Virus Genes (2009) 38:155–170
DOI 10.1007/s11262-008-0297-y

Genomic sequence analysis of the Illinois strain of the *Agrotis ipsilon multiple nucleopolyhedrovirus*

Robert L. Harrison

Received: 3 September 2008 / Accepted: 21 October 2008 / Published online: 18 November 2008
© Springer Science+Business Media, LLC 2008

Abstract The *Agrotis ipsilon multiple nucleopolyhedrovirus* (AgipMNPV) is a group II nucleopolyhedrovirus (NPV) from the black cutworm, *A. ipsilon*, with potential as a biopesticide to control infestations of cutworm larvae. The genome of the Illinois strain of AgipMNPV was completely sequenced. The AgipMNPV genome was 155,122 nt in size and contained 163 open reading frames (ORFs), including 61 ORFs found among all lepidopteran baculoviruses sequenced to date. Phylogenetic inference placed AgipMNPV in a clade with group II NPVs isolated from larvae of *Agrotis* and *Spodoptera* species. Though closely related to the *Agrotis segetum* NPV (AgseNPV), AgipMNPV was found to be missing 15 ORFs present in the AgseNPV genome sequence, including two of the three AgseNPV enhancin genes. Remarkably few polymorphisms were identified in the AgipMNPV sequence even though an uncloned field isolate of this virus was sequenced. A genotype characterized by a 128-bp deletion in the ecdysteroid UDP-glucosyltransferase gene (*egt*) was identified in the AgipMNPV field isolate and among clonal isolates of AgipMNPV. The deletion in *egt* was not associated with differences in budded virus or occluded virus production among AgipMNPV clones in cell culture.

The nucleotide sequence data reported in this article have been submitted to the GenBank nucleotide sequence database and have been assigned the accession number EU839994.

R. L. Harrison (✉)
Invasive Insect Biocontrol and Behavior Laboratory, USDA Agricultural Research Service, Plant Sciences Institute, Building 011A, Room 214, BARC-W, 10300 Baltimore Avenue, Beltsville, MD 20705, USA
e-mail: Robert.L.Harrison@ars.usda.gov

Keywords Baculovirus · Nucleopolyhedrovirus · *Agrotis ipsilon* · Black cutworm · AgipMNPV · Ecdysteroid UDP-glucosyltransferase

Introduction

Baculoviruses are rod-shaped, enveloped DNA viruses of a single family (Baculoviridae) that have been isolated exclusively from arthropods [1]. Most baculoviruses have been identified from insects of the order Lepidoptera (moths and butterflies), though some have been found in mosquitoes (Diptera), sawflies (Hymenoptera), and shrimp (Crustacea). Taxonomic organization of the family Baculoviridae currently places baculoviruses into one of two genera, *Nucleopolyhedrovirus* (NPV) and *Granulovirus* [2]. The life cycles of lepidopteran nucleopolyhedroviruses and granuloviruses (GVs) are essentially identical, with replication resulting in a budded virus phenotype (BV) that spreads infection to other tissues of the host and an occluded virus phenotype (OV) that mediates horizontal transmission within a host population. While the NPVs occlude multiple virions in relatively large polyhedra, GVs occlude one or sometimes two virions in smaller ovoid granules. During assembly of NPV virions, several nucleocapsids often are packaged within a single lipid envelope, while GV virions invariantly consist of a single nucleocapsid per envelope. The NPVs and GVs are also readily distinguished by molecular phylogenetic analysis. Although mosquito and sawfly baculoviruses are currently placed in *Nucleopolyhedrovirus*, a future revision of Baculoviridae is expected to re-distribute these viruses into separate genera [3].

The NPVs are popular tools for expression of genes in insect and mammalian cells [4, 5]. Both NPVs and GVs are

used as environmentally and ecologically safe biopesticides to control agricultural and forestry pests [6, 7].

An NPV has been isolated from the black cutworm, *Agrotis ipsilon* (Noctuidae: Lepidoptera) [8]. *A. ipsilon* is a cosmopolitan pest that feeds on a wide range of plants [9]. The *Agrotis ipsilon multiple nucleopolyhedrovirus* (AgipMNPV) has been explored as a potential insecticide to control black cutworm on corn and turfgrass [10, 11]. Further research is underway to develop AgipMNPV for season-long, multi-year biological control of black cutworm on golf courses [12].

Prior characterization of sequences from AgipMNPV revealed its close relationship with an NPV from the common cutworm, *Agrotis segetum* [13]. The Polish isolate of *A. segetum* NPV-A (AgseNPV-A) exhibits a host range that overlaps with that of AgipMNPV, with each virus able to productively infect *A. ipsilon* and *A. segetum* larvae. AgipMNPV and AgseNPV-A differ somewhat in their virulences towards these two species, with AgipMNPV exhibiting a significantly lower LC_{50} against *A. ipsilon*. The complete genomic sequence of AgseNPV-A was recently reported [14]. This study reports the determination and analysis of the complete genomic sequence of the Illinois strain of AgipMNPV. Comparison with the AgseNPV-A genome confirmed the close relationship between these viruses, but notable differences in open reading frame (ORF) composition were observed.

Methods

Viral DNA isolation

A polyhedra sample of the Illinois strain of AgipMNPV was obtained from Iowa State University. Recently molted 5th instar *A. ipsilon* larvae (from the colony at the Corn Insects and Crop Genetics Research Unit, ARS-USDA, in Ames, IA) were infected *per os* with 8×10^6 polyhedra each. Cadavers were homogenized in 0.5% sodium dodecyl sulfate (SDS) using a Ultra-Turrax T25 mixer (IKA Works, Wilmington, NC). The homogenate was filtered through two layers of cheesecloth and a wire mesh, which were subsequently washed with additional volumes of 0.5% SDS. Polyhedra were pelleted by low-speed centrifugation (750 *g*) for 10 min. Pellets were washed by re-suspension twice in 0.1% SDS and once in 0.5 M NaCl and pelleted by centrifugation after each washing step. Polyhedra were re-suspended in deionized distilled H_2O and solubilized with Na_2CO_3 as previously described [15].

The AgipMNPV OV from solubilized polyhedra was precipitated by mixing with an equal volume of 20% polyethylene glycol (PEG)–1 M NaCl and incubating on ice overnight. The precipitated virions were pelleted by centrifugation (3,000 g) for 15 min. Viral DNA was extracted from the virion pellets by incubation with 10 mM Tris–HCl pH 8.0–10 mM EDTA pH 8.0–0.25% sodium dodecyl sulfate–500 mg/ml protease K for 3 h at 37°C. Viral DNA was recovered by phenol-chloroform extraction and ethanol precipitation, and then subjected to an additional step of purification by CsCl/ethidium bromide equilibrium gradient centrifugation [16].

DNAs from AgipMNPV plaque isolates [17] were isolated from 3rd-passage infected-cell medium using the procedures above, but were not subject to purification by equilibrium gradient centrifugation.

DNA sequencing and analysis

AgipMNPV OV DNA was sheared, cloned, and sequenced as previously described [18]. Sequence data from individual sequencing reactions were assembled and analyzed using Lasergene 7 (DNASTAR). Additional sequencing to address gaps and ambiguities in the genome sequence and to resolve other issues was carried out with custom-designed primers and templates consisting of PCR amplimers. A complete draft of the AgipMNPV genome was obtained with 12.7× coverage.

Open reading frames greater than 50 codons in length that did not overlap larger ORFs by more than 75 nt and were not present in a homologous repeat (*hr*) region were characterized by standard protein–protein BLAST (blast-p; http://www.ncbi.nlm.nih.gov/BLAST/). ORFs not fitting the above criteria that nevertheless possessed homologs in other baculovirus genomes were also selected for characterization. Sequence alignment of 30 core gene amino acid sequences was carried out by CLUSTAL W [19] using the MegAlign program of Lasergene 7. Phylogenetic inference of concatenated alignments containing 15,159 characters per taxon was carried out by minimum evolution (ME) and maximum parsimony (MP) methods using MEGA 4 [20]. Parameters for alignment and phylogenetic inference were as previously described [18]. Taxa analyzed included group II NPVs with completely sequenced genomes [14, 18, 21–35], AcMNPV-C6 [36], and *Cydia pomonella* GV [37].

Quantitation of viral growth

Agrotis ipsilon AiE1611T cells were seeded in six-well plates (0.8×10^6 cells/well) and infected with AgipMNPV clones 1T6, 1T7, and 6T2 at a multiplicity of infection (MOI) of five plaque-forming units (PFU)/cell. Aliquots of infected-cell medium were harvested at 24, 48, and 72 h p. i. and titered by plaque assay [17]. Polyhedra were purified from cells at 96 h p. i. following the procedure of O'Reilly et al. [15] and counted with a hemocytometer. Results were analyzed by two-tailed *t*-test.

Results

Features of the AgipMNPV genome sequence

Assembly of AgipMNPV sequencing data resulted in a genome sequence of 155,122 bp, larger than most baculovirus genomes sequenced to date. In particular, the AgipMNPV genome was 7,508 bp larger than that of AgseNPV-A [14]. Compared to the other 42 baculovirus genomes that have been sequenced, the AgipMNPV genome was smaller than that of *Xestia c-nigrum* (Xecn)GV (178,733 bp; [38]), *Helicoverpa armigera* (Hear)GV (169,794 bp; [39]), *Leucania separata* (Ls) NPV (168,041 bp; [34]), *Lymantria dispar* (Ld)MNPV (161,046 bp; [21]), *Mamestra configurata* (Maco)NPV-B (158,482 bp; [27]), and *Orgyia leucostigma* (Orle)NPV (156,179 bp; accession number NC_010276). The AgipMNPV genome sequence also possessed a relatively high G + C content of 48.6%, lower only than that of LdMNPV (57.5%), *Orgyia pseudotsugata* (Op)MNPV (55.0%; [40]), *Antheraea pernyi* (Anpe)NPV (53.4–53.5%; [41, 42]), *Culex nigripalpus* (Cuni)NPV (50.9%; [43]), and *Choristoneura fumiferana* (Cf)MNPV (50.1; [44]).

Examination of the genome revealed a total of 163 ORFs at least 50 codons in length that exhibited minimal (<75 bp) overlap with larger ORFs or shared significant sequence identity with previously characterized baculovirus ORFs (Table 1, Fig. 1). The ORFs were numbered consecutively with the polyhedrin (*polh*) ORF being designated as the first ORF (*agip1*), and the adenine of the polyhedrin ORF initiation codon was set as the first nucleotide of the genome sequence. As described for other baculoviruses [45], the ORFs were positioned close together on the genome sequence and exhibited no obvious bias in orientation, with 53% of the ORFs in the *polh*-sense orientation. Canonical baculovirus early and late gene promoter motifs were detected upstream of 107 of the ORFs (Table 1). BLAST searches with predicted amino acid sequences revealed that AgipMNPV contained the 30 core genes that have been found in all baculovirus genomes sequenced to date [45, 46] as well as the additional 31 genes found in all lepidopteran baculovirus genomes sequenced to date [3, 34].

