Available online at www.sciencedirect.com

ScienceDirect

Virology 360 (2007) 235–246

VIROLOGY

www.elsevier.com/locate/yviro

Genome sequence and organization of a nucleopolyhedrovirus that infects the tea looper caterpillar, *Ectropis obliqua*

Xiu-Cui Ma [a], Jin-Yan Shang [a], Zhang-Nv Yang [a], Yan-Yuan Bao [a], Qiang Xiao [b], Chuan-Xi Zhang [a,*]

[a] *Institute of Insect Science, Zhejiang University, 268 Kaixuan Road, Hangzhou, 310029, China*
[b] *Tea Research Institute, Chinese Academy of Agricultural Sciences, Hangzhou, 310000, China*

Received 20 September 2006; returned to author for revision 9 October 2006; accepted 11 October 2006
Available online 9 November 2006

Abstract

The complete nucleotide sequence of *Ectropis obliqua* nucleopolyhedrovirus (EcobNPV), which infects the tea looper caterpillar, was determined and analyzed. The double stranded circular genome is composed of 131,204 bp and is 37.6% G+C rich. The analysis predicted 126 putative, minimally overlapping open reading frames (ORFs) with 150 or more nucleotides that together compose 89.8% of the genome. The remaining 10.2% constitute non-coding and three homologous regions. Comparison with previously sequenced baculoviruses indicated that three ORFs were unique to EcobNPV, while the remaining 123 ORFs shared identity with other baculovirus genes. In addition to two *bro* homologues, three other repeat ORFs, including *dbp*, *p26*, and *odv-e66*, were identified. Phylogenetic analysis indicated that each member of the paired ORFs was acquired independently. Gene parity plot analysis and percent identity of gene homologues suggested that EcobNPV is a Group II NPV, although its genomic organization was highly distinct.

Keywords: *Ectropis obliqua*; Nucleopolyhedrovirus; Complete genome; Gene parity plot; Phylogeny

Introduction

The Baculoviridae represent a large and diverse family of rod-shaped viruses with circular, double stranded DNA genomes ranging in size from the 81.7 kb genome of *Neodiprion lecontei* nucleopolyhedrovirus (NeleNPV) to the 178.7 kb genome of *Xestia c-nigrum* GV (XcGV). The baculovirus family includes two genera, the nucleopolyhedroviruses (NPV) and granuloviruses (GV). Lepidoptera NPVs are subdivided into Groups I and II based on their molecular phylogenies (Zanotto et al., 1993). Updated classification has proposed to expand the baculovirus family into four genera: the alphabaculoviruses (Lepidopteran-specific NPV), betabaculoviruses (lepidopteran-specific GV), gammabaculoviruses (hymenopteran-specific NPV), and deltabaculoviruses (dipteran-specific baculovirus) (Jehle et al., 2006). To date, 34 completely sequenced baculovirus genomes have been determined, including 22 alphabaculoviruses, eight betabaculoviruses, three gammabaculoviruses, and one deltabaculovirus based on this proposed scheme.

The tea looper caterpillar, *Ectropis obliqua* (Lepideptera: Geometride), infests tea plants in East Asia and can cause considerable damage (Chen and Huang, 2001). *E. obliqua* nucleopolyhedrovirus (EcobNPV) is a singly embedded NPV (Ma et al., 2006) that is pathogenic to this caterpillar. As a result, EcobNPV has become an important biological insecticide (Hu et al., 1994) that is now commercially available in China.

Restriction maps of EcobNPV have been assembled and 15.5 kb of the *polyhedrin* region is characterized (Ma et al., 2006). In this study, the complete genome of EcobNPV has been sequenced and analyzed in order to better understand how this virus evolved, as well as its molecular mechanisms of infection and replication.

* Corresponding author. Fax: +86 571 86971697.
E-mail address: chxzhang@zju.edu.cn (C.-X. Zhang).

0042-6822/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2006.10.024

Results and discussion

Nucleotide sequence analysis of the EcobNPV genome

The EcobNPV genome consists of 131,204 bp (GenBank accession no.DQ837165), which is similar to the 127.7 kb size predicted by restriction enzyme analysis (Ma et al., 2006), and within the 81.7 and 178.7 kb range for baculovirus genomes. The EcobNPV genome is highly AT-rich, having only 37.6% G+Cs, similar to *Chrysodeixis chalcites* NPV (ChchNPV, 39.1%), *Helicoverpa armigera* NPV (HearNPV, 39.1%), *Helicoverpa zea* SNPV (HzSNPV, 39.1%), *Rachiplusia ou* MNPV (RaouMNPV, 39.1%), and *Adoxophyes honmai* NPV (AdhoNPV, 35.6%), but lower than most other sequenced NPVs.

According to the adopted convention (Vlak and Smith, 1982; Hayakawa et al., 1999; Ijkel et al., 1999), the adenine residue at the translation initiation codon of the *polyhedrin* gene represented the zero point on the EcobNPV physical map, and was designated ORF 1 (Fig. 1 and Table 1). 126 putative ORFs and three homologous regions (hrs) were detected in the EcobNPV genome using computer-assisted analysis to select ORFs starting from the methionine-initiated codon (ATG) and including at least 50 amino acids (aa) having minimal overlap with other ORFs. All 126 ORFs are shown in Table 1 by location, orientation, size, and potential baculovirus homologues. The number of ORFs was similar to other fully sequenced baculoviruses, which range from 89 (NeleNPV) to 181 (XcGV), and was most like AdhoNPV (125). EcobNPV

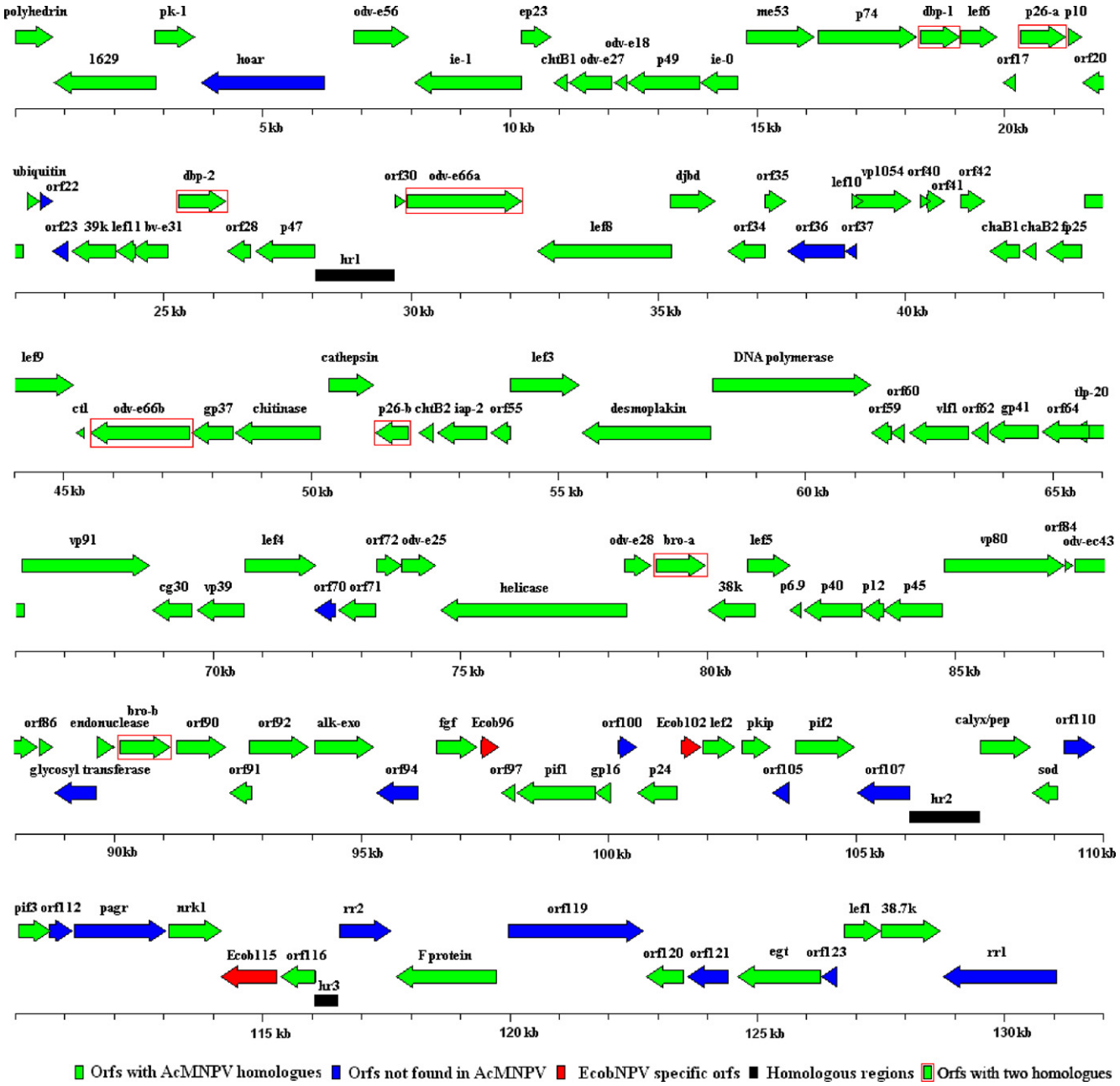

Fig. 1. Linear map of the 126 predicted ORFs for the complete EcobNPV genome.

