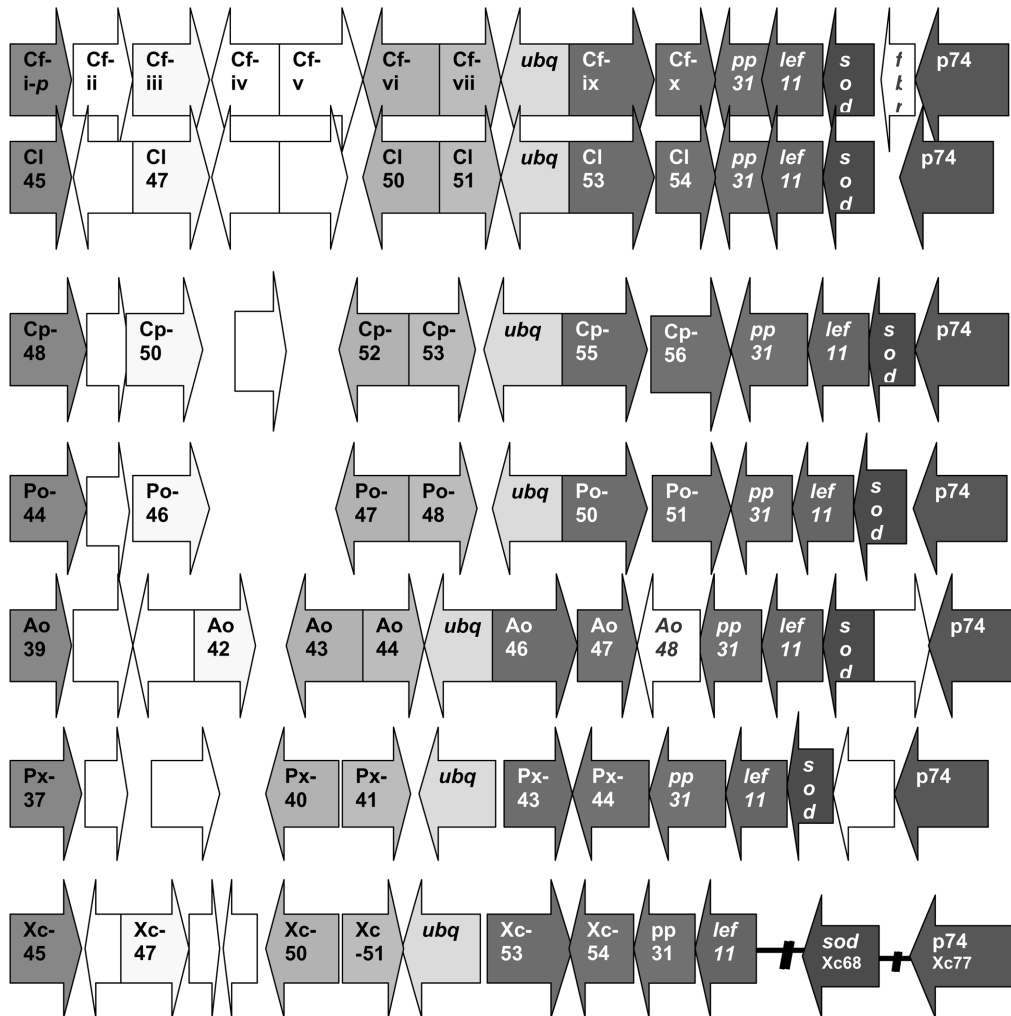


Fig. 1. Schematic illustration of the gene structure and orientation of the *p74*-containing region of the various GVs (i- ChfuGV, ii- ClGV, iii-CpGV, iv-PhopGV, v-AoGV, vi-PxGV, vii-XcGV). The arrows indicate the relative length and the direction of transcription of the ORFs.

have also been suggested. Among the other functions, the role of ubiquitin in controlling protein trafficking within the cell is well established. It has been shown that ubiquitination is essential for the endocytosis of a number of yeast plasma membrane proteins (Hicke, 2001; Pickart, 2001). In most species, there are many genes coding for ubiquitin. However they can be classified into two groups. The first group produces the polyubiquitin molecules that consist of exact head to tail repeats of ubiquitin. The number of repeats varies. Moreover, there are a number of proteins that are evolutionary related to ubiquitin. These include the ubiquitin-like proteins from the baculoviruses proteins (Hicke, 2001; Archibald *et al.*, 2003).

Another important ORF found in the *p74* region is the ORF Cf-xiii, which has the potential to code a 161 amino acid polypeptide homologue to baculoviral copper/zinc superoxide dismutase (SOD). This ORF is located two ORFs upstream of the *p74* on ChfuGV genome. Three evolutionarily distinct

families of the SODs are known, which are classified according to the metal cofactor: Cu/Zn, Fe and Mn types. The Fe and Mn forms are similar in their primary, secondary and tertiary structures, but are different from the Cu/Zn form. Prokaryotes and protists contain Mn, Fe or both types, while most eukaryotic organisms make use of the Cu/Zn type (Van Camp *et al.*, 1990). The ChfuGV SOD homologue along with other baculoviral SOD homologue proteins is also a Cu/Zn type.

