Virus Genes (2008) 36:565–581
DOI 10.1007/s11262-008-0218-0

Genomic sequence analysis of a granulovirus isolated from the Old World bollworm, *Helicoverpa armigera*

Robert L. Harrison · Holly J. R. Popham

Received: 19 December 2007 / Accepted: 28 February 2008 / Published online: 17 April 2008

Abstract The genome of a granulovirus isolated from the Old World bollworm, *Helicoverpa armigera*, was completely sequenced. The size of the *Helicoverpa armigera* granulovirus (HearGV) genome is 169,794 nt containing 179 open reading frames (ORFs), making it the second largest baculovirus genome analyzed to date. The genomes of HearGV and the *Xestia c-nigrum* GV (XecnGV) exhibit extensive sequence similarity and co-linearity, with both genomes containing the same nine homologous regions (*hr*s) with conserved structure and locations and sharing 167 open reading frames (ORFs). Phylogenetic inference and pairwise analysis of Kimura-2-parameter nucleotide distances for the *lef-8*, *lef-9*, and *granulin* genes indicate that HearGV is part of a cluster of granuloviruses typified by XecnGV. The HearGV genome contains all 62 ORFs found in common among other fully sequenced lepidopteran baculovirus genomes, as well as seven ORFs unique to HearGV. In addition, HearGV and XecnGV genomes share 20 ORFs not found among other baculovirus genomes sequenced to date. In addition to possessing ten ORFs with sequence similarity to baculovirus repeated ORFs (*bro*), the HearGV genome contains members of two other gene families with homologues in ascovirus, nucleopolyhedrovirus, and entomopoxvirus genomes. Alignment of the HearGV and XecnGV genome sequences revealed that HearGV is missing approximately 16.6 kbp of XecnGV-homologous sequence and contains approximately 8.2 kbp of sequence not found in the XecnGV genome.

Keywords Baculovirus · Granulovirus · *Helicoverpa armigera* · HearGV · Xestia c-nigrum · XecnGV

The nucleotide sequence data reported in this paper have been submitted to the GenBank nucleotide sequence database and have been assigned the accession number EU255577.

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

H. J. R. Popham
Biological Control of Insects Research Laboratory, USDA, Agricultural Research Service, 1503 S. Providence Road, Columbia, MO 65203, USA
e-mail: Holly.Popham@ars.usda.gov

Introduction

Baculoviruses are DNA viruses of the arthropod-specific virus family Baculoviridae [1]. Members of Baculoviridae have been isolated primarily from the insect order Lepidoptera (moths and butterflies), but some baculoviruses have also been described from mosquitoes, sawflies, and shrimp. Currently, baculoviruses are divided into the genera *Nucleopolyhedrovirus* and *Granulovirus* [2]. Both nucleopolyhedroviruses (NPVs) and granuloviruses (GVs) produce two structurally and functionally distinct types of virions: an occluded form that initiates primary infection of the host insect, and a budded form that spreads infection to other tissues of the infected host. NPVs produce cuboidal polyhedra containing many occluded virions, while GVs produce smaller spheroid occlusions containing a single virion. The GVs have been isolated exclusively from Lepidoptera. GVs cause three distinct types of pathology in infected hosts [3]. Type 1 GV pathology is characterized by an infection limited to the host's midgut and fat body resulting in a relatively slow speed of kill. Type 2 GV

pathology is characterized by infection of most of the host's tissues and a rapid speed of kill. A single GV, the *Harrisina brillians granulovirus*, causes a third type of pathology resulting in an infection that is constrained to the midgut epithelium and that results in the rapid death of the host. Phylogenetic analysis of GV sequences suggests that these different types of GV pathogenesis do not have monophyletic origins [4].

Baculovirus genomes consist of a single large circular double-stranded DNA molecule, ranging in size from 80 to 180 kbp. Several NPV genome sequences have been reported, but to date complete genome sequences have only been reported for six GVs [5–10] with genome sequences for another three GVs (from *Phthorimaea operculella*, *Agrotis segetum*, and *Spodoptera litura*) on file in GenBank.

A GV from the Old World bollworm, *Helicoverpa armigera*, was first described by Whitlock [11]. This virus, *H. armigera* granulovirus (HearGV), kills larvae of *H. armigera* slowly and appears to cause type 1 GV pathology. HearGV exhibits a relatively broad host range in bioassays compared to other GVs [12]. Of particular interest is the interaction between HearGV and other NPVs in mixed infections. The LC_{50} of *Lymantria dispar* multiple nucleopolyhedrovirus (LdMNPV) against *L. dispar* larvae is reduced when LdMNPV polyhedra was mixed with HearGV granules [13], probably due to the action of enhancins in HearGV granules [14]. However, mixed infections of *H. armigera* larvae with *H. armigera* single nucleopolyhedrovirus (HearSNPV) and HearGV results in lower mortalities than infections with either HearSNPV or HearGV alone [15]. In co-infections of *Helicoverpa zea* with HearGV and *H. zea* single nucleopolyhedovirus (HzSNPV), survival times of larvae increase with an increasing dose of HearGV, even when HearGV is applied 36 hr after infection with HzSNPV [16]. Cadavers from co-infected larvae contain HearGV granules instead of HzSNPV polyhedra, suggesting that HearGV inhibits HzSNPV replication.

To add to our knowledge of granulovirus molecular genetics and to acquire preliminary information for studies into HearGV's broad host range and ability to interfere with NPV infection, the complete genome of HearGV was sequenced and analyzed.

Materials and methods

Viral DNA extraction and cloning into fosmid vectors

The HearGV isolate used in previous studies [13, 16] was amplified in *Heliothis virescens* larvae by *per os* infection. Viral particles were isolated by sucrose gradient centrifugation and DNA was subsequently extracted from the purified particles [17]. Because HearGV could not be propagated in cell culture, viral DNA from granules was sheared and cloned into the vector pCC1FOS using the Copy Control Fosmid Library Production Kit (Epicentre Biotechnologies). Fosmid DNA was purified as per kit instructions. Following the characterization of 60 clones by *Eco*RI, *Hind*III, *Bg*II, *Nhe*I, and *Bam*HI restriction endonuclease digest, a complete genomic library of HearGV was chosen consisting of seven overlapping fosmid clones.

DNA sequencing and analysis

The DNA of fosmid clones containing HearGV inserts was sheared and cloned into the pCR-Blunt II-TOPO plasmid vector (Invitrogen) as previously described [18]. White colonies from the cloning procedure were picked, thermally lysed, and inserts amplified by PCR with vector-specific primers as described [18]. The PCR products were precipitated with 20% PEG-2.5 M NaCl to remove excess primers and dNTPs as described [18] and sequenced using nested plasmid vector-specific primers T7 (5′GTAATACG ACTCACTATAGGG-3′) and SP6 (5′-GCTATTTAGGTG ACACTATAG-3′). Reactions were carried out with the Applied Biosytems BigDye Terminator Cycle Sequencing kit with AmpliTaq DNA polymerase, and fragments were electrophoresed on an Applied Biosystems 3100 DNA sequencer.

Contigs were assembled from DNA sequencing runs with the Seqman program of the Lasergene suite (DNASTAR, Inc.). Gaps and ambiguities in the genome sequence were resolved by amplifying the corresponding regions of the sequence from both fosmid and viral DNA by PCR (40 pg DNA/reaction) with custom-designed primers and sequencing the PCR products.

Open reading frames (ORFs) 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 region (*hr*) were identified and selected for further characterization. ORFs with homologues in other baculovirus genomes also were characterized. Predicted amino acid sequence identities were obtained from the results of protein database searches using the standard protein-protein BLAST algorithm (http://www.ncbi.nlm.nih.gov/BLAST/).

For phylogenetic inference, amino acid sequences derived from selected genes were aligned by ClustalW [19] using Gonnet matrices with a gap penalty of 10 and a gap extension penalty of 0.1 for pairwise alignments and 0.2 for multiple alignments. Sequence alignments for different genes were concatenated using BioEdit [20]. The concatenated amino acid alignments were used to construct phylograms with MEGA version 4.0 [21] using minimum evolution (ME) and maximum parsimony (MP) methods. ME and MP trees were sought by using a close-neighbor-interchange heuristic

search, starting with either 1 initial neighbor-joining tree (ME) or 10 initial trees generated by random addition of sequences (MP). For ME trees, Poisson correction distances were estimated with a gamma shape parameter of 2.25. In both cases, the reliability of the trees was tested with bootstrap re-sampling using 1,000 replicates. To calculate Kimura-2-parameter nucleotide distances [22], nucleotide sequences were aligned by ClustalW using IUB matrices, gap penalties of 15 and gap extension penalties of 6.66 for pairwise and multiple alignments. Individual and concatenated nucleotide alignments were used to calculate distances using MEGA version 4.0.

Results and discussion

Characteristics of the HearGV genome sequence

A 13.4X sequence of the HearGV genome was compiled from all sequence data generated. The size of the final draft sequence was 169,794 nt, which makes the HearGV genome the second largest baculovirus genome sequenced after the XecnGV genome, which is 178,733 nt [5]. The third largest baculovirus genome is that of *Leucania separata* nucleopolyhedrovirus, at 168,041 nt [23], followed by *Lymantria dispar* multiple nucleopolyhedrovirus (LdMNPV) at 161,046 nt [24]. The HearGV genome has an A+T content of 59.2%, which is closest to that of XecnGV (59.3%) and *Plutella xylostella* granulovirus (59.3%; [6]). Among GV genomes, only the *Cydia pomonella* granulovirus (CpGV; [7]) has a lower A+T content at 54.8%.

One hundred seventy-nine ORFs were identified that equal or exceed 50 codons in length, had minimal overlap with larger ORFs or shared significant sequence identity with previously characterized baculovirus ORFs (Table 1, Fig. 1). The ORF encoding granulin was designated as the first ORF (*hear1*), and the first nucleotide position for the HearGV genome sequence was set to the adenine of the *granulin* ORF initiation codon. As with other baculovirus genomes, the ORFs were randomly distributed with 90 ORFs in the granulin-sense orientation and 89 in the opposite orientation. Canonical baculovirus early and late gene promoter sequences were associated with 93 of the HearGV ORFs (Table 1).

In regions of the HearGV genome sequence that were covered by more than one fosmid clone, sequence polymorphisms were detected. Amplification and sequencing of selected regions directly from viral DNA where polymorphisms were observed in the assembled sequence confirmed the presence of polymorphic sites in the viral DNA preparation from which the fosmids were constructed. No more than two different nucleotides were observed at polymorphic sites, even in regions covered by more than two fosmid clones. The polymorphisms numbered 333 substitutions and 24 insertions and deletions (indels) ranging from 1 to 71 nt in size. Of the 230 substitutions that occurred in ORFs, 126 were silent, while 104 resulted in amino acid changes. Of the fourteen indels occurring in ORFs, three indels did not alter the reading frame, while eleven caused frameshifts. This result indicated that the HearGV sample used in this study consisted of more than one genotype. Genotypic variation is common among field isolates of baculoviruses [10, 25–27].

