Available online at www.sciencedirect.com

SCIENCE DIRECT

Virology 338 (2005) 209 – 226

VIROLOGY

www.elsevier.com/locate/yviro

Sequence analysis of the complete genome of *Trichoplusia ni* single nucleopolyhedrovirus and the identification of a baculoviral photolyase gene

Leslie G. Willis[a], Robyn Siepp, Taryn M. Stewart[b], Martin A. Erlandson[c,d], David A. Theilmann[a,b,*]

[a]*Pacific Agri-Food Research Centre, Agriculture and Agri-Food Canada, 4200 Highway 97, Summerland, BC, Canada V0H 1Z0*
[b]*Agricultural Sciences, University of British Columbia, AAFC-Saskatoon, SK, Canada S7N 0X2*
[c]*Saskatoon Research Centre, AAFC-Saskatoon, SK, Canada S7N 0X2*
[d]*Department of Applied Microbiology and Food Science, University of Saskatchewan, Canada*

Received 19 March 2005; returned to author for revision 6 April 2005; accepted 25 April 2005
Available online 13 June 2005

Abstract

The genome of the *Trichoplusia ni* single nucleopolyhedrovirus (TnSNPV), a group II NPV which infects the cabbage looper (*T. ni*), has been completely sequenced and analyzed. The TnSNPV DNA genome consists of 134,394 bp and has an overall G + C content of 39%. Gene analysis predicted 144 open reading frames (ORFs) of 150 nucleotides or greater that showed minimal overlap. Comparisons with previously sequenced baculoviruses indicate that 119 TnSNPV ORFs were homologues of previously reported viral gene sequences. Ninety-four TnSNPV ORFs returned an *Autographa californica* multiple NPV (AcMNPV) homologue while 25 ORFs returned poor or no sequence matches with the current databases. A putative photolyase gene was also identified that had highest amino acid identity to the photolyase genes of *Chrysodeixis chalcites* NPV (ChchNPV) (47%) and *Danio rerio* (zebrafish) (40%). In addition unlike all other baculoviruses no obvious homologous repeat (*hr*) sequences were identified. Comparison of the TnSNPV and AcMNPV genomes provides a unique opportunity to examine two baculoviruses that are highly virulent for a common insect host (*T. ni*) yet belong to diverse baculovirus taxonomic groups and possess distinct biological features. In vitro fusion assays demonstrated that the TnSNPV F protein induces membrane fusion and syncytia formation and were compared to syncytia formed by AcMNPV GP64.

Keywords: *Trichoplusia ni* nucleopolyhedrovirus genome; TnSNPV; Baculovirus; Photolyase

Introduction

The cabbage looper (*Trichoplusia ni*) (Lepidoptera: Noctuidae) is becoming a serious pest of the greenhouse industry in the Fraser Valley of British Columbia, Canada, due to resistance to *Bacillus thuringiensis*-based bioinsecticides (Janmaat and Myers, 2003). *T. ni* single nucleopolyhedrovirus (TnSNPV) has been used on a limited trial basis for the control of *T. ni* (Jaques, 1974) and is a potential alternative to *B. thuringiensis*. Baculoviruses with their high rate of infection, high efficacy in control of target species and low impact on non-target organisms are attractive alternative to traditional chemical pesticides.

The Baculoviridae are a diverse family of rod-shaped, occluded viruses that have circular dsDNA genomes ranging in size from 80 to 180 kb. They are divided into two genera: the *Nucleopolyhedrovirus* (NPVs) and the *Granulovirus* (GVs). NPVs typically produce large occlusion bodies (OBs) containing numerous virions while GVs produce smaller granular OBs containing

* Corresponding author. Pacific Agri-Food Research Centre, Agriculture and Agri-Food Canada, 4200 Highway 97, Summerland, BC, Canada V0H 1Z0. Fax: +1 250 494 0755.
E-mail address: TheilmannD@agr.gc.ca (D.A. Theilmann).

0042-6822/$ - see front matter.
doi:10.1016/j.virol.2005.04.041

single virions. NPVs have been further subdivided into two distinct groups (I and II) based on molecular phylogenetic characteristics (Zanotto et al., 1993). Two naturally occurring, but biologically distinct baculoviruses have been identified that infect and kill early instar *T. ni* larvae, TnSNPV and *Autographa californica* multiple nucleopolyhedrovirus (AcMNPV). AcMNPV, the archetype baculovirus, is a group I NPV that exhibits a broad host range across numerous lepidopteran species including *T. ni*. Conversely, TnSNPV exhibits a narrow host range infecting *T. ni* exclusively and is a group II NPV. These two viruses provide a unique opportunity to develop a model system that will identify the shared and unique molecular features found between these two distinct viral systems.

In this study the complete genome of TnSNPV has been sequenced and analyzed to provide a framework for determining the molecular basis for the biological differences between AcMNPV and TnSNPV. The TnSNPV genome comprises 134,394 bp and codes for 144 putative genes having open reading frames of 150 nucleotides or longer. Comparative analysis of these two viruses has shown that they contain many genes that are evolutionarily related, as well as genes such as the F protein, that may play a role in their observed biological differences. Bioinformatic analyses also identified a putative type II cyclobutane pyrimidine dimer-photolyase in the TnsNPV genome, which was characterized by sequence alignment and phylogenetic analysis.

Results

Nucleotide sequence analysis of the TnSNPV genome

A linear map of the 144 predicted open reading frames (ORFs) for the complete TnSNPV genome is illustrated in Fig. 1. Arrows indicate open reading frames and the direction of transcription (left to right clockwise, right to left counterclockwise). Previous estimates for the molecular mass of the TnSNPV genome ranged from 115.5 to 119.2 kbp (Davis and Wood, 1996). Sequence data obtained from the contiguous assembly of 2233 sequences produced a genome size of 134,394 bp (accession no. DQ017380), slightly larger than that of AcMNPV (133,894 bp). The overall G + C content of the entire genome was 39.0%, similar to the content recorded for AcMNPV (41%) (Ayres et al., 1994). Initial ORF prediction carried out using Genemark (Borodovsky and McIninch, 1993) identified 303 putative coding regions. ORFs were defined as methionine-initiated open reading frames encoding putative proteins of more than 50 amino acids with minimal overlap. Homology searches with these putative ORFs were performed using a standalone implementation of BLAST using the tBlastx (translated query vs. translated database), Blastn (nucleotide vs. nucleotide) and Blastp (protein vs. protein) algorithms. A total of 144 putative ORFs, including the 29 core genes found in all baculoviruses (Lauzon et al., 2004) and the 62 genes common among members with lepidopteran hosts (Herniou et al., 2003), were identified.

According to convention, the adenine residue of the start codon of the polyhedrin ORF was designated as the zero point in the genome. Sixty-nine ORFs were found to code in a clockwise (+) orientation with respect to the transcriptional orientation of the polyhedrin gene (ORF 1) and 75 ORFs were found to code in the counter clockwise orientation. Appreciable overlap occurred between 25 ORFs. Table 1 outlines the characteristics of the 144 putative ORFs from TnSNPV. Identity and similarity scores with AcMNPV and non-AcMNPV homologues uncovered by Blast were calculated with ClustalW using the Blosum62 scoring matrix.

Homologous repeat sequences

Variable numbers of homologous region (*hr*) sequences, composed of direct repeats which contain an imperfect palindrome core, have been identified in all of the currently sequenced baculovirus genomes (Kool et al., 1995). These regions range in number from a low of 4 in PlxyGV (Hashimoto et al., 2000) to a high of 13 in LdMNPV (Kuzio et al., 1999). Homologous regions have been shown to enhance RNA polymerase II-mediated transcription of baculovirus early promoters (Guarino and Summers, 1986; Theilmann and Stewart, 1992) and serve as origins of replication in transient replication assays (Ahrens et al., 1995; Kool et al., 1995; Pearson and Rohrmann, 1995; Pearson et al., 1992; Theilmann and Stewart, 1992). Unlike AcMNPV no obvious *hr* sequences were identified in TnSNPV; however, two regions that may function as *hrs* were identified as they contained short homologous repeat sequences with limited similarity (nt 43140–44236 and 125146–125520; Fig. 1). A number of additional sequences were identified in the TnSNPV genome using DotPlot analysis (DNASTAR) that contained a number of short repeat sequences (1233–1347, 2782–2893, 39821–39907, 45290–45342). These short repeat sequences were separated by variable lengths of intervening sequences. No homology between these TnSNPV sequences and other baculovirus *hr* sequences was observed.

Baculovirus-repeated ORFs (bro genes)

Bro genes show sequence similarity to AcMNPV ORF 2 (Ayres et al., 1994) and have been identified with varying frequency in a number of other baculoviruses including 16 in LdMNPV (Kuzio et al., 1999), 5 in BmNPV (Gomi et al., 1999; Kang et al., 1999), 3 in OpMNPV (Ahrens et al., 1997), 3 in HearNPV (Chen et al., 2001) and 7 in XecnGV (Hayakawa et al., 2000). Bro genes have also been found to be associated with regions of viral genome rearrangement