SOD is the best known and perhaps most important of the antioxidant enzymes. SOD is a ubiquitous metalloprotein that catalyzes the destruction (dismutation) of the O_2^- free radical ($2O_2^- + 2H^+ \rightarrow O_2 + H_2O_2$). SOD converts the very harmful free radicals super oxide to the less active peroxide, which can then be further converted into water by other antioxidant enzymes (Petkau *et al.*, 1975; Malmström *et al.*, 1975; Schinina *et al.*, 1989). SOD is widespread in nature and is present in all oxygen-metabolizing cells (Gregory *et al.*,

Table 1. Properties of the ChfuGV ORFs detected in an 8.9 kb *Bam*HI restriction fragment and their related ORFs on the CpGV, PhopGV, AoGV CIGV, PxGV and XcGV genomes

ChfuGV			CpGV			PhopGV			AoGV			CIGV		
ORF	length (nt)	Dir	ORF	length (nt)	Dir	ORF	length (nt)	Dir	ORF	length (nt)	Dir	ORF	length (nt)	Dir
Cf-i-p	-----	>	Cp48	1119	>	Po44	1101	>	Ao39	1128	>	C145	1122	>
Cf-ii	303	>	Cp49	390	<	Po45	204	<	Ao40	417	>	C146	255	<
Cf-iii	1098	>	Cp50	2184	>	Po46	2433	>	Ao42	2628	>	C147	2145	>
Cf-iv	735	<										C148	477	<
Cf-v	621	>	Cp51	639	>							C149	1746	>
Cf-vi	633	<	Cp52	1029	<	Po47	909	<	Ao43	609	<	C150	621	<
Cf-vii	150	>	Cp53	147	>	Po48	144	>	Ao44	147	>	C151	156	>
Cf-viii ubiquitin	288	<	Cp54	ubiquitin 285	<	Po49 ubiquitin	285	<	Ao45	ubiquitin 285	<	C152	ubiquitin 231	<
Cf-ix	1032	>	Cp55	981	>	Po50	1020	>	Ao46	1050	>	C153	1047	>
Cf-x	141	>	Cp56	210	>	Po51	147	>	Ao47	228	>	C154	150	>
									Ao48	360	<			
Cf-xi <i>pp31</i>	756	<	Cp57 <i>pp31</i>	726	<	Po52 <i>pp31</i>	696	<	Ao49 <i>pp31</i>	837	<	C155 <i>pp31</i>	678	<
Cf-xii <i>lef-11</i>	279	<	Cp58 <i>lef-11</i>	405	<	Po53 <i>lef-11</i>	267	<	Ao50 <i>lef-11</i>	285	<	C156 <i>lef-11</i>	285	<
Cf-xiii <i>sod</i>	486	<	Cp59	399	<	Po54 <i>sod</i>	501	<	Ao51 <i>sod</i>	459	<	C157 <i>sod</i>	426	<
Cf-xiv <i>fibrillin</i>	315	<				Ao52	171	>						
Cf-xv P74-p	-----	<	Cp60 p74	2067	<	Po55 p74	1977	<	Ao53 p74	1929	<	C158 p74	2013	<

Most homologous ORFs

Cf-i, Cp48, Po44, Ao39, C145, Px37, Xc45	Cf-vi, Cp52, Po47, Ao43, C150, Px40, Xc49	Cf-vii, Cp53, Po48, Ao44, C151, Px41, Xc51	Cf-viii, Cp54, Po49, Ao45, C152, Px42, Xc52	Cf-xi, Cp57, Po52, Ao49, C155, Px45, Xc55
--	---	--	---	---

1974). The genes coding SOD have been reported in several eukaryotic viruses as well, for example in vertebrate poxviruses, a gene that encodes the homologues of the cellular cupro-zinc superoxide dismutases (Cu-Zn SOD). Sequencing of the *sod*-like genes from various poxviruses has demonstrated that the protein is highly conserved in all the viruses sequenced (Almazan *et al.*, 2001; Cao *et al.*, 2002). The gene encode SOD is present in almost all baculoviruses. The only exception is *Epiphyas postvittana* nucleopolyhedrovirus (EpNPV), which lacks a homologue of the superoxide dismutase (*sod*) gene (Hyink *et al.*, 2002). The role of SOD in baculovirus infections is unclear but considering the function of SOD as an antioxidant enzyme, it is possible that a baculoviral SOD might play a role in the anti-apoptotic process that occurs during the virus infection cycle. Oxidants may be essential biochemical intermediates in the progression of many forms of apoptosis induced by different stimuli.

Exposure of various cell types to oxidative stress-causing agents can induce apoptosis, which can be blocked by a wide range of antioxidants such as SOD (Hockenbery *et al.*, 1993; Buttke and Sandstorm, 1994; Wolf *et al.*, 1994; Atlante *et al.*, 2003).

Another ORF identified (Cf-xi) potentially codes for the protein known as the baculoviral 39k/pp31 protein. This protein is a nuclear matrix associate protein that is essential for the late expression during the baculoviral infection cycle (Gong *et al.*, 1998). The baculoviral 39k/pp31 protein is believed to play a role either at the transcriptional level or in mRNA stabilization. (Lu and Miller, 1995).