Baculovirus genomes generally contain clusters of repeated sequences known as homologous repeat regions (*hr*s). The *hr*s function as enhancers of transcription and origins of replication [28, 29]. While NPV *hr*s typically consist of repeated palindromes, GV *hr*s are more variable in structure. Nine *hr*s were identified in HearGV that consisted of the same kind of repeated sequences as found in the *hr*s of XecnGV [5], with two conserved 10-bp core sequences. These repeats shared the same relative locations on the genome as the XecnGV repeats, and thus were given the same numbering designations. A large proportion of sequence polymorphisms was found in HearGV *hr1*, with 65 substitutions and 6 indels observed among shotgun clones from the fosmids containing this region. While HearGV and XecnGV *hr1*, *hr5a*, and *hr6* contained the same number of repeats, *hr2*, *hr3*, *hr4*, *hr5*, *hr7*, and *hr8* differed between the two viruses by one or two repeats. Differences in the number of repeats among the *hr*s of closely related viruses have been documented previously for several pairs of closely related NPVs [18, 25, 30–32].

Relationships with other granuloviruses

Roelvink and co-workers [14] previously sequenced a 3,213-bp region from a *H. armigera* granulovirus. This sequence is 100% identical with three 1-nt gaps to nt 146202–149415 of the HearGV genome reported here. BLAST searches with HearGV nucleotide sequences and ORF predicted amino acid sequences revealed that HearGV also was closely related to XecnGV. To examine the relationship of HearGV to other GVs, phylogenetic trees were inferred from a set of concatenated aligned partial amino acid sequences of granulin and late expression factors 8 and 9 (lef-8 and lef-9; Fig. 2) for HearGV, other completely sequenced GVs, *Autographa californica* multiple nucleopolyhedrovirus (AcMNPV-C6, [33]) and partial sequences from other GVs [4, 34]. This analysis confirmed the close relationship of HearGV and XecnGV. HearGV and XecnGV were placed in a clade of closely related GVs isolated from Lepidoptera of family Noctuidae, including *Autographa gamma* GV, *Hoplodrina ambigua* GV, *Euxoa ochrogaster* GV, and *Scotogramma trifolii* GV (Fig. 2).

Table 1 Features of the HearGV genome

ORF/other feature	Name	Position[a]	aa (Da)[b]	Promoter motifs[c]	Comparison with other viruses					
					XecnGV		CpGV		PlxyGV	
					ORF (size)/hr	% ID (range)	ORF (size)	% ID (range)[d]	ORF (size)	% ID (range)[d]
1	*granulin*	1 → 747	248 (29204)	L	*xc1* (248)	100 (248/248)	*cp1* (248)	87.1 (216/248)	*px4* (248)	86.7 (215/248)
2	*orf1629* (capsid)	794 ← 1495	233 (26504)	L	*xc2* (231)	96.6 (225/233)	*cp2* (174)	54.3 (19/35)	*px5* (131)	NSS
3	*pk-1*	1428 → 2327	299 (35707)	L	*xc3* (302)	98.0 (296/302)	*cp3* (279)	42.2 (122/289)	*px6* (302)	55.1 (156/283)
4		2361 ← 3335	324 (38859)	L	*xc4* (324)	96.3 (312/324)				
5	*p10*	3457 → 3711	84 (9046)	L	*xc5* (84)	98.8 (83/84)			*px2* (83)	34.7 (25/72)
6		3746 ← 4309	187 (21934)		*xc7* (187)	99.5 (186/187)	*cp4* (188)	35.7 (65/182)	*px8* (187)	25.1 (46/183)
7		4299 → 4559	86 (10278)	L	*xc8* (86)	97.7 (84/86)			*px9* (93)	33.3 (26/78)
8	*ie-1*	4566 ← 6023	485 (56662)		*xc9* (484)	97.3 (472/485)	*cp7* (488)	25.7 (131/510)	*px10* (484)	31.3 (119/380)
9	*ac146*	6055 → 6645	196 (21760)		*xc10* (196)	98.0 (192/196)	*cp8* (192)	33.7 (65/193)	*px11* (180)	30.7 (51/166)
10	*ac145*	6669 ← 6968	99 (11567)	L	*xc11* (99)	99.0 (98/99)	*cp9* (101)	52.0 (52/100)	*px12* (99)	54.1 (53/98)
11	*odv-e18*	6987 ← 7238	83 (8986)	L	*xc12* (83)	98.8 (82/83)	*cp14* (84)	69.5 (57/82)	*px13* (83)	68.8 (33/48)
12	*p49*	7242 ← 8603	453 (52683)	L	*xc13* (453)	97.1 (440/453)	*cp15* (457)	41.6 (192/461)	*px14* (446)	40.6 (186/458)
13		8675 ← 9364	229 (25821)		*xc14* (229)	98.7 (226/229)			*px15* (216)	24.4 (52/213)
14	*odv-e56*	9381 ← 10442	353 (38505)	L	*xc15* (353)	87.5 (309/353)	*cp18* (355)	58.2 (203/349)	*px16* (351)	50.9 (179/352)
15	*ac29*	10472 → 10687	71 (8471)		*xc16* (71)	97.2 (69/71)	*cp19* (75)	34.6 (18/52)	*px17* (58)	NSS
16	*pep*	10707 ← 11270	187 (21477)	L	*xc17* (187)	98.9 (185/187)	*cp20* (235)	43.0 (99/230)	*px20* (235)	38.2 (91/238)
17	*pep*	11359 ← 11820	153 (17244)	L	*xc18* (153)	99.3 (152/153)	*cp23* (152)	52.9 (74/140)	*px23* (131)	48.1 (63/131)
18	*pep/p10*	11841 → 13001	386 (40761)	L	*xc19* (386)	99.7 (385/386)	*cp22* (347)	49.9 (192/385)	*px21* (320)	45.9 (175/381)
19		13063 → 13341	92 (10248)	L	*xc20* (91)	84.8 (78/92)	*cp9* (101)	29.3 (22/75)		
20	*p94*	13428 → 15902	824 (95835)	L	*xc21* (826)	95.5 (789/826)				
21	*sprT*	15957 ← 16436	159 (18011)							
22		16638 ← 17663	341 (39797)		*xc23* (350)	85.1 (303/356)				
23		17737 ← 18111	124 (13969)	L	*xc24* (124)	96.8 (120/124)				
24		19361 → 20770	469 (53422)	L	*xc25* (449)	95.3 (428/449)	*cp29* (304)	24.0 (46/192)	*px24* (338)	24.4 (40/164)
25		21293 → 22405	370 (42916)		*xc26* (370)	94.1 (348/370)	*cp30* (181)	NSS	*px25* (431)	20.7 (79/382)
26	*efp*	22479 → 24224	581 (67488)		*xc27* (599)	96.0 (556/579)	*cp31* (601)	32.3 (195/603)	*px26* (544)	33.3 (193/579)
27		24381 → 25364	327 (37466)		*xc28* (323)	91.7 (288/314)				
28		25361 ← 26083	240 (27854)	L	*xc29* (170)	97.6 (165/169)	*cp33* (294)	27.9 (79/283)	*px28* (210)	33.6 (81/241)
29		26100 ← 26678	192 (22150)	L	*xc30* (128)	95.7 (111/116)	*cp34* (229)	45.8 (87/190)		
					xc31 (85)	91.2 (73/80)				
30	*pif-3*	26708 → 27295	195 (22240)		*xc32* (195)	99.0 (193/195)	*cp35* (199)	44.4 (84/189)	*px29* (181)	49.7 (85/171)
31		27302 → 27595	97 (10816)	L	*xc33* (97)	92.9 (91/98)	*cp38* (113)	43.9 (18/41)		

Table 1 continued

ORF/other feature	Name	Position[a]	aa (Da)[b]	Promoter motifs[c]	Comparison with other viruses					
					XecnGV		CpGV		PlxyGV	
					ORF (size)/hr	% ID (range)	ORF (size)	% ID (range)[d]	ORF (size)	% ID (range)[d]
32		27621 → 27962	113 (13449)	L	*xc34* (113)	100 (113/113)	*cp39* (105)	49.4 (44/89)	*px31* (129)	44.1 (41/93)
33	*lef-2*	27984 → 28553	189 (22186)		*xc35* (189)	98.9 (186/188)	*cp41* (171)	41.2 (68/165)	*px32* (270)	44.5 (73/164)
34		28556 → 28825	89 (10384)		*xc36* (89)	100 (89/89)	*cp42* (82)	38.2 (26/68)		
35		28829 ← 29344	171 (20217)		*xc37* (175)	91.4 (160/175)				
36		29386 ← 29721	111 (12872)		*xc38* (111)	93.7 (104/111)				
37		29848 ← 30297	149 (17631)		*xc39* (149)	98.0 (146/149)	*cp45* (154)	19.6 (28/143)	*px34* (133)	NSS
38	*mmp*	30398 ← 32188	596 (68500)		*xc40* (469)	95.3 (447/469)	*cp46* (545)	28.8 (156/541)	*px35* (382)	36.0 (149/414)
39		32272 ← 33417	381 (45478)		*xc42* (379)	93.7 (357/381)				
40	*p13*	33475 → 34308	277 (32454)	L	*xc43* (277)	99.3 (275/277)	*cp47* (269)	49.4 (131/265)	*px36* (263)	51.1 (135/264)
41		34561 → 35046	161 (18810)	E	*xc44* (148)	84.8 (128/151)				
42	*pif-2*	35064 → 36230	388 (44968)	L	*xc45* (388)	98.7 (383/388)	*cp48* (372)	54.0 (203/376)	*px37* (368)	52.3 (195/373)
	hr1 (5 repeats)	36236–36849			*hr1*(5 repeats)					
43		36862 ← 37110	82 (9791)	L	*xc46* (81)	84.1 (69/82)	*cp49* (129)	31.2 (24/77)		
44		37130 → 40969	1279 (144763)	L	*xc47* (220)	86.3 (177/205)	*cp50* (727)	24.6 (97/394)	*px38* (153)	37.3 (28/75)
					xc48 (839)	61.1 (563/921)				
45	*ac106*	40956 ← 41729	257 (30333)	L	*xc50* (272)	85.7 (233/272)	*cp52* (342)	53.7 (101/188)	*px40* (206)	76.6 (141/184)
46	*ac110*	41774 → 41935	53 (6040)		*xc51* (53)	98.1 (52/53)	*cp53* (48)	48.6 (17/35)	*px41* (53)	43.1 (22/51)
47	*v-ubi*	41975 ← 42208	77 (8704)	L	*xc52* (77)	98.7 (76/77)	*cp54* (94)	84.2 (64/76)	*px42* (114)	88.2 (67/76)
48	*ac109*	42289 → 43350	353 (40139)	L	*xc53* (353)	99.7 (352/353)	*cp55* (326)	38.3 (125/326)	*px43* (414)	40.2 (143/356)
49		43373 → 43705	110 (12529)	L	*xc54* (110)	98.2 (108/110)	*cp56* (69)	38.1 (16/42)	*px44* (130)	IS
50	*pp31*	43764 ← 44645	293 (33410)	L	*xc55* (295)	98.0 (289/295)	*cp57* (241)	23.9 (55/230)	*px45* (252)	45.6 (113/248)
51	*lef-11*	44608 ← 44916	102 (11716)		*xc56* (102)	97.1 (99/102)	*cp58* (134)	50.0 (52/104)	*px46* (96)	57.3 (51/89)
	hr2 (4 repeats)	45070–45563			*hr2* (5 repeats)					
52		45644 → 46510	288 (33746)	E, L	*xc57* (278)	91.6 (252/275)				
53		46596 ← 48314	572 (66565)							
54	*bro-a*	48536 ← 50086	516 (59879)		*xc60* (484)	68.1 (325/477)				
55	*bro-b*	50131 ← 50310	59 (6830)		*xc60* (484)	80.0 (36/45)				
56		50427 ← 50954	175 (20791)	L						
57		51166 ← 51624	152 (17757)		*xc61* (455)	94.0 (141/150)				
58	*bro-c*	51741 ← 52397	218 (25405)	L	*xc62* (211)	54.7 (122/223)				
59		52581 ← 54635	684 (79704)		*xc64* (689)	83.9 (578/689)				
60		54682 ← 56124	480 (58314)	L	*xc65* (480)	93.8 (450/480)				