ORFs had an average length of 954 bp, with ORF 73 (*helicase*) being the largest (3753 bp) and ORF 47 (*ctl*, conotoxin-like protein) being the smallest (159 bp). The 126 predicted ORFs encoded 39,620 aa. The total coding sequence and intergenic regions were 117,873 and 13,331 bp, and represented 89.8% and 10.2% of the genome, respectively. The three hrs were distributed along the genome with sizes ranging from 493 to 1615 bp, and their total sequence was 3541 bp, accounting for 2.7% of the genome. 17 ORFs overlapped with adjacent ORFs by between 2 and 389 bp, with a total number of 987 bp.

Of the 126 ORFs identified in EcobNPV, 37 (29%) possessed a consensus early promoter motif (a TATA box followed by a CAGT or CATT motif 20 to 40 bp downstream and up to 180 bp upstream of the initiation codon), 75 (60%) only contained a late promoter motif ((A/T/G)TAAG up to 180 bp upstream of the initiation codon), and 22 (17%) had early and late promoter motifs, which might allow transcription during both the early and late stages of infection. 35 ORFs lacked any recognizable consensus early or late promoter motifs up to 180 bp upstream of ATG. 58 ORFs (46%) were directed clockwise and 68 ORFs (54%) were directed counterclockwise according to the transcription orientation of the *polyhedrin* gene.

Genome organization and gene homology of EcobNPV as compared to other baculoviruses

The EcobNPV genome was compared with those from *Autographa californica* MNPV (AcMNPV) (Ayres et al., 1994), *Lymantria dispar* MNPV (LdMNPV) (Kuzio et al., 1999), *Mamestra configurata* NPV-B (MacoNPV-B) (Li et al., 2002a,b), *Spodoptera exigua* MNPV (SeMNPV) (Ijkel et al., 1999), ChchNPV (Van Oers et al., 2005), and XcGV (Hayakawa et al., 1999) (Tables 1 and 2). All 62 lepidopteran baculovirus genes, including the 30 core baculovirus genes (Herniou et al., 2003), were found in the EcobNPV genome (Fig. 1 and Table 1). Of the 126 EcobNPV ORFs identified, 123 had either an assigned function or homologues in other baculoviruses. EcobNPV shared the largest number of homologues with MacoNPV-B (118), accounting for 94% of the total ORFs. 104 (83%) EcobNPV ORFs had homologues in AcMNPV and 113 (90%) had homologues in SeMNPV. A lower percentage of homologues (63%) were found in XcGV (Tables 1 and 2). The mean ORF amino acid identity between EcobNPV and the Group II baculoviruses, SeMNPV, MacoNPV-B, and ChchNPV was similar (40–42%), and higher than between EcobNPV and AcMNPV (34%), supporting the phylogenetic relationship between Group I and Group II NPVs. Further evidence that EcobNPV is a Group II NPV was provided by the presence of the baculovirus F protein gene (Ecob118), and absence of the Group I NPV specific *gp64* gene (Kuzio et al., 1999; Ijkel et al., 1999; Chen et al., 2001). This finding is in agreement with the classification that was given to EcobNPV based on phylogenetic analysis (Ma et al., 2006). *Polyhedrin* was the most conserved ORF within all six NPVs (EcobNPV shared 85%, 82%, 94%, 90% and 93% amino acid identity with AcMNPV, LdMNPV, MacoNPV-B, SeMNPV, and ChchNPV, respectively), followed by *ubiquitin* (75%, 79%, 84%, 85%, and 77% identity).

Of the 126 EcobNPV ORFs, GenBank BLAST results showed that 114 ORFs have their best matched homologues in Group II NPV members, including AgseNPV, MacoNPV, LdMNPV, SeMNPV, TnNPV, ChchNPV, AdhoNPV and HearNPV. Five ORFs have the highest similarity to the homologues in Group I NPVs, including 3 (Ecob28, 31, 67) from AcMNPV and 2 (Ecob50 and 107) from CfDEFNPV, while 3 ORFs (Ecob48, 53, and 88) were found to be best matched with that of XcGV (Table 1). In addition to 3 unique EcobNPV ORFs, Ecob79 (p6.9) was not matched with any proteins in GenBank by using the BLAST software. However, the homologues of p6.9 were found in all sequenced baculoviruses due to the presence of polyarginine and polyserine. Ecob79 potentially encodes a protamine-like basic DNA binding protein (BDBP), which is characterized by being rich in arginine (39%) and serine (19%).

The genomic organization of EcobNPV was examined using the gene parity plot to reveal gene order and similarity by finding ORFs from other baculoviruses that share homology with EcobNPV (Hu et al., 1998) (Fig. 2). Comparison of the relative gene order between EcobNPV and AcMNPV, LdMNPV, MacoNPV-B, SeMNPV, ChchNPV, or XcGV, revealed 14 gene clusters that were conserved in all seven baculoviruses (Fig. 2 and Table 1). The presence of these clusters may have resulted from physical constraint that prevented their separation (Herniou et al., 2003). At the level of gene content, arrangement, and homology, EcobNPV was more closely related to MacoNPV-B, SeMNPV, and ChchNPV than AcMNPV and XcGV (Table 1 and Fig. 2). 58 ORFs were unique to AcMNPV (*y*-axis) and 22 ORFs were unique to EcobNPV (*x*-axis) (Fig. 2A). The amino acid sequence identities of most ORFs that EcobNPV and AcMNPV shared in common were below 50%. In contrast, 27 ORFs were unique to SeMNPV, and 13 ORFs were unique to EcobNPV. Only 8 and 12 ORFs present in EcobNPV were not found in MacoNPV-B and ChchNPV, respectively.

Compared to MacoNPV-B and SeMNPV, most of the EcobNPV genome was inversely oriented relative to the *polyhedrin* gene. While a considerable amount of co-linearity was observed between EcobNPV and the Group II NPVs, 22 EcobNPV genes were not co-linear with SeMNPV homologues, including Ecob13 (Se7), Ecob14 (Se131), Ecob27 (Se126), Ecob29 (Se115), Ecob48 (Se57/114), Ecob49 (Se25), Ecob50 (Se19), Ecob51 (Se16), Ecob52 (Se87/129), Ecob53 (Se96/134), Ecob54 (Se88), Ecob87 (Se56), Ecob99 (Se9), Ecob101 (Se10), Ecob103 (Se12), Ecob106 (Se35), Ecob108 (Se46), Ecob114 (Se54), Ecob116 (Se53), Ecob117 (Se45), Ecob118 (Se8), and Ecob126 (Se139) (Table 1 and Fig. 2). Most of these ORFs were located in two high variable regions of the EcobNPV genome. The ‘right’ part of this genome (ORF 99-118) displayed a high degree of gene scrambling in the gene parity plot analysis. Unlike other baculoviruses described previously (Ijkel et al., 1999; Chen et al., 2001), the ‘central’ region (ORF48-54), including two duplicate ORFs, *odv-e66* and *p26*, showed considerable gene translocations (Fig. 2). The