Another conserved baculoviral gene homologue that is present in this region of the ChfuGV genome is a putative late expression factor (*lef-11*). The baculoviral *lef-11* gene has been reported to be a gene of a great importance for the baculoviral late gene expression and is essential for viral DNA

Table 1. Continued

PxGV			XcGV		
ORF	length (nt)	Dir	ORF	length (nt)	Dir
Px37	1107	>	Xc45	1167	>
			Xc46	246	<
Px38	462	>	Xc47	663	>
Px39	1131	>	Xc48	2520	>
Px40	621	<	Xc49	546	<
			Xc50	819	<
Px41	162	>	Xc51	162	>
Px42	345	<	Xc52	234	<
ubiquitin			ubiquitin		
Px43	1245	>	Xc53	1062	>
Px44	393	<	Xc54	333	>
Px45	759	<	Xc55	888	<
<i>pp31</i>			<i>pp31</i>		
Px46	291	<	Xc56	309	<
lef-11			lef-11		
Px47	462	<	Xc68	462	<
sod			sod		
Px48	156	<			
Px49			Xc77	2133	<
p74	1737	<	P74		

Most homologous ORFs:	Cf-xii, Cp58, Po53, Ao50, Cl56, Px46, Xc56	Cf-xiii, Cp59, Po54, Ao51, Cl57, Px47, Xc68	Cf-xv, Cp60, Po55, Ao53, Cl58, Px49, Xc77
-----------------------	--	---	---

replication during the infection cycle to support the optimal levels of the transient expression from the baculoviral late promoter (Lin *et al.*, 2001; Lin and Blissard 2002). The ORF encodes the ChfuGV LEF-11 homologue, which is situated three ORFs upstream to ChfuGV p74, and has the potential to encode 92 amino acids. Phylogenetic analysis of the LEF-11 proteins separated the GVs from the NPVs, and demonstrates that ChfuGV LEF-11 is closely related to the *Phthorimaea operculella* granulovirus (PhopGV), the *Cryptophlebia leucotreta* granulovirus (CIGV), and the *Cydia pomonella* granulovirus (CpGV) (Fig. 2).

Sequence analysis of the p74 gene region also revealed the presence of an ORF (Cf-xiv) containing 315 nucleotides with a homology to the baculoviral *fibrillin* proteins. This protein owes its name to its fibrillar structure (van Oers and Vlak, 1997), and is located adjacent to the p74 gene, which encodes a highly conserved viral envelope protein. The presumed ChfuGV *fibrillin* gene codes a 104 amino acid protein with a molecular weight equal to 11811 Da. To date, several *fibrillin* genes have been identified in different baculoviruses. These

include *Autographa californica* multiple nucleopolyhedrovirus (AcMNPV) (Kuzio *et al.*, 1984), *Orgyia pseudotsugata* multicapsid nucleopolyhedrovirus (OpMNPV) (Leisy *et al.*, 1986), *Bombyx mori* nucleopolyhedrovirus (BmNPV) (Hu *et al.*, 1994), *Perina nuda* nucleopolyhedrovirus (PnNPV) (Genbank/EMBL accession no. AAC13879), *Spodoptera exigua* nucleopolyhedrovirus (SeNPV) (Zuidema *et al.*, 1993), *Choristoneura fumiferana* nucleopolyhedrovirus (CfNPV) (Wilson *et al.*, 1995), *Hyphantria cunea* nucleopolyhedrovirus (HcNPV) (Lee *et al.*, 1996a), *Lymantria dispar* nucleopolyhedrovirus (LdNPV) (Kuzio *et al.*, 1999), *Spodoptera litura* nucleopolyhedrovirus (SINPV) (Faktor *et al.*, 1997; Wei *et al.*, 1998), *Mamestra configurata* nucleopolyhedrovirus (McNPV) (Li *et al.*, 1997; Li *et al.*, 2002), *Trichoplusia ni* nucleopolyhedrovirus (TnNPV) (Fielding and Davison, 2000; Fielding *et al.*, 2002), *Buzura suppressaria* nucleopolyhedrovirus (BsNPV) (van Oers *et al.*, 1998), *Helicoverpa armigera* nucleopolyhedrovirus (HaNPV) (Wang *et al.*, 2001), *Helicoverpa zea* nucleopolyhedrovirus (HzNPV) (Chen *et al.*, 2002), *Anticarsia gemmatilis* nucleopolyhedrovirus (Razuck *et al.*, 2002), EpNPV (Genbank/EMBL accession no. AAK85684) and *Rachiplusia ou* nucleopolyhedrovirus (RoNPV) (Genbank/EMBL accession no. AAN28152).

The baculoviral *fibrillin* gene product along with the baculoviral occlusion matrix protein (granulin or polyhedrin) is generally produced in substantial quantities throughout the late stages of the baculoviral infection cycle. Studies have implicated *fibrillin* (p10) in nuclear disintegration, occlusion body formation and the release of the occlusion body (Rohrmann, 1986; Williams *et al.*, 1989; van Oers *et al.*, 1993; Lee *et al.*, 1996b). It has also been reported that *fibrillin* is not a structural protein but it can be found in fibrillar structures in the cytoplasm of the infected insect cells (Vlak *et al.*, 1988; van Oers *et al.*, 1994). Several studies have shown that the *fibrillin* genes in the AcMNPV, OpMNPV, CfMNPV, SeNPV, and SINPV are located adjacent to the p74 genes on the genome of these viruses (Leisly *et al.*, 1986; Zuidema *et al.*, 1993; Ayers *et al.*, 1994; Wilson *et al.*, 1995; Ahrens *et al.*, 1997; Faktor *et al.*, 1997).