Table 1 continued

ORF/other feature	Name	Position[a]	aa (Da)[b]	Promoter motifs[c]	Comparison with other viruses					
					XecnGV		CpGV		PlxyGV	
					ORF (size)/hr	% ID (range)	ORF (size)	% ID (range)[d]	ORF (size)	% ID (range)[d]
61		56235 ← 56891	218 (24799)		*xc66* (218)	92.2 (201/218)	*cp64* (230)	NSS		
62	*he65*	57174 ← 58817	547 (63998)		*xc67* (568)	95.2 (519/545)				
63	*sod*	58940 ← 59404	154 (16243)	L	*xc68* (153)	96.7 (148/153)	*cp59* (132)	60.2 (77/128)	*px47* (153)	57.1 (84/147)
64		59472 ← 59642	56 (6509)	L	*xc70* (56)	98.2 (55/56)				
	hr3 (4 repeats)	59721–60295			*hr3* (6 repeats)					
65		60395 → 61009	204 (23420)	L	*xc71* (356)	92.2 (190/206)				
66		61078 ← 62106	342 (38150)	L	*xc72* (342)	96.8 (331/342)				
67		62474 → 63574	366 (41777)		*xc73* (458)	96.4 (353/366)				
68		63634 ← 64749	371 (43853)	L	*xc74* (368)	84.1 (311/370)				
69		64805 → 65098	97 (11615)	L	*xc75* (97)	97.9 (95/97)	*cp65* (79)	35.6 (32/90)		
	hr4 (7 repeats)	65114–65953			*hr4* (6 repeats)					
70		67077 ← 67256	59 (6591)							
71	*bro-d*	67316 → 68137	273 (31633)		*xc76* (273)	96.0 (262/273)				
72	*p74*	68177 → 70309	710 (80722)	E, L	*xc77* (710)	99.2 (704/710)	*cp60* (688)	43.1 (304/706)	*px49* (578)	44.3 (266/600)
73		70306 ← 70620	104 (12031)	L	*xc77a* (103)	96.0 (99/103)	*cp62* (189)	51.6 (32/62)	*px50* (135)	43.2 (41/95)
74	*p47*	70696 → 71880	394 (46139)		*xc78* (394)	99.5 (392/394)	*cp68* (460)	56.1 (222/396)	*px51* (386)	57.5 (223/388)
75		71909 ← 72424	171 (20451)							
76		72480 ← 72692	70 (8410)							
77	*ac38*	72982 → 73659	225 (27144)	L	*xc79* (225)	99.6 (224/225)	*cp69* (220)	63.6 (140/220)	*px52* (207)	73.3 (154/210)
78		73652 → 74200	182 (20346)		*xc80* (182)	99.5 (181/182)	*cp71* (203)	50.6 (78/154)	*px53* (159)	53.6 (81/151)
79	*p38.7*	74227 ← 74781	184 (21164)	L	*xc81* (184)	94.0 (173/184)	*cp73* (198)	28.3 (32/113)	*px54* (150)	36.5 (38/104)
80	*lef-1*	74782 ← 75498	238 (27913)		*xc82* (238)	98.7 (235/238)	*cp74* (235)	50.6 (120/237)	*px55* (251)	50.2 (119/237)
81		75567 → 76115	182 (19287)	L	*xc83* (182)	95.6 (173/181)				
82	*pif-1*	76132 → 77754	540 (61347)	L	*xc84* (540)	98.5 (532/540)	*cp75* (538)	45.2 (245/542)	*px7* (536)	45.7 (247/540)
83	*fgf-1*	77793 ← 78491	232 (26593)		*xc85* (232)	94.8 (220/232)	*cp76* (195)	27.7 (51/184)	*px56* (221)	31.9 (68/213)
84		78543 ← 78890	115 (13608)		*xc86* (115)	95.7 (110/115)	*cp77* (103)	NSS	*px57* (100)	NSS
85		79022 → 79516	164 (18684)	L	*xc87* (164)	91.5 (150/164)	*cp79* (156)	29.3 (46/157)	*px59* (79)	27.6 (21/76)
86	*lef-6*	79513 ← 79812	99 (11741)		*xc88* (99)	94.9 (94/99)	*cp80* (101)	45.1 (32/71)	*px60* (86)	40.3 (25/62)
87	*dbp*	79874 ← 80707	277 (31635)		*xc89* (277)	97.8 (271/277)	*cp81* (290)	26.6 (51/192)	*px61* (263)	31.0 (79/255)
88		80816 ← 81028	70 (8762)		*xc89a* (70)	98.0 (69/70)	*cp82* (306)	31.2 (15/48)		
89		80989 ← 81729	246 (28949)		*xc90* (246)	95.1 (234/246)				
90	*p45*	81728 → 82846	372 (43903)	L	*xc91* (372)	98.9 (368/372)	*cp83* (439)	54.4 (209/384)	*px63* (377)	54.0 (198/367)

Table 1 continued

ORF/other feature	Name	Position[a]	aa (Da)[b]	Promoter motifs[c]	Comparison with other viruses					
					XecnGV		CpGV		PlxyGV	
					ORF (size)/hr	% ID (range)	ORF (size)	% ID (range)[d]	ORF (size)	% ID (range)[d]
91	*p12*	82855 ← 83220	121 (13093)	L	*xc92* (120)	98.3 (119/121)	*cp84* (109)	40.8 (49/120)	*px64* (96)	45.6 (36/79)
92	*p40*	83272 → 84387	371 (42390)	L	*xc93* (372)	97.6 (363/372)	*cp85* (380)	50.1 (184/367)	*px66* (366)	47.4 (170/359)
93	*p6.9*	84447 → 84623	58 (7209)	L	*xc94* (60)	96.7 (58/60)	*cp86* (49)	59.2 (29/49)	*px67* (56)	58.6 (34/58)
94	*lef-5*	84664 ← 85470	268 (30720)		*xc95* (245)	98.7 (227/230)	*cp87* (242)	55.9 (132/236)	*px69* (247)	51.4 (133/259)
95	*38K*	85399 → 86304	301 (36043)	L	*xc96* (301)	97.7 (294/301)	*cp88* (343)	47.5 (141/297)	*px70* (340)	49.5 (143/289)
96	*ac96*	86321 ← 86794	157 (18023)	L	*xc97* (157)	98.7 (155/157)	*cp89* (161)	46.8 (73/156)	*px71* (161)	35.3 (55/156)
97	*helicase*	86793 → 90275	1160 (136126)		*xc98* (1159)	98.5 (1143/1160)	*cp90* (1131)	37.4 (436/1167)	*px72* (1124)	49.2 (572/1163)
98	*odv-e25*	90350 ← 91012	220 (24213)	L	*xc99* (220)	98.6 (217/220)	*cp91* (213)	65.9 (145/220)	*px74* (214)	61.6 (135/219)
99	*ac93*	91062 ← 91538	158 (18301)	L	*xc100* (122)	95.7 (112/117)	*cp92* (161)	41.5 (66/159)	*px75* (156)	45.2 (70/155)
100	*p33*	91636 → 92394	252 (30167)	L	*xc101* (251)	95.2 (239/251)	*cp93* (251)	53.8 (135/251)	*px76* (250)	56.1 (138/246)
101	*bro-e*	92422 ← 93201	259 (28651)		*xc109* (308)	68.9 (175/254)				
102	*bro-f*	93206 ← 93712	168 (19463)		*xc109* (308)	69.7 (23/33)				
103	*chaB*	93827 ← 94090	87 (10160)	L	*xc102* (87)	86.9 (73/84)				
104		94128 ← 94352	74 (8756)	L						
105	*chitinase*	94507 → 96264	585 (64934)	L	*xc103* (594)	98.6 (554/562)	*cp10* (594)	62.3 (344/552)		
106		96261 ← 96563	100 (11106)	L	*xc105* (91)	45.3 (43/95)	*cp9* (101)	30.9 (25/81)	*px12* (98)	28.9 (24/83)
107		96578 ← 96862	94 (10457)	L	*xc105* (91)	91.5 (86/94)	*cp9* (101)	33.3 (26/78)	*px12* (98)	27.5 (22/80)
108		96989 → 97324	111 (12235)	L	*xc106* (90)	92.9 (65/70)				
109	*gp37*	97393 → 98127	244 (27733)	L	*xc107* (244)	93.0 (227/244)	*cp13* (251)	43.9 (107/244)		
110		98227 → 98694	155 (17356)		*xc108* (152)	87.7 (136/155)				
111		98624 → 98818	64 (7410)							
112	*lef-4*	98843 ← 100195	450 (52209)		*xc110* (447)	95.1 (424/446)	*cp95* (480)	42.6 (205/481)	*px78* (432)	43.8 (194/443)
113	*vp39*	100244 → 101233	329 (37186)	L	*xc111* (329)	97.6 (321/329)	*cp96* (285)	36.7 (116/316)	*px79* (320)	42.1 (141/335)
114	*odv-ec27*	101309 → 102175	288 (33917)		*xc112* (288)	98.6 (284/288)	*cp97* (288)	44.3 (125/282)	*px80* (287)	50.5 (143/283)
115		102346 → 102537	63 (7516)	L						
116		102490 ← 103608	372 (44472)	L	*xc113* (373)	98.1 (365/372)	*cp99* (394)	26.8 (99/370)	*px82* (340)	28.9 (102/353)
	hr5 (2 repeats)	103708–103973			*hr3* (3 repeats)					
117	*bro-g*	104088 ← 105236	382 (44487)		*xc114* (427)	98.7 (377/382)				
118		105369 → 106529	386 (44996)		*xc115* (397)	97.2 (375/386)				
119		106629 → 107000	123 (14524)	L	*xc116* (123)	98.4 (121/123)	*cp100* (108)	35.8 (38/106)	*px83* (69)	44.3 (27/61)
120		107041 ← 107586	181 (21394)	E, L	*xc117* (181)	95.0 (171/180)				