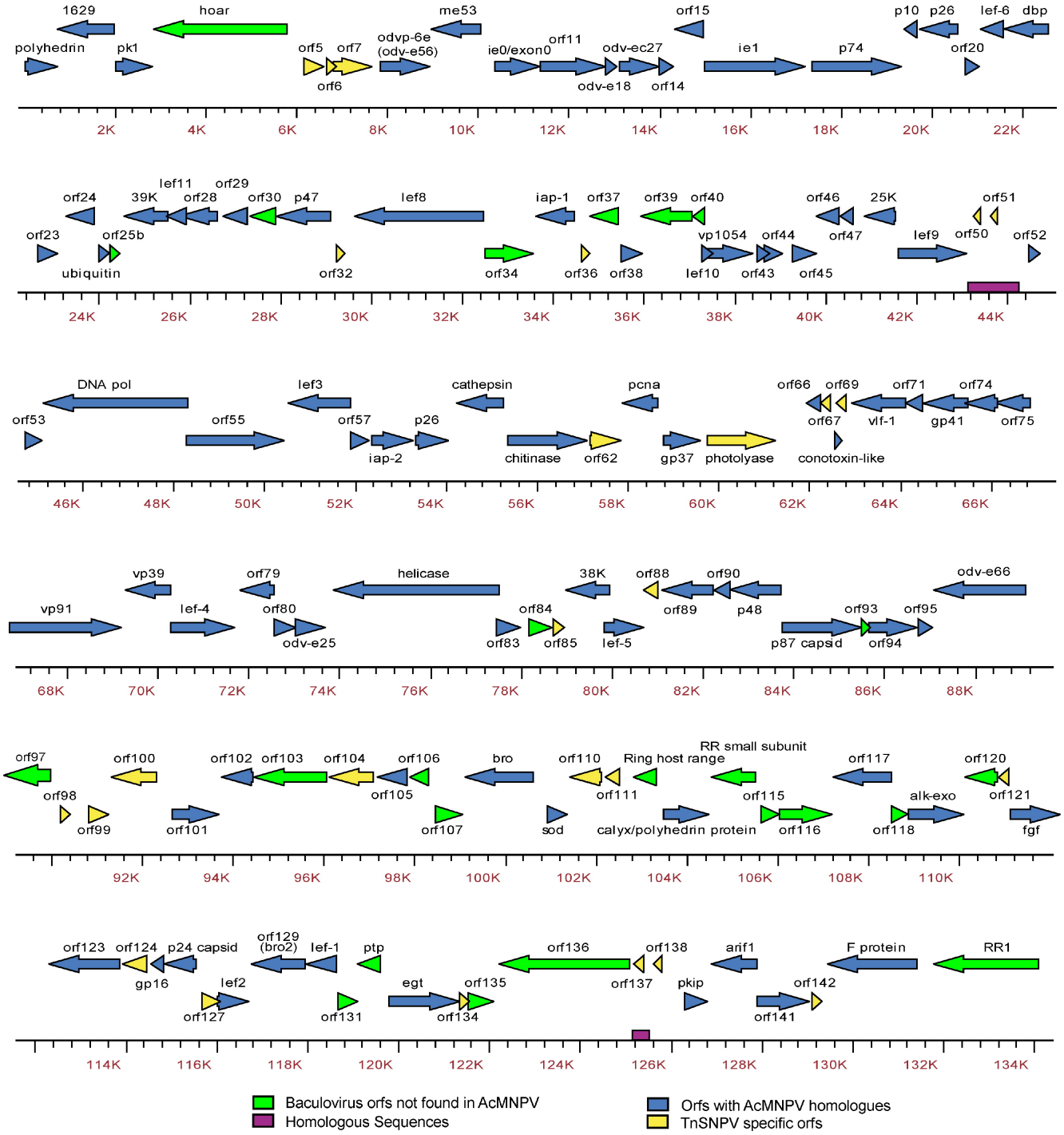

Fig. 1. Linear map of the 144 predicted ORFs for the complete TnSNPV genome. Arrows indicate open reading frames and the direction of transcription. ORFs were defined as methionine-initiated open reading frames encoding putative proteins of more than 50 amino acids with minimal overlap. According to convention, the adenine residue of the start codon of the polyhedrin gene was designated as the zero point in the genome. Sixty-nine ORFs were found to code in a clockwise (+) orientation with respect to the transcriptional orientation of the polyhedrin gene (ORF 1) and 75 ORFs were found to code in the counter clockwise orientation.

but the molecular basis for this is unknown (Li et al., 2002a, 2005). Two baculovirus-repeated ORFs (*bro* genes) were identified in TnSNPV (ORF 108 and 129) and were designated *bro1* to *bro2*, respectively, based on their order in the genome.

Genes associated with virion structure

The 15 structural protein genes conserved in all baculoviruses sequenced to date (Hayakawa et al., 2000) are also present in TnSNPV. These structural protein genes

Table 1
Summary of the main characteristics of the TnSNPV genome

1	2	3	4	5	6	7	8	9	10	11	12	13	14
Position	Strand	ORF	Promoter	aa	AcMNPV homologue	Ac ORF	% ID	% Sim	aa	Best non-AcMNPV blast hit	% ID	% Sim	aa
1…741	+	1	L	246	**polyhedrin**	8	91	96	245	ThorNPV polyhedrin	99	100	246
738…1979	–	2	None	413	1629	9	21	37	543	SeMNPV 1629	26	44	462
2003…2821	+	3	L	272	pk-1	10	40	59	272	SeMNPV pk-1	50	66	295
2868…5771	–	4	None	967	–					SeMNPV hoar	17	31	828
6156…6593	+	5	L, E	145	–					–			
6667…6879	+	6	None	70	–					–			
6809…7657	+	7	L, E	282	–					–			
7845…8921	+	8	E	358	**odvp-6e**	148	53	68	376	LdMNPV odvp-6e	56	72	356
8951…10063	–	9	L	370	me53	139	21	39	450	SeMNPV me53	38	55	390
10378…11358	+	10	L (E)	326	ie-0/exon0	141	25	42	261	MacoNPV-A ie0/exon0	35	52	234
11375…12802	+	11	L	275	**orf-142**	142	52	69	477	MacoNPV-B orf166	61	77	461
12812…13057	+	12	L	81	odv-e18	143	46	54	62	HearNPV odv-e18	71	75	81
13098…13988	+	13	L, E	298	**odv-ec27**	144	49	67	290	MacoNPV-B odv-ec27	65	81	278
13997…14278	+	14	L	93	orf-145	145	38	48	77	SeMNPV orf134	61	78	93
14321…14944	–	15	L	207	orf-146	146	32	54	201	MacoNPV-A orf163	43	64	191
14983…17184	+	16	E	733	ie-1	147	25	42	581	SA TnSNPV ie-1	76	85	739
										MacoNPV-A ie-1	30	48	607
17340…19325	+	17	L	661	**p74**	138	54	71	645	AdhoNPVp74	57	74	671
19375…19650	–	18	L	91	p10	137	27	44	94	SA TnSNPV p10	91	93	88
										SeMNPV p10	54	71	88
19716…20567	–	19	L	283	p26	136	30	44	240	SA TnSNPV p26	77	84	280
										MacoNPV-B p26	52	70	266
20717…21016	+	20	L	81	orf -29	29	26	49	71	SA TnSNPV Seorf128 like	68	72	71
										MacoNPV-A orf157	56	67	81
21061…21546	–	21	L	161	lef6	28	23	40	173	SA TnSNPV lef6	78	84	158
										MacoNPV-B lef6	36	52	141
21558…22559	–	22	L	333	dbp	25	26	45	316	HearNPV G4 dbp	35	55	323
22641…23069	+	23	L, E	142	orf-26	26	23	39	129	MacoNPV-A orf154	32	54	119
23257…23886	–	24	L	209	orf-34	34	29	48	215	HzSNPV orf27	44	54	255
23991…24224	+	25	L	77	**v-ubiquitin**	35	68	84	77	SeMNPV v-ubiquitin	50	55	129
24224…24460	+	25b	L	78	–					MacoNPV-B orf149	28	54	65
24548…25501	–	26	None	317	39K/pp31	36	28	46	275	SeMNPV 39K / pp31	37	54	317
25476…25934	–	27	None	152	lef11	37	25	37	112	MacoNPV-B lef-11	45	57	124
25769…26617	–	28	L	282	orf-38	38	43	55	216	MacoNPV-A orf147	58	67	230
26727…27248	–	29	L	173	orf-63	63	23	48	155	BmNPV orf51	27	49	155
27313…27894	–	30	L	193	–					SeMNPV orf117	23	43	191
27903…29105	–	31	None	400	**p47**	40	52	69	401	SeMNPV p47	62	78	400
29221…29409	+	32	L	62	–					–			
29637…32474	–	33	None	945	**lef8**	50	58	72	876	LdMNPV lef8	61	74	874
32501…33556	+	34	E	351	–					MacoNPV-A J Domain-like protein	23	43	388
33633…34461	–	35	E	275	iap1	27	30	46	286	BmNPV iap	35	48	346
34613…34792	+	36	L	59	–					–			
34789…35442	–	37	None	217	–					LdMNPV orf53	20	32	300

35503...35958	+	38	L	151	orf-53	53	42	62	139	SeMNPV orf108	55	74	137
35924...37054	–	39	L	376	–					SeMNPV orf107	36	52	344
37086...37319	–	40	L	77	–					MacoNPV-B orf134	46	69	75
37279...37506	+	41	L	75	lef10	53.5	43	68	78	MacoNPV-A lef10	61	80	75
37367...38374	+	42	None	335	orf-54	54	41	58	365	MacoNPV-A vp1054	61	81	336
38497...38772	+	43	L	91	orf-55	55	35	54	73	MacoNPV-B orf131	42	54	69
38648...39037	+	44	L, E	129	orf-56	56	22	32	84	SeMNPV orf103	31	53	93
39273...39788	+	45	None	171	orf-57	57	41	61	161	MacoNPV-B orf130	43	67	159
39790...40296	–	46	L	168	orf-59	59	22	32	69	SeMNPV orf101	42	58	195
40317...40592	–	47	L	91	orf-60	60	41	61	87	SeMNPV orf100	56	73	89
40839...41525	–	48	L	228	**25K protein**	61	49	66	212	SeMNPV orf98	65	74	195
41600...43090	+	49	None	496	**lef9**	62	66	80	516	MacoNPV-B lef9	76	88	497
43259...43417	–	50	None	52	–					–			
43621...43776	–	51	None	51	–					–			
44452...44709	+	52	L	85	orf-76	76	41	76	84	MacoNPV-A orf117	79	93	85
44716...45105	+	53	L	129	orf-75	75	26	53	133	MacoNPV-B orf115	62	84	129
45127...48285	–	54	E	1052	DNA-pol	65	44	60	984	SeMNPV dnapol	62	75	1063
48284...50422	+	55	None	712	orf-66	66	23	43	808	SeMNPV orf92	32	53	704
50529...51899	–	56	None	456	lef3	67	21	43	385	SeMNPV lef3	32	50	422
51898...52296	+	57	None	132	22.3 kDa protein	68	31	45	192	SeMNPV orf90	58	77	133
52346...53245	+	58	L	299	iap-2	71	15	31	286	MacoNPV-A iap-2	45	58	253
53308...54027	+	59	None	239	p26	136	22	37	240	SeMNPVp26	40	58	250
54210...55244	–	60	L	344	**cathepsin**	127	53	71	323	MacoNPV-B cathepsin	59	75	337
55359...57098	+	61	L	579	**chitinase**	126	64	76	551	MacoNPV-A chitinase	65	78	562
57167...57832	+	62	None	221	–					–			
57867...58652	–	63	None	261	pcna	49	25	46	256	*Xenopus laevis* pcna	28	50	261
58769...59587	+	64	L	272	**gp37**	64	49	64	302	MacoNPV-B gp37	68	79	262
59735...61243	+	65	E	525	–					ChchNPV photolyase	37	54	516
61927...62220	–	66	None	97	orf-111	111	24	34	67	ChchNPV Acorf-111 homolog	56	60	76
62243...62455	–	67	None	70	–					–			
62534...62707	+	68	None	57	conotoxin-like	3	12	25	53	MacoNPV-B conotoxin-like	17	32	50
62577...62789	–	69	None	70						–			
62925...64103	–	70	L	392	**vlf-1**	77	66	84	379	AdhoMNPV vlf-1	71	84	389
64100...64477	–	71	L	125	orf-78	78	31	50	109	HezeSNPV orf74	43	65	110
64503...65474	–	72	L	323	gp41	80	43	58	354	SpliNPV gp41	61	74	337
65404...66132	–	74	None	242	orf-81	81	40	58	233	MacoNPV-B orf102	52	65	240
66026...66862	–	75	None	278	tlp	82	23	36	180	LdMNPV orf82	39	51	223
66732...69194	+	76	L, E	820	p95 capsid	83	41	59	847	MacoNPV-A vp91 capsid	50	66	809
69283...70290	–	77	L	335	vp39	89	40	56	347	SeMNPV vp39	59	75	326
70265...71680	+	78	None	471	lef-4	90	46	61	464	MacoNPV-B lef-4	59	76	454
71791...72546	–	79	None	251	**orf-92**	92	51	71	259	SeMNPV orf73	65	83	252
72548...73030	+	80	L, E	160	**orf-93**	93	50	64	161	MacoNPV-B orf94	76	86	161
73027...73689	+	81	L	220	odv-e25	94	43	66	228	MacoNPV-B odv-e25	78	90	216
73849...77502	–	82	L	1217	helicase	95	41	61	1221	LdMNPV helicase	49	68	1218
77459...77977	+	83	None	172	orf-96	96	45	65	173	MacoNPV-B orf91	65	84	172
78157...78672	+	84	None	171	–					CpGV orf119	32	52	162
78706...78945	+	85	None	79	–					–			
78973...79929	–	86	None	318	38K	98	40	62	320	MacoNPV-B orf87	59	74	300
79822...80694	+	87	None	290	lef-5	99	48	63	265	MacoNPV-B lef-5	62	78	273