Table 1
ORFs identified in EcobNPV

ORF	Name	Position	Promoter	Length (aa)	Homologous ORF#/amino acid identity (%)						Best matched baculovirus ORF, BLAST score (bits)
					AcMNPV	LdMNPV	Maco-B	SeMNPV	Chch	XcGV	
1	polyhedrin	1>741	L	246	8(85)	1(82)	1(94)	1(90)	1(93)	1(53)	OpS ph, 468
2	orf1629	773<2818	None	681	9(16)	2(16)	2(15)	2(15)	2(18)	2(18)	Ld2, 43.5
3	pk-1	2817>3608	L	263	10(33)	3(41)	3(42)	3(45)	3 (45)	3(26)	Agse3, 246
4	hoar	3741<6227	E	828	–	–	4(15)	4(18)	4(16)	–	Hear4, 142
5	odv-e56	6847>7938	L	363	148(50)	14(59)	6(46)	6 (48)	7(52)	15(40)	Ld14, 359
6	ie-1	8050<10,215	E	721	147(22)	15(28)	161(30)	132(28)	16(22)	9(13)	MacoB161, 324
7	ep23	10,224>10,826	None	200	146(26)	16(26)	162(31)	133(26)	15(32)	10(20)	MacoB162, 113
8	chtB1	10,851<11,129	L	92	145(46)	17(59)	163(54)	134(54)	14(55)	11(36)	Adho21, 127
9	odv-e27	11,168<12,031	L	287	144(39)	18(59)	164(51)	135(51)	13(51)	112(21)	Ld18, 266
10	odv-e18	12,090<12,347	L	85	143(55)	19(69)	165(47)	136(53)	12(57)	12(40)	Ld19, 43.1
11	p49	12,362<13,810	L	482	142(44)	20(52)	166(53)	137(52)	11(58)	13(26)	Chch11, 549
12	ie-0	13,830<14,588	L	252	141(26)	21(32)	167(36)	138(33)	10(36)	14(12)	Chch10, 187
13	me53	14,813>16,165	L	450	139(14)	23(33)	7(25)	7 (21)	8(25)	180(15)	Ld23, 224
14	p74	16,243>18,201	L	652	138(55)	27(62)	159(55)	131(58)	17(59)	77(31)	Ld27, 767
15	dbp-1	18,319>19,116	E	265	25(11)	37(12)/47(16)	154(14)	126(19)	22(16)	89(14)	Hz24, 87.8
16	lef6	19,141>19,866	L	241	28(17)	38(23)	155(26)	127(24)	21(26)	88(18)	Tn21, 71.6
17	Unknown	19,957<20,196	None	79	29(31)	39(25)	156(32)	128(30)	20(37)	16(17)	Chch20, 50.8
18	p26-a	20,358>21,260	L	300	136(23)	40(23)	157(39)	129(37)/87(16)	19(36)/63(13)	–	Agse142, 233
19	p10	21,308>21,571	L	87	137(21)	41(53)	158(61)	130(62)	18(60)	5(32)	Agse143, 86.7
20	Unknown	21,582<22,151	L	189	34(24)	42(37)	152(46)	124(45)	25(49)	–	Agse136, 184
21	ubiquitin	22,219>22,458	E, L	79	35(75)	43(73)	151(84)	123(85)	26(77)	52(77)	Splt32, 115
22	Unknown	22,488>22,715	E	75	–	–	150(25)	122(17)	27(27)	–	Adho15, 36.6
23	Unknown	22,755<23,045	None	–	–	–	–	121(16)	–	–	Agse133, 32.3
24	39k	23,136<24,017	E, L	293	36(32)	44(37)	149(28)	120(27)	28(28)	55(13)	Agse132, 154
25	lef11	24,046<24,411	None	121	37(23)	45(33)	148(39)	119(39)	29(38)	56(33)	Ld45, 79.3
26	bv-e31	24,342<25,073	E, L	243	38(52)	46(49)	147(59)	118(54)	30(49)	79(31)	MacoA148, 271
27	dbp-2	25,281>26,228	L	315	25(20)	47(33)/37(13)	154(18)	126(20)	22(19)	89(14)	Ld47, 198
28	Unknown	26,289<26,741	E	151	63(24)	117(13)	–	–	31(24)	–	Ac63, 51.2
29	p47 hr1	26,868<28,049 28,050–29,664	E, L	393	40(52)	48(59)	144(58)	115(57)	33(59)	78(39)	MacoA145, 455
30	Unknown	29,665>29,850	L	61	43(23)	–	142(33)	113(36)	35(33)	–	Agse124, 42
31	odv-e66a	29,896>32,211	E, L	771	46(63)	131(31)	143(31)	114(19)/57(19)	101(35)	149(36)	Ac46, 819
32	lef8	32,565<35,240	None	891	50(60)	51(64)	140(65)	112(61)	37(63)	148(46)	Hear38, 1169
33	djbd	35,261>36,160	E, L	299	51(13)	–	139(14)	111(13)	38(14)	–	Adho48, 55.5
34	Unknown	36,384<37,136	E	250	52(15)	53(16)	137(23)	109(17)	40(21)	–	Chch40, 60.5
35	Unknown	37,189>37,608	E, L	139	53(40)	54(53)	136(52)	108(49)	41(47)	171(16)	Ld54, 167
36	Unknown	37,605<38,747	E, L	380	–	55(14)	135(16)	107(18)	42(17)	–	Hear45, 54.7
37	Unknown	38,762<38,977	L	71	–	–	134(28)	–	43(24)	–	Hear46, 312
38	lef-10	38,949>39,170	L	73	53a(37)	56(36)	133(44)	106(40)	44(48)	174(27)	Tn41, 65.9
39	vp1054	39,037>40,137	None	366	54(35)	57(49)	132(50)	105(49)	45(47)	175(29)	Ld57, 364
40	Unknown	40,319>40,516	None	65	55(28)	58(31)	131(42)	104(51)	46(45)	–	Se104, 53.9
41	Unknown	40,458>40,805	E, L	115	56(18)	–	130(26)	103(33)	47(31)	–	Se103, 57.8
42	Unknown	41,143>41,628	E	161	57(32)	60(44)	129(37)	102(37)	48(41)	–	Ld60, 150
43	chaB1	41,688<42,281	L	197	59(35)	61(28)	128(30)	101(31)	49(37)	–	Tn46, 88.6
44	chaB2	42,348<42,611	L	87	60(35)	62(41)	127(39)	100(40)	50(37)	102(20)	Hz53, 60.8
45	fp25	42,833<43,540	None	235	61(48)	63(24)	124(64)	98(63)	51(54)	140(27)	Agse106, 231
46	lef9	43,643>45,157	None	504	62(65)	64(13)	123(69)	97(70)	52 (73)	139(53)	Chch52, 701
47	ctl	45,223<45,384	E, L	53	3(40)	149(43)	106(62)	–	74(63)	127(54)	MacoA107, 53.1
48	odv-e66b	45,516<47,513	L	665	46(32)	131(46)	77(50)/143(27)	57 (31)/114(26)	101(50)	149(56)	XcGV149, 774
49	gp37	47,558<48,391	L	277	64(44)	68(53)	32(58)	25 (56)	67(55)	107(39)	MacoB32, 332
50	chitinase	48,434<50,170	L	578	126(67)	70(64)	19(63)	19(61)	65(66)	103(56)	CfDEF121, 760
51	cathepsin	50,337>51,236	E, L	299	127(68)	78(66)	28(57)	16(55)	64(52)	58(45)	Busu cath, 485
52	p26-b	51,281<51,949	E	222	136(15)	40(14)	108(29)	87(31)/129(14)	63(26)	–	Agse94, 134
53	chtB2	52,170<52,451	L	93	150(16)/145(14)	30(15)/17(15)	163(14)/38(20)	96(12)/134(17)	14(15)	20(25)	XcGV20, 46.2
54	iap-2	52,530<53,528	E, L	332	71(22)	79(37)	109(34)	88(31)	62(30)	137(13)	Chch62, 137
55	Unknown	53,608<54,015	E	135	68(326)	80(45)	111(52)	90(48)	61(51)	135(23)	Agse97, 134
56	lef3	54,014>55,324	None	436	67(18)	81(27)	112(24)	91(29)	60(27)	134(12)	Agse98, 196
57	desmoplakin	55,451<58,054	L	867	66(15)	82(14)	113(16)	92(19)	59(17)	133(14)	Se92, 97.1
58	DNA pol	58,053>61,250	E	1065	65(38)	83(49)	114(52)	93(50)	58(50)	132(28)	MacoA115, 1049

Table 1 (*continued*)