Table 1 continued

ORF/other feature	Name	Position[a]	aa (Da)[b]	Promoter motifs[c]	Comparison with other viruses					
					XecnGV		CpGV		PlxyGV	
					ORF (size)/hr	% ID (range)	ORF (size)	% ID (range)[d]	ORF (size)	% ID (range)[d]
121	*vp91 capsid*	107674 ← 109884	736 (82402)	L	*xc118* (741)	97.4 (722/741)	*cp101* (665)	33.3 (247/742)	*px84* (533)	38.5 (122/317)
	hr5a (1 repeat)	108748–108874			*hr5a* (1 repeat)					
122	*tlp20*	109850 → 110329	159 (17921)		*xc119* (161)	96.2 (153/159)	*cp102* (216)	38.0 (16/42)	*px85* (139)	29.0 (45/155)
123	*ac81*	110350 → 110913	187 (21504)	L	*xc120* (187)	99.5 (186/187)	*cp103* (191)	57.9 (103/178)	*px86* (191)	61.2 (112/183)
124	*gp41*	110972 → 111844	290 (33079)	L	*xc121* (290)	99.7 (289/290)	*cp104* (289)	50.7 (147/290)	*px87* (283)	44.8 (130/290)
125	*ac78*	111908 → 112219	103 (11606)	L	*xc122* (103)	96.1 (99/103)	*cp105* (111)	25.4 (17/67)	*px88* (89)	33.0 (34/103)
126	*vlf-1*	112200 → 113324	374 (43388)	L	*xc123* (373)	98.7 (368/373)	*cp106* (378)	54.2 (202/373)	*px89* (346)	56.0 (191/341)
127		113325 ← 113867	180 (21821)		*xc124* (180)	96.1 (173/180)			*px90* (176)	37.1 (59/159)
128	*ac76*	113907 → 114164	85 (9708)		*xc125* (85)	98.8 (84/85)	*cp107* (84)	57.6 (49/85)	*px91* (81)	61.7 (50/81)
129	*ac75*	114224 → 114661	145 (17066)	L	*xc126* (145)	99.3 (144/145)	*cp108* (148)	36.7 (54/147)	*px92* (145)	34.7 (51/147)
130	*ctl*	114642 → 114830	62 (7030)		*xc127* (52)	93.9 (46/49)				
131		114849 ← 115280	143 (16594)		*xc128* (143)	95.8 (137/143)				
132	*lef-7*	115437 → 116375	312 (36130)	E	*xc129* (312)	91.7 (286/312)				
133	*bro-h*	116623 → 118014	463 (53569)		*xc131* (442)	92.8 (410/442)				
134	*dnapol*	118070 ← 121366	1098 (127087)		*xc132* (1098)	98.9 (1086/1098)	*cp111* (1051)	52.2 (547/1048)	*px93* (979)	52.9 (540/1001)
135	*desmoplakin*	121365 → 123350	661 (76336)		*xc133* (661)	96.7 (639/661)	*cp112* (718)	20.9 (142/678)	*px94* (651)	25.6 (173/675)
136	*lef-3*	123420 ← 124475	351 (40980)	L	*xc134* (163)	96.9 (340/351)	*cp113* (353)	27.2 (98/360)	*px95* (297)	27.2 (82/302)
137	*ac68*	124444 → 124854	136 (16254)		*xc135* (120)	99.1 (114/115)	*cp114* (126)	41.9 (54/129)	*px96* (128)	41.3 (50/121)
138		124917 → 125432	171 (20036)		*xc136* (171)	97.l (166/171)			*px97* (194)	23.2 (44/190)
139	*iap-5*	125496 → 126350	284 (33009)		*xc137* (285)	97.9 (279/285)	*cp116* (275)	32.1 (92/287)	*px98* (281)	38.2 (99/259)
140	*lef-9*	126527 → 128008	493 (56997)		*xc139* (493)	99.6 (491/493)	*cp117* (499)	63.0 (310/492)	*px99* (494)	64.0 (317/495)
141	*fp*	128053 → 128496	147 (16909)	L	*xc140* (147)	97.3 (143/147)	*cp118* (161)	50.3 (75/149)	*px100* (138)	52.6 (71/135)
142	*ligase*	128497 ← 130080	527 (60635)		*xc141* (527)	98.9 (521/527)	*cp120* (570)	42.8 (244/570)	*px101* (523)	42.7 (229/536)
143		130294 → 130560	88 (10305)		*xc142* (88)	98.9 (87/88)	*cp121* (71)	NSS	*px102* (61)	28.8 (15/52)
144		130622 → 130822	66 (7457)		*xc143* (66)	100 (66/66)	*cp122* (66)	34.7 (17/49)	*px103* (66)	30.4 (21/69)
145	*fgf-2*	130857 ← 132086	409 (46745)		*xc144* (409)	96.1 (393/409)	*cp123* (400)	30.6 (90/294)	*px104* (396)	31.0 (97/313)
146	*alk-exo*	132098 → 133471	457 (53070)		*xc145* (457)	96.7 (442/457)	*cp125* (398)	49.0 (214/437)	*px106* (378)	45.3 (174/384)
147	*helicase-2*	133541 → 134908	455 (52330)	E	*xc146* (455)	96.0 (437/455)	*cp126* (457)	50.9 (136/267)	*px107* (436)	52.5 (232/442)
148		135003 → 135998	331 (39636)		*xc147* (309)	83.6 (230/275)				
149	*lef-8*	136034 ← 138613	859 (99648)		*xc148* (859)	98.5 (846/859)	*cp131* (873)	61.1 (540/884)	*px109* (838)	64.4 (549/853)
150	*odv-e66*	138673 ← 140679	668 (75178)	L	*xc149* (668)	98.8 (660/668)			*px30* (682)	44.4 (275/619)

Table 1 continued

ORF/other feature	Name	Position[a]	aa (Da)[b]	Promoter motifs[c]	Comparison with other viruses					
					XecnGV		CpGV		PlxyGV	
					ORF (size)/hr	% ID (range)	ORF (size)	% ID (range)[d]	ORF (size)	% ID (range)[d]
	hr6 (4 repeats)	140724–141219			*hr6* (4 repeats)					
151	*enhancin-1*	141226 ← 143697	823 (93996)	L	*xc150* (824)	96.8 (798/824)				
152	*enhancin-2*	143852 ← 146449	865 (98848)	L	*xc152* (867)	96.4 (836/867)				
153	*enhancin-3*	146473 → 149181	902 (104792)	L	*xc154* (898)	86.6 (781/902)				
154		149239 → 150408	389 (44497)	L	*xc155* (391)	96.1 (365/380)				
155		150451 → 150810	119 (13906)		*xc157* (59)	96.6 (57/59)				
156		150866 ← 151321	151 (17335)		*xc158* (150)	88.0 (132/150)				
157		151555 → 153285	576 (66655)							
158	*bro-i*	153414 ← 154211	265 (29404)		*xc159* (408)	93.6 (248/265)				
159	*bro-j*	154226 ← 154426	66 (7246)		*xc159* (408)	67.5 (27/40)				
160		154776 → 156236	486 (56201)		*xc161* (486)	96.3 (468/486)				
161		156265 ← 156825	186 (21999)		*xc162* (186)	97.8 (182/186)				
	hr7 (3 repeats)	156913–157292			*hr7* (5 repeats)					
162		157480 ← 157701	73 (8746)		*xc164* (53)	51.9 (28/54)				
163		157627 → 157983	118 (12916)	L	*xc165* (118)	97.5 (115/118)	*cp132* (131)	NSS	*px110* (114)	NSS
164	*enhancin-4*	158015 ← 160585	856 (97834)	L	*xc166* (856)	95.9 (821/856)				
165		160814 → 161146	110 (13356)		*xc167* (105)	97.1 (101/104)				
166		161278 → 161865	195 (22374)	E	*xc168* (198)	84.8 (168/198)				
167		161908 ← 162342	144 (17056)	L	*xc169* (144)	93.8 (135/144)	*cp119* (162)	45.9 (67/146)	*px65* (152)	21.9 (23/105)
	hr8 (4 repeats)	162476–162976			*hr8* (5 repeats)					
168		162983 ← 163174	63 (7254)		*xc170* (63)	96.8 (61/63)	*cp133* (62)	55.6 (20/36)		
169	*ac53*	163161 → 163580	139 (16062)		*xc171* (139)	98.6 (137/139)	*cp134* (133)	38.7 (53/137)	*px112* (139)	46.7 (50/107)
170		163584 ← 164720	378 (43667)	L	*xc172* (378)	96.6 (365/378)	*cp135* (354)	26.8 (60/224)	*px113* (198)	NSS
171		164745 ← 164948	67 (7860)	L	*xc173* (67)	100 (67/67)	*cp136* (67)	NSS	*px114* (109)	NSS
172	*lef-10*	164926 → 165138	70 (7609)	L	*xc174* (70)	98.6 (69/70)	*cp137* (89)	34.3 (24/70)		
173	*vp1054*	165017 → 165988	323 (37407)	L	*xc175* (323)	98.1 (317/323)	*cp138* (332)	47.6 (151/317)	*px115* (311)	44.7 (139/311)
174		166044 → 166256	70 (8139)		*xc176* (70)	100 (70/70)			*px116* (59)	43.3 (26/60)
175		166331 → 166690	119 (13527)		*xc177* (119)	96.6 (115/119)				
176	*fgf-3*	166734 → 167789	351 (38509)	L	*xc178* (332)	88.2 (313/355)	*cp140* (347)	26.2 (84/321)	*px117* (249)	26.5 (63/238)
177		167908 → 168501	197 (23513)		*xc179* (197)	94.9 (187/197)				
178	*me53*	168467 → 169444	325 (38177)		*xc180* (325)	97.8 (318/325)	*cp143* (303)	34.0 (105/309)	*px120* (308)	40.6 (127/313)

Table 1 continued

ORF/other feature	Name	Position[a]	aa (Da)[b]	Promoter motifs[c]	Comparison with other viruses					
					XecnGV		CpGV		PlxyGV	
					ORF (size)/hr	% ID (range)	ORF (size)	% ID (range)[d]	ORF (size)	% ID (range)[d]
179		169447 → 169776	109 (12360)		*xc181* (110)	96.3 (105/109)				

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

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

[c] Promoter motifs present upstream of ORF. E: 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 25–35 bp. L: Late promoter motif (A/T/G)TAAG 120 bp upstream of start codon

[d] NSS: No significant sequence similarity detected by BLASTp; the presence of the homologue was inferred from comparisons among other GV genomes

These viruses, along with XecnGV, are considered to be isolates of the same virus species [34].