(continued on next page)

Table 1 (*continued*)

1	2	3	4	5	6	7	8	9	10	11	12	13	14
Position	Strand	ORF	Promoter	aa	AcMNPV homologue	Ac ORF	% ID	% Sim	aa	Best non-AcMNPV blast hit	% ID	% Sim	aa
80688...80996	–	88	None	102	–	100	34	34	55	–			
81081...82211	–	89	L	376	orf-101	101	44	64	361	SeMNPV orf64	62	79	388
82234...82584	–	90	L	116	orf-102	102	25	53	122	SeMNPV orf63	47	69	106
82571...83707	–	91	L	378	p48	103	48	64	387	MacoNPV-A orf083	71	85	377
83744...85480	+	92	None	578	orf-104	104	21	37	691	SeMNPV orf61 p87capsid	31	52	556
85477...85662	+	93	None	61	–					HearNPV orf96	45	63	58
85646...86716	+	94	L, E	356	orf-109	109	48	64	390	SeMNPV orf59	64	83	356
86751...87035	+	95	L	94	orf-108	108	31	46	105	SpliNPV orf97	32	53	112
87071...89104	–	96	L	677	odv-e66	46	39	56	704	MacoNPV-B odv-e66a	61	77	672
89189...90097	–	97	L	302	–					SeMNPV orf56	54	67	283
90185...90406	+	98	None	73	–					–			
90819...91250	+	99	L	143	–					–			
91327...92313	–	100	None	328	–					–			
92645...93688	+	101	E	347	putative histidinol-phosphatase	33	15	26	182	SeMNPV orf54	34	58	364
93755...94435	–	102	L	226	orf-106	106	14	24	110	MacoNPV-B orf70	67	80	216
94478...96055	–	103	L	525	–					SeMNPV orf52	27	48	529
96116...97075	–	104	L	319	–					–			
97175...97825	–	105	L	216	orf-115	115	37	54	204	MacoNPV-B orf67	43	61	203
97900...98298	–	106	E	132	–					MacoNPV-B orf66	22	43	121
98449...99036	+	107	L	195	–					MacoNPV-A orf120	35	55	181
99125...100597	–	108	None	490	bro1	2	20	34	329	XecnNPV orf60 bro-a	57	75	484
100906...101361	+	109	L	151	**sod**	31	72	87	151	AdhoNPV sod	74	87	156
101416...102087	–	110	None	223	–					–			
102212...102499	–	111	E	95	–					–			
102811...103329	–	112	None	172	–					LSDV LD140 RING finger host range	15	30	240
103461...104468	+	113	L	335	polyhedrin calyx protein	131	27	40	252	SeMNPV polyhedrin calyx protein	60	73	335
104541...105497	–	114	None	318	–					MacoNPV-A ribonucleotide reductase small subunit	64	81	313
105618...106016	+	115	L, E	132	–					MacoNPV-A orf57	36	55	135

106039…107169	+	116	L, E	376	–					OpMNPV orf113	42	63	361
107205…108506	–	117	L	433	orf-18	18	26	43	433	MacoNPV-A orf56	42	60	389
108508…108870	+	118	L	120	–					MacoNPV-B orf59	36	56	112
108878…110089	+	119	None	403	alkaline exonuclease	133	38	52	419	HearNPV alkaline exonuclease	43	62	429
110104…110844	–	120	L	246	–					MacoNPV-A orf53	33	57	238
110877…111083	–	121	None	68	–					–			
111116…112216	+	122	E	366	fibroblast growth factor	32	16	32	181	MacoNPV-A fibroblast growth factor	26	48	373
112297…113853	–	123	L	518	orf-119	119	47	64	530	AdhoNPV orf117	50	69	531
113902…114459	–	124	None	185	–					–			
114531…114818	–	125	L	95	gp16	130	32	55	106	SeMNPV gp16	47	78	94
114829…115545	–	126	L	238	p24 capsid protein	129	39	54	198	SeMNPVp24 capsid	56	79	248
115674…116078	+	127	L	134	–					–			
116005…116682	+	128	None	225	lef-2	6	42	60	210	MacoNPV-A lef-2	45	65	211
116745…117932	–	129	L	395	bro2 (bro-like C-terminal domain)	13	24	43	327	MacoNPV-B orf31	39	60	349
117972…118616	–	130	None	214	lef-1	14	35	50	266	MacoNPV-B lef-1	53	68	215
118656…119099	+	131	L	147	–					SeMNPV orf15	24	43	154
119101…119595	–	132	None	164	–					MacoNPV-B ptp2	39	57	179
119766…121334	+	133	None	522	egt	15	47	65	506	SeMNPV egt	66	83	523
121336…121548	+	134	None	70	–					–			
121536…122081	+	135	None	181	–					SfMNPV gi:32307379	34	70	174
122214…125078	–	136	None	954	–					MacoNPV-A orf42	32	52	848
125161…125376	–	137	None	71	–					–			
125609…125806	–	138	L	65	–					–			
126301…126804	+	139	L	167	pkip	24	29	47	169	SeMNPV pkip	47	63	164
126898…127887	–	140	None	329	actin rearrangement inducing factor	21	21	36	319	MacoNPV-A arif-1	25	38	290
127882…129030	+	141	L	382	**orf-22**	22	59	75	382	SeMNPV orf35	59	73	413
129104…129301	+	142	None	65	–					–			
129428…131413	–	143	L, E	661	orf-23 (Fusion)	23	19	40	690	SeMNPV Fusion	46	68	665
131779…134097	–	144	None	772	–					*Ecotropis obliqua* ribonucleotide reductase 1	54	70	760

Columns 1–5 show the predicted ORF nucleotide coding positions, direction of coding clockwise (+) or counterclockwise (−), ORF number, predicted late (L) and/or early (E) promoters and the predicted number of amino acids for each protein. No identifiable promoter is indicated by "none". Columns 6–10 show the predicted AcMNPV homologue, the AcMNPV homologue ORF number, the Blosum62 percent identity and percent similarity scores of the TnSNPV ORF and its AcMNPV homologue and the number of amino acids in the AcMNPV homologue. Columns 11–14 show the best non-AcMNPV blast hit for the predicted TnSNPV ORF, the Blosum62 percent identity and percent similarity scores of the TnSNPV ORF and its non-AcMNPV homologue and the number of amino acids in the non-AcMNPV homologue. AcMNPV homologues that have higher than average homology to TnSNPV ORFs are in bold. For TnSNPV *ie1*, *p10*, *p26* and *lef6* the homology with the South African (SA) isolate of a TnSNPV as well as the next best blast hit is shown (columns 11–14).

include those coding for capsid-associated proteins *vp*39 (ORF 77), *vp*91 (ORF 76) and *odv-ec*27 (ORF 13); basic DNA binding protein *p*6.9 (ORF 22); ODV envelope proteins *odv-e*18 (ORF 12), *odv-e*25 (ORF 81), *odvp-6e* (ORF 8) and *odv-e66* (ORF 96); ODV-associated proteins tegument protein *gp*41 (ORF 73) and *p*74 (ORF 17); genes for proteins associated with the occlusion body, *polyhedrin* (ORF 1) and p10 (ORF 18); as well as *vp*1054 (ORF 42) and *pk*-1 (ORF 3). An Ld130 or fusion protein homologue was found at ORF 143. Interestingly, TnSNPV polyhedrin (ORF 1) was nearly 100% identical to the *polyhedrin* gene of an NPV isolated from *Thysanoplusia orichalcea* L. (Lepidoptera: Noctuidae) (ThorNPV) in Indonesia (Cheng and Carner, 2000). The two genes differed at amino acid 123 where a valine was present in TnSNPV and an isoleucine in ThorNPV. Comparison of TnSNPV REN profiles with published restriction fragment profile of ThorNPV showed that there is extensive variation between the profiles of the two genomes indicating they are not the same virus.

Regulation of gene expression

Ten genes were previously identified as being required for maximal late gene expression in AcMNPV (Lu and Miller, 1995). Nine of these genes were identified in the TnSNPV genome: *lef*-4 (ORF 78), *lef*-5 (ORF 87), *lef*-6 (ORF 21), *lef*-8 (ORF 33), *lef*-9 (ORF 49), *lef*-10 (ORF 41), *lef*-11 (ORF 27), *39k* (ORF 26) and *p47* (ORF 31). A *lef*-12 homologue was not identified. Similarly, no *lef*-12 homologue was identified from the analysis of the group II *Mamestra configurata* NPV (MacoNPV) genome (Li et al., 2002b).

Inhibitors of apoptosis

The baculovirus p35-like genes were the first baculovirus gene family shown to be involved in the inhibition of apoptosis (Clem, 1997). A second family of inhibitors of apoptosis (*iap*) has been found in all members of the Baculoviridae sequenced to date. Apoptotic inhibition has been rescued in AcMNPV *p35* deletion mutants with a variety of baculovirus *iap* homologues (Seshagiri and Miller, 1997). Two TnSNPV ORFs (35 and 58) show homology to *iap* genes. ORF 35 shows strong homology with numerous baculovirus homologues as well as AcMNPV *iap-1*. ORF 58 shows strong homology with AcMNPV *iap-2*. No *p35* or AcMNPV *iap-4* homologues were identified in TnSNPV.