ORF	Name	Position	Promoter	Length (aa)	Homologous ORF#/amino acid identity (%)						Best matched baculovirus ORF, BLAST score (bits)
					AcMNPV	LdMNPV	Maco-B	SeMNPV	Chch	XcGV	
59	Unknown	61,287<61,679	E, L	130	75(19)	84(47)	115(40)	94(40)	57(41)	126(13)	Ld84, 115
60	Unknown	61,687<61,944	L	85	76(39)	85(79)	116(75)	95(69)	56(78)	125(35)	Chch56, 102
61	vlf1	62,055<63,251	E, L	398	77(65)	86(71)	105(60)	82(63)	76(68)	123(25)	Ld86, 531
62	Unknown	63,310<63,642	L	110	78(28)	87(29)	104(35)	81(37)	77(32)	122(20)	Hz74, 55.5
63	gp41	63,665<64,681	None	338	80(11)	88(10)	103(10)	80(10)	78(10)	121(11)	Ld88, 315
64	Unknown	64,773<65,720	L	315	81(48)	81(11)	102(52)	79(52)	79(52)	120(39)	Tn74, 217
65	tlp-20	65,332<66,186	None	284	82(26)	82(13)	101(42)	78(39)	80(38)	119(16)	Se78, 123
66	vp91	66,116>68,677	L	853	83(36)	91(35)	100(40)	77(41)	81(38)	118(14)	Agse88, 684
67	cg30	68,761<69,570	None	269	88(21)	–	99(15)	76(14)	–	–	Ac56, 58.9
68	vp39	69,664<70,623	L	639	89(33)	92(48)	98(40)	75(44)	82(44)	111(23)	Ld92, 299
69	lef4	70,625>72,040	None	471	90(39)	93(44)	97(45)	74(46)	83(50)	110(25)	Tn78, 429
70	Unknown	72,045<72,458	L	137	–	–	96(23)	–	118(18)	–	MacoB96, 54.3
71	Unknown	72,524<73,279	None	251	92(49)	94(53)	95(62)	73(66)	84(61)	101(34)	Se73, 338
72	Unknown	73,285>73,782	E, L	165	93(44)	95(64)	94(65)	72(64)	85(63)	100(30)	Agse82, 194
73	odv-e25	73,784>74,464	L	226	94(37)	96(63)	93(60)	71(59)	86(64)	99(38)	Chch86, 224
74	helicase	74,596<78,351	L	1251	95(36)	97(49)	92(49)	70(49)	87(49)	98(20)	Agse80, 1138
75	odv-e28	78,308>78,829	None	173	96(49)	98(59)	91(64)	69(68)	88(59)	97(29)	Agse79, 166
76	bro-a	78,936>79,916	None	326	2(15)	154(58)	20(52)	–	55(65)	109(58)	Ha60, 439
77	38k	80,002<80,931	L	309	98(38)	99(47)	87(50)	67(50)	91(47)	96(34)	MacoB87, 325
78	lef5	80,785>81,651	E, L	288	99(46)	100(52)	86(58)	66(55)	92(62)	95(36)	Chch92, 277
79	p6.9	81,648<81,863	L	71	100(47)	101(51)	85(52)	65(44)	93(52)	94(35)	–
80	p40	81,926<83,083	L	385	101(34)	102(46)	84(46)	64(46)	94(50)	93(19)	Chch94, 356
81	p12	83,116<83,538	L	140	102(21)	103(34)	83(25)	63(26)	95 (27)	92(13)	Ld103, 62
82	p45	83,531<84,724	L	397	103(31)	104(58)	82(54)	62(56)	96(55)	91(27)	Ld104, 464
83	vp80	84,764>87,193	L	809	104(12)	105(18)	81(18)	61(17)	97 (20)	–	Chch97, 114
84	Unknown	87,201>87,366	None	55	110(29)	106(55)	80(62)	60(47)	98(56)	51(36)	MacoB80, 48.1
85	odv-ec43	87,396>88,469	L	357	109(47)	107(56)	79(47)	59(50)	99(50)	53(27)	Adho82, 440
86	Unknown	88,512>88,772	E, L	86	108(27)	108(35)	78(40)	58(45)	100(42)	–	Se58, 43.1
87	glycosyl transferase	88,769<89,605	L	278	–	–	75(51)	56(55)	102(46)	–	Se56, 320
88	endonuclease	89,682>90,011	L	109	79(24)	–	15(24)	–	–	75(35)	XcGV75, 58.2
89	bro-b	90,125>91,138	None	337	2(24)	153(22)	121(64)	–	55(16)	109(16)	MacoA122, 415
90	Unknown	91,287>92,273	None	328	112(39)	109(29)	–	–	–	147(30)	Ld109, 159
91	Unknown	92,319<92,747	E, L	142	19(25)	159(25)	49(27)	42(25)	126(32)	–	Tn118, 72.8
92	Unknown	92,758>93,939	None	393	18(13)	158(13)	50(31)	43(32)	125(24)	–	Agse45, 241
93	alk-exo	94,071>95,249	None	392	133(32)	157(35)	48(33)	41(35)	127(33)	145(30)	Ld157, 285
94	Unknown	95,283<96,125	None	280	–	–	47(17)	40(25)	128(28)		Tn120, 82.4
95	fgf	96,527>97,345	L	272	32(22)	156(19)	46(24)	38(19)	130(21)	144(12)	Ld156, 116
96	Ecob96	97,442>97,792	None	116	–	–	–	–	–	–	–
97	Unknown	97,815<98,084	L	89	120(20)	155a(27)	45(26)	37(29)	–	–	MacoA50, 34.3
98	pif1	98,111<99,712	L	533	119(51)	155(43)	44(49)	36(47)	131 (49)	84(27)	Adho117, 641
99	gp16	99,757<100,053	L	98	130(36)	–	10(45)	9 (45)	133(43)	–	Tn125, 82
100	Unknown	100,207>100,593	L	128	–	135(20)	–	–	–	–	Ha68, 41.2
101	p24	100,605<101,387	L	260	129(40)	–	11(46)	10(50)	134(51)	80(18)	Tn126, 223
102	Ecob102	101,493>101,882	E, L	129	–	–	–	–	–	–	–
103	lef2	101,931>102,566	None	211	6(33)	137(37)	13(42)	12(40)	136(42)	35(19)	Orerlef2, 190
104	pkip	102,722>103,291	L	189	24(17)	110(24)	40(35)	32(38)	146(34)	–	Agse33, 81.6
105	Unknown	103,357<103,689	E	110	–	111(24)	41(29)	33(26)	–	–	Agse34, 73.6
106	pif2	103,804>104,982	E, L	392	22(65)	119(62)	43(65)	35(68)	148(64)	45(47)	Agse36, 575
107	Unknown hr2	105,045<106,112 106,113–107,545	None	355	–	–	–	–	124(36)	–	CfDEF108, 323
108	calyx/pep	107,546>108,538	L	330	131(19)	136(45)	61(47)	46(41)	121(49)	–	Chch121, 246
109	sod	108,597<109,091	None	164	31(58)	145(56)	65(59)	48(60)	115(60)	68(50)	Busu sod, 203
110	Unknown	109,226>109,837	E	203	–	144(16)	66(13)	49(16)	111(15)	–	Clbi101, 36.2
111	pif3	110,046>110,672	E, L	208	115(42)	143(47)	67(54)	50(52)	110(37)	32(34)	Se50, 219
112	Unknown	110,659>111,117	None	152	–	142(15)	68(16)	51(20)	–	–	Agse57, 34.7
113	pagr	111,160>113,010	L	616	–	141(16)	69(15)	52(16)	108(17)		Hear100, 110
114	nrk1	113,078>114,145	None	355	33(24)	138(30)	71(40)	54(41)	106(35)	–	Se54, 290
115	Ecob115	114,142<115,266	E	374	–	–	–	–	–	–	–
116	Unknown hr3	115,346<116,038 116,039–116,531	L	230	106(59)	140(64)	70(61)	53(58)	107(69)	50(38)	Chch107, 278

(continued on next page)

Table 1 (*continued*)

ORF	Name	Position	Promoter	Length (aa)	Homologous ORF#/amino acid identity (%)						Best matched baculovirus ORF, BLAST score (bits)
					AcMNPV	LdMNPV	Maco-B	SeMNPV	Chch	XcGV	
117	rr2	116,532>117,557	E	341	–	120(46)	52(60)	45(59)	122(64)	–	Agse47, 378
118	F protein	117,681<119,711	E, L	676	23(14)	130(44)	8 (31)	8(32)	117(16)	27(20)	Clbi129, 628
119	Unknown	119,952>122,669	None	905	–	129(26)	37(20)	30(21)	143(20)	–	Ld129, 451
120	Unknown	122,736<123,491	None	251	17(19)	128(22)	36(18)	29(24)	–	–	Splt2, 103
121	Unknown	123,563<124,381	E	272	–	127(18)	35(17)	28(18)	142(24)	–	Agse29, 42
122	egt	124,585<126,267	None	560	15(39)	125(43)	34(54)	27(52)	141(58)	–	Chch141, 632
123	Unknown	126,275<126,598	L	107	–	124(39)	29(35)	15(29)	139(27)	–	Eups1, 141
124	lef1	126,735>127,463	None	242	14(35)	123(44)	30(46)	14(44)	138(44)	82(30)	Ld123, 233
125	38.7 k (bro)	127,485>128,669	E	394	13(20)	122(21)	31(33)	13(30)	137(25)	81(14)	MacoB31, 216
126	rr1	128,745<131,027	None	760	–	148(14)	168(54)	139(52)	151(55)	–	Eups rr1, 1066

Note. Putative EcobNPV ORFs are listed in column 1 along with the gene homologues designated in column 2. Column 3 indicates ORF location and transcriptional direction on the EcobNPV genome. Column 4 indicates the presence of early (E) and/or late (L) promoters located upstream of the start codon of each ORF. E indicates a TATA sequence followed by a CAGT or CATT mRNA start site sequence 20–40 nucleotides downstream, within 180 bp upstream of the start codon. L indicates the presence of a (A/T/G)TAAG sequence. The number of amino acids is listed in columns 5. Columns 6–11 list the homologue ORF and % aa identity in parentheses from AcMNPV, LdMNPV, MacoNPV-B, SeMNPV, ChchNPV and XcGV, respectively.

gene parity plots revealed that gene order is not well conserved between EcobNPV and other Group II NPVs, which is distinct from Group I NPVs that all share similar gene arrangements (Hu et al., 1998; Ijkel et al., 1999; Chen et al., 2001). Indeed, Group II clustering is far less stable than Group I clustering based on single gene phylogenetic analysis (Hu et al., 1997; Chen et al., 1997).