Jehle et al. [4] proposed a criterion for distinguishing among baculovirus species using nucleotide distances (in base substitutions per site) calculated with the Kimura-2-parameter (K-2-P) method for the three marker genes used in Fig. 2 for phylogenetic analysis (*granulin*, *lef-8*, and *lef-9*). Under this criterion, viruses with distances of <0.015 for single or concatenated sequences are considered to belong to the same species, while viruses with distances >0.05 are considered to be different species. For distances from 0.015 to 0.05, additional information is required to make a decision about species boundaries. K-2-P distances were calculated for the clade of viruses containing HearGV, TnGV, and the XecnGV isolates. The pairwise distances between HearGV and the XecnGV isolates for *lef-8* and *lef-9* ranged from 0.015–0.021 to 0.015–0.022, respectively (Table 2). For the more strongly conserved granulin gene, pairwise distances ranged from 0.010 to 0.012. The pairwise distances for the concatenated sequences ranged from 0.010 to 0.017. For all four sets of data, the nucleotide distances between HearGV and the XecnGV isolates were larger than the pairwise distances among the XecnGV isolates themselves (Table 2). However, the results of this analysis do not lend themselves to a straightforward conclusion about the taxonomic position of HearGV with respect to XecnGV and other XecnGV variants.

Gene content and order

Gene-parity plot analysis [35] revealed a strong degree of co-linearity between HearGV and other GVs (data not shown). PlxyGV and CpGV are missing many of the ORFs found in HearGV, which is expected given the significantly smaller sizes of the genomes of these GVs. Comparison of HearGV with AcMNPV-C6 revealed that the order of some ORFs in these NPVs was conserved between the two viruses, but the orientation of a large proportion of these ORFs was inverted relative to the polyhedrin gene. These results are similar to those of previous analyses of GV and NPV gene order [6–10].

All the 62 genes common among lepidopteran baculovirus genome sequences as of 2006 [2] were found in the HearGV genome. Seven HearGV ORFs have no discernible homologues in baculovirus genomes sequenced to date (Table 3). Four of these ORFs (*hear70*, *hear104*, *hear111*, and *hear115*) are small (<75 codons) with either no BLAST hits or hits with relatively high *E*-values (0.47 or higher). The other three ORFs (*hear56*, *hear75*, and *hear76*) have homologues in other families of insect DNA viruses.

The predicted amino acid sequence for *hear56* produced BLAST hits with orf76 of *Spodoptera frugiperda* ascovirus 1a (SfAV-1a; [36]) and orf95 of *H. virescens* ascovirus 3e

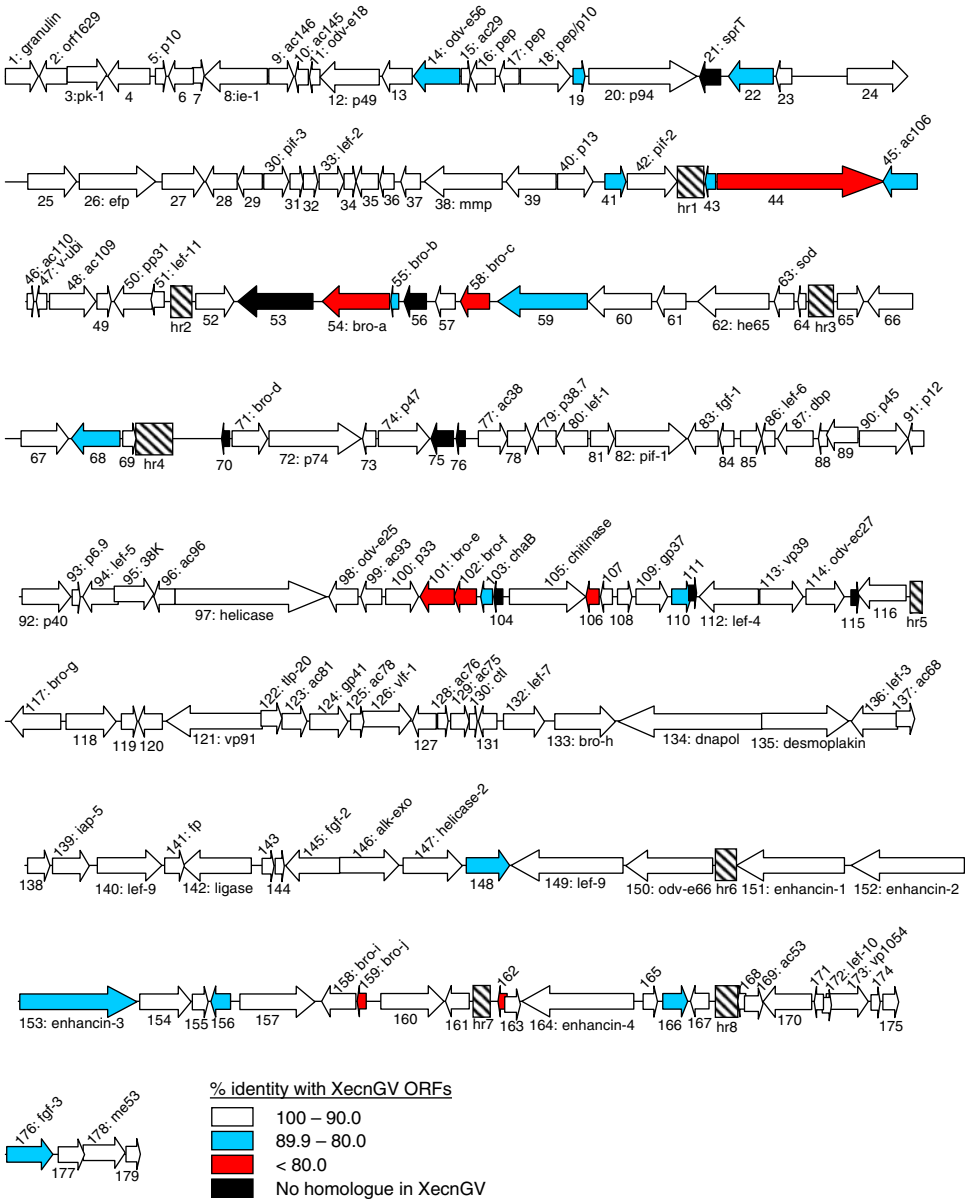

Fig. 1 Map of the ORFs and other features of the HearGV genome. ORFs are represented by arrows, with the position and direction of the arrow indicating ORF position and orientation. The number of each ORF is displayed, with the name of the ORF following a colon. Homologous repeat regions (*hr*s) are represented by hatched boxes. The shading of the ORFs indicates the degree of predicted amino acid sequence similarity to XecnGV homologues

(HvAV-3e; [37]). The role of these ORFs in ascovirus biology is currently unknown. ORFs *hear75* and *hear76* showed sequence similarity to members of the polydnavirus Rep gene family in ichnoviruses of *Hyposoter fugitivus*, *Hyposoter didymator*, *Tranosema rostrales*, and *Campoletis sonorensis*. Polydnavirus Rep genes contain one or more copies of an imperfectly conserved 540-bp sequence [38, 39]. Though they are present in large numbers in ichnovirus genomes and expressed in both the wasp and parasitized larval host, the function of their gene products is unknown [40, 41].

Another HearGV ORF of interest is *hear21*, a homologue of the *sprT* family of putative metallopeptidases thought to be involved in transcriptional elongation [42]. SprT homologues are also found in *Trichoplusia ni* ascovirus 2c (TnAV-2c; [43]) and in *Mamestra configurata* multiple nucleopolyhedroviruses A and B (MacoNPV-A and MacoNPV-B; [44, 45]).

There are 20 ORFs that HearGV and XecnGV have in common which are not present in other baculovirus genomes sequenced to date (Table 4). One of these ORFs (*hear108*) exhibits significant sequence similarity to ascovirus ORFs. The other ORFs either exhibit no significant similarity to other amino acid sequences or have BLAST hits with modest *E*-values (0.0008–0.0096) to uncharacterized sequences. Canonical baculovirus early and late gene promoter motifs are associated with ten of these ORFs. Several of these ORFs are located next to or near *hr3* and *hr7*.

Other HearGV ORFs exhibit interesting patterns of similarity with genes from other viruses. After *xc21*, the next two