Although the gene between ChfuGV and other baculoviruses is conserved, comparative sequence analysis of the various *fibrillin* proteins have revealed significant deviations in the amino acid sequences (data not shown). However, a relatively higher level of similarity in their secondary structures, such as the existence of a coiled-coil region in the N-terminal, can be found in the majority of the *fibrillin* proteins. This similarity suggests the existence of a parallel function for the baculoviral *fibrillin* protein homologues. The gene homology might not only provide information on the relatedness amongst baculoviral *fibrillin* homologues but also the location of the *fibrillin* genes. Comparative studies revealed that the arrangement of p74 and the *fibrillin* genes in ChfuGV was identical to those of the AcMNPV, OpMNPV, SeNPV, TnNPV, CfNPV, McNPV,

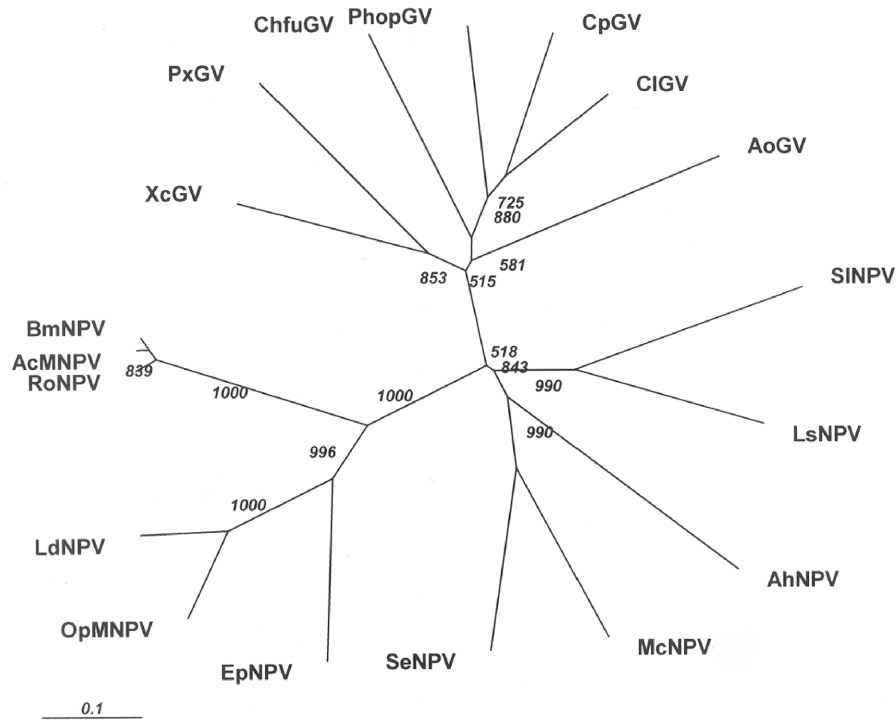


Fig. 2. Phylogenetic analysis of the conceptual protein product of the baculoviral late expression factor 11 (*lef-11*) genes. The unrooted tree was produced using the Neighbor joining (NJ) method with PAUP4.0b4a. The bootstrap values (1000 replicates) are shown. The tree was constructed according to the following settings: 1- uninformative characters ignored. 2- Branch-and bound search options: an initial upper bound: compute via stepwise; b keep minimal only; c collapse zero-length branches ON; MULPARS ON; e addition sequences: furthest. Accession numbers for the sequences used are as follow: ChfuGV AAN77198; CpGV AAK70718; CIGV AAQ21651; PhopGV AAM70251; XcGV AAF05170; AoGv AAP85687; PxGV AAG27344; McNPV AAM95134; SeNPV AAF33648; AhNPV AP006270; SINPV AAL01720; RoNPV AAN28128; LsNPV AB009614; AcMNPV AAA66667; BmNPV AAC63712; LdNPV AAC70230; OpMNPV BAA02637; EpNPV AAK85585.

SINPV, EpNPV, and RoNPV (data not shown). The arrangement of the p74 and *fibrillin* genes in the LdNPV was different, even though they were located within the same region of the genome and maintained a similar direction of the transcription compared with the abovementioned baculoviruses.

The ChfuGV *fibrillin* homologue contains a 12-fold EP (glutamic acid and proline) dipeptide repeat within an E and P rich region at its conceptual C-terminal. These EP repeats have also been observed in the C-terminal region of the *Trypanosoma sp.* procyclins. Procyclins, which are the product of a small multigenic family, are glycosyl phosphatidylinositol-anchored proteins that contain the characteristic amino acid repeats at the C-terminal (either EP or GPEET) (Ruepp *et al.*, 1999; Rangarajan *et al.*, 2000; Acosta-Serrano *et al.*, 2001; Butikofer *et al.*, 2002). It was suggested that one function of the protease-resistant C-terminal is to protect the parasite surface from the digestive enzymes in the insect gut.

Compared with other *fibrillin* proteins, ChfuGV *fibrillin* and SINPV *fibrillin* are the longest with 104 amino acids. A previous study reported that the SINPV *fibrillin* additional amino acid residues reside within a defined region. This

region is well known for its potential to shape a coiled-coil configuration. Therefore, SINPV *fibrillin* was predicted to form a longer coiled-coil region (Faktor *et al.*, 1997). Analysis on the ChfuGV *fibrillin* protein sequence using diverse multiple alignment tools revealed that the extra amino acid residues of the ChfuGV *fibrillin* protein, was mainly concentrated inside the EP rich region (from 71 to 91) and at the terminal region of the protein, instead of being located within the coiled-coil region, as was reported for SINPV *fibrillin*. The result of the phylogenetic analysis on *fibrillin* protein species showed that ChfuGV *fibrillin* and the EpNPV, AgNPV PnNPV and OpMNPV *fibrillin* proteins share a common ancestor (data not shown).