Gene order is poorly conserved in baculoviruses. A core cluster of four genes, including *helicase*, *lef-5*, Ac96, and *38k* (Ac98), was thought to be present in all baculoviruses based on the sequencing of 13 baculovirus genomes (Herniou et al., 2003). In the EcobNPV genome, a *bro-a* gene is inserted after *helicase* and Ac96 and before *38k* and *lef-5*, however, which separates these genes into two clusters. Genes from other completely sequenced genomes were also found between these two clusters, such as BmNPV ORF 80 (*bro-b*) and 81 (*bro-c*), *Orgyia pseudotsugata* MNPV (OpMNPV) ORF 98, SeMNPV ORF 68, MacoNPV-A ORF 89 (*enhancin*), 90 (*bro-c*), 91, and 92, MacoNPV-B ORF 88 (*vef*), 89 (*bro-d*), 90, and 91, *Choristoneura fumiferana* MNPV (CfMNPV) ORF 89 and 90, *Agrotis segetum* NPV (AgseNPV) ORF 75 (*vef-1*), 76 (*vef-2*), 77 (*bro-b*), and 78, *Trichoplusia ni* SNPV (TnSNPV) ORF 84 and 85, and ChchNPV ORF 89 and 90. Thus, the four gene cluster is not well conserved, and may have diverged into two clusters during baculovirus evolution. The conservation of these two clusters in all completely sequenced baculovirus genomes may suggest that the physical structure of these regions is essential for virus replication or gene overlapping, as suggested previously (Herniou et al., 2003; Garcia-Maruniak et al., 2004). Thus, the continuity of the gene cluster would be maintained through evolution of the baculovirus lineages.

Unique EcobNPV ORFs

The Blast search found no homologues in the GenBank to Ecob96 and Ecob115, while a mouse *espin* gene that encodes a multifunctional actin cytoskeletal regulatory protein was found to show a low similarity (BLAST score=33.5 bits) with Ecob102. Thus, the three ORFs (Ecob96, Ecob102, Ecob115) in the EcobNPV genome were unique in sequenced baculoviruses. These ORFs encode peptides with 116, 129, and 374 aa, respectively. Baculovirus early or late promoter motifs, as well as putative polyadenylation signals, were found upstream of these ORFs indicating that they may represent expressed genes. The potential functions of these ORFs are worth further exploration.

Homologous regions (hrs)

Homologous regions have been found in most NPV genomes except ChchNPV (Van Oers et al., 2005) and TnSNPV (Willis et al., 2005). In the EcobNPV genome, three hrs were identified

Table 2
Characteristics of several baculovirus genomes

Characteristic	EcobSNPV	AcMNPV	LdMNPV	MacoNPV-B	SeMNPV	ChchNPV	XcGV
Size (bp)	131 204	133 894	161 046	158 482	135 611	149 622	178 733
G + C content (mol%)		40.7	57.5	40.0	43.7	39.1	40.7
Total no. ORFs	126	156	163	168	139	151	181
No. hrs	3	8	13	4	6	-	9
Mean amino acid identity (%) with EcobSNPV	-	34	39	41	40	42	29
No. homologues in EcobSNPV	-	104	108	118	113	114	81
References		Ayres et al. (1994)	Kuzio et al. (1999)	Li et al. (2002a)	Ijkel et al. (1999)	van Oers et al. (2005)	Hayakawa et al. (1999)

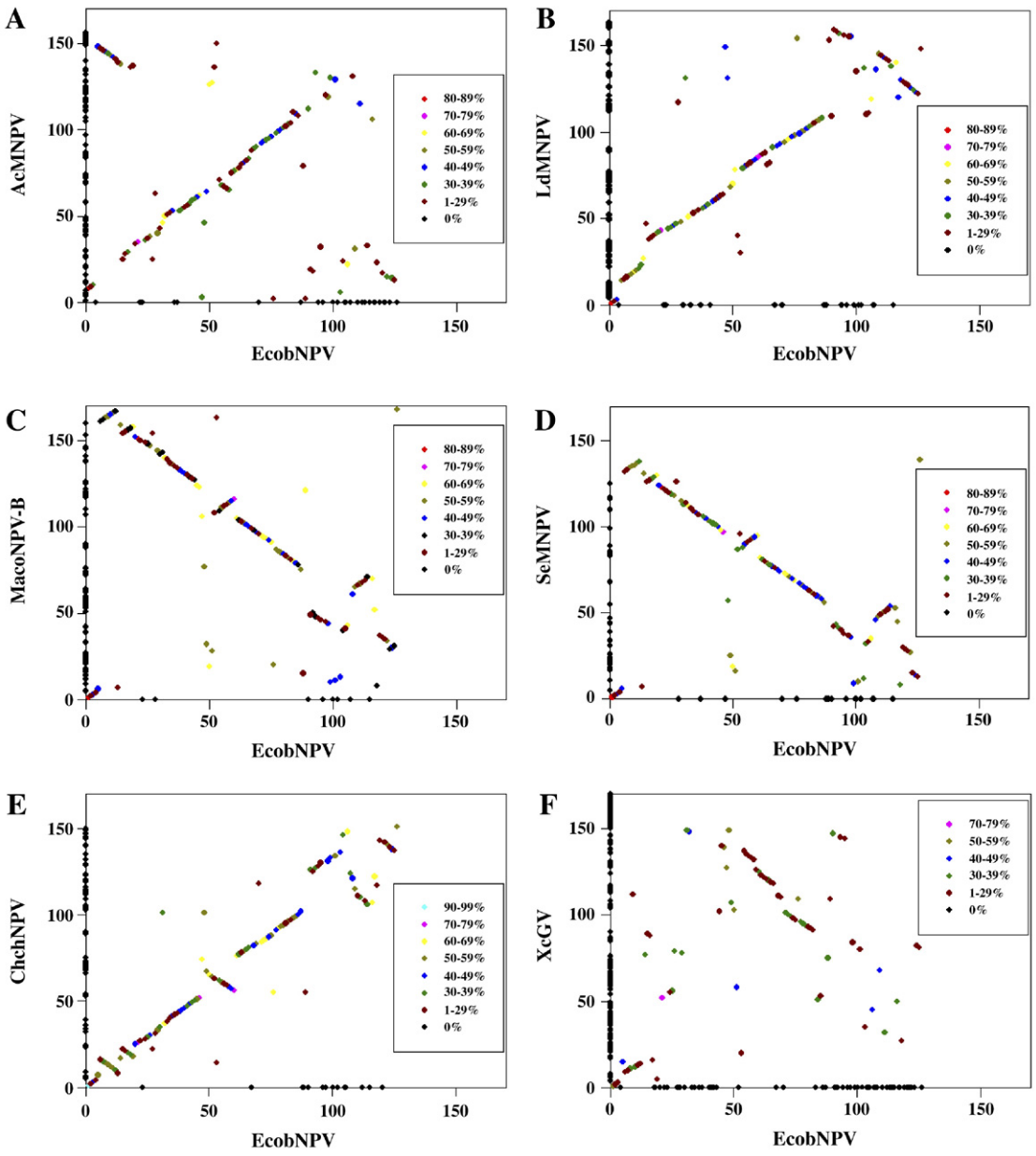

Fig. 2. Gene parity plot comparison of EcobNPV with AcMNPV (A), LdMNPV (B), MacoNPV-B (C), SeMNPV (D), ChchNPV (E) and XcGV (F). Homologous genes are plotted based on their relative location in the genomes using the method of Hu et al. (1998) by the sigma Plot 10.0 software. Those without homologues are aligned on the *x* or *y* axis, respectively. Each dot represents an ORF and level of identity is indicated by dot color.

and dispersed unevenly along the genome in AT rich intergenic regions, with the sizes of 1615 (hr1), 1433 (hr2) and 493 bp (hr3), respectively. The genes that flanked each hr were not conserved among baculoviruses. The combined size of the three hrs in EcobNPV was 3541 bp, or 2.7% of the genome. EcobNPV and *Cryptophlebia leucotreta* GV (CrleGV) have the smallest number of hrs of all baculoviruses, which range from three (CrleGV) to 17 (*Spodoptera litura* MNPV, SpltMNPV).