Auxiliary genes

Auxiliary genes are defined as being non-essential for viral replication, but they provide a selective advantage to the virus in terms of its epizoological relationships with insect hosts. Baculovirus genomes can often contain a large variety of auxiliary genes. These genes include *proliferating cell nuclear antigen* (*pcna*), *protein tyrosine phosphatase* (*ptp*), *ubiquitin*, *p10*, *superoxide dismutases* (*sod*), *conotoxin-like peptide* (*ctl*), *cathepsin*, *chitinase*, *ecdysteroid UDP-glucosyltransferase* (*egt*), *fibroblast growth factor* (*fgf*), *actin rearrangement-inducing factor-1* (*arif-1*), *viral enhancing factor* (*vef*) and *protein kinase* (*pk*) *1* and *2*. Homologues for all these genes except *vef* and *pk2* were found in the TnSNPV genome.

Transcriptional regulation

Putative early and late promoters were identified by searching 125 bp upstream of each ORF ATG start codon for either the TATA with a CAKT capsite (located approximately 25–40 bp downstream) or the DTAAG motif. Table 1 indicates L and or E for the presence of late and early promoters, respectively. A total of 166 putative early and late promoters were identified. In many cases two putative promoters could be identified for a single open reading frame. Eight ORFS (2, 6, 26, 27, 34, 67, 132 and 137) displayed a classic late or early promoter motif outside the designated 125 bp upstream region. Two additional ORFs (62 and 75) both contained late promoter motifs 3′ of the ATG start site. A promoter motif could not be identified 5′ of the putative ATG start site on the remaining 46 ORFs. It seems likely that additional, as yet unidentified, promoter sequences might exist within the TnSNPV genome, similar to other baculoviruses.

Homology between TnSNPV, AcMNPV and other baculoviruses

TnSNPV (narrow host range) and AcMNPV (broad host range) can co-exist simultaneously within *T. ni* and yet display the very different genetic properties of type II and type I NPVs, respectively. To identify genes that may be related to the biological differences and similarities of these viruses, it is helpful to identify both homologous and non-homologous gene sequences between the two viruses. Additionally, the levels of homology with other single and multiple nucleopolyhedrovirus genes may also help contribute to a better understanding of evolutionary relationships and host range.

The key properties of the TnSNPV genome are outlined in Table 1, which shows homology to AcMNPV genes and also the best hit to a non-AcMNPV gene. In order to assess their relative protein identity and similarity with the AcMNPV genome, each TnSNPV ORF was aligned pairwise with its AcMNPV homologue. The percent identity and similarities ranged from a high of 91% and 96% for AcMNPV polyhedrin to a low of 12% and 25% for the conotoxin-like homolog (Table 1).

Putative AcMNPV homologues were reported for 94 (65%) of the putative TnSNPV ORFs. The remaining 50 putative TnSNPV ORFs did not produce a significant hit

with the AcMNPV genome. Many of these ORFs did however return significant hits with other non-AcMNPV genomes.

Of particular interest among the non-baculovirus homologs identified was a putative class II cyclobutane pyrimidine dimer (CPD) photolyase gene (TnSNPV *orf65*) (Table 1, Fig. 1). This is the first photolyase gene reported to date in any completely sequenced baculovirus genome. Results from partial sequencing have also recently identified a baculovirus photolyase gene in the *Chrysodeixis chalcites* NPV (ChchNPV) (van Oers et al., 2004). Twenty-five ORFs (5, 6, 7, 30, 32, 36, 50, 51, 62, 67, 85, 98, 99, 100, 104, 110, 111, 121, 124, 127, 131, 134, 137, 138 and 142) reported very poor or no homology with any known protein in the current databases.

A gene parity plot was used to compare the gene order in TnSNPV and AcMNPV. Closely related viruses are expected to display a near colinear arrangement of genes, which decreases with increasing divergence (Hu et al., 1998). The gene parity plot (Fig. 2) revealed a considerable level of colinearity between AcMNPV and TnSNPV. Eight regions (TnSNPV ORF 1–3; 9–16; 24–28; 41–49; 54–58; 70–76; 77–83; and 86–92) had clusters of 3 or more ORFs forming a linear arrangement. It is also apparent that in general these clusters contained genes having the highest level of identity, suggesting that there has been selection pressure on amino sequence and gene order or grouping. However, the gene with the second highest identity, ORF 108 *superoxide dismutase* (SOD), was not found in a colinear cluster. SOD is a highly conserved baculovirus gene and is found in most baculovirus genomes sequenced to date (Herniou et al., 2001, 2003; Tomalski et al., 1991).

TnSNPV F protein

Envelope fusion proteins are essential for the spread of BV during baculovirus infection. AcMNPV and other group I NPVs contain a gene that encodes a GP64 homologue, while type II NPVs utilize envelope fusion proteins related to the LdMNPV LD130 or SE8 proteins, also known as Fusion (F) proteins. The two proteins are distinctly different in structure and sequence, yet both mediate membrane fusion and are required for virus attachment and cell entry. The TnSNPV genome was found to encode a typical group II NPV F protein (Tn143) and does not encode a homologue of GP64. As both TnSNPV and AcMNPV are highly virulent for *T. ni*, a comparison of the TnSNPV F protein and AcMNPV GP64 may shed light on the relationship between these two fusion proteins and their roles in baculovirus infection and host range.

An alignment with homologous F proteins Tn143, Se8 and the non-functional Ac23 is shown in Fig. 3. The predicted signal peptides, transmembrane domains and conserved cysteine residues for all three proteins are shown. Tn143 shares greater sequence identity (46%) and similarity (68%) with Se8 compared to Ac23 (19% and 40%, respectively) and this homology is also evident in the location of key cysteines in the three proteins. The F protein for AcMNPV also includes a 13 amino acid leader sequence not found in SeMNPV or TnSNPV, which may potentially be incorrect. A second methionine, found at position 14,

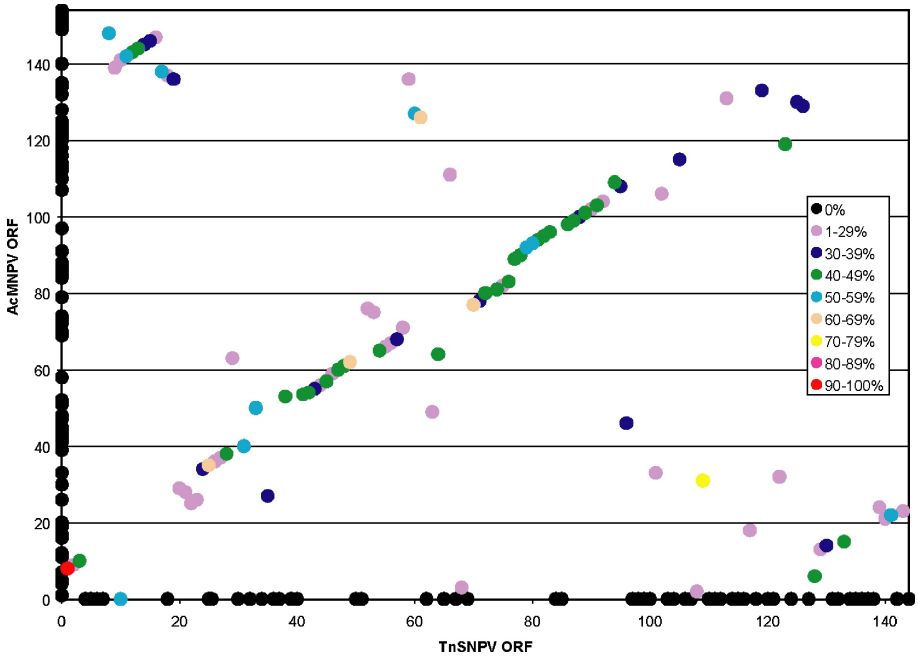

Fig. 2. Gene parity analysis of the genome organization of AcMNPV and TnSNPV. The genomes are represented on the horizontal (TnSNPV) and vertical (AcMNPV) axes with the position of each ORF relative to the axis plotted. Genes present only in one genome are plotted along the vertical or horizontal axes, respectively. Level of identity is indicated by dot color.

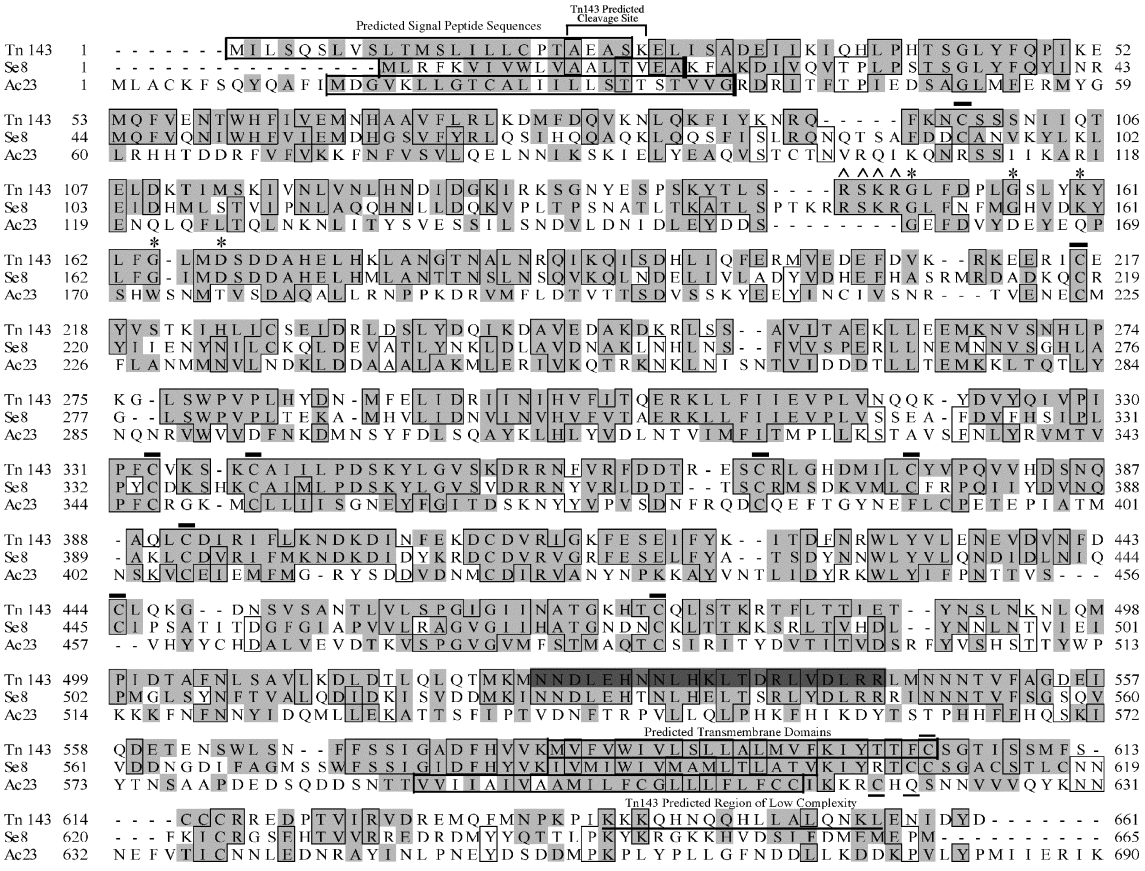

Fig. 3. ClustalW alignment of the TnSNPV (Tn143), SeMNPV (Se8) and AcMNPV (Ac23) fusion protein homologs. Shaded regions indicate regions of 66% or greater similarity. Narrow boxed regions indicate regions of complete identity. Heavy boxed regions indicate predicted signal peptides (N-terminal end) and transmembrane domains (C-terminal end). Conserved cysteine residues are overlined (Pearson et al., 2000). The inverted square bracket indicates the predicted cleavage site [AEA-SK] for the Tn143 signal peptide, ^^^^ predicted furin cleavage site consensus sequence; * furin conserved amino acid residues. The underlined Tn143 residues indicate a predicted region of low complexity as predicted using ClustalW (Thompson et al., 1994). The dark shaded box indicates the Tn143 predicted coiled coil region. Distinct domains were predicted for the putative TnSNPV F protein using the Simple Modular Architecture Research Tool (SMART) (Letunic et al., 2002; Nielsen et al., 1997; Schultz et al., 1998).