An ORF coding a putative baculoviral envelope protein (p74) is located at the extremity of the clone (Cf-xv-p). This protein has been already identified in a number of other baculoviruses including *Adoxophyes orana* granulovirus (AoGV) (Wormleaton *et al.*, 2003), CIGV (Lange and Jehle, 2003), PhopGV (Accession no. NC-004062), CpGV (Luque *et al.*, 2001), *Plutella xylostella* granulovirus (PxGV) (Hashimoto *et al.*, 2000), *Xestia c-nigrum* granulovirus (XcGV) (Hayakawa *et al.*, 1999), AcMNPV (Ayres *et al.*,

1994), OpMNPV (Ahrens *et al.*, 1997), CfNPV (Hill *et al.*, 1993), LdNPV (Kuzio *et al.*, 1999), BmNPV (Gomi *et al.*, 1999), SINPV (Pang *et al.*, 2001), SeNPV (Ijkel *et al.*, 1999), HaNPV (Zhang and Wu 2001), McNPV (Li *et al.*, 2002), and *Culex nigripalpus* baculovirus (CnBV)(Afonso *et al.*, 2001). The baculoviral p74 proteins are believed to be highly conserved ODV envelope proteins that might play a major role in initiating a viral infection (Slack *et al.*, 2001). It was suggested that the N terminal of AcMNPV p74 is located outside of the ODV envelope while the C terminal acts as a transmembrane anchor (Faulkner *et al.*, 1997).

The results of a study on the ChfuGV p74 protein supported the suggestion that the N terminal of the p74 protein is a flexible structure with a higher number of turns and surface loops. The multiple alignment of the ChfuGV p74 protein with those of the other BVs revealed that several glycine and proline residues were preserved in almost all the baculoviral p74 proteins (Rashidan *et al.*, 2003). Another remarkable aspect of the ChfuGV p74 protein is the existence of an 8 cysteine residue cluster, six of those cysteine residues were shown to be conserved in all of the known p74 homologues (Rashidan *et al.*, 2003). The high level of preservation of cysteine, glycine and proline residues in all baculoviral p74 proteins could mean that the p74 protein plays an important role in the life cycle of all baculoviruses and mutations on these preserved residues might ultimately lead to the exclusion of the mutated virus from the evolutionary pathway.

Traditionally, the sequence of the occlusion body matrix proteins, granulin and polyhedrin, has been used to determine the phylogenetic relationship in the baculoviridae family. However, the soundness of this approach was initially questioned by Herniou *et al.* (2001). In our latest paper (Rashidan *et al.*, 2004), it was demonstrated that the GVs could be grouped in 2 distinct groups: Group I (ChfuGV, CpGV, PhopGV and AoGV) and Group II (XcGV, PxGV, LoGV and TnGV). These results in the current study confirm the previously reported results in that the CIGV complete genome sequencing assisted in adding another member to the group I, GVs.

The ChfuGV conserved proteins are most closely related to those of the CpGV, PhopGV and AoGV. The ChfuGV along with the CpGV, PhopGV, AoGV and CIGV share a high degree of gene order preservation within the genome region studied (Fig. 1). The GVs and NPVs showed different patterns in their gene order in the p74 regions (data not shown).

Acknowledgments The authors wish to acknowledge Dr. Anne Bruneau, Department of Biology, University of Montreal for her assistance in the phylogenetic studies, and Dr. David Morse, Department of Biology, University of Montreal for his scientific discussions. This research was funded by the Ministère des Ressources Naturelles du Québec Grant number 0316-240S to CG.

References

- Acosta-Serrano, A., Vassella, E., Liniger, M., Kunz Renggli, C., Brun, R., Roditi, I. and Englund, P. T. (2001) The surface coat of procyclic *Trypanosoma brucei*: programmed expression and proteolytic cleavage of procyclin in the tsetse fly. *Proc. Natl. Acad. Sci. USA* **98**, 1513-1518.
- 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., 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.
- Almazan, F., Tschärke, D. C. and Smith, G. L. (2001) The vaccinia virus superoxide dismutase-like protein (A45R) is a virion component that is nonessential for virus replication. *J. Virol.* **75**, 7018-7029.
- Altshul, S. F., Gish, W., Miller, W., Meyers, E. W. and Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403-410.
- Archibald, J. M., Longet, D., Pawlowski, J. and Keeling, P. J. (2003) A novel polyubiquitin structure in cercozoa and foraminifera: evidence for a new eukaryotic supergroup. *Mol. Biol. Evol.* **20**, 62-66.
- Atlante, A., Bobba, A., Calissano, P., Passarella, S. and Marra, E. (2003) The apoptosis/necrosis transition in cerebellar granule cells depends on the mutual relationship of the antioxidant and the proteolytic systems which regulate ROS production and cytochrome C release en route to death. *J. Neurochem.* **84**, 960-971.
- 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.
- Bah, A., Bergeron, J., Arella, M., Lucarotti, C. J. and Guertin, C. (1997) Identification and sequence analyses of the granulin gene of *Choristoneura fumiferana* granulovirus. *Arch. Virol.* **142**, 1577-1584.
- Blissard, G. W. and Rohrmann, G. F. (1990) Baculovirus diversity and molecular biology. *Annu. Rev. Entomol.* **35**, 127-155.
- Blissard, G. W. (1996) Baculovirus-insect cell interactions. *Cytotechnology* **20**, 73-93.
- Buttke, T. M. and Sandstrom, P. A. (1994) Oxidative stress as a mediator of apoptosis. *Immunol. Today* **15**, 710.
- Butikofer, P., Vassella, E., Boschung, M., Renggli, C. K., Brun, R., Pearson, T. W. and Roditi, I. (2002) Glycosylphosphatidylinositol-anchored surface molecules of *Trypanosoma congolense* insect forms are developmentally regulated in the tsetse fly. *Mol. Biochem. Parasitol.* **119**, 7-16.
- Cao, J. X., Teoh, M. L., Moon, M., McFadden, G. and Evans, D. H. (2002) Leporipoxvirus Cu-Zn superoxide dismutase homologs inhibit cellular superoxide dismutase, but are not essential for virus replication or virulence. *Virology* **296**, 125-135.
- Chen, X., Zhang, W.-J., Wong, J., Chun, G., Lu, A., McCutchen, B. F., Presnail, J. K., Herrmann, R., Dolan, M., Tingey, S., Hu, Z.-H. and Vlak, J. M. (2002) Comparative analysis of the