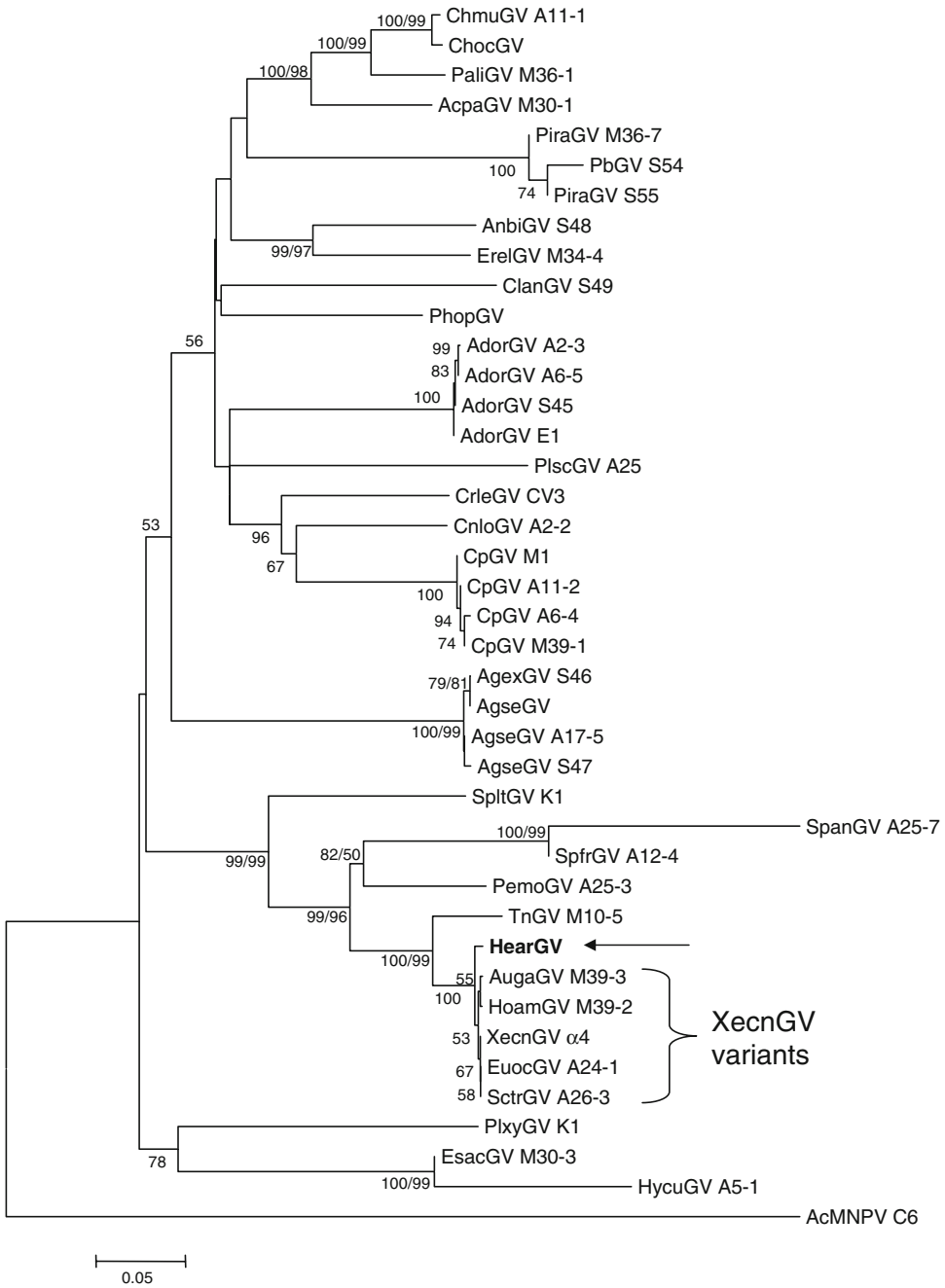

Fig. 2 Phylogenetic analysis of concatenated amino acid sequence alignments, showing bootstrap values >50% for ME and MP trees at each node (ME/MP). The location of HearGV (bold) is indicated by an arrow. A consensus ME phylogram of concatenated partial *polh*, *lef-8*, and *lef-9* sequence alignments for the following completely sequenced granuloviruses and other granuloviruses described in [34] and [4]: *Choristoneura murinana* GV (ChmuGV A11-1); *Choristoneura occidentalis* GV [10]; *Pandemis limitata* GV (PaliGV M36-1); *Amelia pallorana* GV (AcpaGV M30-1); *Pieris rapae* GV (PiraGV isolates M36-7 and S55); *Pieris brassicae* (PbGV S54); *Andraca bipunctata* GV (AnbiGV S48); *Erinnyis ello* GV (ErelGV M34-4); *Clostera anachoreta* GV (ClanGV S49); *Phthorimaea operculella* GV (accession number NC_004062); *Adoxophyes orana* GV (AdorGV isolates E1 [8], A2-3, A6-5, and S45); *Plathypena scabra* GV (PlscGV A25); *Cryptophlebia leucotreta* GV (CrleGV CV3 [9]); *Cnephasia longana* GV (CnloGV A2-2); *Cydia pomonella* GV (CpGV isolates M1 [7], A11-2, A6-4, and M39-1); *Agrotis exclamationis* GV (AgexGV S46); *Agrotis segetum* GV (AgseGV complete sequence accession number NC_005839 and isolates A17-5 and S47); *Spodoptera litura* GV (SpltGV K1, accession number NC_009503), *Spodoptera androgea* GV (SpanGV A25-7); *Spodoptera frugiperda* GV (SpfrGV A12-4); *Peridroma morpontora* GV (PemoGV A25-3); *Trichoplusia ni* GV (TnGV M10-5); *Autographa gamma* GV (AugaGV M39-3); *Hoplodrina ambigua* GV (HoamGV M39-2); *Xestia c-nigrum* GV (XecnGV α4 [5]); *Euxoa ochrogaster* GV (EuocGV A24-1); *Scotogramma trifolii* GV (SctrGV A26-3); *Plutella xylostella* GV (PlxyGV K1 [6]); *Estigmene acrea* GV (EsacGV M30-3); *Hyphantria cunea* GV (HycuGV A5-1); *Autographa californica* MNPV (AcMNPV C6 [33])

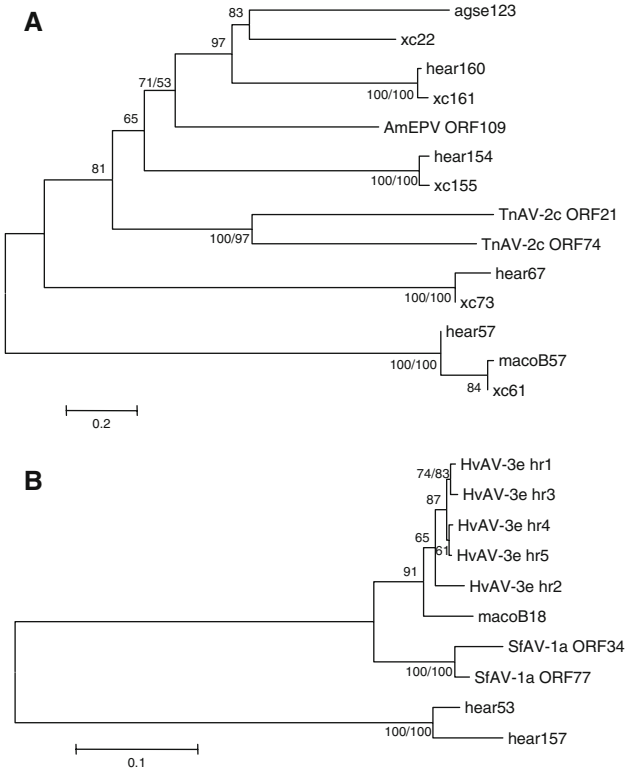

Fig. 3 Phylogenetic analysis of HearGV gene family amino acid sequences. Consensus ME phylograms inferred from the alignment of (**a**) *hear57*/*hear67*/*hear154*/*hear160* and homologous ORFs, and (**b**) *hear53*/*hear157* and homologous ORFs, with bootstrap values >50% shown at interior branches for ME and MP analysis (ME/MP) where they occur. The sequences used derive from XecnGV (*xc22*, *xc61*, *xc73*, *xc155*, and *xc161* [5]), *Amsacta moorei* entomopoxvirus (AmEPV ORF109 [52]), *Trichoplusia ni* ascovirus 2c (TnAV-2c ORF21 and ORF74; [43]), *Mamestra configurata* NPV-B (*macoB57* and *macoB18*; [45]), *Heliothis virescens* ascovirus 3e (HvAV-3e hr1 through hr5; [37]), and *Spodoptera frugiperda* ascovirus 1a (SfAV-1a ORF34 and ORF77; [36])

top BLAST hits for ORF *hear20*, a homologue of the NPV gene *p94*, are ORFs 3004 (accession number AAV98008, 49% identity) and 3003 (accession number AAV98006, 44.7% identity) of *Cotesia plutellae* bracovirus. ORFs *hear27*, *hear36*, and *hear89* have homologues only in XecnGV and *Spodoptera litura* granulovirus (SpltGV, accession number NC_009503), while *hear148* homologues are found in SpltGV and a selection of NPVs. Among granulovirus genomes, ORFs *hear39*, *hear52*, *hear81*, *hear131*, and *hear132* (*lef-7*) only have homologues in XecnGV, but homologues of these ORFs also occur in a subset of NPVs.

Multigene families

Multigene families, or groups of related genes, occur in a number of large DNA viruses. One of the most widespread multigene families is the baculovirus repeated ORF (*bro*) family, found in many invertebrate DNA viruses [46].

Table 2 Estimates of pairwise nucleotide distances (base substitutions/site) of the nucleotide sequences of (A) *lef-9* and *lef-8* fragments and of (B) *granulin* and concatenated *granulin*/*lef-8*/*lef-9* fragments of HearGV, TnGV, and XecnGV variants

(A)

lef-8	lef-9						
	1	2	3	4	5	6	7
1. AugaGV M39-3	–	0.007	**0.019**	0.004	0.011	0.011	0.184
2. EuocGV A24-1	0.015	–	**0.015**	0.004	0.000	0.000	0.176
3. HearGV	**0.021**	**0.018**	–	**0.022**	**0.015**	**0.022**	**0.189**
4. HoamGV M39-2	0.003	0.012	**0.018**	–	0.007	0.007	0.189
5. SctrGV A26-3	0.015	0.000	**0.018**	0.012	–	0.007	0.184
6. XecnGV α4	0.015	0.000	**0.018**	0.012	0.000	–	0.184
7. TnGV M10-5	0.194	0.188	**0.184**	0.190	0.188	0.188	–

(B)

granulin/lef-8/lef-9	granulin						
	1	2	3	4	5	6	7
1. AugaGV M39-3	–	0.002	**0.012**	0.000	0.004	0.002	0.118
2. EuocGV A24-1	0.009	–	**0.010**	0.002	0.002	0.000	0.116
3. HearGV	**0.017**	**0.015**	**–**	**0.012**	**0.012**	**0.010**	**0.118**
4. HoamGV M39-2	0.002	0.007	**0.016**	–	0.004	0.002	0.118
5. SctrGV A26-3	0.010	0.001	**0.015**	0.008	–	0.002	0.113
6. XecnGV α4	0.009	0.000	**0.015**	0.007	0.001	–	0.116
7. TnGV M10-5	0.163	0.159	**0.159**	0.162	0.158	0.159	–

Pairwise distances between HearGV and another GV are in bold

Gene expression, nucleic acid binding activity, nucleosome association, and protein localization and trafficking have been characterized for some NPV *bro* genes and proteins [46–50], but the functions of *bro* gene products in the baculovirus life cycle are not known with precision. The number of *bro* genes in a genome varies from virus to virus. The XecnGV has eight *bro*-homologous sequences, including one ORF (*xc62*) not previously identified as a *bro* family member. HearGV has 10 *bro*-homologous sequences, including homologues for XecnGV *bro-a* (*xc60*), *xc62*, *bro-b* (*xc76*), *bro-c* (*xc109*), *bro-d* (*xc114*), *bro-f* (*xc131*), and *bro-g* (*xc159*). In HearGV, there are three adjacent pairs of *bro* ORFs (*hear54* and *hear55*; *hear101* and *hear102*; and *hear158* and *hear159*). ORF *hear55* occurs in an approximately 1.3-kbp insertion of novel sequence (not found in the XecnGV genome) that also includes *hear56*. Although the *hear55* coding sequence aligns with aa 179–223 of XecnGV *bro xc60*, the top BLAST hit for *hear55* is MacoNPV-A *bro-c*. The ORFs *hear101* and *hear102* are contained within an approximately 1.5-kbp insertion of novel sequence, and both align with different portions of *xc109*. An approximately 1.3-kbp sequence that would

Table 3 ORFs unique to HearGV

ORF	Features
hear56	Homologues in *Heliothis virescens* AV-3e (ORF 94; $E = 2.98261 \times 10^{-16}$) and *Spodoptera frugipeda* AV-1a (ORF 76; $E = 2.4456 \times 10^{-10}$); sequence missing from XecnGV
hear70	ORF truncated to 3 codons in XecnGV sequence
hear75	Sequence similarity with ichnovirus Rep genes (closest match: *Hyposoter fugitivus* ichnovirus repeat element protein-d3.2, $E = 1.9573 \times 10^{-09}$); sequence missing from XecnGV
hear76	Sequence similarity with ichnovirus Rep genes (closest match: *Campoletis sonorensis* ichnovirus repeat element protein BHv 0.9, $E = 0.0016663$); sequence missing from XecnGV
hear104	ORF truncated to 45 codons in XecnGV sequence
hear111	ORF truncated to 28 codons in XecnGV sequence
hear115	ORF truncated to 36 codons in XecnGV sequence