would agree more closely with the predicted start sites of both SeMNPV and TnSNPV. Removal of these additional 13 amino acids does not, however, significantly change the alignment of the F proteins (data not shown).

GP64 homologs are typically highly related (>74% amino acid sequence identity) and found only in group I NPVs. Conversely, F proteins are considerably less homologous (20–40% amino acid sequence identity) (Pearson and Rohrmann, 2002) and are typically found in both NPVs and GVs. Tn143 exhibits 46% identity with Se8, the most similar homologue. Tn143 also has low homology (18% identity) with the *env* gene of TED, an insect retrovirus known to be integrated into the *T. ni* genome (Friesen et al., 1986) and capable of transposition into the baculovirus genome (Friesen and Nissen, 1990; Ozers and Friesen, 1996).

Phylogenetic analysis was performed on the predicted amino acid residues of 22 putative baculovirus F proteins using the neighbor-joining method (Saitou and Nei, 1987) (Fig. 4). The resulting cladogram shows a clear separation between groups I and II NPVs and concurs with the positioning of TnSNPV as a group II baculovirus (Herniou et al., 2003; Zanotto et al., 1993). TnSNPV is contained within a clade with the other group II viruses MacoNPV-A,-B and SeMNPV. The close relationship between these three viruses was also mirrored by many other TnSNPV genes (Table 1).

To functionally compare TnSNPV F protein (Tn143) with AcMNPV GP64, we constructed several TnSNPV F protein expression vectors for use in transient transfection assays to compare syncytia formation (Blissard and Wenz, 1992). Two constructs were designed, ptntTnF and pie2TnF, that expressed Tn143 under the control of its native promoter or the highly expressed OpMNPV *ie2* early promoter (Hegedus et al., 1998), respectively. AcMNPV *gp64* expression was under control of the native AcMNPV promoter (Blissard and Wenz, 1992). Syncytia formation was monitored in the *T. ni* cell line Tn5B-1 (Hi-5) using previously described methods (Blissard and Wenz, 1992). Transfected Hi-5 cells were examined and scored for the formation of multi-nucleated syncytia by phase contrast microscopy after a short exposure to low pH media. Hi-5

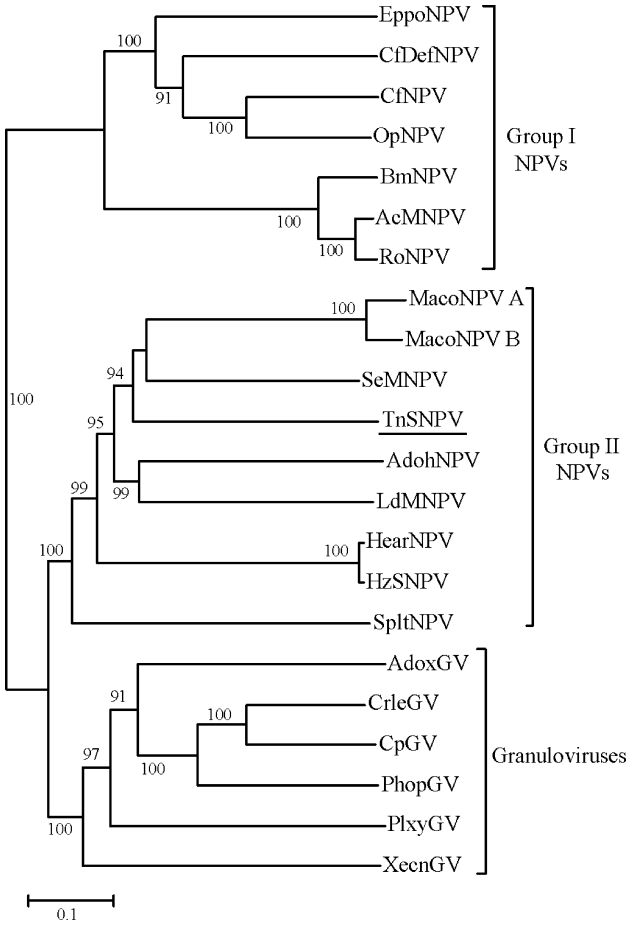

Fig. 4. Phylogenetic analysis using the predicted amino acid residues from 22 baculovirus F proteins. Sequences were aligned with a gap penalty of 10, an extend gap penalty of 0.05 and a delay of divergence of 40%. Phylogenetic analysis was performed using the neighbor-joining method (Saitou and Nei, 1987) with branch lengths and 1000 bootstrap replicates. The same topology was reflected in analysis of other highly conserved genes including structural proteins of the occlusion-derived virus envelope (data not shown). Scale represents 0.1 substitutions per site for the trees' branch length.

cells transfected with pAcgp64 developed, as expected, numerous large syncytia after exposure to reduced pH conditions (Fig. 5). In comparison, Hi-5 cells transfected with Tn143 under the native TnSNPV or the OpMNPV *ie2* promoter developed many smaller syncytia after exposure to reduced pH conditions (Fig. 5). No significant difference between the two Tn143 expression constructs was observed. Control constructs and mock-transfected cells exposed to acidic pH conditions exhibited no signs of cell fusion.

Photolyase

Blast homology searches revealed that *Tn65* encodes a putative class II cyclobutane pyrimidine dimer (CPD)-photolyase. This predicted photolyase protein shares very high homology with *photolyase* genes from more than 27 different organisms including: *Drosophila melanogaster*, Fowlpox virus, *Monodelphis domestica* (short-tailed opossum), Myxoma virus, *Potorous tridactylus* (long-nosed potoroo), *Oryzias latipes* (Japanese medaka), *Danio rerio* (zebrafish), *Carassius auratus* (goldfish) and the recently discovered *photolyase* of *C. chalcites* NPV. *Tn65 photolyase* encodes a 502-aa protein with a predicted molecular weight of 58.6 kDa and has been tentatively identified as an early gene product based upon promoter sequence locations. Two putative overlapping early promoters (containing TATA and CAKT sequences with a capsite approximately 26 and 28 bp downstream) are located 41 bp upstream from the putative ATG start position.

ClustalW pairwise alignment scores of the type II CPD photolyase were calculated and identity and similarity scores ranged from 23% to 45% and 37% to 57%, respectively. Among the proteins compared, the ChchNPV *photolyase* generated the highest homology scores (45% identity, 57% similarity) while the zebrafish (*D. rerio*) generated the second highest homology scores (37% identity, 54% similarity). A phylogenetic analysis using the predicted amino acid residues from 15 class II CPD photolyase genes and TnSNPV *orf65* is shown in Fig. 6. The alignment of the five most similar genes, with the location of the photolyase domain that binds the light harvesting cofactor and the FAD binding domains, is shown in Fig. 7.

Discussion

TnSNPV has been recognized as a useful biological insecticide for several decades and has been used successfully in numerous field trials (Jaques, 1974). In addition, the virus has been reported to have been isolated from locations around the world (del Rincon-Castro and Ibarra, 1997; Fielding and Davison, 1999). To enable the further development of this virus as a viral insecticide and to further understand the biology of this virus we have sequenced the entire TnSNPV genome.

The genome size of TnSNPV was determined to be 134,394 bp, larger than previously predicted by restriction fragment analyses (Davis and Wood, 1996; del Rincon-Castro and Ibarra, 1997). Sequence analysis predicted 144 open reading frames based on methionine-initiated ORFs encoding putative proteins of more than 50 amino acids and a minimal overlap with adjacent ORFs (Table 1). BLAST searches and comparisons with previously sequenced baculoviruses indicate that 119 ORFs were homologues of known baculovirus genes. The 144 ORFs identified were densely arranged with minimal intergenic distances. Their distribution with respect to transcriptional orientation was random and evenly divided between clockwise and counterclockwise directions (Fig. 1). Twenty-five ORFs showed no significant homology to any genes in the current databases.