Sequence analysis confirmed that the three hrs comprised two apparent domains with perfect or near-perfect 24 bp palindromes and 37 bp flanking repeats at the head/end of one or both sides of the palindromes (Fig. 3). Hr1 had 19 palindromes and 22 flanking repeats, hr2 had 19 palindromes and 19 flanking repeats, and hr3 had five palindromes and five flanking repeats. The hrs shared significant intragenomic sequence homology, even though they had no homology with other sequenced baculovirus genomes. The palindromes had 92% to 100% sequence conservation, but the flanking repeat sequence was much more variable, with conservation ranging from 51% to 100%. Sequence alignment of the consensus regions of the two domains in each hr is shown in Fig. 3. Since hrs share higher similarity within a virus strain than between species, the amplification process appears to be tightly linked to functional conservation.

It is interesting to note that hr2 and hr3 are present at the "right" portion of the genome, which represents the highly variable regions. The hypervariable regions are also located roughly around the hr region in other baculoviruses, suggesting a possible

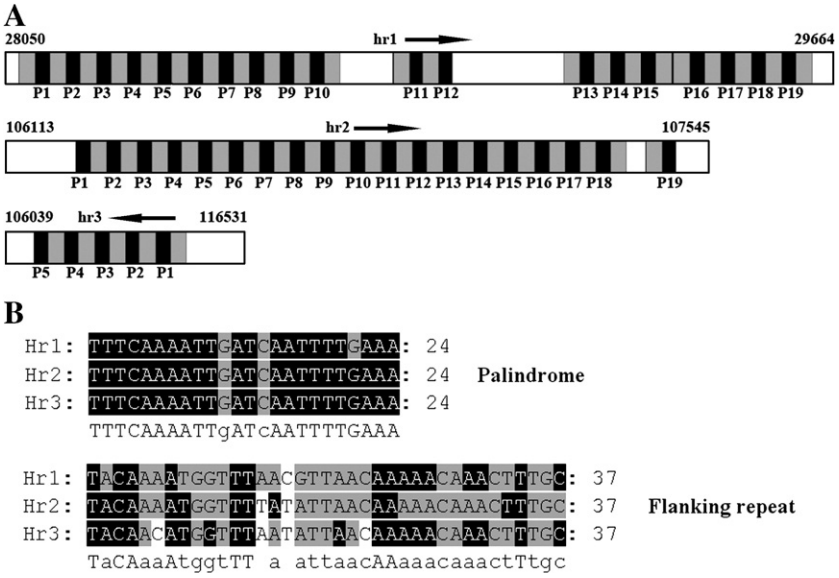

Fig. 3. EcobNPV homologous regions. (A) Organization of repeats in each hr, ■ represents the core palindrome, ▒ represents flanking repeat sequence, and □ represents the sequences between palindrome and flanking repeat. (B) Sequence alignment of core palindrome and flanking repeat sequence. A deduced consensus sequence of two regions from each hr was used for this alignment. Conserved sequences are indicated with different shading: black indicates 100% conservation, gray >60% conservation and no shading <60% conservation.

role for hrs in recombination events, consistent with the fact that hrs may be related to the recombination and rearrangement within or between baculovirus genomes since hrs have been identified as enhancer elements for gene expression and act as origins of DNA replication (Leisy and Rohrmann, 1993).

Baculovirus-repeated ORFs (bro genes)

The occurrence of *bro* genes is a common feature in baculoviruses (Kuzio et al., 1999). *Bro* genes are present in a number of baculoviruses, with 1–16 copies related to AcMNPV ORF 2. We identified two *bro* genes in EcobNPV (ORFs 76 and 89), which were named *bro-a* and *bro-b* based on their order of appearance in the genome. *Bro-a* shared between 15% and 65% identity with its homologues in AcMNPV, LdMNPV, MacoNPV-B, and ChchNPV, whereas *bro-b* shared 16% to 64% identity with its homologues in these baculoviruses (Table 1). No *bro* gene homologues exist in the SeMNPV genome (Ijkel et al., 1999), while seven *bro* genes are present in the XcGV genome (Hayakawa et al., 1999). EcobNPV *bro-a* was located within a conserved cluster of four genes in the baculovirus genome described by Herniou et al. (2003), which produced the two gene clusters, *helicase* and Ac96, and *38k* and *lef5*. *Bro-a* shared only 14% identity with *bro-b* in the EcobNPV genome. While *bro-a* belonged to the Group II bro genes, *bro-b* shared the highest similarity with Group I *bro* genes, based on LdMNPV BRO protein classification (Kuzio et al., 1999). A single-stranded DNA (ssDNA) binding motif found in BRO proteins (Zemskov et al., 2000) was conserved at the N-terminus of *bro-a* and *bro-b*. *Bro* genes are associated with regions of viral genome rearrangement (Li et al., 2002a, 2005; Willis et al., 2005). BmNPV BRO proteins have nucleic acid binding activity that influences host DNA replication/transcription (Zemskov et al., 2000), and recently, BRO proteins are shown to function as nucleocytoplasmic shuttling proteins that utilize the CRM1-mediated nuclear export pathway (Kang et al., 2006).

EcobNPV genes with two homologues

The three genes, *odv-e66*, *p26*, and *dbp* had two copies in the EcobNPV genome. *Odv-e66* encoded an occlusion-derived virion (ODV) protein that helps assemble the envelope of occlusion-derived baculoviruses. The two copies of *odv-e66*, Ecob31 (*odv-e66a*) and Ecob48 (*odv-e66b*), encoded 771 and 665 aa proteins, respectively, and their identity was 33%. There were also two copies of *odv-e66* in SeMNPV (Se57 and Se114), MacoNPV-A (ORFs 78 and 144), and MacoNPV-B (ORFs 77 and 143) (Ijkel et al., 1999; Li et al., 2002a, 2002b). Ecob31 and AcMNPV *odv-e66* shared 63% identity, which was higher than the other Group II NPVs, while Ecob48 shared more similarity with XcGV (56%) and Group II NPVs (50% with MacoNPV and ChchNPV) than AcMNPV (32%). The phylogenetic tree shown in Fig. 4 suggests that Ecob31 (*odv-e66a*) was acquired from a second source that was more closely related to Group I NPVs than the Ecob48 copy (*odv-e66b*).

Like other Group II NPVs, including SeMNPV, AgseNPV, MacoNPV-A and B, ChchNPV, and TnNPV, EcobNPV had two copies of *p26* (ORFs 18 and 52). Two copies of *p26* are also present in a single Group I NPV, CfMNPV (Cf7 and Cf128) (de Jong et al., 2005). Ecob18 (*p26a*) and 52 (*p26b*) shared only 18% and 15% identity with the AcMNPV *p26* gene, respectively. Ecob18 shared 37% identity with Se129 and only 16% identity with Se87, while Ecob52 shared 31% identity with Se87 and only 14% identity with Se129. The same differences were observed between EcobNPV and AgseNPV, MacoNPV, ChchNPV, and TnNPV. The genomic location of Ecob18 between Ac29 and the *p10* gene was also conserved for