- complete genome sequences of *Helicoverpa zea* and *Helicoverpa armigera* single-nucleocapsid nucleopolyhedroviruses. *J. Gen. Virol.* **83**, 673-684.
- Cheng, C. H., Liu, S. M., Chow, T. Y., Hsiao, Y. Y., Wang, D. P., Huang, J. J. and Chen, H. H. (2002) Analysis of the Complete Genome Sequence of the Hz-1 Virus Suggests that it is related to members of the Baculoviridae. *J. Virol.* **76**, 9024-9034.
- Faktor, O., Toister-Achituv, M., Nahum, O. and Kamensky, B. (1997) The p10 gene of *Spodoptera littoralis* nucleopolyhedrovirus: nucleotide sequence, transcriptional analysis and unique gene organization in the p10 locus. *J. Gen. Virol.* **78**, 2119-2128.
- Fielding, B. C. and Davison, S. (2000) Identification and characterization of the *Trichoplusia ni* single capsid nuclear polyhedrosis virus p10 gene. *Virus Genes* **20**, 189-192.
- Faulkner, P., Kuzio, J., Williams, G. V. and Wilson, J. A. (1997) Analysis of p74, a PDV envelope protein of *Autographa californica* nucleopolyhedrovirus required for occlusion body infectivity in vivo. *J. Gen. Virol.* **78**, 3091-3100.
- Fielding, B. C., Khan, S., Wang, W., Kruger, C., Abrahams, R. and Davison, S. P. (2002) The genetic organization of a 2,966 basepair DNA fragment of a single capsid nucleopolyhedrovirus isolated from *Trichoplusia ni*. *Virus Genes* **25**, 35-43.
- Forté, A. J., Guertin, C. and Cabana, J. (1999) Pathogenicity of a granulovirus towards *Choristoneura fumiferana*. *Canadian Entomologist* **131**, 725-727.
- Gomi, S., Majima, K. and Maeda, S. (1999) Sequence analysis of the genome of *Bombyx mori* nucleopolyhedrovirus. *J. Gen. Virol.* **80**, 1323-1337.
- Gong, M., Jin, J. and Guarino, L. A. (1998) Mapping of ORF121, a factor that activates baculovirus early gene expression. *Virology* **244**, 495-503.
- Gregory, E., Goscin, S. and Fridovich, I. (1974) Superoxide Dismutase and Oxygen Toxicity in a Eukaryote, *J. Appl. Bacteriol.* **117**, 456-462.
- Hashimoto, Y., Hayakawa, T., Ueno, Y., Fujita, T., Sano, Y. and Matsumoto, T. (2000) Sequence analysis of the *Plutella xylostella* granulovirus genome. *Virology* **275**, 358-372.
- Hayakawa, T., Ko, R., Okano, K., Seong, S. I., Goto, C. and Maeda, S. (1999) Sequence analysis of the *Xestia c-nigrum* granulovirus genome. *Virology* **262**, 277-297.
- Herniou, E. A., Luque, T., Chen, X., Vlask, 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.
- Hess, R. T. and Falcon, L. A. (1987) Temporal events in the invasion of the codling moth, *Cydia pomonella*, by a granulosis virus: an electron microscope study. *J. Invertebr. Pathol.* **50**, 85-92.
- Hicke, L. (2001) A new ticket for entry into budding vesicles: ubiquitin. *Cell* **106**, 527-530.
- Hill, J. E., Kuzio, J., Wilson, J. A., MacKinnon, E. A. and Faulkner, P. (1993) Nucleotide sequence of the p74 gene of a baculovirus pathogenic to the spruce budworm, *Choristoneura fumiferana* multicapsid nuclear polyhedrosis virus. *Biochim. Biophys. Acta* **1172**, 187-189.
- Hockenbery, D. M., Oltvai, Z. N., Yin, X. M., Millman, C. L. and Korsmeyer, S. J. (1993) Bcl-2 functions in an antioxidant pathway to prevent apoptosis. *Cell* **75**, 241-251.