Sequence data have been reported for small regions of a South-African isolate of an SNPV from *T. ni*, which

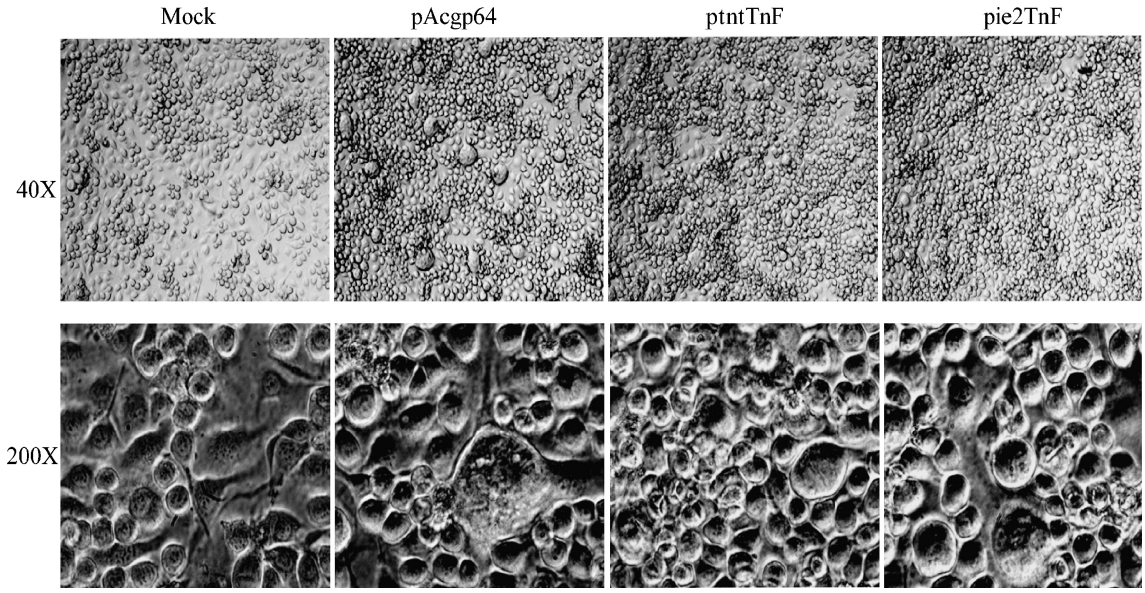

Fig. 5. Fusion assay analysis of TnSNPV F protein and comparison to AcMNPV GP64 in transfected *T. ni* Tn5b-1 cells. Cells were transfected with constructs expressing TnSNPV F and GP64 and exposed to low pH media. The ptntTnF and pie2TnF constructs express the predicted TnSNPV F fusion protein under the control of the native TnSNPV and OpMNPV *ie2* promoter, respectively. The *gp64* construct expresses GP64 under the control of the native AcMNPV promoter. Cells were incubated for 30 min in media (pH 5.0), washed twice with normal pH media and examined for syncytia formation by light microscopy at 40× and 200× magnifications. Mock control cells were transfected with constructs expressing GFP.

included coding sequences for *ie1* (ORF 16), *polyhedrin* (ORF 1), *p26* (ORF 19), *lef-6* (ORF 121), *SeMNPV orf128* homolog (ORF 120) and *p10* (ORF 18). Interestingly, comparisons of the predicted proteins of these genes with the TnSNPV sequences reported in this study show significant divergence (Table 1). The South African NPV isolated from *T. ni* would therefore appear to represent a closely related but separate viral species.

Comparison of the TnSNPV genes with the nr GenBank database showed that the gene with the highest homology

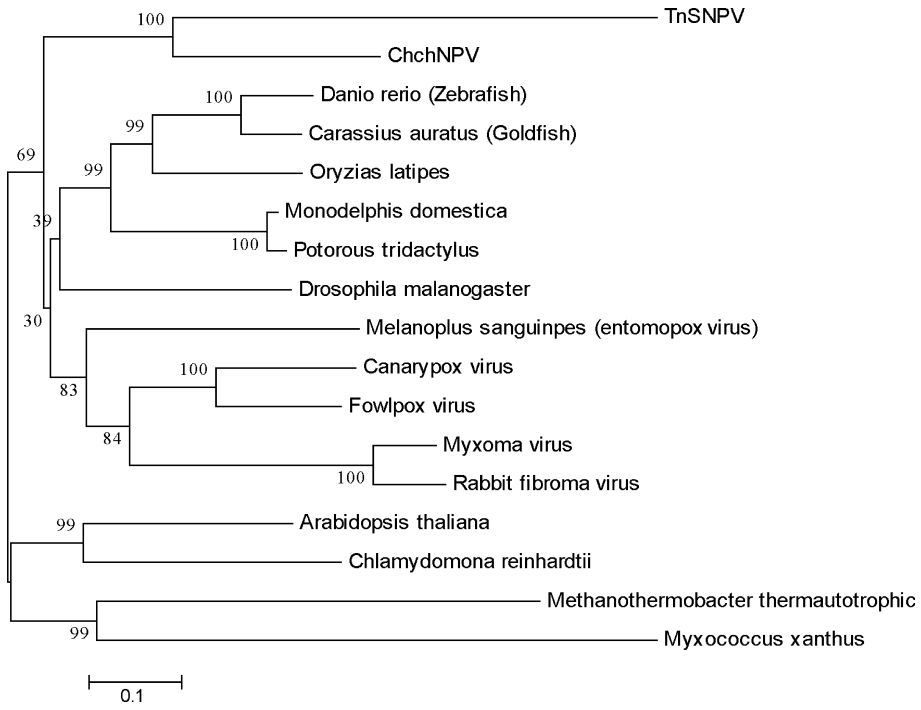

Fig. 6. Phylogenetic analysis using the predicted amino acid residues from 15 class II CPD photolyase genes and TnSNPV *orf65*. Sequences were aligned with ClustalW using a gap penalty of 10, an extend gap penalty of 0.05 and a delay of divergence of 40%. Phylogenetic analysis was performed using the neighbor-joining method (Saitou and Nei, 1987) and branch numbers represent bootstrap scores (%) of 1000 replicates. Scale represents 0.1 substitutions per site for the trees' branch length.

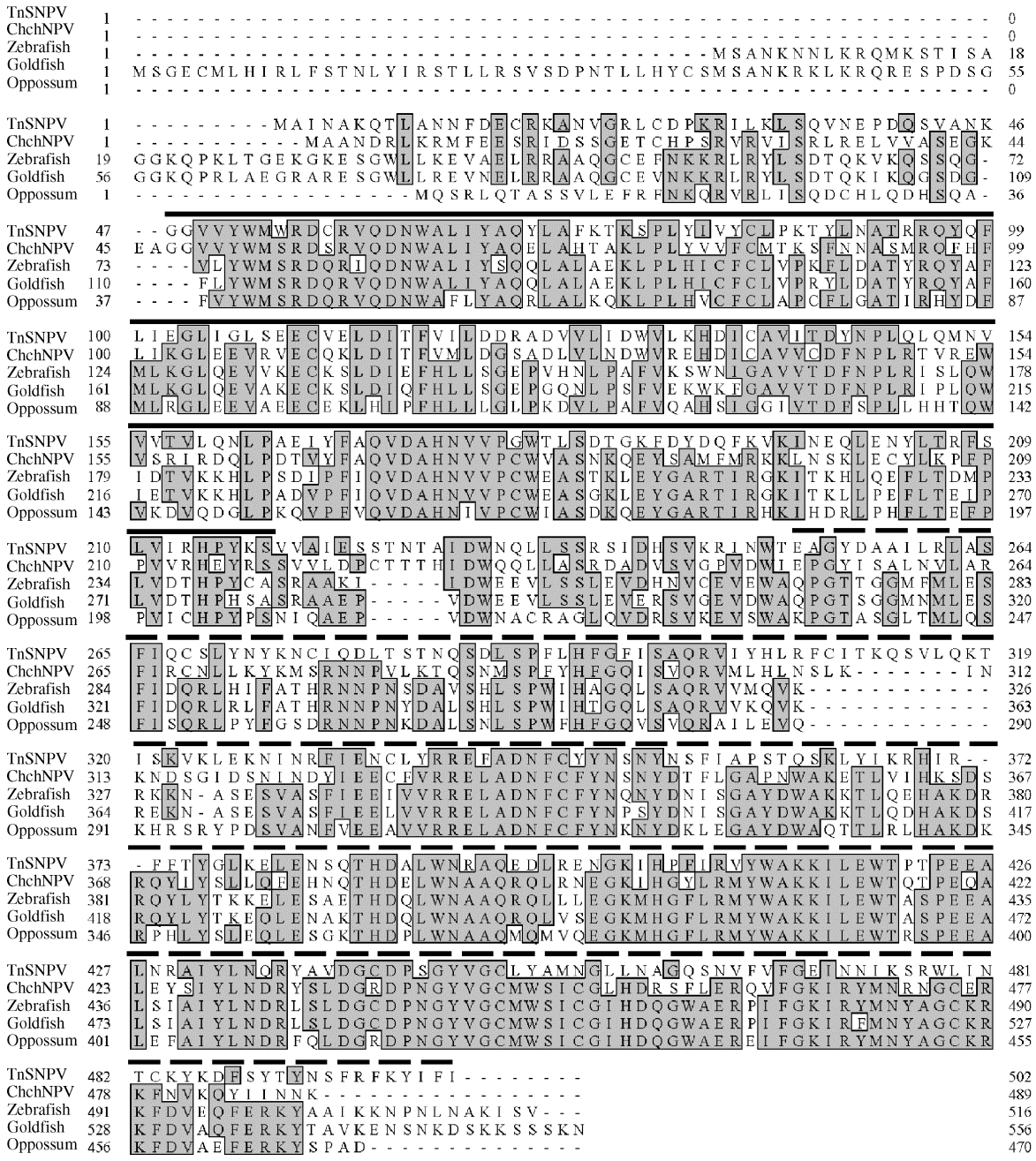

Fig. 7. ClustalW alignment of TnSNPV photolyase (ORF 65) with class II CPD photolyases from ChchNPV, zebrafish (*D. rerio*), goldfish (*C. auratus*) and short-tailed opossum (*M. domestica*). Shaded boxed regions indicate regions of 60% or greater identity. The conserved photolyase domain that binds the light harvesting cofactor (solid overline; pfam00875) and the FAD binding domains (dashed overline; pfam03441) were identified by comparison with the conserved domain database (Marchler-Bauer et al., 2003).

was *polyhedrin* from ThorNPV, which had only a single conservative amino acid difference. Interestingly, ThorNPV exhibits a tetrahedral polyhedra which is quite distinct from the more common TnSNPV polyhedral shape. This suggests that the single amino acid 123 may be responsible for the altered polyhedra shape, a supposition that could be tested by mutating the TnSNPV gene. If such a mutation did not alter polyhedra shape, it would suggest that other viral genes may be involved in determining polyhedral structure.

Comparison of TnSNPV and AcMNPV

TnSNPV and AcMNPV represent two separate lineages of the NPV genus (groups I and II) yet both are highly virulent for *T. ni*. We were interested in elucidating the similarities and differences between these two genomes in order to extend our understanding of the molecular differences in their pathology. These two viruses have diverged significantly. They possess a total of 94 genes in common with an average amino acid identity and similarity of 37%

and 53%, respectively (Fig. 2, Table 1). In contrast the average identity and similarity of the non-AcMNPV genes with the highest homology to each ORF of the TnSNPV genome was 49% and 65%. Not surprisingly, the predicted AcMNPV protein with the highest identity was polyhedrin (91%), and the second highest identity (72%) was superoxide dismutase (SOD). SOD is not conserved in all baculovirus genomes sequenced to date and has been classified as a potentially beneficial but non-essential gene (Herniou et al., 2003). It is therefore surprising that it ranks as one of the most highly conserved genes between TnSNPV and AcMNPV. Although no function has been attributed to this gene in any baculovirus, it is assumed that it is required for reducing oxidative stress in infected cells. The very high homology between the TnSNPV and the AcMNPV SODs may imply some type of host-specific SOD function.