Relationships with other baculoviruses

BLAST analyses of AgipMNPV ORFs generally revealed strong similarities between AgipMNPV sequences and those of AgseMNPV-A, *Spodoptera exigua* (Se)MNPV [22], *Spodoptera frugiperda* (Sf)MNPV [18, 35] and the NPVs characterized from *M. configurata* (MacoNPV-A and -B; [26, 27]). A more comprehensive examination of the relationship of AgipMNPV to other group II NPVs was carried out by phylogenetic inference with concatenated alignments of the 30 baculovirus core genes from completely sequenced group II NPVs as well as AcMNPV-C6 [36] and *Cydia pomonella granulovirus* (CpGV; [37]) (Fig. 2). Both ME and MP trees generated by this analysis contained the clade consisting of *Agrotis* and *Spodoptera* NPVs that had been reported in previous analyses. This clade in turn was part of a larger clade containing the MacoNPV-A and -B viruses. These relationships enjoyed strong bootstrap support in both ME and MP phylograms. Strong bootstrap support in both ME and MP phylograms also was observed for some of the terminal clades containing two or three NPVs. Aside from these terminal clades, relationships among group II NPVs were characterized by long branches in the phylogram, indicating a profound degree of divergence in this group of baculoviruses.

The close relationships among the *Agrotis* and *Spodoptera* NPVs could also be observed by the strongly conserved order of homologous ORFs along the genomes of these viruses, as assessed by gene-parity plot analysis (data not shown; [47]). Exceptions to this general co-linearity could be seen in the region between the *cathepsin* and *gp37* genes, in which the order and orientation of some orthologous ORFs (*he65*, *agip24*/*agse22*, *chitinase*) were reversed when comparing viruses of *Agrotis* and *Spodoptera* (Fig. 3). Evidence for rearrangement in this region also has been observed in the genomes of MacoNPV-A and -B [27].

The genome sequence reported in this study exhibited 100% nucleotide sequence identities with seven prior nucleotide sequence entries in GenBank for individual AgipMNPV genes. Most of these entries (accession numbers AY519204-AY519206, AY136483, and AY136484) likely derive from the Illinois strain of AgipMNPV. One partial polyhedrin sequence (accession number DQ014542) derives from a Kentucky strain of AgipMNPV. This sequence, along with the matching sequence from the Illinois strain deposited by the same group (accession number DQ014543), differed from the AgipMNPV genome sequence at the 3′ terminal nucleotide of the sequence file.

Homologous repeat regions

In baculovirus genomes, the larger intergenic spaces are often occupied by regions of repeated sequences referred to as *homologous regions* or *hr*s [45]. The *hr*s are thought to function as enhancers and origins of replication [48, 49]. In NPVs, the repeat units that make up *hr*s are usually imperfect palindromes.

Examination of the AgipMNPV intergenic spaces revealed the presence of *hr*s with palindromic repeats

Table 1 Features of the AgipMNPV genome

ORF/other feature	*Name*	Position[a]	aa (Da)[b]	Promoter motifs[c]	Comparison with other viruses					
					AgseNPV		SeMNPV		AcMNPV	
					ORF (size)/hr	% ID (range)	ORF (size)	% ID (range)	ORF (size)	% ID (range)
1	*polh*	1–>741	246 (29029)	L	*agse1* (246)	94.7 (233/246)	*se1* (246)	92.7 (228/246)	*ac8* (245)	86.4 (209/242)
2	*orf1629*	833<–2392	519 (56833)		*agse2* (448)	49.8 (264/530)	*se2* (462)	46.5 (244/525)	*ac9* (543)	24 (119/488)
3	*pk-1*	2391–>3212	274 (32086)	E	*agse3* (264)	83.3 (220/264)	*se3* (295)	71.6 (192/268)	*ac10* (272)	42.2 (114/270)
4	*hoar*	3304<–5517	738 (82427)	E	*agse4* (704)	39.4 (305/774)		33.0 (262/793)		
5		5474<–5710	78 (8778)	L						
6		5964<–6200	78 (9056)							
7		6130<–6501	123 (14349)	E					*ac152* (92)	59.4 (19/32)
8		6337–>7485	382 (42357)							
9	*odv-e56*	7577–>8683	368 (40402)	L	*agse6* (369)	71.7 (263/367)	*se6* (371)	64.1 (237/370)	*ac148* (376)	53.8 (184/342)
10	*me53*	9034–>10095	353 (41183)	L	*agse7* (349)	82.2 (290/353)	*se7* (390)	61.1 (212/347)	*ac139* (363)	27.5 (96/349)
11		10229–>10399	56 (6791)							
12	*efp/ld130*	10892–>12937	681 (78432)	L	*agse8* (670)	65.5 (448/684)	*se8* (656)	64.4 (439/682)	*ac23* (690)	23.5 (117/498)
13		13119<–14087	322 (37356)	C	*agse9* (335)	47.9 (156/326)				
14	*gp16*	14177<–14467	96 (10791)	L	*agse10* (96)	78.1 (75/96)	*se9* (94)	59.6 (56/94)	*ac130* (106)	34.4 (32/93)
15	*p24*	14494<–15243	249 (27101)	L	*agse11* (232)	68.8 (165/240)	*se10* (248)	64.7 (163/252)	*ac129* (198)	36.2 (81/224)
16		15325–>15669	114 (13337)	L	*agse12* (106)	53.0 (61/115)	*se11* (105)	43.9 (50/114)		
17	*lef-2*	15626–>16273	215 (24604)		*agse13* (210)	68.9 (146/212)	*se12* (209)	64.2 (136/212)	*ac6* (210)	40.6 (86/212)
18		16346<–17095	249 (28186)	L	*agse103* (133)	35.5 (38/107)	*se96* (113)	56.8 (25/44)	*ac150* (99)	43.5 (20/46)
19	*bro-a*	17351–>18832	493 (57518)							
20		18920<–19564	214 (24698)	L						
hr1 (5 repeats)		19778–20012								
21	*38.7 k*	20047<–21174	375 (42352)	L	*agse16* (367)	64.9 (248/382)	*se13* (363)	57.8 (216/374)	*ac13* (327)	28.6 (74/259)
22	*lef-1*	21176<–21832	218 (25259)		*agse17* (215)	75.7 (159/210)	*se14* (216)	69.5 (146/210)	*ac14* (266)	45.3 (97/214)
23	*v-cath*	22234<–23328	364 (41272)		*agse19* (343)	85.4 (292/342)	*se16* (337)	87.2 (294/337)	*ac127* (323)	53.1 (179/337)
24		23378<–24727	449 (49689)	C	*agse22* (449)	62.2 (281/452)	*se22* (103)	51.1 (45/88)		
							se23 (211)	43.3 (94/217)		
							se24 (72)	62.9 (44/70)		
25	*he65*	25055–>26677	540 (62787)		*agse20* (245)	74.6 (182/244)			*ac105* (553)	37.8 (221/58)
26		26813<–27259	148 (17368)				*se21* (98)	52.2 (48/92)		
hr1a (1 repeat)		27390–27429								
27	*chitinase*	27524–>29266	580 (64009)	L	*agse23* (582)	84.9 (478/563)	*se19* (572)	83.3 (464/557)	*ac126* (551)	66.0 (359/544)
28		29435–>30193	252 (29436)		*agse25* (247)	28.4 (71/250)				
29		30236<–30538	100 (11740)	E, L					*ac79* (104)	48.4 (44/91)

Table 1 continued

ORF/other feature	*Name*	Position[a]	aa (Da)[b]	Promoter motifs[c]	Comparison with other viruses					
					AgseNPV		SeMNPV		AcMNPV	
					ORF (size)/hr	% ID (range)	ORF (size)	% ID (range)	ORF (size)	% ID (range)
30	*gp37*	30583–>31368	261 (29969)	L	*agse26* (260)	77.3 (197/255)	*se25* (267)	77.1 (205/266)	*ac64* (203)	57.3 (137/239)
31	*ptp-2*	31381<–31875	164 (19069)	L	*agse27* (170)	64.0 (105/164)	*se26* (165)	59.0 (98/166)		
32	*egt*	32020–>33600	526 (60202)		*agse28* (523)	83.0 (439/529)	*se27* (523)	76.9 (407/529)	*ac15* (506)	47.3 (246/520)
33		33766–>34365	199 (22997)		*agse29* (186)	59.4 (107/180)	*se28* (190)	47.1 (89/189)		
34		34378–>35049	223 (25806)		*agse30* (214)	59.6 (130/218)	*se29* (213)	49.1 (105/214)	*ac17* (164)	30.1 (25/83)
35		35126<–37819	897 (103409)	L	*agse31* (853)	52.2 (473/906)	*se30* (886)	51.5 (469/911)		
36		37934–>38305	123 (13708)	E, L	*agse32* (186)	33.0 (35/106)	*se31* (241)	28.6 (36/126)	*ac4* (151)	NSS
37		38352–>38879	175 (20394)	L	*agse32* (186)	28.9 (57/197)	*se31* (241)	44.6 (37/83)	*ac4* (151)	NSS
38	*pkip-1*	38993–>39490	165 (19524)	L	*agse33* (166)	74.1 (123/166)	*se32* (408)	61.6 (98/159)	*ac24* (169)	23.7 (32/135)
39		39655<–40719	354 (41490)							
40		40908<–41240	110 (12326)		*agse34* (112)	61.3 (68/111)	*se33* (112)	50.9 (57/112)		
41	*arif-1*	41245<–42063	272 (30521)		*agse35* (290)	58.3 (155/266)	*se34* (281)	51.6 (147/285)	*ac20-ac21* (417)	26 (53/201)
42	*pif-2*	41972–>43207	411 (46732)		*agse36* (400)	86.5 (346/400)	*se35* (413)	82.6 (341/413)	*ac22* (382)	60.2 (228/379)
43	*pif-1*	43233–>44870	545 (62093)	E	*agse37* (527)	71.0 (389/548)	*se36* (526)	57.0 (311/546)	*ac119* (530)	45.9 (253/551)
44		44903–>45166	87 (10169)	E	*agse38* (80)	63.8 (51/80)	*se37* (80)	56.2 (45/80)	*ac120* (82)	28.6 (22/77)
45	*fgf*	45239<–46375	378 (41759)	E	*agse39* (373)	46.2 (176/381)	*se38* (404)	41.5 (170/410)	*ac32* (181)	25.7 (29/113)
46		46734–>47471	245 (28329)	L	*agse40* (257)	61.1 (149/244)	*se40* (241)	53.1 (121/228)		
47	*alk-exo*	47523<–48755	410 (47115)	L	*agse41* (402)	66.4 (271/408)	*se41* (413)	54.0 (221/409)	*ac133* (419)	39.8 (168/422)
hr2 (4 repeats)		48797–49004			*hr2* (3 repeats)					
48		49097–>49321	74 (8659)							
49		49851<–50186	111 (12598)	L	*agse44* (118)	80.4 (86/107)	*se42* (81)	64.5 (49/76)	*ac19* (108)	NSS
50		50200–>51357	385 (45226)	L	*agse45* (380)	74.9 (287/383)	*se43* (280)	68.1 (263/386)	*ac18* (353)	25.7 (100/389)
51		51460<–51846	128 (15133)	E	*agse46* (152)	78.1 (100/128)	*se44* (140)	59.4 (63/106)		
52	*rr2B*	51948–>52889	313 (36472)		*agse47* (364)	89.1 (279/313)	*se45* (313)	76.7 (240/313)		
53		53029–>54354	441 (50935)	L	*agse48* (143)	39.5 (155/392)				
54	*pep*	54549<–55682	377 (42719)	L	*agse49* (341)	70.6 (266/377)	*se46* (335)	65.8 (250/380)	*ac131* (252)	32.8 (120/366)
55		55879<–56394	171 (19541)		*agse51* (97)	78.6 (77/98)	*se47* (103)	62.1 (64/103)	*ac117* (95)	31.7 (32/101)
56		56638<–57060	140 (16008)	L	*agse53* (132)	30.2 (39/129)				
57	*sod*	57099<–57557	152 (15999)		*agse54* (114)	83.4 (126/151)	*se48* (151)	79.1 (117/148)	*ac31* (151)	71.1 (106/149)
58		57629–>57994	121 (13146)	E	*agse55* (134)	61.1 (66/108)	*se49* (130)	21.5 (23/107)		
59	*pif-3*	58030–>58722	230 (25052)		*agse56* (207)	80.9 (165/204)	*se50* (214)	72.2 (151/209)	*ac115* (204)	48.2 (96/199)
60		58724–>59182	152 (17596)	E	*agse57* (168)	58.3 (88/151)	*se51* (142)	44.6 (62/139)		