- Hyink, O., Dellow, R. A., Olsen, M. J., Caradoc-Davies, K. M., Drake, K., Herniou, E. A., Cory, J. S., O'Reilly, D. R. and Ward, V. K. (2002) Whole genome analysis of the Epiphyas postvittana nucleopolyhedrovirus. *J. Gen. Virol.* **83**, 957-971.
- Hu, N. T., Lu, Y. F., Hashimoto, Y., Maeda, S. and Hou, R. F. (1994) The p10 gene of natural isolates of *Bombyx mori* nuclear polyhedrosis virus encodes a truncated protein with an M(r) of 7700. *J. Gen. Virol.* **75**, 2085-2088.
- Ijkel, W. F., van Strien, E. A., Heldens, J. G., Broer, R., Zuidema, D., Goldbach, R. W. and Vlask, J. M. (1999) Sequence and organization of the *Spodoptera exigua* multicapsid nucleopolyhedrovirus genome. *J. Gen. Virol.* **80**, 3289-3304.
- 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.
- Lange, M. and Jehle, J. A. (2003) The genome of the *Cryptophlebia leucotreta* granulovirus. *Virology* **317**, 220-236.
- Lee, H. H., Park, S. and Cha, S. C. (1996a) Nucleotide sequence analyses of p10 gene and its promoter of *Hyphantria cunea* nuclear polyhedrosis virus strain HL-2. *Taehan Paerosu Hakhoe Chi* **26**, 131-137.
- Lee, S. Y., Poloumienko, A., Belfry, S., Qu, X., Chen, W., MacAfee, N., Morin, B., Lucarotti, C. and Krause, M. (1996b) A common pathway for p10 and calyx proteins in progressive stages of polyhedron envelope assembly in AcMNPV-infected *Spodoptera frugiperda* larvae. *Arch. Virol.* **141**, 1247-1258.
- Leisy, D. J., Rohrmann, G. F., Nesson, M. and Beaudreau, G. S. (1986) Nucleotide sequencing and transcriptional mapping of the *Orygia pseudotsugata* multicapsid nuclear polyhedrosis virus p10 gene. *Virology* **153**, 157-167.
- Lin, G. and Blissard, G. W. (2002) Analysis of an *Autographa californica* nucleopolyhedrovirus *lef-11* knockout: LEF-11 is essential for viral DNA replication. *Virol.* **76**, 2770-2779.
- Lin, G., Slack, J. M. and Blissard, G. W. (2001) Expression and localization of LEF-11 in *Autographa californica* nucleopolyhedrovirus-infected Sf9 cells. *J. Gen. Virol.* **82**, 2289-2294.
- Li, Q., Donly, C., Li, L., Willis, L. G., Theilmann, D. A. and Erlanson, M. (2002) Sequence and organization of the *Mamestra configurata* nucleopolyhedrovirus genome. *Virology* **294**, 106-121.
- Lu, A. and Miller, L. K. (1995) The roles of eighteen baculovirus late expression factor genes in transcription and DNA replication. *J. Virol.* **69**, 975-982.
- Luque, T., Finch, R., Crook, N., O'Reilly, D. R. and Winstanley, D. (2001) The complete sequence of the *Cydia pomonella* granulovirus genome. *J. Gen. Virol.* **82**, 2531-2547.
- Malmström, B. G., Andréasson, L. E. and Reinhammar, B. (1975) Super-oxide dismutase; in *The Enzymes*, Boyer, P. D. (ed), pp. 507-579, Academic Press, New York, USA.
- Miller, L. K. (2001) Insect viruses; in *Fields Virology*, B. N. Fields et al. (eds), pp. 533-585, Lippincott-Raven publishers, New York, USA.
- Pang, Y., Yu, J., Wang, L., Hu, X., Bao, W., Li, G., Chen, C., Han, H., Hu, S. and Yang, H. (2001) Sequence analysis of the *Spodoptera litura* multicapsid nucleopolyhedrovirus genome. *Virology* **287**, 391-404.
- Pearson, W. R. and Lipman, D. J. (1988) Improved Tools for Biological Sequence Comparison. *PNAS* **85**, 2444-2448.
- Petkau, A., Chelack, W., Pleskach, S., Meeker, B. and Brady, C. (1975) Radioprotection of mice by superoxide dismutase.