Table 4 ORFs unique to HearGV and XecnGV

HearGV ORF/*hr*	Position[a]	aa (Da)[b]	Promoter motifs[c]	XecnGV homologue (size)/*hr*	% ID (range)
hear4	2361 ← 3335	324 (38859)	L	*xc4* (324)	96.3 (312/324)
hear 22	16638 ← 17663	341 (39797)		*xc23* (350)	85.1 (303/356)
hear 23	17737 ← 18111	124 (13969)	L	*xc24* (124)	96.8 (120/124)
hear 35	28829 ← 29344	171 (20217)		*xc37* (175)	91.4 (160/175)
hear 41	34561 → 35046	161 (18810)	E	*xc44* (148)	84.8 (128/151)
hear 64	59472 ← 59642	56 (6509)	L	*xc70* (56)	98.2 (55/56)
hr3 (4 repeats)	59721–60295			*hr3* (6 repeats)	
hear 65	60395 → 61009	204 (23420)	L	*xc71* (356)	92.2 (190/206)
hear 66	61078 ← 62106	342 (38150)	L	*xc72* (342)	96.8 (331/342)
hear 68	63634 ← 64749	371 (43853)	L	*xc74* (368)	84.1 (311/370)
hear 108	96989 → 97324	111 (12235)	L	*xc106* (90)	92.9 (65/70)
hear 110	98227 → 98694	155 (17356)		*xc108* (152)	87.7 (136/155)
hear 118	105369 → 106529	386 (44996)		*xc115* (397)	97.2 (375/386)
hear 120	107041 ← 107586	181 (21394)	E, L	*xc117* (181)	95.0 (171/180)
hear 161	156265 ← 156825	186 (21999)		*xc162* (186)	97.8 (182/186)
hr7 (3 repeats)	156913–157292			*hr7* (5 repeats)	
hear 162	157480 ← 157701	73 (8746)		*xc164* (53)	51.9 (28/54)
hear 165	160814 → 161146	110 (13356)		*xc167* (105)	97.1 (101/104)
hear 166	161278 → 161865	195 (22374)	E	*xc168* (198)	84.8 (168/198)
hear 175	166331 → 166690	119 (13527)		*xc177* (119)	96.6 (115/119)
hear 177	167908 → 168501	197 (23513)		*xc179* (197)	94.9 (187/197)
hear 179	169447 → 169776	109 (12360)		*xc181* (110)	96.3 (105/109)

[a] Direction of ORF in the HearGV 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. E: 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 25 - 35 bp. L: Late promoter motif (A/T/G)TAAG 120 bp upstream of start codon

contain the *xc109* homologue is missing from where it should occur in the HearGV genome, suggesting that the position of *hear101* and *hear102* is a consequence of re-arrangement within the HearGV genome. Indels in the region of HearGV corresponding to *xc159* resulted in the division of this ORF into *hear158* and *hear159*. Of the three pairs of *bro* sequences, only *hear54* has an intact Bro-N motif which is involved in DNA binding [50, 51].

HearGV contained homologues to a group of four XecnGV ORFs (*xc61*, *xc73*, *xc155*, and *xc161*) that are part of a five-gene family with significant sequence similarity to each other [5]. Homologues to various members of this family also occur in another granulovirus (AgseGV), an NPV (MacoNPV-B), an ascovirus (TnAV-2c), and an entomopoxvirus (*Amsacta moorei* entomopoxvirus (AmE-PV; [52]). In ME and MP phylograms, the ascovirus and

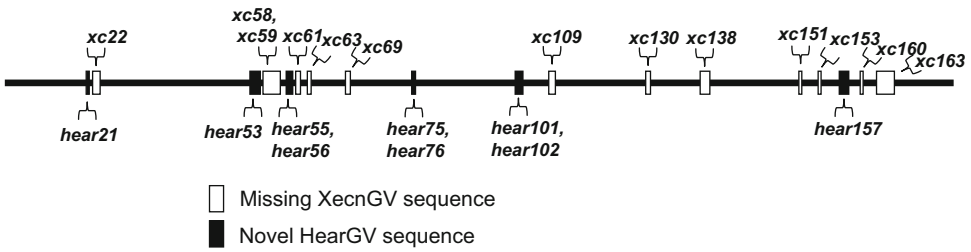

Fig. 4 Map of deletions and insertions in the HearGV genome relative to the XecnGV genome. The relative positions and sizes of insertions and deletions are indicated by boxes on a linear representation of the HearGV genome. Regions of XecnGV sequence that are missing from the homologous location in the HearGV genome are denoted by white boxes, while regions of novel HearGV sequence not found in the homologous location in the XecnGV genome are denoted by black boxes. XecnGV ORFs that are located in the XecnGV sequences missing from HearGV are indicated, as are HearGV ORFs present in the novel HearGV sequences. In the case of *xc63*, a HearGV homologue is present but truncated by 300 codons due to the deletion of XecnGV-homologous sequence in this region

granulovirus members of this family form distinct clades (Fig. 3a). The MacoNPV-B ORF *macoB57* was found to be more closely related to *xc61* than the HearGV homologue. Deletions in the HearGV genome removed sequence containing the HearGV homologue for *xc22* and 300 C-terminal *xc61*-homologous codons from *hear57*. An additional pair of homologous ORFs, *hear155* and *xc157*, also exhibits a low degree of sequence similarity to members of this gene family. No conserved domains that would suggest a function were detected among the genes in this family.

Open reading frames *hear53* and *hear157* have 92% sequence identity with each other by BLAST. These ORFs contain a hypothetical DNA binding domain (transposase_35) normally found in the C-terminus of many transposases. Five homologues for these genes are present in *H. virescens* ascovirus 3e (HvAV-3e; [37]). These ORFs constitute the homologous repeat regions for HvAV-3e. Two homologues are also found in the *Spodoptera frugiperda* ascovirus 1a (SfAV-1a; [36]) genome, and a single homologue is found in MacoNPV-B. No homologues for these ORFs are present in the XecnGV genome. Phylogenetic analysis divided the ascovirus and HearGV members of this gene family into separate clades (Fig. 3b). The MacoNPV-B homologue, *macoB18*, was grouped either with the HvAV-3e ORFs (by minimum evolution) or the SfAV-1a ORFs (by maximum parsimony; data not shown). However, *macoB18* is only 301 codons in length and aligns with the C-terminus of the ascovirus ORFs. In contrast, the HearGV homologues, at 572 and 576 codons, align with most of the length of the ascovirus ORFs. A high concentration of polymorphisms (83 substitutions and six indels) occurs in *hear53*.

Comparison with XecnGV

Excluding gaps inserted to optimize the alignment, the overall nucleotide sequence identity between the HearGV and XecnGV genomes is 94.8%. The smaller size of the HearGV genome relative to the XecnGV genome is due to the reduced number of repeats in the *hr*s shared by the two viruses (Table 1) and a loss of approximately 16.6 kbp of XecnGV-homologous sequence from the non-*hr* regions of HearGV, counterbalanced by the occurrence of approximately 8.2 kbp of novel sequence not found in XecnGV (Fig. 4). Of the 179 ORFs in HearGV, 167 have homologues in the XecnGV genome. Two of these ORFs include homologues for two previously uncharacterized XecnGV ORFs (*xc77a* and *xc89a*) that also occur in other GVs (Table 1). There are differences in the promoter motif composition of 17 of the ORFs that HearGV and XecnGV have in common. Eight HearGV ORFs have a late or early promoter motif within 120 bp of the start of the ORF, while the corresponding XecnGV homologues are lacking these motifs. Nine XecnGV ORFs are preceded by early or late promoter motifs not present in the corresponding HearGV homologues. In addition to conservation of ORF content and *hr* placement, three relatively large (>500 bp) non-*hr* intergenic regions are conserved in the HearGV and XecnGV genomes. These intergenic regions lie between ORF pairs *hear23*/*xc24* and *hear24*/*xc25*, *hear24*/*xc25* and *hear25*/*xc26*, and *hr4* and *hear70*/*xc76*.

Two XecnGV ORFs were fused into a single ORF in HearGV. A two-nt insertion at HearGV nt 26447–26448 relative to XecnGV resulted in a frameshift and fusion of sequences homologous to XecnGV ORFs *xc30* and *xc31* into a single ORF, *hear29*. A one-nt insertion at HearGV nt 37744 relative to XecnGV caused a frameshift enabling the *xc47*-homologous ORF to extend past a stop codon into the *xc48*-homologous ORF, leading to a fusion of these ORFs into *hear44*.

Fourteen XecnGV ORFs (*xc6*, *xc22*, *xc49*, *xc58*, *xc59*, *xc63*, *xc69*, *xc104*, *xc130*, *xc138*, *xc151*, *xc153*, *xc160*, and *xc163*) are not represented by homologues in the HearGV genome (Table 5). In addition, homologues for the XecnGV ORFs *xc41* and *xc156* are not listed among the ORFs of HearGV in Table 1. These sequences are present in the HearGV genome, but they are nested within larger

Table 5 XecnGV ORFs missing from HearGV

XecnGV ORF	Features
xc6	Missing initiation codon in HearGV
xc22	Sequence missing from HearGV
xc49	Missing initiation codon in HearGV
xc58	*v-cath* (*ac127*) homologue; sequence missing from HearGV
xc59	*Chrysodeixis chalcites* NPV ORF *chch133* homologue; sequence missing from HearGV
xc63	*Mamestra configurata* NPV-A ORF 29/NPV-B ORF 24 homologue; sequence missing from HearGV
xc69	Sequence missing from HearGV
xc104	Missing initiation codon in HearGV
xc130	XecnGV *bro-e*; sequence missing from HearGV
xc138	*Chrysodeixis chalcites* NPV ORF *chch133* homologue; sequence missing from HearGV
xc151	Homologues in *Mamestra configurata* NPV-A & -B, *Agrotis segetum* NPV and GV; sequence missing from HearGV
xc153	Sequence missing from HearGV
xc160	*Autographa californica* MNPV *ac111* homologue; sequence missing from HearGV
xc163	Sequence missing from HearGV

neighboring ORFs (*hear38* and *hear155*, respectively). Of the missing ORFs, five (*xc6*, *xc63*, *xc104*, *xc153*, and *xc160*) are relatively small (<100 codons). The remaining ORFs are >150 codons in size. The absence from HearGV of homologues for *xc22*, *xc58*, *xc59*, *xc63*, *xc69*, *xc130*, *xc138*, *xc151*, *xc153*, *xc160*, and *xc163* was due to the deletion of XecnGV-homologous sequences containing those ORFs from the HearGV genome (Fig. 4). Homologues for *xc6*, *xc49*, and *xc104* were absent from the HearGV genome due to substitutions leading to the loss of the initiation codons for those ORFs. Two of the missing ORFs (*xc59* and *xc138*) share significant sequence similarity with each other and with ORF 133 from the *Chrysodeixis chalcites* NPV genome [53]. ORF *xc63* has homologues in MacoNPV-A and -B, while ORFs *xc151* and *xc160* have homologues in several NPVs. ORF *xc58* encodes a viral cathepsin L present in most baculoviruses that is involved in the liquefaction of host tissues and disintegration of the host cuticle often seen with baculovirus infections of larvae [54, 55]. Of nine completely sequenced granulovirus genomes, four (XecnGV, CrleGV, AgseGV, and CpGV) contain *cathepsin* genes, while five (PhopGV, PlxyGV, AdorGV, SpltGV, and ChocGV) do not possess a *cathepsin* gene.