Table 1 continued

ORF/other feature	*Name*	Position[a]	aa (Da)[b]	Promoter motifs[c]	Comparison with other viruses					
					AgseNPV		SeMNPV		AcMNPV	
					ORF (size)/hr	% ID (range)	ORF (size)	% ID (range)	ORF (size)	% ID (range)
61		59280–>60731	483 (55737)	L	*agse58* (480)	74.1 (352/475)	*se52* (529)	57.4 (291/507)		
62	*ac106/107*	60772–>61464	230 (26154)	L	*agse59* (216)	83.0 (191/230)	*se53* (222)	76.5 (176/230)	*ac106–ac107* (243)	57 (118/204)
63	*hisP*	61541<–62617	358 (41250)	C	*agse60* (362)	68.8 (249/362)	*se54* (264)	64.9 (233/359)	*ac33* (182)	29.8 (42/141)
hr3 (7 repeats)		62751–63172			*hr3* (3 repeats)					
64	*dUTPase*	63232–>63660	142 (15583)		*agse62* (142)	79.6 (113/142)	*se55* (143)	72.7 (104/143)		
65		63816–>64400	194 (22700)							
66		64375–>64542	55 (6614)	E						
67		64481–>64633	50 (6037)	E						
68	*bro-b*	64584–>65579	331 (37101)							
69		65755–>66354	199 (23860)	C	*agse63* (195)	77.8 (154/198)				
70	*p13*	66486–>67316	276 (32307)	L	*agse64* (274)	87.1 (237/272)	*se56* (283)	74.8 (205/274)		
71	*odv-e66*	67416–>69506	696 (78914)	L			*se57* (723)	50.5 (378/749)	*ac46* (704)	39.1 (255/652)
72		69496<–69858	120 (13253)	E, L	*agse65* (109)	64.4 (76/118)	*se58* (114)	67.6 (75/111)	*ac108* (105)	36.5 (27/74)
73	*odv-ec43*	69864<–70934	356 (41399)		*agse66* (356)	95.2 (339/356)	*se59* (356)	90.7 (323/356)	*ac109* (390)	44.2 (174/394)
74	*ac110*	70918<–71100	60 (7137)	E	*agse67* (59)	90.0 (54/60)	*se60* (59)	76.4 (42/55)	*ac110* (56)	38.0 (19/50)
75	*vp80/p87*	71097<–72908	603 (67609)	E	*agse68* (582)	58.6 (365/623)	*se61* (556)	56.1 (338/602)	*ac104* (691)	22.8 (105/460)
76	*p45*	72945–>74078	377 (43838)	L	*agse69* (377)	88.3 (333/377)	*se62* (375)	80.9 (305/377)	*ac103* (387)	53.1 (203/382)
77	*p12*	74068–>74388	106 (11715)	L	*agse70* (113)	78.6 (81/103)	*se63* (106)	77.1 (81/105)	*ac102* (122)	28.8 (30/104)
78	*p40*	74420–>75571	383 (42750)	E, L	*agse71* (369)	80.9 (301/372)	*se64* (388)	78.8 (308/391)	*ac101* (361)	45.8 (176/384)
79	*p6.9*	75656–>75913	85 (9945)	L	*agse72* (80)	70.2 (59/84)	*se65* (75)	73.2 (60/82)	*ac100* (55)	54 (29/53)
80	*lef-5*	75910<–76734	274 (32168)		*agse73* (272)	89.3 (242/271)	*se66* (279)	82.1 (224/273)	*ac99* (265)	54.9 (130/237)
81	*38k*	76636–>77535	299 (35811)	L	*agse74* (264)	85.9 (226/263)	*se67* (300)	77.6 (228/294)	*ac98* (320)	45.8 (142/310)
82	*vef-2*	77751–>80444	897 (102626)	L	*agse76* (883)	54.3 (482/887)				
83	*bro-c*	80488<–81528	346 (39429)		*agse77* (350)	78.4 (276/352)			*ac2* (328)	53.1 (172/324)
84		81597<–82031	144 (17146)		*agse78* (147)	87.1 (128/147)				
85	*odv-e28*	82093<–82602	169 (19365)	E	*agse79* (173)	91.1 (154/169)	*se69* (170)	80.4 (135/168)	*ac96* (173)	52.6 (90/171)
86	*helicase*	82571–>86203	1210 (141532)	E, L	*agse80* (1213)	84.9 (1031/1214)	*se70* (1222)	76.3 (939/1230)	*ac95* (1221)	41.4 (512/1236)
87	*odv-e25*	86335<–86988	217 (24485)	E, L	*agse81* (216)	88.9 (193/217)	*se71* (216)	88.0 (191/217)	*ac94* (228)	46.4 (104/224)
88	*ac93*	86985<–87461	158 (18128)		*agse82* (158)	94.3 (149/158)	*se72* (157)	78.1 (121/155)	*ac93* (161)	50.3 (80/159)
89	*p33*	87468–>88226	252 (30658)	E	*agse83* (252)	93.3 (235/252)	*se73* (252)	87.7 (221/252)	*ac92* (259)	52.5 (136/259)
90		88334–>88969	211 (24334)	L	*agse84* (207)	43.5 (84/193)				

Table 1 continued

ORF/other feature	*Name*	Position[a]	aa (Da)[b]	Promoter motifs[c]	Comparison with other viruses					
					AgseNPV		SeMNPV		AcMNPV	
					ORF (size)/hr	% ID (range)	ORF (size)	% ID (range)	ORF (size)	% ID (range)
91	*lef-4*	89042<–90487	481 (54957)		*agse85* (457)	70.6 (338/479)	*se74* (466)	62.4 (302/484)	*ac90* (464)	44.0 (217/493)
92	*vp39*	90486–>91469	327 (36841)	E, L	*agse86* (207)	89.3 (292/327)	*se75* (326)	80.3 (261/325)	*ac89* (347)	44.0 (136/309)
93	*cg30*	91697–>93442	581 (64422)		*agse87* (453)	29.1 (120/412)	*se76* (461)	25.4 (116/457)	*ac88* (264)	24 (27/111)
94	*vp91*	93606<–96047	813 (91894)	L	*agse88* (778)	72.1 (575/798)	*se77* (813)	61.2 (502/820)	*ac83* (847)	40.3 (345/857)
95	*telokinin-like*	96016–>96648	210 (22920)	L	*agse89* (187)	68.1 (143/210)	*se78* (196)	62.6 (134/214)	*ac82* (180)	31.0 (52/168)
96	*ac81*	96422–>97201	259 (29839)	L	*agse90* (191)	79.3 (149/188)	*se79* (240)	68.1 (171/251)	*ac81* (233)	51.5 (118/229)
97	*gp41*	97164–>98171	335 (37333)	L	*agse91* (297)	94.6 (281/297)	*se80* (331)	88.1 (290/329)	*ac80* (409)	54.5 (169/310)
98	*ac78*	98168–>98533	121 (13648)	E, L	*agse92* (124)	63.6 (82/129)	*se81* (127)	57.4 (74/129)	*ac78* (109)	35.7 (40/112)
99	*vlf-1*	98535–>99671	378 (44463)	L	*agse93* (382)	92.9 (355/382)	*se82* (372)	91.4 (340/372)	*ac77* (379)	64.3 (238/370)
100	*p26*	99760<–100500	246 (28009)	E	*agse94* (244)	74.8 (184/246)	*se87* (248)	66.8 (167/250)		
101	*iap-2*	100594<–101478	294 (34032)		*agse95* (259)	67.8 (181/267)	*se88* (317)	57.8 (175/303)	*ac71* (249)	36.0 (91/253)
102		101351<–102223	290 (33360)		*agse96* (280)	71.1 (199/280)	*se89* (299)	65.4 (189/289)	*ac69* (262)	45.3 (121/267)
103	*ac68*	102195<–102572	125 (14416)	E	*agse97* (122)	84.7 (100/118)	*se90* (133)	79.0 (94/119)	*ac68* (192)	47.7 (53/111)
104	*lef-3*	102571–>103764	397 (45429)	E	*agse98* (422)	63.2 (263/416)	*se91* (422)	55.2 (232/420)	*ac67* (385)	28.7 (111/387)
105	*desmoplakin*	103937<–106147	736 (83150)	E	*agse99* (695)	58.7 (437/744)	*se92* (704)	45.7 (344/752)	*ac66* (808)	21.0 (174/828)
106	*dnapol*	106146–>109277	1043 (120824)		*agse100* (1025)	81.0 (846/1045)	*se93* (1063)	77.3 (815/1054)	*ac65* (984)	45.3 (458/1010)
107	*ac75*	109337<–109726	129 (15238)	L	*agse101* (125)	93.0 (120/129)	*se94* (129)	81.4 (105/129)	*ac75* (133)	26.6 (33/124)
108	*ac76*	109733<–109990	85 (9829)	L	*agse102* (85)	89.4 (76/85)	*se95* (85)	89.4 (76/85)	*ac76* (84)	43.0 (37/86)
109		110097–>110732	211 (24090)	L	*agse103* (133)	55.0 (44/80)	*se96* (113)	38.7 (43/111)	*ac150* (99)	40 (21/52)
110		110729<–110893	54 (6127)	L						
111		110835–>111512	225 (26507)		*agse104* (221)	71.4 (160/224)				
112	*bro-d*	111550<–112611	353 (40970)							
113	*lef-9*	112730<–114256	508 (58241)		*agse105* (514)	89.1 (450/505)	*se97* (495)	86.1 (428/497)	*ac62* (516)	65.1 (317/487)
114	*fp25k*	114333–>114944	203 (23319)	L	*agse106* (195)	96.3 (183/190)	*se98* (195)	91.1 (174/191)	*ac61* (214)	62.8 (118/188)
115	*p94*	115080–>117740	886 (104164)				*se99* (719)	24.0 (161/671)	*ac134* (803)	38.2 (313/819)
116	*bro-e*	117824–>118306	160 (18508)		*agse107* (163)	72.9 (113/155)				
117		118344–>118655	103 (12110)	E, L	*agse108* (86)	75.0 (72/96)	*se100* (89)	73.7 (73/99)	*ac60* (87)	43.2 (35/81)
118		118671–>119261	196 (22718)	L	*agse109* (174)	50.0 (100/200)	*se101* (195)	46.2 (98/212)	*ac58-ac59* (172)	35.5 (59/166)
119		119254<–119736	160 (18791)	C	*agse110* (160)	69.4 (109/157)	*se102* (178)	66.7 (108/162)	*ac57* (161)	42.8 (68/159)
120		119976<–120257	93 (10485)	L	*agse111* (77)	54.5 (30/55)	*se103* (93)	56.5 (52/92)	*ac56* (84)	34.5 (20/58)
121		120199<–120567	122 (13982)		*agse112* (69)	73.9 (51/69)	*se104* (67)	63.2 (43/68)	*ac55* (73)	42.0 (29/69)
122	*vp1054*	120521<–121546	341 (39290)		*agse113* (337)	77.5 (262/338)	*se105* (346)	76.4 (262/343)	*ac54* (365)	39.8 (144/362)