- Biochem. Biophys. Res. Commun.* **65**, 886-893.
- Pickart, C. M. (2001) Mechanisms underlying ubiquitination. *Ann. Rev. Biochem.* **70**, 503-533.
- Rangarajan, D., Harvey, T. I. and Barry, J. D. (2000) Characterization of the loci encoding the glutamic acid and alanine rich protein of *Trypanosoma congolense*. *Mol. Biochem. Parasitol.* **105**, 281-290.
- Rashidan, K. K., Nassoury, N., Giannopoulos, P. N. and Guertin, C. (2002) Identification and characterization of a conserved baculoviral structural protein ODVP-6E/ODV-E56 from *Choristoneura fumiferana* granulovirus. *J. Bioch. Mol. Biol.* **35**, 595-603.
- Rashidan, K. K., Nassoury, N., Tazi, S., Giannopoulos, P. N. and Guertin, C. (2003) *Choristoneura fumiferana* Granulovirus p74 protein, a highly conserved baculoviral envelope protein. *J. Biochem. Mol.* **36**, 475-487.
- Rashidan, K. K., Nassoury, N., Giannopoulos, P. N., Mauffette, Y. and Guertin, C. (2004). Identification, Characterization and Phylogenetic Analysis of Conserved Genes within the *odvp-6e/odv-e56* Gene Region of *Choristoneura fumiferana* Granulovirus. *J. Bioch. Mol. Biol.* **37**, 206-212.
- Razuck, F. B., Ribeiro, B., Vargas, J. H., Wolff, J. L. and Ribeiro, B. M. (2002) Characterization of the p10 gene region of *Anticarsia gemmatalis* nucleopolyhedrovirus. *Virus Genes* **24**, 243-247.
- Ruepp, S., Kurath, U., Renggli, C. K., Brun, R. and Roditi, I. (1999) Glutamic acid/alanine-rich protein from *Trypanosoma congolense* is the functional equivalent of 'EP' procyclin from *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **98**, 151-156.
- Rohrmann, G. F. (1986) Polyhedrin structure. *J. Gen. Virol.* **67**, 1499-1513.
- Schinina, M. E., Barra, D., Bossa, F., Calabrese, L., Montesano, L., Carri, M. T., Mariottini, P., Amaldi, F. and Rotilio, G. (1989) Primary structure from amino acid and cDNA sequences of two Cu, Zn superoxide dismutase variants from *Xenopus laevis*. *Arch. Biochem. Biophys.* **272**, 507-515.
- Slack, J. M., Dougherty, E. M. and Lawrence, S. D. (2001) A study of the *Autographa californica* multiple nucleopolyhedrovirus ODV envelope protein p74 using a GFP tag. *J. Gen. Virol.* **82**, 2279-2287.
- Swofford, D. L. (2000) *PAUP Version 4: Phylogenetic Analysis Using Parsimony*, Sinauer, Sunderland, USA.
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. and Higgins, D. G. (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **24**, 4876-4882.
- Van Camp, W., Bowler, C., Villaruel, R., Tsang, E. W., Van Montagu, M. and Inze, D. (1990) Characterization of iron superoxide dismutase cDNAs from plants obtained by genetic complementation in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **87**, 9903-9907.
- van Oers, M. M., Flipsen, J. T., Reusken, C. B., Sliwinsky, E. L., Goldbach, R. W. and Vlak, J. M. (1993) Functional domains of the p10 protein of *Autographa californica* nuclear polyhedrosis virus. *J. Gen. Virol.* **74**, 563-574.
- van Oers, M. M., Flipsen, J. T., Reusken, C. B. and Vlak, J. M. (1994) Specificity of baculovirus p10 functions. *Virology* **200**, 513-523.
- van Oers, M. M. and Vlak, J. M. (1997) The baculovirus 10-kDa protein. *J. Invertebr. Pathol.* **70**, 1-17.
- van Oers, M. M., Hu, Z., Arif, B. M., van Strien, E. A., van Lent, J. W. and Vlak, J. M. (1998) The single-nucleocapsid nucleopolyhedrovirus of *Buzura suppressaria* encodes a P10 protein. *J. Gen. Virol.* **79**, 1553-1562.
- Vlak, J. M., Klinkenberg, F. A., Zaal, K. J., Usmany, M., Klinge-roode, E. C., Geervliet, J. B., Roosien, J. and van Lent, J. W. (1988) Functional studies on the p10 gene of *Autographa californica* nuclear polyhedrosis virus using a recombinant expressing a p10-beta-galactosidase fusion gene. *J. Gen. Virol.* **69**, 765-776.
- Wang, H., Chen, X., Vlak, J. M. and Hu, Z. (2001) Sequence and transcriptional analysis of the p10 gene of *Helicoverpa armigera* single nucleocapsid polyhedrovirus. *Bing Du Xue Bao* **17**, 81-86.
- Wei, Y. J., Long, Q. X., Chen, S. W. and Wang, X. Z. (1998) Nucleotide sequence and characterization of the p10 gene of *Spodoptera litura* nuclear polyhedrosis virus. *Sheng Wu Hua Xue Yu Sheng Wu Wu Li Xue Bao (Shanghai)* **30**, 550-555.
- Williams, G. V., Rohel, D. Z., Kuzio, J. and Faulkner, P. (1989) A cytopathological investigation of *Autographa californica* nuclear polyhedrosis virus p10 gene function using insertion/deletion mutants. *J. Gen. Virol.* **70**, 187-202.
- Wilson, J. A., Hill, J. E., Kuzio, J. and Faulkner, P. (1995) Characterization of the baculovirus *Choristoneura fumiferana* multicapsid nuclear polyhedrosis virus p10 gene indicates that the polypeptide contains a coiled-coil domain. *J. Gen. Virol.* **76**, 2923-2932.
- Wolfe, J. T., Ross, D. and Cohen, G. M. (1994) A role for metals and free radicals in the induction of apoptosis in thymocytes. *FEBS Lett.* **352**, 58-62.
- Wormleaton, S., Kuzio J and Winstanley D. (2003) The complete sequence of the *Adoxophyes orana* granulovirus genome. *Virology* **311**, 350-365.
- Zhang, C. X. and Wu, J. C. (2001) Genome Structure and the p10 Gene of the *Helicoverpa armigera* Nucleopolyhedrovirus. *Sheng Wu Hua Hsueh Yu Sheng Wu Wu Li Hsueh Pao* **33**, 179-184.
- Zuidema, D., van Oers, M. M., van Strien, E. A., Caballero, P. C., Kloek, E. J., Goldbach, R. W. and Vlak, J. M. (1993) Nucleotide sequence and transcriptional analysis of the p10 gene of *Spodoptera exigua* nuclear polyhedrosis virus. *J. Gen. Virol.* **74**, 1017-1024.