Many of the genes in common between TnSNPV and AcMNPV could be involved in the adaptation to the host, *T. ni*. On average the identity between TnSNPV and AcMNPV proteins was quite low compared to the best non-AcMNPV BLAST hit. To identify common genes that are potentially required for infection of *T. ni*, we identified those TnSNPV and AcMNPV genes that had higher percent identity scores than the average best non-AcMNPV identity score (48%) (Table 1). Of the 18 genes examined that met this criteria, 10 are found in all baculoviruses (*odvp-6e/odv-e56*, *Acorf142*, *odv-ec27*, *p74*, *p47*, *lef8*, *lef9*, *vlf-1*, *Acorf-92*, *Acorf-22*) and four are found in all lepidopteran baculoviruses (*polyhedrin*, *ubiquitin*, *fp25k*, *Acorf-93*) (Herniou et al., 2003). The other four genes (*cathepsin-like protein*, *chitinase*, *gp37* and *sod*) are only found in a subset of viruses. The relatively high degree of conservation between these particular TnSNPV and AcMNPV genes suggests that these genes may be important for infection of the common host *T. ni*. Interestingly, four of these genes are known to be required for late gene transcription (*p47*, *lef-8*, *lef*-9 and *vlf-1*), six are structural genes (*polyhedrin*, *odvp-6e/odv-e56*, *fp25k*, *odv-ec27*, *p74*, *gp37*) and three are known or thought to provide non-essential auxiliary functions (*sod*, *cathepsin-like protein*, *chitinase*). Of the structural proteins, ODV-EC27 and P74 are known to be required for oral infectivity, which suggests that they interact with host proteins for infection. It is unknown if the core late RNA polymerase proteins have host-specific interactions. Cathepsin and chitinase require intimate interaction with the host in order to disrupt the exoskeleton, produce liquefaction and release occlusion bodies. Saville et al. (2002) recently showed that deletion of the AcMNPV cathepsin ER retention motif (KDEL) resulted in a virus that had increased virulence for *T. ni*, suggesting that this protein can impact host specificity. In addition to the genes described above it is also possible that non-homologous genes in the two genomes could be essential for the infection of *T. ni* for each respective virus as alternate infection pathways may be utilized.

Photolyase

Photoreactivation is the process of reversing the harmful effects of ultraviolet light (200–300 nm) upon exposure to blue light (350–450 nm) and subsequent repair by a photo reactivating enzyme called DNA photolyase (Sancar, 2003). Cyclobutane pyrimidine dimers (CPDs) are the most frequently observed form of DNA damage produced by shortwave ultraviolet (UV-C) light and account for approximately 70–80% of lesions in vivo. CPD photolyases can be separated into two subclasses (I and II) based on amino acid sequence similarity (Kanai et al., 1997; Yasui et al., 1994). The TnSNPV *photolyase* is a type II CPD photolyase and is similar to the recently reported ChchNPV *photolyase* (van Oers et al., 2004). Interestingly the most similar non-baculoviral proteins are the photolyases from zebrafish and goldfish (Figs. 6 and 7).

Several poxviruses are known to encode type II CPD photolyases and functional analysis of myxoma and shope fibroma proteins has demonstrated that they are catalytically active and effective in eliminating lethal photoproducts in wild-type viruses (Bennett et al., 2003; Srinivasan et al., 2001). Shope fibroma and myxoma virus CPD photolyase genes are highly conserved and share 85% amino acid identity over 445 residues (Bennett et al., 2003). The TnSNPV identity scores with other type II CPD photolyase genes (including ChchNPV) were relatively low, ranging from 23 to 37. Phylogenetic analysis did not place TnSNPV photolyase in a clade with any of the known *photolyases* and its origin remains unclear.

Baculoviruses are typically exposed to large amounts of ultraviolet light both inside and outside their insect hosts. Therefore, an adaptive mechanism for dealing with the UV-induced damage incurred during normal transmission would be beneficial to the overall fitness of the virus. The identification of *photolyase* in both TnSNPV and ChchNPV suggests that some baculoviruses have developed mechanisms to deal with UV light. However, only two photolyase genes have been identified out of the 30 baculovirus genomes that have been sequenced to date. Therefore, most baculoviruses either do not require *photolyase* or have developed alternate methods to repair UV damage.

Synergism between TnSNPV and AcMNPV during viral infection of *T. ni* has been reported (Lara-Reyna et al., 2003) and clearly an opportunity for gene transfer between these two viruses would exist. An enzyme such as TnSNPV *photolyase* should provide an increase in fitness by enabling genomic DNA damage repair and it would not require the formation of a functional protein complex such as that observed with helicase (Bideshi and Federici, 2000). However, AcMNPV does not contain a photolyase and it must be concluded that selection for the incorporation of this gene is not favored under normal circumstances. Occlusion-derived virus (ODV) of NPVs which initiate the primary midgut infection can contain virions that

incorporate single (S) or multiple (M) nucleocapsids per envelope. TnSNPV is an S phenotype and therefore fewer genomes per virion may have a lower tolerance for UV inactivation and require additional UV-repair provided by *photolyase*. AcMNPV may also use alternative genes such as PCNA (proliferating cell nuclear antigen) (O'Reilly et al., 1989) to combat the damaging effects of UV light. Interestingly, incorporation of an algal virus pyrimidine dimer-specific glycosylase (cv-PDG) into the AcMNPV genome enhanced UV protection for BV but not for ODV, which is the form that is most exposed to UV light (Petrik et al., 2003). However, this DNA repair enzyme decreased the dose and time required for lethality relative to wild-type virus in *Spodoptera frugiperda* larvae but not in *T. ni* larvae.

F protein

As TnSNPV is a group II NPV we were very interested in comparing the activity of the fusion protein (Tn143) with that of AcMNPV GP64 in cells derived from their common host *T. ni* (Fig. 5). The F protein of TnSNPV has numerous features common to viral fusion proteins including a signal peptide, a transmembrane domain and a peptide cleavage site. In addition, the protein contains numerous conserved cysteine residues and a putative coiled coil region (Figs. 3 and 4). Envelope proteins from retroviruses, paramyxoviruses, filoviruses and orthomyxoviruses have all been characterized by similar features (Eckert and Kim, 2001; Skehel and Wiley, 2000). Fusion assays showed that Tn143 is a functional F protein capable of generating numerous syncytia. Although these syncytia were significantly smaller than those formed by GP64 (Fig. 5), they were similar to those produced by other functional F proteins (Westenberg et al., 2004). AcMNPV GP64 null viruses can be rescued by F proteins from *Lymantria distria* MNPV and SeMNPV. However, virus production by viruses pseudotyped with the F protein from either SeMNPV or LdMNPV was approximately 2 log units lower than the AcMNPV-GP64 control viruses (Lung et al., 2002). TnSNPV and AcMNPV are both highly virulent for *T. ni* yet they use completely different classes of fusion proteins. It would be very interesting to determine if the TnSNPV F protein could rescue an AcMNPV GP64 null virus more efficiently than the SeMNPV and LdMNPV F proteins, which infect different hosts.

All group I NPVs, including AcMNPV, utilize GP64 for cell-to-cell transmission of infection (Monsma et al., 1996) but also contain a non-functional F protein homologue (AcMNPV Ac23; Fig. 3 and Table 1), which has been shown to be a pathogenicity factor (Lung et al., 2003). It has been suggested that GP64 is a recent addition to group I genomes and that it has displaced the envelope fusion function of the F protein (Pearson et al., 2000). The fluid nature of insect virus genomes (Herniou et al., 2003) and the local proximity of TnSNPV and AcMNPV in co-infections provides an opportunity for the horizontal transfer of *gp64* to TnSNPV. Transfer would seem to be highly likely if these genes were modular and conferred some fitness advantage. However, the transfer of *gp64* has not taken place and experiments involving simultaneous infection with AcMNPV and TnSNPV suggest that the two viruses are equally fit during intra-host competition and that viral recombination does not occur (Milks et al., 2001). Co-infections are known to occur in naturally infected insects collected in greenhouses in British Columbia (M. Erlandson, unpublished data), suggesting that both F protein and GP64-based infection strategies can be simultaneously successful.

A number of factors may be preventing TnSNPV from incorporating *gp64* into its genome. Broad host range multinucleocapsid viruses like AcMNPV may use *gp64* as a mechanism to quickly transmit budded virus to tracheal cells before virus replication is complete (Washburn et al., 2003). This early transmission may provide less of an advantage for an S phenotype virus such as TnSNPV, which has fewer nucleocapsids for early packaging. Sequencing the TnSNPV genome has revealed the considerable differences in genomic content between these two viruses. These differences, and the complex interplay of genetic elements associated with them, may simply be too great to permit the exchange of the key components such as the *fusion* gene or *gp64*.

Sequencing the genome of TnSNPV provides a complete genetic database that will permit us to compare and contrast the pathogenic mechanisms of groups I and II viruses for the host *T. ni.* In addition, it will permit the further development of TnSNPV as a biocontrol agent for *T. ni*, which is becoming a significant economic problem in greenhouses in Canada (Janmaat and Myers, 2003).

Materials and methods

Viral DNA isolation

TnSNPV was prepared by feeding fourth instar larvae on 16 mm^3 plugs of artificial diet inoculated with approximately 500 PIBs, adding fresh diet as required and allowing the infection to proceed for 5 days. Larvae were then collected and TnSNPV genomic DNA was extracted from purified polyhedra as previously described (Erlandson, 1990).

TnSNPV DNA shotgun cloning and sequencing

Cloning and initial sequence assembly was carried out by Greenomics (Wageningen, The Netherlands). Six-fold coverage was obtained using a shotgun sequencing strategy. Total TnSNPV genomic DNA was sheared by nebulization into fragments with an average size of 1200 bp and blunt end repaired using Pfu DNA polymerase (Stratagene) according to manufacturer's protocols. Viral DNA frag-

ments were then size fractionated using gel electrophoresis and ligated into the *Eco*RV restriction site of pBluescript SK (Stratagene). Ligation products were transformed into *E. coli* XL2-Blue competent cells (Stratagene). Recombinant clones were picked randomly from the resulting transformants and DNA templates for sequencing were further prepared using a Qiagen BioRobot 9600. Individual clones were sequenced using the ABI Prism Big Dye Terminator Cycle Sequencing Kit (Perkin Elmer) with FS AmpliTaq DNA polymerase (Perkin Elmer) and analyzed using an ABI 3700 DNA Analyzer.

Sequence analysis

Initial base calling and assembly was carried out using the PHRED base caller and PHRAP assembler software (Bonfield et al., 1995; Ewing and Green, 1998). The GAP4 interface and its features were then used for editing and sequence finishing. Consensus calculations with a quality cut-off value of 40 were performed from within GAP4 using a probabilistic consensus algorithm based on expected error rates outputted by PHRED. Sequencing of PCR products was used to fill any remaining gaps in the sequence. Sequence traces were manually edited to correct errors in automated base calls. Final contiguous sequence assembly was repeated using DNASTAR (Lasergene) and Sequencher (Gene Codes Corporation). After a completed sequence had been assembled, each base call within the assembly was manually inspected.