Table 1 continued

ORF/other feature	*Name*	Position[a]	aa (Da)[b]	Promoter motifs[c]	Comparison with other viruses					
					AgseNPV		SeMNPV		AcMNPV	
					ORF (size)/hr	% ID (range)	ORF (size)	% ID (range)	ORF (size)	% ID (range)
123	*lef-10*	121407<–121640	77 (8066)	E, L	*agse114* (76)	76.4 (55/72)	*se106* (77)	73.4 (58/79)	*ac53a* (78)	49.3 (33/67)
124		121621–>121827	68 (8591)	L	*agse115* (80)	71.2 (47/66)				
125		121847–>122875	342 (38490)	L	*agse116* (326)	62.6 (214/342)	*se107* (344)	53.4 (189/354)		
126	*ac53*	122885<–123298	137 (16315)	L	*agse117* (121)	71.1 (86/121)	*se108* (137)	67.9 (93/137)	*ac53* (139)	45.6 (62/136)
127		123368–>123961	197 (23314)	C	*agse118* (169)	47.4 (91/192)	*se109* (162)	35.7 (41/115)	*ac52* (123)	23.1 (49/212)
hr4 (7 repeats)		124013–124448			*hr4* (3 repeats)					
128	*iap-3*	124657–>125475	272 (30869)	L	*agse119* (271)	62.9 (176/280)	*se110* (313)	46.0 (133/289)		
129	*bjdp*	125539<–126690	383 (43962)	E, L	*agse120* (397)	59.1 (227/384)	*se111* (415)	42.1 (170/404)	*ac51* (318)	20.5 (60/292)
130	*lef-8*	126711–>129353	880 (101696)		*agse121* (882)	87.5 (773/883)	*se112* (906)	80.1 (726/906)	*ac50* (876)	61.8 (545/882)
131		129462<–129872	136 (15675)	E	*agse122* (133)	36.9 (38/103)				
132		130089<–130292	67 (8129)	L	*agse124* (68)	71.6 (48/67)	*se113* (57)	66.0 (33/50)	*ac43* (77)	35.5 (22/62)
133	*odv-e66b*	130469<–132541	690 (78532)	E, L	*agse125* (678)	64.6 (446/690)	*se114* (685)	56.8 (400/704)		
134	*p47*	132594–>133787	397 (46488)	L	*agse126* (401)	83.3 (334/401)	*se115* (400)	81.5 (326/400)	*ac40* (401)	54.6 (212/388)
135		133801<–134784	327 (37406)		*agse127* (338)	54.8 (182/332)				
hr5 (2 repeats)		134977–135069			*hr5 (3 repeats)*					
136		135028–>135726	232 (26830)		*agse129* (173)	64.4 (114/177)	*se117* (191)	36.8 (67/182)		
137	*ac38*	135795–>136496	233 (27849)	C, E, L	*agse130* (225)	91.6 (206/225)	*se118* (261)	83.6 (179/214)	*ac38* (216)	62.6 (127/203)
138	*lef-11*	136427–>136798	123 (14117)		*agse131* (125)	73.6 (89/121)	*se119* (103)	68.9 (71/103)	*ac37* (112)	37.0 (34/92)
139	*pp31/39 k*	136764–>137693	309 (34587)		*agse132* (292)	67.3 (208/309)	*se120* (317)	53.0 (174/328)	*ac36* (275)	36.0 (112/311)
140		137773–>138069	98 (12117)	E	*agse133* (94)	47.3 (44/93)	*se121* (96)	35.5 (27/76)		
141		138184<–138372	62 (6991)		*agse134* (69)	60.3 (38/63)	*se122* (69)	47.7 (31/65)		
142	*v-ubi*	138366<–138614	82 (9402)	L	*agse135* (78)	93.4 (71/76)	*se123* (80)	93.4 (71/76)	*ac35* (77)	75 (57/76)
143		138694–>139227	177 (20854)	L	*agse136* (178)	83.1 (147/177)	*se124* (187)	72.3 (133/184)	*ac34* (215)	34.6 (63/182)
144		139488<–139664	58 (6753)							
hr6 (4 repeats)		139797–140019								
145		140028<–140399	123 (14025)	L	*agse138* (123)	63.4 (78/123)	*se125* (135)	61.7 (71/115)	*ac26* (129)	39.6 (42/106)
146	*dbp1*	140488–>141465	325 (37723)		*agse139* (312)	70.3 (230/327)	*se126* (328)	56.4 (189/335)	*ac25* (316)	28.2 (87/309)
147	*lef-6*	141483–>141989	168 (19425)	L	*agse140* (168)	57.5 (100/174)	*se127* (163)	47.0 (79/168)	*ac28* (173)	30.8 (52/169)
148	*ac29*	142059<–142307	82 (95900)	L	*agse141* (77)	80.5 (66/82)	*se128* (136)	68.6 (59/86)	*ac29* (71)	27.5 (22/80)
149	*p26b*	142420–>143238	272 (30519)	L	*agse142* (275)	75.8 (194/256)	*se129* (278)	67.2 (180/268)	*ac136* (240)	34.3 (82/239)
150	*p10*	143271–>143567	98 (10627)	L	*agse143* (87)	65.4 (51/78)	*se130* (88)	67.1 (49/73)	*ac137* (94)	NSS
151	*p74*	143680<–145620	646 (73367)	E, L	*agse144* (644)	78.5 (507/646)	*se131* (653)	72.3 (472/653)	*ac138* (645)	57.3 (368/642)
152		145762–>146025	87 (10200)		*agse145* (83)	56.0 (47/84)				