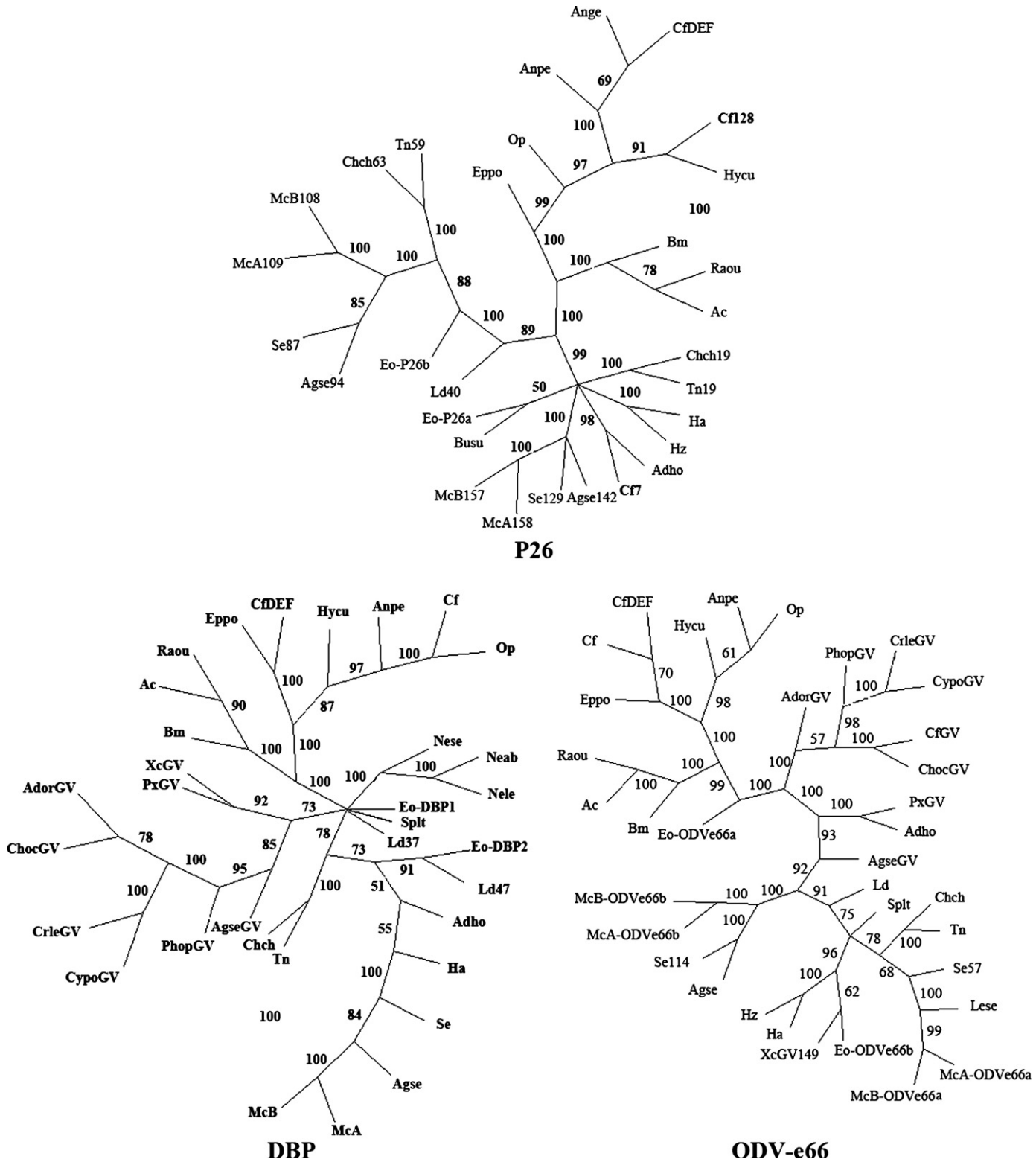

Fig. 4. Phylogenetic trees of baculovirus P26, DBP and ODV-e66. The bootstrap 50% majority-rule consensus trees were made with the Distance method (PAUP, Version 4.0) using multiple alignments of amino acid sequences. Statistical support for each node was evaluated by bootstrap analysis with 1000 replicates. The data sources and abbreviations of Baculovirus names are given in the text.

its homologues, Se129, Agse142, MacoB157, Chch19, and Tn19, while the location of homologues with higher identity to Ecob52 were more divergent in these baculoviruses. Thus it is likely that the two copies of *p26* were acquired from different sources, a result supported by the phylogenetic tree shown in Fig. 4.

EcobNPV is the second baculovirus identified to contain two copies of DNA-binding protein (*dbp*), the first being LdMNPV

(Kuzio et al., 1999). The identity of these two EcobNPV ORFs (Ecob15 and 27), however, was only 17%. Ecob15 (*dbp-1*) also shared low identity with its homologues in LdMNPV, Ld37 (12%) and Ld47 (16%). Ecob27 (*dbp-2*) shared low identity with Ld37 (16%) and slightly more similarity to Ld47 (33%). The phylogenetic tree of 33 *dbp* homologues indicated that Ecob27 was likely acquired from the same source as Ld47 during evolution (Fig. 4). DBP has been identified as an early gene product that can unwind double-stranded DNA, bind preferentially to single-stranded DNA, and colocalize with IE-1 and LEF-3 (Mikhailov et al., 1998; Okano et al., 1999).

In this report, we have analyzed the complete EcobNPV genome sequence and compared it with other baculovirus genomes. So far, three ORFs were found to be unique to EcobNPV and two copies of *bro*, *odv-e66*, *p26*, and *dbp* were identified. The two copies of each gene may have different sources. Based on the percent identity of gene homologues, the phylogeny of particular genes (Ma et al., 2006), and gene parity plots, EcobNPV was most closely related to Group II NPV. Although certain genes that were clustered in the genomes of EcobNPV and other Group II NPVs, the gene arrangement in EcobNPV was quite different from that in LdMNPV, MacoNPV-B, SeMNPV and ChchNPV. Comparison of EcobNPV with these Group II NPVs showed extensive genomic translocations in addition to cluster inversions. Thus, the genomic organization of EcobNPV was highly distinct. The sequence information presented here should provide further insight into baculovirus evolution and host specificity.

Materials and methods

Virus DNA isolation

The EcobNPV was originally isolated from the tea looper, *E. oblique*, in the Anhui Province of China. The 2nd instar *E. obliqua* individuals were infected with a low mortality dose of wild-type EcobNPV and virus DNA from each diseased single larva were extracted and examined with restriction endonucleases, respectively. A genotype (Strain A1) whose genomic DNA showed no submolar band when digested with restriction endonucleases was selected and used in this study. The viruses were propagated in fourth instar larvae and occlusion bodies were purified by sucrose-gradient centrifugation. Viral genomic DNA was extracted from purified occlusions as previously described (O'Reilly et al., 1992).

EcobNPV DNA cloning and sequencing

Purified genomic DNA was sheared by ultrasonication into fragments of 1 to 1.5 kb. The ends of each random fragment were repaired using the large fragment of T4 DNA polymerase (Klenow), according to the manufacturer's protocols. Viral DNA fragments were then cloned into pUC19. The ligation products were transformed into *Escherichia coli* JM109 competent cells (Promega). DNA templates for sequencing were prepared from over 2000 clones, and sequencing was performed using the ABI PRISM TM 3700 DNA Analyser and Big Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin Elmer). The combined sequence generated from these clones represented a tenfold genomic coverage. Sequencing of PCR products was used to fill the remaining gaps in the sequence.

Sequence analysis

Genomic DNA composition and ORFs were analyzed using the Wisconsin Genetic Computer Group program and Genetyx-win (Software Development Co. Ltd, Japan). ORFs encoding more than 50 aa were designated as putative genes. Relevant ORFs were checked for maximum alignment with known baculovirus gene homologues from GenBank. Homologous repeat regions were detected using the Search Direct, Inverted and Complementary Repeat programs of Genetyx-win (Parameters: minilength 20, maxilength 160, matching percentage of sites 75). DNA and protein comparisons with entries in the GenBank were performed using the BLAST and Genetyx program. Gene parity plot analysis was performed to compare gene order in the EcobNPV genome with that from other baculoviruses using SigmaPlot 10.0 software, as described by Hu et al. (1998). In this analysis, both shared and non-shared genes were included. Percent similarity/identity scores were obtained using DNASTAR's Clustal method with default conditions. Multiple amino acid sequence alignments were performed using ClustalX (1.81) software. Following alignments, the bootstrap 50% majority-rule consensus trees were made using the distance method on PAUP∗ 4.0 beta 10 version (Swofford, 2003). Statistical support for each node was evaluated by bootstrap analysis with 1000 replicates. The following genomes were used to identify gene conservation and/or phylogenetic analysis: *Adoxophyes orana* GV (AdorGV; GenBank accession no. NC_005038; Wormleaton et al., 2003), *C. fumiferana* GV (CfGV; GenBank accession no. AF517768), *C. leucotreta* GV (CrleGV; GenBank accession no. NC_005068; Lange and Jehle, 2003), *Cydia pomonella* GV (CypoGV; GenBank accession no. NC_002816; Luque et al., 2001), *Phthorimaea operculella* GV (PhopGV; GenBank accession no. NC_004062; Taha et al., 2000), *Plutella xylostella* GV (PxGV; GenBank accession no. NC_002593; Hashimoto et al., 2000), *X. c-nigrum* GV (XcGV; GenBank accession no. NC_002331; Hayakawa et al., 1999), *A. segetum* GV (AgseGV; GenBank accession no. NC_005839), *Choristoneura occidentalis* GV (ChocGV; GenBank accession no. NC_008168; Escasa et al., 2006), *A. honmai* NPV (AdhoNPV; GenBank accession no. NC_004690; Nakai et al., 2003), *A. segetum* NPV (AgseNPV; GenBank accession no. NC_007921; Jakubowska et al., 2006), *Antheraea pernyi* NPV (AnpeNPV; GenBank accession no. NC_008035), *A. californica* NPV (AcNPV; GenBank accession no. NC_001623; Ayres et al., 1994), *Bombyx mori* NPV (BmNPV; GenBank accession no. NC_001962; Gomi et al., 1999), *C. fumiferana* defective MNPV (CfDEFMNPV; GenBank accession no. NC_005137; Lauzon et al., 2005), *C. fumiferana* MNPV (CfMNPV; GenBank accession no. NC_004778; de Jong et al., 2005), *C. chalcites* NPV (ChchNPV; GenBank accession no. NC_007151; Van Oers et al., 2005), *E.*