The HearGV genome contains 10 ORFs (*sprT (hear21)*, *hear53*, *hear56*, *hear70*, *hear75*, *hear76*, *hear104*, *hear111*, *hear115*, and *hear157*) with no homologues listed for XecnGV. Six of these ORFs (*hear21*, *hear53*, *hear56*, *hear75*, *hear76*, and *hear157*) are contained within insertions of novel sequence in the HearGV genome that are not present in XecnGV (Fig. 4). Homologous sequences for *hear70*, *hear111*, and *hear115* exist in the XecnGV genome, but are shorter than 50 codons in length due to substitutions and indels resulting in the appearance of stop codons in the XecnGV sequences. A substitution in XecnGV eliminates the start codon for *hear104*.

Conclusions

Whole-genome analyses of baculovirus phylogeny have revealed that baculovirus genomes in general exhibit a tremendous degree of "genomic plasticity" [56, 57]. This genetic fluidity is also on display in the HearGV genome and its relationship to the XecnGV genome. Although HearGV and XecnGV share a high degree of nucleotide sequence identity, numerous re-arrangements in the form of insertions and deletions have occurred in HearGV and XecnGV since they diverged. A number of ORFs unique to either XecnGV or HearGV individually, or unique to both viruses, are located next to or near homologous regions (Tables 1 and 4), further highlighting the role that these sequences play in gene loss and acquisition events [58]. Many genes in HearGV and XecnGV are represented by homologues in the genomes of nucleopolyhedrovirues, ascoviruses, entomopoxviruses, and polydnaviruses, suggesting the acquisition or exchange of genes by non-homologous recombination with viruses other than GVs.

Results of pairwise nucleotide distances for three marker genes suggest that HearGV either belongs to the XecnGV species cluster or is very closely related to XecnGV. In addition to the four variants of XecnGV identified by Lange et al. [34], restriction endonuclease analysis by Goto et al. [59] previously identified additional GVs from four other noctuid hosts (*Hydraecia amurensis*, *Celaena leucostigma*, *Aletia pallens*, and *Pseudaletia separata*) that likely are also variants of XecnGV. These data suggest that the virus species group that includes HearGV and XecnGV possess a very

broad host range, not only among granuloviruses, but among baculoviruses in general. The ORFs unique to HearGV and XecnGV (Table 4), as well as other ORFs shared by these GVs that do not occur in other GVs, may account for the broad host range exhibited by this virus group.

Acknowledgments The authors wish to thank Tad Sonstegard (USDA-ARS, Bovine Functional Genomics Laboratory) for the use of his GeneMachines Hydroshear, Jing Hu (USDA-ARS, Invasive Insect Biocontrol and Behavior Laboratory) and Stephen Rehner (USDA-ARS, Systematic Mycology and Microbiology Laboratory) for assistance with sequencing, and Sandra Brandt (USDA-ARS Biological Control of Insects Research Laboratory) for assistance with fosmid cloning**.** 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. D.A. Theilmann, G.W. Blissard, B. Bonning, J.A. Jehle, R. O'Reilly D, 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
2. 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)
3. B.A. Federici, in *The Baculoviruses*, ed. by L.K. Miller (Plenum Press, New York, 1997), pp. 33–59
4. J.A. Jehle, M. Lange, H. Wang, Z. Hu, Y. Wang, R. Hauschild, Virology **346**, 180–193 (2006)
5. T. Hayakawa, R. Ko, K. Okano, S.I. Seong, C. Goto, S. Maeda, Virology **262**, 277–297 (1999)
6. Y. Hashimoto, T. Hayakawa, Y. Ueno, T. Fujita, Y. Sano, T. Matsumoto, Virology **275**, 358–372 (2000)
7. T. Luque, R. Finch, N. Crook, D.R. O'Reilly, D. Winstanley, J. Gen. Virol. **82**, 2531–2547 (2001)
8. S. Wormleaton, J. Kuzio, D. Winstanley, Virology **311**, 350–365 (2003)
9. M. Lange, J.A. Jehle, Virology **317**, 220–236 (2003)
10. S.R. Escasa, H.A.M. Lauzon, A.C. Mathur, P.J. Krell, B.M. Arif, J. Gen. Virol. **87**, 1917–1933 (2006)
11. V.H. Whitlock, J. Invertebr. Pathol. **23**, 70–75 (1974)
12. J.J. Hamm, Environ. Entomol. **11**, 159–160 (1982)
13. M. Shapiro, J. Econ. Entomol. **93**, 1633–1637 (2000)
14. P.W. Roelvink, B.G. Corsaro, R.R. Granados, J. Gen. Virol. **76**, 2693–2705 (1995)
15. V.H. Whitlock, J. Invertebr. Pathol. **29**, 297–303 (1977)
16. K.J. Hackett, A. Boore, C. Deming, E. Buckley, M. Camp, M. Shapiro, J. Invertebr. Pathol. **75**, 99–106 (2000)
17. D.R. O'Reilly, L.K. Miller, V.A. Luckow, *Baculovirus Expression Vectors* (W.H. Freeman and Company, New York, 1992)
18. R.L. Harrison, D.E. Lynn, Virus Genes **35**, 857–873 (2007)
19. J.D. Thompson, D.G. Higgins, T.J. Gibson, Nucleic Acids Res. **22**, 4673–4680 (1994)
20. T.A. Hall, Nucleic Acids Symp. Ser. **41**, 95–98 (1999)
21. K. Tamura, J. Dudley, M. Nei, S. Kumar, Mol. Biol. Evol. **24**, 1596–1599 (2007)
22. M. Kimura, J. Mol. Evol. **16**, 111–120 (1980)
23. H. Xiao, Y. Qi, Virus Genes **35**, 845–856 (2007)
24. 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)
25. L. Li, Q. Li, L.G. Willis, M. Erlandson, D.A. Theilmann, C. Donly, J. Gen. Virol. **86**, 91–105 (2005)
26. J.S. Cory, J.H. Myers, Annu. Rev. Ecol. Evol. Syst. **34**, 239–272 (2003)
27. J.S. Cory, B.M. Green, R.K. Paul, F. Hunter-Fujita, J. Invertebr. Pathol. **89**, 101–111 (2005)
28. R.D. Possee, G.F. Rohrmann, in *The Baculoviruses*, ed. by L.K. Miller (Plenum, New York, 1997), pp. 109–140
29. S. Hilton, D. Winstanley, J. Gen. Virol. **88**, 1496–1504 (2007)
30. 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)
31. R.L. Harrison, B.C. Bonning, J. Gen. Virol. **84**, 1827–1842 (2003)
32. C.X. Zhang, X.C. Ma, Z.J. Guo, Virology **333**, 190–199 (2005)
33. M.D. Ayres, S.C. Howard, J. Kuzio, M. Lopez-Ferber, R.D. Possee, Virology **202**, 586–605 (1994)
34. M. Lange, H. Wang, H. Zhihong, J.A. Jehle, Virology **325**, 36–47 (2004)
35. 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)
36. D.K. Bideshi, M.V. Demattei, F. Rouleux-Bonnin, K. Stasiak, Y. Tan, S. Bigot, Y. Bigot, B.A. Federici, J. Virol. **80**, 11791–11805 (2006)
37. S. Asgari, J. Davis, D. Wood, P. Wilson, A. McGrath, J. Gen. Virol. **88**, 1120–1132 (2007)
38. D.A. Theilmann, M.D. Summers, J. Virol. **61**, 2589–2598 (1987)
39. D.A. Theilmann, M.D. Summers, Virology **167**, 329–341 (1988)
40. R.S. Hilgarth, B.A. Webb, J. Gen. Virol. **83**, 2393–2402 (2002)
41. K. Tanaka, R. Lapointe, W.E. Barney, A.M. Makkay, D. Stoltz, M. Cusson, B.A. Webb, Virology **363**, 26–35 (2007)
42. C.P. Ponting, Nucleic Acids Res. **30**, 3643–3652 (2002)
43. L. Wang, J. Xue, C.P. Seaborn, B.M. Arif, X.W. Cheng, Virology **354**, 167–177 (2006)
44. Q. Li, C. Donly, L. Li, L.G. Willis, D.A. Theilmann, M. Erlandson, Virology **294**, 106–121 (2002)
45. L. Li, C. Donly, Q. Li, L.G. Willis, B.A. Keddie, M.A. Erlandson, D.A. Theilmann, Virology **297**, 226–244 (2002)
46. D.K. Bideshi, S. Renault, K. Stasiak, B.A. Federici, Y. Bigot, J. Gen. Virol. **84**, 2531–2544 (2003)
47. W. Kang, M. Kurihara, S. Matsumoto, Virology **350**, 184–191 (2006)
48. W. Kang, M. Suzuki, E. Zemskov, K. Okano, S. Maeda, J. Virol. **73**, 10339–10345 (1999)
49. W.K. Kang, N. Imai, M. Suzuki, M. Iwanaga, S. Matsumoto, E.A. Zemskov, Arch. Virol. **148**, 99–113 (2003)
50. E.A. Zemskov, W. Kang, S. Maeda, J. Virol. **74**, 6784–6789 (2000)
51. L.M. Iyer, E.V. Koonin, L. Aravind, Genome Biol. **3**, RESEARCH0012 (2002)
52. A.L. Bawden, K.J. Glassberg, J. Diggans, R. Shaw, W. Farmerie, R.W. Moyer, Virology **274**, 120–139 (2000)
53. 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)
54. J.M. Slack, J. Kuzio, P. Faulkner, J. Gen. Virol. **76**, 1091–1098 (1995)
55. R.E. Hawtin, T. Zarkowska, K. Arnold, C.J. Thomas, G.W. Gooday, L.A. King, J.A. Kuzio, R.D. Possee, Virology **238**, 243–253 (1997)
56. E.A. Herniou, T. Luque, X. Chen, J.M. Vlak, D. Winstanley, J.S. Cory, D.R. O'Reilly, J. Virol. **75**, 8117–8126 (2001)
57. E.A. Herniou, J.A. Olszewski, J.S. Cory, D.R. O'Reilly, Annu. Rev. Entomol. **48**, 211–234 (2003)
58. 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)
59. C. Goto, Y. Minobe, T. Iizuka, J. Gen. Virol. **73**, 1491–1497 (1992)