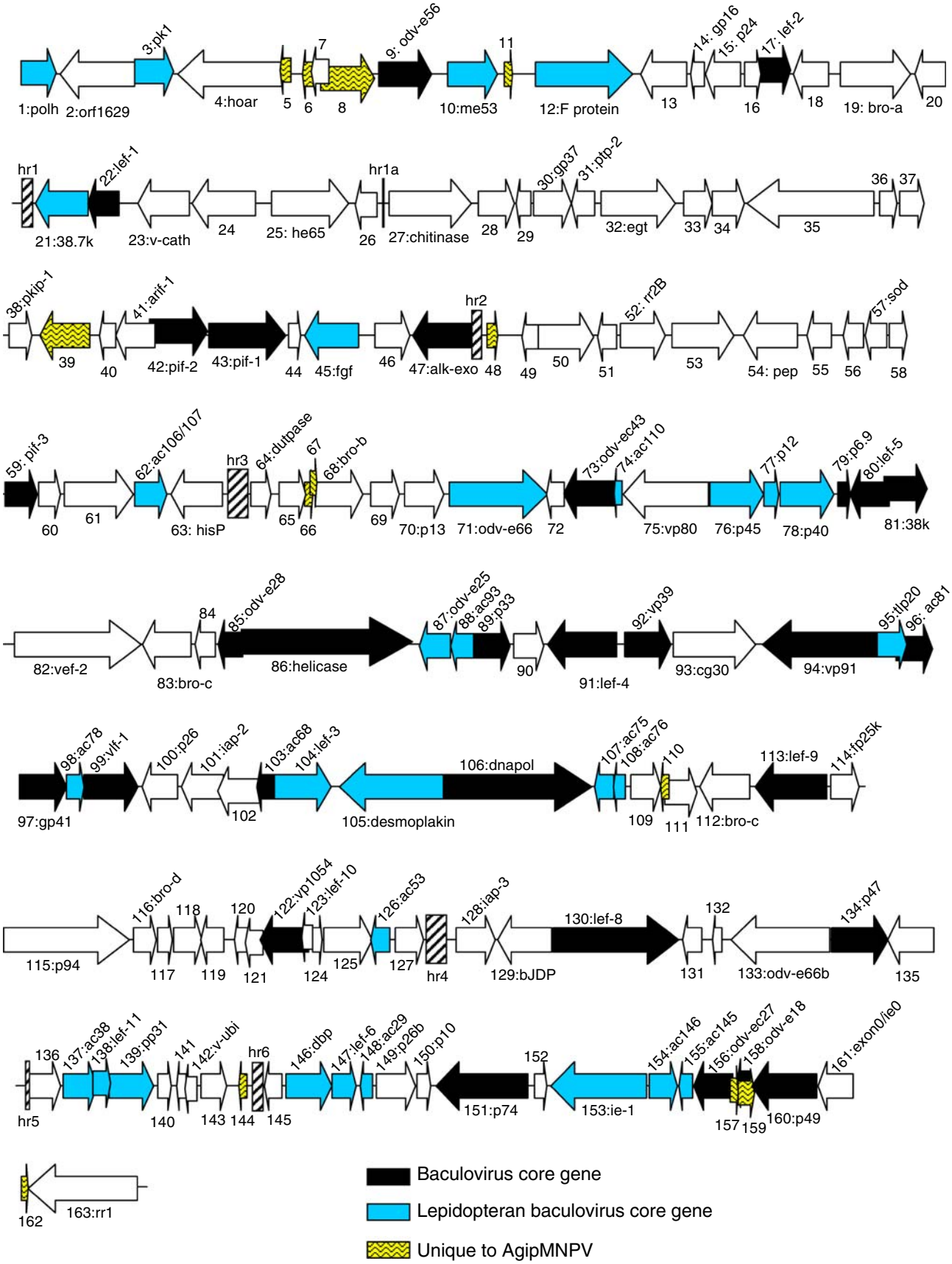
Table 1 continued

ORF/other feature	*Name*	Position[a]	aa (Da)[b]	Promoter motifs[c]	Comparison with other viruses					
					AgseNPV		SeMNPV		AcMNPV	
					ORF (size)/hr	% ID (range)	ORF (size)	% ID (range)	ORF (size)	% ID (range)
153	*ie-1*	146103<–148130	675 (77485)		*agse146* (661)	58.5 (393/672)	*se132* (714)	47.2 (339/718)	*ac147* (582)	33.8 (134/397)
154	*ac146*	148181–>148777	198 (21871)	E, L	*agse147* (191)	66.2 (131/198)	*se133* (200)	60.6 (120/198)	*ac146* (201)	31.5 (63/200)
155	*ac145*	148817<–149095	92 (10674)	L	*agse148* (92)	87.0 (80/92)	*se134* (92)	78.3 (72/92)	*ac145* (77)	39.0 (30/77)
156	*odv-ec27*	149135<–149965	276 (32101)	L	*agse149* (279)	88.5 (246/278)	*se135* (281)	85.0 (238/280)	*ac144* (290)	54.0 (157/291)
157		149913–>150077	54 (6042)							
158	*odv-e18*	150049<–150330	93 (9581)	E, L	*agse150* (77)	62.4 (58/93)	*se136* (80)	63.4 (59/93)	*ac143* (62)	58.2 (39/67)
159		150077–>150406	109 (11983)							
160	*p49*	150355<–151737	460 (54137)	L	*agse151* (460)	92.8 (427/460)	*se137* (460)	85.5 (394/461)	*ac142* (477)	49.7 (230/463)
161	*ie-0/exon0*	151752<–152486	244 (28406)	L	*agse152* (239)	75.9 (173/228)	*se138* (244)	73.2 (167/228)	*ac141* (261)	31.5 (70/222)
162		152464–>152619	51 (6035)							
163		152603<–154921	772 (87217)	C	*agse153* (779)	78.4 (611/779)	*se139* (869)	61.9 (471/761)		

[a] Direction of ORF in the AgipMNPV genome is indicated by the arrow

[b] Number of amino acids encoded by ORF and mass in daltons

[c] Promoter motifs present upstream of ORF. C: Cap site (initiator) CA(G/T)T 120 bp upstream of start codon, preceded by a TATA box TATA(A/T)A(A/T) within 40 bp. E: early promoter motif CGTGC, 210 bp upstream of start codon. L: Late promoter motif (A/T/G)TAAG 120 bp upstream of start codon


1:polh
2:orf1629
3:pk1
4:hoar
5
6
7
8
9: odv-e56
10:me53
11
12:F protein
13
14: gp16
15: p24
16
17: lef-2
18
19: bro-a
20
hr1
21:38.7k
22:lef-1
23:v-cath
24
25: he65
26
hr1a
27:chitinase
28
29
30:gp37
31:ptp-2
32:egt
33
34
35
36
37
38:pklp-1
39
40
41:arif-1
42:pif-2
43:pif-1
44
45:fgf
46
47:alk-exo
hr2
48
49
50
51
52: rr2B
53
54: pep
55
56
57:sod
58
59: pif-3
60
61
62:ac106/107
63: hisP
hr3
64:dutpase
65
66
67
68:bro-b
69
70:p13
71:odv-e66
72
73:odv-ec43
74:ac110
75:vp80
76:p45
77:p12
78:p40
79:p6.9
80:lef-5
81:38k
82:vef-2
83:bro-c
84
85:odv-e28
86:helicase
87:odv-e25
88:ac93
89:p33
90
91:lef-4
92:vp39
93:cg30
94:vp91
95:tlp20
96: ac81
97:gp41
98:ac78
99:vlf-1
100:p26
101:iap-2
102
103:ac68
104:lef-3
105:desmoplakin
106:dnapol
107:ac75
108:ac76
109
110
111
112:bro-c
113:lef-9
114:fp25k
115:p94
116:bro-d
117
118
119
120
121
122:vp1054
123:lef-10
124
125
126:ac53
127
hr4
128:iap-3
129:bJDP
130:lef-8
131
132
133:odv-e66b
134:p47
135
hr5
136
137:ac38
138:lef-11
139:pp31
140
141
142:v-ubi
143
144
hr6
145
146:dbp
147:lef-6
148:ac29
149:p26b
150:p10
151:p74
152
153:ie-1
154:ac146
155:ac145
156:odv-ec27
157
158:odv-e18
159
160:p49
161:exon0/ie0
162
163:rr1
Baculovirus core gene
Lepidopteran baculovirus core gene
Unique to AgipMNPV

◀**Fig. 1** Map of the ORFs and other features of the AgipMNPV genome. ORFs are represented by arrows, with the position and direction of the arrow indicating ORF position and orientation. The number of each ORF is displayed, with the name of the ORF following a colon. Homologous repeat regions (hrs) are represented by hatched boxes

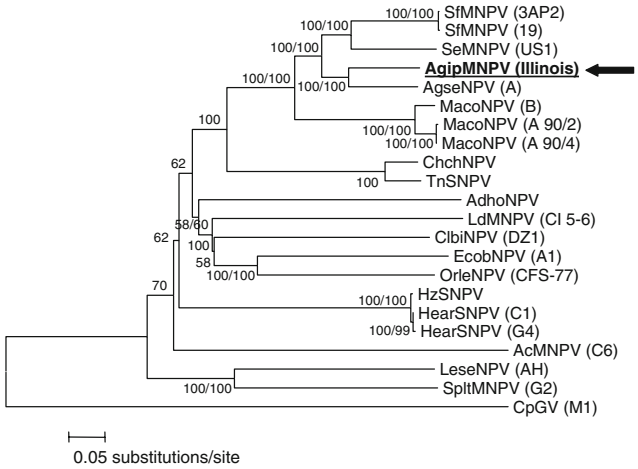


Fig. 2 Phylogenetic inference of concatenated amino acid sequence alignments of the 30 baculovirus core genes, showing bootstrap values >50% for ME and MP trees at each node (ME/MP). The location of AgipMNPV (bold) is indicated by an arrow. Shown is a consensus ME phylogram of concatenated baculovirus core gene sequence alignments for the following completely sequenced NPVs: *Spodoptera* frugiperda (SfMNPV) isolates 3AP2 and 19, *Spodoptera exigua* MNPV (SeMNPV-US1), *Agrotis segetum* NPV (AgseNPV-A), *Mamestra configurata* NPV (MacoNPV) isolates A-90/2, A-90/4, and B; *Chrysodeixis chalcites* NPV (ChChNPV), *Trichoplusia ni* SNPV (TnSNPV), *Adoxophyes honmai* NPV (AdhoNPV), *Lymantria dispar* MNPV (LdMNPV-Cl5-6), *Clanis bilineata* NPV (ClbiNPV-DZ1, accession no. DQ504428), *Ectropis oblique* NPV (EcobNPV-A1), *Orgyia leucostigma* NPV (OrleNPV-CFS-77, accession no. NC_010276), *Helicoverpa zea* SNPV (HzSNPV), *Helicoverpa armigera* NPV (HearNPV) isolates C1 and G4, *Autographa californica* MNPV (AcMNPV-C6), *Leucania separata* NPV (LeseNPV-AH), *Spodoptera litura* NPV (SpltNPV-G2), and *Cydia pomonella* GV (CpGV-M1)

similar to those found in AgseNPV-A (Fig. 4). Seven *hr*s were identified, including one (*hr1a*) consisting of a single repeat unit. The positions of AgipMNPV *hr2*, *hr3*, *hr4*, and *hr5* were conserved with those of *hr2*, *hr3*, *hr4*, and *hr5* of AgseNPV-A (Table 1). The unit repeats of the AgipMNPV *hr*s were 40–48 bp and exhibited a high degree of sequence similarity with each other (Fig. 4a). The consensus sequence of the AgipMNPV *hr* repeats in turn shared a remarkably high degree of sequence similarity with the consensus repeats of the other viruses in the *Agrotis-Spodoptera* clade (Fig. 4b).

Gene content

The 163 ORFs identified in AgipMNPV include 62 ORFs present in all lepidopteran NPVs and 138 homologs of ORFs in AgseNPV. The ORFs *agip53*, *agip56*, and *agip69* are homologous to ORFs that were previously identified as being unique to AgseNPV-A (*agse48*, *agse53*, and *agse63*, respectively). In addition, the AgipMNPV genome contains 23 ORFs with no homologs in AgseNPV (Table 2). Ten of these ORFs are homologous to other baculovirus ORFs, including a second *odv-e66* gene [50] which is orthologous to the second *odv-e66* ORFs found in some of the other group II NPVs. The other 13 ORFs are unique to AgipMNPV and exhibited either no similarity to other sequences in Genbank or BLAST matches with modest E-values. Of the unique ORFs, *agip8* (382 codons), *agip39* (354 codons), and *agip159* (109 codons) are relatively large, while the others are less than 100 codons in size. Promoter motifs are associated with four of the unique ORFs.

AgipMNPV is missing 15 ORFs present in AgseNPV-A. Seven of these ORFs are unique to AgseNPV-A, and the other eight are homologs to other baculovirus genes. Among the AgseNPV-A genes missing from AgipMNPV is *lef-7*, which encodes a late gene expression factor required for optimal levels of gene expression and replication in a cell line-specific fashion [51, 52]. Among completely sequenced group II NPVs, *lef-7* only occurs in AgseNPV, SfMNPV, SeMNPV, and MacoNPV-A. Also missing from AgipMNPV are two of the three enhancin genes (*vef-1* and *vef-3*) present in AgseNPV.

The AgipMNPV genome was found to contain four pairs of duplicated ORFs. Two of these pairs are homologs of *odv-e66* (*agip71* and *agip133*) and *p26* (*agip100* and *agip149*) genes that also occur twice in other group II NPV genomes. As reported for duplicated *odv-e66* and *p26* ORFs in other group II NPVs, the AgipMNPV ORFs in these pairs appear to derive from independent lineages. The ORFs *agip18* and *agip109* are homologs of the AcMNPV *ac150* subtype of the 11 K-like peptide family [53]. These two ORFs only share 32% amino acid sequence identity with each other. Phylogenetic inference places *agip109* in a clade containing *agse103* and homologs from SfMNPV and SeMNPV (*sf95* and *sf96*, respectively), while *agip18* occurs in a clade containing homologs from *Xestia c-nigrum granulovirus* (XecnGV; *xecn151*) and MacoNPV-A (*macoA19* from the 90/2 and 90/4 isolates of this virus; data not shown). These results suggest that *agip18* and *agip109* are also from separate lineages. The AgipMNPV genome was also found to possess a single homolog of the 11 K-like peptide *ac145* subtype. Finally, ORFs *agip36* and *agip37* both appear to be homologs of AgseMNPV ORF *agse32*. ORF *agip37* exhibited a greater degree of sequence similarity with *agse32*, and phylogenetic inference placed *agip37* in a clade with *agse32* and homologs from the *Spodoptera* NPVs (data not shown).