obliqua SNPV (EcobNPV; GenBank accession no. DQ837165), *Epiphyas postvittana* NPV (EppoNPV; GenBank accession no. NC_003083; Hyink et al., 2002), *H. armigera* NPV-C1 (HaSNPV-C1; GenBank accession no. NC_003094; Zhang et al., 2005), *H. zea* SNPV (HzSNPV; GenBank accession no. NC_003049; Chen et al., 2002), *Hyphantria cunea* NPV (HycuNPV; GenBank accession no. NC_007767; Ikeda et al., 2006), *Leucania separata* NPV (LeseNPV; GenBank accession no. AB009613), *L. dispar* MNPV (Ld MNPV; GenBank accession no. NC_001973; Kuzio et al., 1999), *M. configurata* NPV-A (McNPV-A; GenBank accession no. NC_003529; Li et al., 2002b), *M. configurata* NPV-B (McNPV-B; GenBank accession no. NC_004117; Li et al., 2002a), *O. pseudotsugata* MNPV (OpMNPV; GenBank accession no. NC_001875; Ahrens et al., 1997), *R. ou* MNPV (RaouMNPV; GenBank accession no. NC_004323; Harrison and Bonning, 2003), *S. exigua* MNPV (SeMNPV; GenBank accession no. NC_002169; Ijkel et al., 1999), *S. litura* MNPV (SpltMNPV; GenBank accession no. NC_003102; Pang et al., 2001), *T. ni* SNPV (TnSNPV; GenBank accession no. NC_007383; Willis et al., 2004), *Neodiprion abietis* NPV (NeabNPV; GenBank accession no. NC_008252; Duffy et al., 2006), *Neodiprion lecontii* NPV (NeleNPV; GenBank accession no. NC_005906; Lauzon et al., 2004), *Neodiprion sertifer* NPV (NeseNPV; GenBank accession no. NC_005905; Garcia-Maruniak et al., 2004).

Acknowledgments

This project was supported by the grants from the Foundation for the Author of National Excellent Doctoral Dissertation of P. R. China (200055), the National Basic Research Program of China (2003CB114400) and the National Natural Science Foundation of China (30271008 and 30570074).

References

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

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

Chen, Y.G., Huang, M.D., 2001. Bioloical control progress of *Ectropis obliqua* Prout. Natural Enemies Insects 23, 181–184.

Chen, X., Hu, Z., Jehle, J.A., Zhang, Y., Vlak, J.M., 1997. Analysis of the ecdysteroid UDP-glucosyltransferase gene of *Heliothis armigera* single-nucleocapsid baculovirus. Virus Genes 15, 219–225.

Chen, X., IJkel, W.F., Tarchini, R., Sun, X., Sandbrink, H., Wang, H., Peters, S., Zuidema, D., Lankhorst, R.K., Vlak, J.M., Hu, Z., 2001. The sequence of the *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus genome. J. Gen. Virol. 82, 241–257.

Chen, X., Zhang, W.J., Wong, J., Chun, G., Lu, A., McCutchen, B.F., Presnail, J.K., Hermann, R., Dolan, M., Tingey, S., Hu, Z.H., Vlak, J.M., 2002. Comparative analysis of the complete genome sequences of *Helicoverpa zea* and *Helicoverpa armigera* single nucleocapsid nucleopolyhedroviruses. J. Gen. Virol. 83, 673–684.

de Jong, J.G., Lauzon, H.A., Dominy, C., Poloumienko, A., Carstens, E.B., Arif, B.M., Krell, P.J., 2005. Analysis of the *Choristoneura fumiferana* nucleopolyhedrovirus genome. J. Gen. Virol. 86, 929–943.

Duffy, S.P., Young, A.M., Morin, B., Lucarotti, C.J., Koop, B.F., Levin, D.B., 2006. Sequence analysis and organization of the *Neodiprion abietis* nucleopolyhedrovirus genome. J. Virol. 80, 6952–6963.

Escasa, S.R., Lauzon, H.A., Mathur, A.C., Krell, P.J., Arif, B.M., 2006. Sequence analysis of the *Choristoneura occidentalis* granulovirus genome. J. Gen. Virol. 87, 1917–1933.

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

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

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

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

Hayakawa, T., Ko, R., Okano, K., Seong, S.I., Goto, C., Maeda, S., 1999. Sequence analysis of the *Xestia c-nigrum* granulovirus genome. Virology 262, 277–297.

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

Hu, C., Zhu, J.Q., Ye, G.Y., Hong, J., 1994. *Ectropis oblique* Prout, a Serious Geometrid Pest of Tea Bush in East China. Shanghai Scientific and Technical Publishers, Shanghai, p. 217.

Hu, Z.H., Broer, R., Westerlaken, J., Martens, J.W., Jin, F., Jehle, J.A., Wang, L.M., Vlak, J.M., 1997. Characterization of the ecdysteroid UDP-glucosyltransferase gene of a single nucleocapsid nucleopolyhedrovirus of *Buzura suppressaria*. Virus Res. 47, 91–97.

Hu, Z.H., Arif, B.M., Jin, F., Martens, J.W., Chen, X.W., Sun, J.S., Zuidema, D., Goldbach, R.W., Vlak, J.M., 1998. Distinct gene arrangement in the *Buzura suppressaria* single-nucleocapsid nucleopolyhedrovirus genome. J. Gen. Virol. 79, 2841–2851.

Hyink, O., Dellow, R.A., Olsen, M.J., Caradoc-Davies, K.M., Drake, K., Herniou, E.A., Cory, J., O'Reilly, D.R., Ward, V.K., 2002. Whole genome analysis of the *Epiphyas postvittana* nucleopolyhedrovirus. J. Gen. Virol. 83, 957–971.

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

Ikeda, M., Shikata, M., Shirata, N., Chaeychomsri, S., Kobayashi, M., 2006. Gene organization and complete sequence of the *Hyphantria cunea* nucleopolyhedrovirus genome. J. Gen. Virol. 87, 2549–2562.

Jakubowska, A.K., Peters, S.A., Ziemnicka, J., Vlak, J.M., van Oers, M.M., 2006. Genome sequence of an enhancin gene-rich nucleopolyhedrovirus (NPV) from *Agrotis segetum*: collinearity with *Spodoptera exigua* multiple NPV. J. Gen. Virol. 87, 537–551.

Jehle, J.A., Blissard, G.W., Bonning, B.C., Cory, J.S., Herniou, E.A., Rohrmann, G.F., Theilmann, D.A., Thiem, S.M., Vlak, J.M., 2006. On the classification and nomenclature of baculoviruses: a proposal for revision. Arch. Virol. 151, 1257–1266.

Kang, W., Kurihara, M., Matsumoto, S., 2006. The BRO proteins of *Bombyx mori* nucleopolyhedrovirus are nucleocytoplasmic shuttling proteins that utilize the CRM1-mediated nuclear export pathway. Virology 350, 184–191.

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

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

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

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

Leisy, D.J., Rohrmann, G.F., 1993. Characterization of the replication of plasmids containing *hr* sequences in baculovirus-infected *Spodoptera frugiperda* cells. Virology 196, 722–730.

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

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

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

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

Ma, X.C., Xu, H.J., Tang, M.J., Xiao, Q., Hong, J., Zhang, C.X., 2006. Morphological, phylogenetic and biological characteristics of *Ectropis obliqua* single-nucleocapsid nucleopolyhedrovirus. J. Microbiol. 44, 77–82.

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

Nakai, M., Goto, C., Kang, W., Shikata, M., Luque, T., Kunimi, Y., 2003. Genome sequence and organization of a nucleopolyhedrovirus isolated from the smaller tea tortrix, *Adoxophyes honmai*. Virology 316, 171–183.

Okano, K., Mikhailov, V.S., Maeda, S., 1999. Colocalization of baculovirus IE-1 and two DNA-binding proteins, DBP and LEF-3, to viral replication factories. J. Virol. 73, 110–119.

O'Reilly, D.R., Miller, L.K., Luckow, V.A., 1992. Baculovirus Expression Vectors: A Laboratory Manual. W.H. Freeman, New York, USA.

Pang, Y., Yu, J.X., Wang, L.H., Hu, X., Bao, W., Li, G., Chen, C., Han, H., Hu, S., Yang, H., 2001. Sequence analysis of the *Spodoptera litura* multicapsid nucleopolyhedrovirus genome. Virology 287, 391–404.

Swofford, D.L., 2003. PAUP∗: Phylogenetic Analysis Using Parsimony (∗ and Other Methods). Sinaur Associates, Sunderland, MA. version 4.0b10.

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

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

Vlak, J.M., Smith, G.E., 1982. Orientation of the genome of *Autographa californica* nuclear polyhedrosis virus: a proposal. J. Virol. 41, 1118–1121.

Willis, L.G., Seipp, R., Stewart, T.M., Erlandson, M.A., Theilmann, D.A., 2005. Sequence analysis of the complete genome of *Trichoplusia ni* single nucleopolyhedrovirus and the identification of a baculoviral photolyase gene. Virology 344, 209–226.

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

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

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

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