The AgipMNPV genome sequence also contains five *baculovirus-repeated ORF*s (*bro*s), members of a widespread multigene family found in many invertebrate

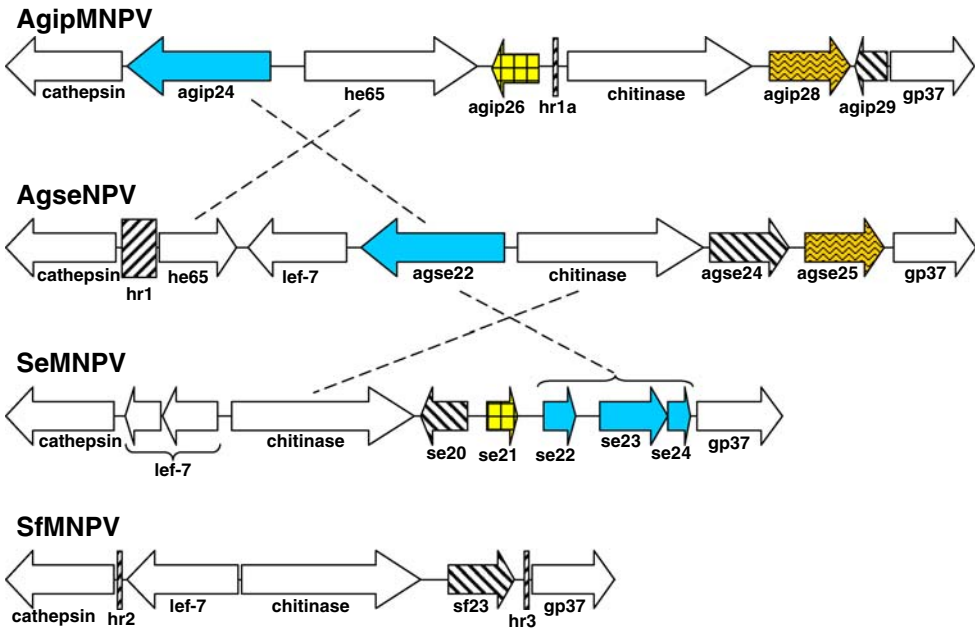


Fig. 3 ORF arrangement and orientation among NPVs of the *Agrotis-Spodotpera* group in a region encompassed by the *cathepsin* and *gp37* genes. Arrows with the same label or shade represent orthologous ORFs. Hatched arrows represent ORFs unique to individual genomes with the *Agrotis-Spodotpera* NPV group. Hatched boxes represent homologous regions (*hr*s)

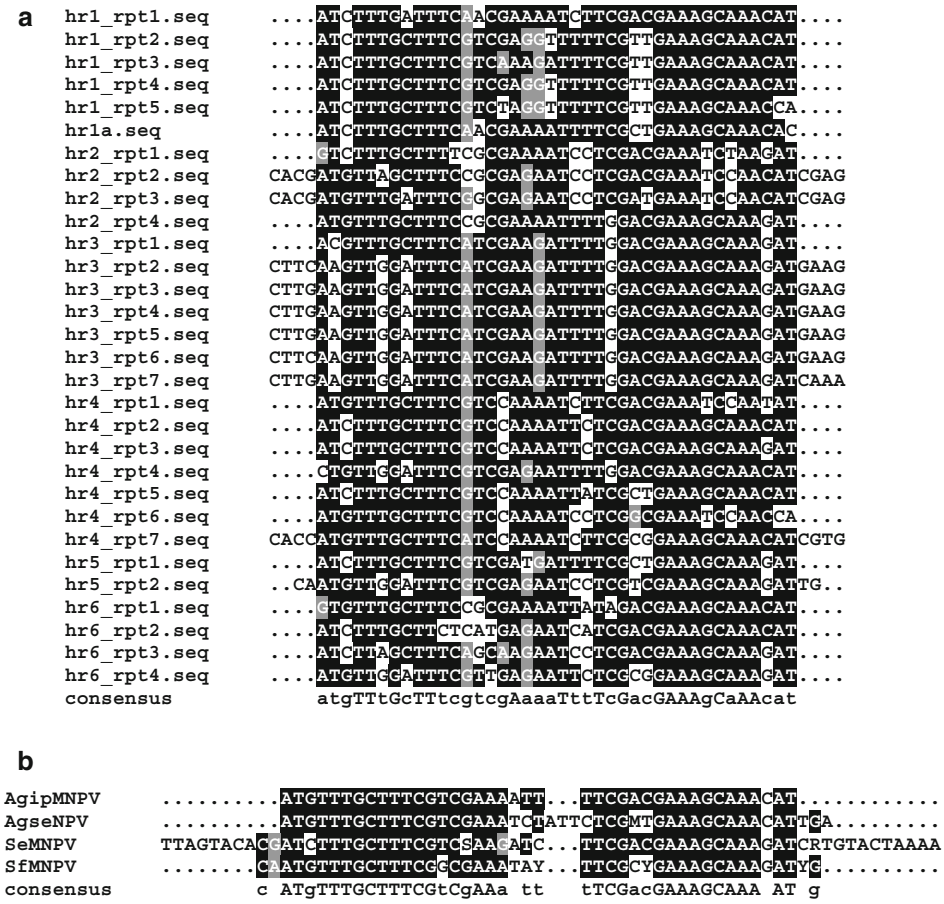


Fig. 4 Alignment of homologous region (*hr*) palindromic repeats. **a** Alignment of individual palindromic repeats from AgipMNPV *hr*s. Identical nucleotides occupying >50% of aligned positions are shaded in black, and nucleotides of the same class as conserved nucleotides (containing either a purine or pyrimidine base) are shaded in gray. Nucleotides in the consensus sequence are denoted by uppercase letters for positions in the alignment with completely identical residues, and lowercase letters for positions in the alignment with a majority of identical residues. **b** Alignment of palindromic repeat consensus sequences from AgipMNPV, AgseNPV, SeMNPV, SfMNPV-3AP2. IUB nucleotide symbols: M = A or C; R = G or A; S = C or G; and Y = C or T

viruses [54]. Of the five *bro* genes in AgipMNPV, only *bro-b* (*agip68*), *bro-c* (*agip83*), and *bro-e* (*agip116*) are orthologues of *bro* genes in the AgseNPV-A genome. The best BLAST matches for AgipMNPV *bro-a* (*agip19*) and *bro-d* (*agip112*) are LeseNPV *bro-c* (*lese52*) and MacoNPV-B *bro-e* (*macoB121*). All AgipMNPV *bro* genes

Table 2 AgipMNPV and AgseNPV ORFs with no orthologues in each other's genomes

ORFs		Features
AgipMNPV ORFs not found in AgseNPV	*agip5*, *agip6*, *agip8*, *agip11*, *agip39*, *agip48*, *agip66*, *agip67*, *agip110*, *agip144*, *agip157*, *agip159*, *agip162*	Unique to AgipMNPV
	agip7	Similar to AcMNPV *ac152*. Best match: MacoNPV-A *macoA8*
	agip19	AgipMNPV *bro-a*
	agip20	Best matches: HearGV *hear156* and LdMNPV *hrf-1*
	agip26	Similar to SeMNPV *se21*. Best match: MacoNPV-B *macoB21*
	agip29	Similar to AcMNPV *ac79*. Best match: MacoNPV-B *macoB15*
	agip65	Similar to CpGV *cp109* (best match), *cp64*, and HearSNPV *hear134*
	agip68	AgipMNPV *bro-b*
	agip71	AgipMNPV *odv-e66*
	agip112	AgipMNPV *bro-d*
	agip115	AcMNPV *p94* homolog. Best match: MacoNPV-A *macoA126*
AgseNPV ORFs not found in AgipMNPV	*agse5*, *agse14*, *agse15*, *agse42*, *agse43*, *agse61*, *agse137*	Unique to AgseNPV
	agse18	SeMNPV *se15* homolog
	agse21	*lef-7* homolog
	agse24	TnSNPV *tn62* homolog
	agse50	AgseNPV *bro-a*
	agse52	MacoNPV-B *macoB63* homolog
	agse75	AgseNPV *vef-1*
	agse123	AgseNPV *bro-d*
	agse128	AgseNPV *vef-3*

except *bro-e* possess an intact Bro-N DNA binding domain [55, 56].

Genetic variation in the AgipMNPV Illinois isolate

Genotypic variation is common among field isolates of baculoviruses [57, 58]. Though the AgipMNPV genome sequence presented in this study was derived from an uncloned field isolate, there were remarkably few nucleotide sequence polymorphisms detected during sequencing. An AT-rich region between ORFs *agip53* and *agip54* (*pep*) contains a variable number of repeats of the dinucleotide sequence AT, ranging from 36 to 44 repeats with a majority of shotgun clones containing 39 copies of the repeat. Three single-nucleotide polymorphisms (SNPs) were also detected in the ORFs *agip76* (*p45*), *agip96* (*ac81*), and *agip112* (*bro-d*). While the SNP in *agip96* is silent, the SNPs in *agip76* and *agip96* result in amino acid substitutions.

AgipMNPV genotype with a deleted *egt* gene

The baculovirus *egt* gene encodes an ecdysteroid UDP-glucosyltransferase. This enzyme inactivates the insect developmental hormone ecdysone by transferring a glucose or galactose from a nucleotide-sugar donor to a hydroxyl group on the ecdysone [59, 60]. Inactivation of ecdysone by EGT during infection is thought to increase the yield of progeny polyhedra by delaying or blocking larval development, which prevents the cessation of larval feeding that takes place just prior to molting [61–63]. The *egt* gene is not required for baculovirus infection or replication at the molecular and cellular level.

In an initial assembly of the AgipMNPV sequence, the *egt* gene was found to be split into two ORFs. Only the N-terminal half of the upstream *egt* ORF exhibited sequence similarity with other NPV EGT sequences. To confirm this arrangement in AgipMNPV, the corresponding region of the *egt* gene was amplified by PCR from OV DNA and from the BV DNA of 20 AgipMNPV plaque isolates generated during a previous study [17]. PCR and sequencing of amplimers revealed two genotypes in the OV DNA from AgipMNPV polyhedra (Fig. 5): a genotype with a single intact *egt* ORF, and a genotype bearing a 128-bp deletion (nt 32287–32414). The deletion resulted in the split *egt* gene observed earlier, causing a frameshift and premature termination of the *egt* ORF followed by an ORF encoding the C-terminal remainder of EGT. These two

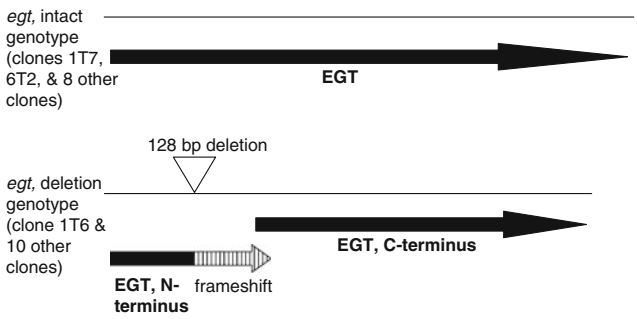


Fig. 5 Structural variation in the ecydysteroid UDP-glucosyltransferase (*egt*) gene in different genotypes of AgipMNPV. The location of a deletion in the *egt* ORF leading to a frameshift (deletion genotype, hatched arrow) in genotypes detected in the Illinois field isolate and clonal isolates of AgipMNPV is indicated

genotypes also were detected among the AgipMNPV plaque isolates. A deletion identical to that observed in OV was present in 10 of the 20 isolates (AgipMNPV-1T1, 1T2, 1T4, 1T6, 1T9, 1T10, 6T1, 6T6, 6T8, and 6T9). Nine of the AgipMNPV plaque isolates (1T3, 1T5, 1T7, 1T8, 6T2, 6T3, 6T5, 6T7, and 6T10) had an intact, undeleted *egt* gene. One isolate (6T4) contained both genotypes.

Deletions removing some or all of the *egt* coding sequence have been observed in tissue culture stocks of AcMNPV [64, 65], the Z isolate of *Antheraea pernyi* nucleopolyhedrovirus [42], and in genotypes from two separate populations of SfMNPV [18, 66]. With both Nicaraguan and American SfMNPV populations, *egt* deletion genotypes represented 85–100% of plaque isolates derived from infected larvae, yet were far less prevalent or even undetectable when viral DNA from larvae was analyzed directly by PCR or restriction endonuclease digest [18, 66]. These results suggest that *egt* deletions may confer a growth advantage in tissue culture. To assess this possibility with the *egt* deletion genotype detected in AgipMNPV, one-step growth curves were generated with plaque isolates containing deleted (AgipMNPV-1T6) and intact (AgipMNPV-1T7 and -6T2) *egt* genes. No correlation between *egt* mutation and progeny virus production was observed. The three AgipMNPV isolates produced roughly equivalent quantities of BV upon infection of AiE1611T cells (Fig. 6a), with the *egt* deletion genotype (1T6) producing moderately larger quantities of BV than clone 6T2 at 24 h p i (2.4-fold; $P = 0.044$) and clone 1T7 at 72 h p i (2.9-fold; $P = 0.0027$). AgipMNPV-6T2 produced approximately 2- to 3-fold more polyhedra/cell than the other two clones (Fig. 6b; $p \leq 0.0002$). Since mutations in the *fp25 k* gene also can account for differences in BV and polyhedra production [67, 68], the *fp25 k* genes of the three isolates were amplified and sequenced. No mutations were detected in the *fp25 k* genes of the three isolates.

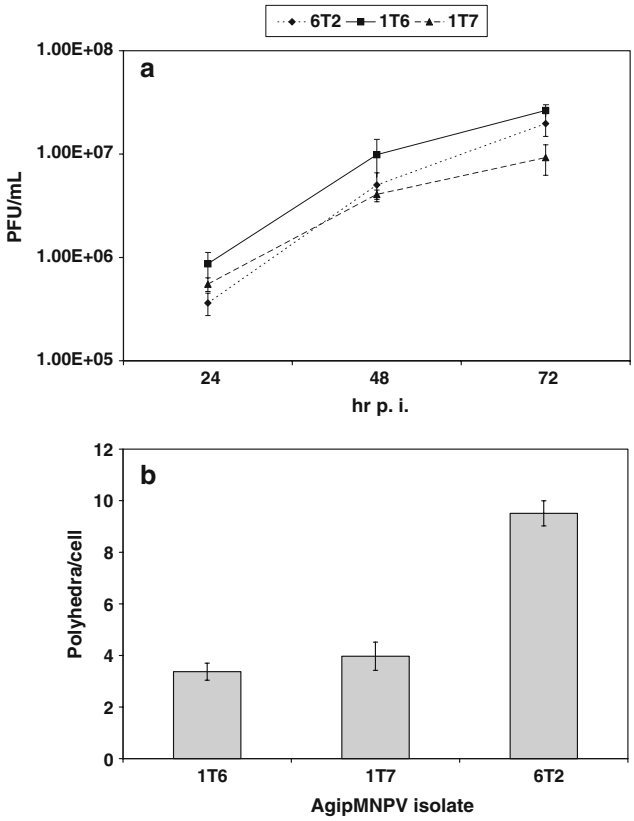


Fig. 6 Virus production by AgipMPNPV plaque isolates with intact or deleted *egt* genes. **a** Titers of budded virus produced by AiE1611T cells infected with AgipMNPV clone 6T2 (◆), 1T6 (■), and 1T7 (▲), displayed as average plaque-forming units (pfu)/ml of three replicate samples/time point. **b** Polyhedra produced by AiE1611T cells infected with AgipMNPV-1T6, -1T7, and -6T2. Average polyhedra from three replicate infections/virus are shown. For both (**a**) and (**b**), error bars represent one standard deviation

Discussion

AgipMNPV is part of a well-defined clade of group II NPVs that includes viruses from *Agrotis* and *Spodoptera* host species. The presence of *hr*s with a core repeat of the sequence TTTGCTTT(N_{18-21})AAAGCAAA appears to be diagnostic for NPVs of this clade. Among the NPVs of this group that have been sequenced, AgipMNPV is most closely related to AgseNPV-A. The two viruses nevertheless have diverged significantly, with an average ORF amino acid sequence identity of 70.6% (± 15.9%), and clearly are distinct virus species. While *A. ipsilon* is found in both the United States and Europe, AgipMNPV isolates have only been identified in the United States. *A. segetum*, in contrast, is not believed to be in the United States, and NPVs isolated from *A. segetum* originate only from Europe. The resulting geographic separation and adaptation to different hosts likely accounts for the degree of divergence between AgipMNPV and AgseNPV-A.

The AgseNPV-A genome is unusual in having three enhancin genes. The LdMNPV genome possesses two enhancin genes, both of which are necessary for optimal virulence against *L. dispar* larvae [69, 70]. While AgipMNPV only retains one of the three enhancin genes possessed by AgseNPV, it nevertheless exhibits a 9.4-fold lower LC_{50} against *A. ipsilon* larvae and an LC_{50} against *A. segetum* larvae that is only moderately (2.7-fold) higher than that of AgseNPV-A. This observation suggests that the presence and expression of all three AgseNPV-A enhancins only may be required for optimal potency against *A. segetum*, and also that the two additional enhancins in AgseNPV-A only make a modest contribution to virulence against *A. segetum* larvae. The AgipMNPV genome has an extra copy of an *ac150* orthologue.

Inactivating *ac150* in AcMNPV results in significantly reduced virulence against three different hosts [71], and appears to work synergistically to increase virulence against *Heliothis virescens* larvae [72]. The extra *ac150* gene in AgipMNPV may account for its lower LC_{50} against *A. ipsilon* larvae. Experiments with recombinant AgipMNPV or AgseNPV-A with inactivated or additional enhancins and *ac150* genes may clarify the role that these genes play in host range differences reported for these viruses.

Acknowledgements The author wishes to thank Tad Sonstegard (USDA-ARS, Bovine Functional Genomics Laboratory, Beltsville, MD) for the use of his GeneMachines Hydroshear, Jing Hu and Dan Rowley (USDA-ARS, Invasive Insect Biocontrol and Behavior Laboratory, Beltsville, MD) for assistance with sequencing, and anonymous referees for a critical review of the manuscript. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture.

References

1. B·C. Bonning, in *Comprehensive Molecular Insect Science*, ed. by L. Gilbert, K. Iatrou, S·S. Gill (Elsevier, Amsterdam; Boston, 2005), pp. 233–269
2. D.A. Theilmann, G.W. Blissard, B. Bonning, J.A. Jehle, D.R. O'Reilly, G.F. Rohrmann, S. Thiem, J.M. Vlak. in *Virus Taxonomy-Classification and Nomenclature of Viruses 8th Report of the International Committee on the Taxonomy of Viruses*, ed. by C.X. Fauquet, M.A. Mayo, J. Maniloff, U. Desselberger, L.A. Ball (Elsevier, Amsterdam, 2005), pp. 177–185
3. J.A. Jehle, G.W. Blissard, B.C. Bonning, J.S. Cory, E.A. Herniou, G.F. Rohrmann, D.A. Theilmann, S.M. Thiem, J.M. Vlak, Arch. Virol. **151**, 1257–1266 (2006). doi:10.1007/s00705-006-0763-6
4. T.A. Kost, J.P. Condreay, D.L. Jarvis, Nat. Biotechnol. **23**, 567–575 (2005). doi:10.1038/nbt1095
5. M.D. Summers, Adv. Virus Res. **68**, 3–73 (2006). doi:10.1016/S0065-3527(06)68001-9
6. F. Moscardi, Annu. Rev. Entomol. **44**, 257–289 (1999). doi:10.1146/annurev.ento.44.1.257
7. N. van Beek, D.C. Davis, Methods Mol. Biol. **388**, 367–378 (2007). doi:10.1007/978-1-59745-457-5-19
8. A.J. Boughton, R.L. Harrison, L.C. Lewis, B.C. Bonning, J. Invertebr. Pathol. **74**, 289–294 (1999). doi:10.1006/jipa.1999.4901
9. R.W. Rings, F.J. Arnold, B.A. Johnson, B.C. Bonning, Bull. Entomol. Soc. Am. **21**, 229–234 (1975)
10. A.J. Boughton, L.C. Lewis, J. Econ. Entomol. **94**, 1045–1052 (2001)
11. C.A. Prater, C.T. Redmond, W. Barney, B.C. Bonning, D.A. Potter, J. Econ. Entomol. **99**, 1129–1137 (2006)
12. D.A. Potter, A. Bixby, Green Sect. Rec. **46**, 11–13 (2008)
13. S. El-Salamouny, M. Lange, M. Jutzi, J. Huber, J.A. Jehle, J. Invertebr. Pathol. **84**, 75–82 (2003). doi:10.1016/j.jip.2003.08.005
14. A.K. Jakubowska, S.A. Peters, J. Ziemnicka, J.M. Vlak, M.M. van Oers, J. Gen. Virol. **87**, 537–551 (2006). doi:10.1099/vir.0.81461-0
15. D.R. O'Reilly, L.K. Miller, V.A. Luckow, *Baculovirus Expression Vectors* (W·H. Freeman and Company, New York, 1992)
16. J.M. Slack, S.D. Lawrence, Methods Cell Sci. **24**, 155–163 (2002). doi:10.1023/A:1024413604663
17. R.L. Harrison, D.E. Lynn, J. Invertebr. Pathol. **99**, 28–34 (2008). doi:10.1016/j.jip.2008.02.015
18. R.L. Harrison, B. Puttler, H.J. Popham, J. Gen. Virol. **89**, 775–790 (2008). doi:10.1099/vir.0.83566-0
19. J.D. Thompson, D.G. Higgins, T.J. Gibson, Nucleic Acids Res. **22**, 4673–4680 (1994). doi:10.1093/nar/22.22.4673
20. K. Tamura, J. Dudley, M. Nei, S. Kumar, Mol. Biol. Evol. **24**, 1596–1599 (2007). doi:10.1093/molbev/msm092
21. J. Kuzio, M.N. Pearson, S.H. Harwood, C.J. Funk, J.T. Evans, J.M. Slavicek, G.F. Rohrmann, Virology **253**, 17–34 (1999). doi:10.1006/viro.1998.9469
22. W.F. IJkel, E.A. van Strien, J.G. Heldens, R. Broer, D. Zuidema, R.W. Goldbach, J.M. Vlak, J. Gen. Virol. **80**, 3289–3304 (1999)
23. Y. Pang, J. Yu, L. Wang, X. Hu, W. Bao, G. Li, C. Chen, H. Han, S. Hu, H. Yang, Virology **287**, 391–404 (2001). doi:10.1006/viro.2001.1056
24. X. Chen, I.J. WF, R. Tarchini, X. Sun, H. Sandbrink, H. Wang, S. Peters, D. Zuidema, R.K. Lankhorst, J.M. Vlak, Z. Hu, J. Gen. Virol. **82**, 241–257 (2001)
25. X. Chen, W.J. Zhang, J. Wong, G. Chun, A. Lu, B.F. McCutchen, J.K. Presnail, R. Herrmann, M. Dolan, S. Tingey, Z.H. Hu, J.M. Vlak, J. Gen. Virol. **83**, 673–684 (2002)
26. Q. Li, C. Donly, L. Li, L.G. Willis, D.A. Theilmann, M. Erlandson, Virology **294**, 106–121 (2002). doi:10.1006/viro.2001.1313
27. L. Li, C. Donly, Q. Li, L.G. Willis, B.A. Keddie, M.A. Erlandson, D.A. Theilmann, Virology **297**, 226–244 (2002). doi:10.1006/viro.2002.1411
28. M. Nakai, C. Goto, W. Kang, M. Shikata, T. Luque, Y. Kunimi, Virology **316**, 171–183 (2003). doi:10.1016/j.virol.2003.08.002
29. L.G. Willis, R. Seipp, T.M. Stewart, M.A. Erlandson, D.A. Theilmann, Virology **338**, 209–226 (2005). doi:10.1016/j.virol.2005.04.041
30. C.X. Zhang, X.C. Ma, Z.J. Guo, Virology **333**, 190–199 (2005). doi:10.1016/j.virol.2004.12.028
31. L. Li, Q. Li, L.G. Willis, M. Erlandson, D.A. Theilmann, C. Donly, J. Gen. Virol. **86**, 91–105 (2005). doi:10.1099/vir.0.80488-0
32. M.M. van Oers, M.H. Abma-Henkens, E.A. Herniou, J.C. de Groot, S. Peters, J.M. Vlak, J. Gen. Virol. **86**, 2069–2080 (2005). doi:10.1099/vir.0.80964-0
33. X.C. Ma, J.Y. Shang, Z.N. Yang, Y.Y. Bao, Q. Xiao, C.X. Zhang, Virology **360**, 235–246 (2007). doi:10.1016/j.virol.2006.10.024
34. H. Xiao, Y. Qi, Virus Genes **35**, 845–856 (2007). doi:10.1007/s11262-007-0106-z
35. J.L. Wolff, F.H. Valicente, R. Martins, J.V. Oliveira, P.M. Zanotto, J. Gen. Virol. **89**, 1202–1211 (2008). doi:10.1099/vir.0.83581-0
36. M.D. Ayres, S.C. Howard, J. Kuzio, M. Lopez-Ferber, R.D. Possee, Virology **202**, 586–605 (1994). doi:10.1006/viro.1994.1380
37. T. Luque, R. Finch, N. Crook, D.R. O'Reilly, D. Winstanley, J. Gen. Virol. **82**, 2531–2547 (2001)
38. T. Hayakawa, R. Ko, K. Okano, S.I. Seong, C. Goto, S. Maeda, Virology **262**, 277–297 (1999). doi:10.1006/viro.1999.9894
39. R.L. Harrison, H.J. Popham, Virus Genes **36**, 565–581 (2008). doi:10.1007/s11262-008-0218-0
40. C.H. Ahrens, R.L. Russell, C.J. Funk, J.T. Evans, S.H. Harwood, G.F. Rohrmann, Virology **229**, 381–399 (1997). doi:10.1006/viro.1997.8448
41. Q. Fan, S. Li, L. Wang, B. Zhang, B. Ye, Z. Zhao, L. Cui, Virology **366**, 304–315 (2007). doi:10.1016/j.virol.2007.04.027
42. Z.M. Nie, Z.F. Zhang, D. Wang, P.A. He, C.Y. Jiang, L. Song, F. Chen, J. Xu, L. Yang, L.L. Yu, J. Chen, Z.B. Lv, J.J. Lu, X.F. Wu, Y.Z. Zhang, BMC Genomics **8**, 248 (2007). doi:10.1186/1471-2164-8-248

43. C.L. Afonso, E.R. Tulman, Z. Lu, C.A. Balinsky, B.A. Moser, J.J. Becnel, D.L. Rock, G.F. Kutish, J. Virol. **75**, 11157–11165 (2001). doi:10.1128/JVI.75.22.11157-11165.2001
44. J.G. de Jong, H.A. Lauzon, C. Dominy, A. Poloumienko, E.B. Carstens, B.M. Arif, P.J. Krell, J. Gen. Virol. **86**, 929–943 (2005). doi:10.1099/vir.0.80490-0
45. M.M. van Oers, J.M. Vlak, Curr. Drug Targets **8**, 1051–1068 (2007). doi:10.2174/138945007782151333
46. C.B. McCarthy, D.A. Theilmann, Virology **375**, 277–291 (2008). doi:10.1016/j.virol.2008.01.039
47. Z.H. Hu, B.M. Arif, F. Jin, J.W. Martens, X.W. Chen, J.S. Sun, D. Zuidema, R.W. Goldbach, J.M. Vlak, J. Gen. Virol. **79**, 2841–2851 (1998)
48. S. Hilton, D. Winstanley, J. Gen. Virol. **88**, 1496–1504 (2007). doi:10.1099/vir.0.82760-0
49. R.D. Possee, G.F. Rohrmann, in *The Baculoviruses*, ed. by L.K. Miller (Plenum, New York, 1997), pp. 109–140
50. T. Hong, S.C. Braunagel, M.D. Summers, Virology **204**, 210–222 (1994). doi:10.1006/viro.1994.1525
51. C.J. Chen, S.M. Thiem, Virology **227**, 88–95 (1997). doi:10.1006/viro.1996.8341
52. A. Lu, L.K. Miller, J. Virol. **69**, 6265–6272 (1995)
53. D. Dall, T. Luque, D. O'Reilly, Bioessays **23**, 184–193 (2001). doi:10.1002/1521-1878(200102)23:2<184::AID-BIES1026>3.0.CO;2-H
54. D.K. Bideshi, S. Renault, K. Stasiak, B.A. Federici, Y. Bigot, J. Gen. Virol. **84**, 2531–2544 (2003). doi:10.1099/vir.0.19256-0
55. L.M. Iyer, E.V. Koonin, L. Aravind, Genome Biol. **3**, RESEARCH0012 (2002)
56. E.A. Zemskov, W. Kang, S. Maeda, J. Virol. **74**, 6784–6789 (2000). doi:10.1128/JVI.74.15.6784-6789.2000
57. J.S. Cory, B.M. Green, R.K. Paul, F. Hunter-Fujita, J. Invertebr. Pathol. **89**, 101–111 (2005). doi:10.1016/j.jip.2005.03.008
58. J.S. Cory, J.H. Myers, Annu. Rev. Ecol. Evol. Syst. **34**, 239–272 (2003). doi:10.1146/annurev.ecolsys.34.011802.132402
59. D.R. O'Reilly, M.R. Brown, L.K. Miller, Insect Biochem. Mol. Biol. **22**, 313–320 (1992). doi:10.1016/0965-1748(92)90069-Q
60. D.R. O'Reilly, O.W. Howarth, H.H. Rees, L.K. Miller, Insect Biochem. **21**, 795–801 (1991). doi:10.1016/0020-1790(91)90121-T
61. D.R. O'Reilly, R.S. Hails, T.J. Kelly, J. Invertebr. Pathol. **72**, 269–275 (1998). doi:10.1006/jipa.1998.4785
62. D.R. O'Reilly, L.K. Miller, Science **245**, 1110–1112 (1989). doi:10.1126/science.2505387
63. K.R. Wilson, D.R. O'Reilly, R.S. Hails, J.S. Cory, Biol. Control **19**, 57–63 (2000). doi:10.1006/bcon.2000.0841
64. S. Kumar, L.K. Miller, Virus Res. **7**, 335–349 (1987). doi:10.1016/0168-1702(87)90047-5
65. D.R. O'Reilly, A.L. Passarelli, I.F. Goldman, L.K. Miller, J. Gen. Virol. **71**, 1029–1037 (1990). doi:10.1099/0022-1317-71-5-1029
66. O. Simon, T. Williams, M. Lopez-Ferber, P. Caballero, Appl. Environ. Microbiol. **70**, 5579–5588 (2004). doi:10.1128/AEM.70.9.5579-5588.2004
67. B. Beames, M.D. Summers, Virology **168**, 344–353 (1989). doi:10.1016/0042-6822(89)90275-4
68. R.L. Harrison, M.D. Summers, J. Gen. Virol. **76**, 1451–1459 (1995). doi:10.1099/0022-1317-76-6-1451
69. D.S. Bischoff, J.M. Slavicek, J. Virol. **71**, 8133–8140 (1997)
70. H.J. Popham, D.S. Bischoff, J.M. Slavicek, J. Virol. **75**, 8639–8648 (2001). doi:10.1128/JVI.75.18.8639-8648.2001
71. J.H. Zhang, T. Ohkawa, J.O. Washburn, L.E. Volkman, J. Gen. Virol. **86**, 1619–1627 (2005). doi:10.1099/vir.0.80930-0
72. R. Lapointe, H.J. Popham, U. Straschil, D. Goulding, D.R. O'Reilly, J.A. Olszewski, J. Virol. **78**, 6439–6448 (2004). doi:10.1128/JVI.78.12.6439-6448.2004