The TnSNPV genome is generally homologous with other baculovirus genes and open reading frames (ORFs) were easily identified using a combination of Genemark, Blast analyses and visual inspection. Gene homology searches were performed using the current downloadable GenBank database and standalone implementations of the Blastx, Blastn and Phi-blast algorithms (Altschul et al., 1990). Sequence alignments and percent similarity/identify scores were obtained using ClustalW (Thompson et al., 1994). Phylogenies were calculated using ClustalW alignments and the neighbor-joining method with DSGene (Accelrys) and Mega2 (Kumar et al., 2001).

Cultured cell lines

T. ni (Tn5b-1) cells were cultured in TC100 medium (Invitrogen) supplemented with 10% fetal bovine serum (FBS), gentamicin (50 μg/ml) (Invitrogen) at 27 °C in sterile 6-well plates (NUNC) at a density of 1×10^6 cells/well.

F protein expression vector construction

PCR was used to amplify the TnSNPV F protein (ORF 143) from wild-type TnSNPV genomic DNA. Three primers were synthesized. Primer 407 (5′-TGG TG**A AGC** ***TT****A CGA ACG CAA ACA ACA****A***) included a *Hin*dIII site (bold) upstream of the A in the ATG start site for ORF 143 (bold) (genome positions 131413–131428 italicized). Primer 409 (5′-A**GA ATT C***GT GAA ATG CCG CAG AAA ACT*) contains the terminal stop codon of F protein ORF and additional 3′ polyA signal regions (genome positions 129263–129282 italicized). This primer pair produced a 2184-bp product that was directionally cloned into p2ZOp2E (Hegedus et al., 1998) under the OpMNPV IE2 promoter and named pie2TnF. A third primer was constructed to amplify the TnSNPV protein F and its native late and early gene promoters. Primer 408 (5′-AAC **AAG CTT** *CGT GGG TTG TGA GAG C*) (genome positions 131791–131806 italicized) included a *Hin*dIII site (bold) was used with 409 (above) to generate a 2560-bp product that was directionally cloned into p2ZeoKS (Pfeifer et al., 1997) and named ptntTnF. The resulting constructs pie2TnF and ptntTnF along with a plasmid expressing the AcMNPV GP64 (pAcgp64) were used in transient transfections under various pH conditions to evaluate syncytia formation. A plasmid that constitutively expressed green fluorescent protein (GFP) was used to determine transfection efficiency.

Cell fusion assay

Two micrograms of each plasmid was transfected into monolayers of Tn5b-1 or Sf-9 cells in 6-well tissue culture plates using lipofectin. Lipofectin was produced as described (Campbell, 1995). Cells were washed once with 1.5 ml Graces' medium after 4 h of incubation at 27 °C and allowed to grow in TC100 medium for 48 h. At this time the pH was reduced in treatment wells by replacing the media with low pH Graces' (pH 5.0) supplemented with 10% FBS. Cells were incubated 30 min in pH 5.0 and then washed twice with normal (pH 6.1) Graces' media and allowed to grow 4 h in Graces' media supplemented with 10% FBS. Control wells were washed and treated with normal pH Graces' and incubated in Graces' supplemented with 10% FBS.

Quantification was performed by photographing wells and comparing the relative amount of membrane fusion present in control and treatment wells. GFP fluorescence from the control plasmid was used to estimate overall transfection efficiency.

Acknowledgments

This study was supported in part by the Biocontrol Network, which is funded by the Natural Sciences and Engineering Research Council of Canada.

References

Ahrens, C.H., Pearson, M.N., Rohrmann, G.F., 1995. Identification and characterization of a second putative origin of DNA replication in a baculovirus of *Orgyia pseudotsugata*. Virology 207 (2), 572–576.

Ahrens, C.H., Russell, R.L., 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 (2), 381–399.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215 (3), 403–410.

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 (2), 586–605.

Bennett, C.J., Webb, M., Willer, D.O., Evans, D.H., 2003. Genetic and phylogenetic characterization of the type II cyclobutane pyrimidine dimer photolyases encoded by Leporipoxviruses. Virology 315 (1), 10–19.

Bideshi, D.K., Federici, B.A., 2000. DNA-independent ATPase activity of the *Trichoplusia ni* granulovirus DNA helicase. J. Gen. Virol. 81 (Pt. 6), 1601–1604.

Blissard, G.W., Wenz, J.R., 1992. Baculovirus gp64 envelope glycoprotein is sufficient to mediate pH-dependent membrane fusion. J. Virol. 66 (11), 6829–6835.

Bonfield, J.K., Smith, K., Staden, R., 1995. A new DNA sequence assembly program. Nucleic Acids Res. 23 (24), 4992–4999.

Borodovsky, M., McIninch, J., 1993. Recognition of genes in DNA sequence with ambiguities. Biosystems 30 (1–3), 161–171.

Campbell, M.J., 1995. Lipofection reagents prepared by a simple ethanol injection technique. BioTechniques 18 (6), 1027–1032.

Chen, X., WF, I.J., 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 (1), 241–257.

Cheng, X.W., Carner, G.R., 2000. Characterization of a single-nucleocapsid nucleopolyhedrovirus of *Thysanoplusia orichalcea* L. (Lepidoptera: Noctuidae) from Indonesia. J. Invertebr. Pathol. 75 (4), 279–287.

Clem, R.J., 1997. Regulation of programmed cell death by baculoviruses. In: Miller, L.K. (Ed.), The Baculoviruses. Plenum, New York.

Davis, T.R., Wood, H.A., 1996. In vitro characterization of a *Trichoplusia ni* single nucleocapsid nuclear polyhedrosis virus. J. Gen. Virol. 77 (9), 2303–2310.

del Rincon-Castro, M.C., Ibarra, J.E., 1997. Genotypic divergence of three single nuclear polyhedrosis virus (SNPV) stranins from the cabbage looper, *Trichoplusia ni*. Biochem. Syst. Ecol. 25 (4), 287–295.

Eckert, D.M., Kim, P.S., 2001. Mechanisms of viral membrane fusion and its inhibition. Annu. Rev. Biochem. 70, 777–810.

Erlandson, M., 1990. Biological and biochemical comparison of *Mamestra configurata* and *Mamestra brassicae* nuclear polyhedrosis virus isolates pathogenic for the bertha armyworm, *Mamestra configurata* (Lepidoptera: Noctuidae). Invertebr. Pathol. 56, 47–56.

Ewing, B., Green, P., 1998. Base-calling of automated sequencer traces using phred: II. Error probabilities. Genome Res. 8 (3), 186–194.

Fielding, B.C., Davison, S., 1999. The characterization and phylogenetic relationship of the *Trichoplusia ni* single capsid nuclear polyhedrosis virus polyhedrin gene. Virus Genes 19 (1), 67–72.

Friesen, P.D., Nissen, M.S., 1990. Gene organization and transcription of TED, a lepidopteran retrotransposon integrated within the baculovirus genome. Mol. Cell. Biol. 10 (6), 3067–3077.

Friesen, P.D., Rice, W.C., Miller, D.W., Miller, L.K., 1986. Bidirectional transcription from a solo long terminal repeat of the retrotransposon TED: symmetrical RNA start sites. Mol. Cell. Biol. 6 (5), 1599–1607.

Gomi, S., Majima, K., Maeda, S., 1999. Sequence analysis of the genome of *Bombyx mori* nucleopolyhedrovirus. J. Gen. Virol. 80 (5), 1323–1337.

Guarino, L.A., Summers, M.D., 1986. Functional mapping of a *trans*-activating gene required for expression of a baculovirus delayed-early gene. J. Virol. 57 (2), 563–571.

Hashimoto, Y., Hayakawa, T., Ueno, Y., Fujita, T., Sano, Y., Matsumoto, T., 2000. Sequence analysis of the *Plutella xylostella* granulovirus genome. Virology 275 (2), 358–372.

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

Hegedus, D.D., Pfeifer, T.A., Hendry, J., Theilmann, D.A., Grigliatti, T.A., 1998. A series of broad host range shuttle vectors for constitutive and inducible expression of heterologous proteins in insect cell lines. Gene 207 (2), 241–249.

Herniou, E.A., Luque, T., Chen, X., Vlak, J.M., Winstanley, D., Cory, J.S., O'Reilly, D.R., 2001. Use of whole genome sequence data to infer baculovirus phylogeny. J. Virol. 75 (17), 8117–8126.

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, 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 (11), 2841–2851.

Janmaat, A.F., Myers, J., 2003. Rapid evolution and the cost of resistance to *Bacillus thuringiensis* in greenhouse populations of cabbage loopers, *Trichoplusia ni*. Proc. R. Soc. Lond. B Biol. Sci. 270 (1530), 2263–2270.

Jaques, R.P., 1974. Occurrence and accumulation of viruses of *Trichoplusia ni* in treated field plots. J. Invertebr. Pathol. 23 (2), 140–152.

Kanai, S., Kikuno, R., Toh, H., Ryo, H., Todo, T., 1997. Molecular evolution of the photolyase-blue-light photoreceptor family. J. Mol. Evol. 45 (5), 535–548.

Kang, W., Suzuki, M., Zemskov, E., Okano, K., Maeda, S., 1999. Characterization of baculovirus repeated open reading frames (bro) in *Bombyx mori* nucleopolyhedrovirus. J. Virol. 73 (12), 10339–10345.

Kool, M., Ahrens, C.H., Vlak, J.M., Rohrmann, G.F., 1995. Replication of baculovirus DNA. J. Gen. Virol. 76 (9), 2103–2118.

Kumar, S., Tamura, K., Jakobsen, I.B., Nei, M., 2001. MEGA2: molecular evolutionary genetics analysis software. Bioinformatics 17 (12), 1244–1245.

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 (1), 17–34.

Lara-Reyna, J., Del Rincon-Castro, M.C., Ibarra, J.E., 2003. Synergism between the nucleopolyhedroviruses of *Autographa californica* and *Trichoplusia ni*. Acta Virol. 47 (3), 189–194.

Lauzon, H.A., 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 (13), 7023–7035.

Letunic, I., Goodstadt, L., Dickens, N.J., Doerks, T., Schultz, J., Mott, R., Ciccarelli, F., Copley, R.R., Ponting, C.P., Bork, P., 2002. Recent improvements to the SMART domain-based sequence annotation resource. Nucleic Acids Res. 30 (1), 242–244.

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 (2), 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 (1), 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 (1), 91–105.

Lu, A., Miller, L.K., 1995. The roles of eighteen baculovirus late expression factor genes in transcription and DNA replication. J. Virol. 69 (2), 975–982.

Lung, O., Westenberg, M., Vlak, J.M., Zuidema, D., Blissard, G.W., 2002. Pseudotyping *Autographa californica* multicapsid nucleopolyhedrovirus (AcMNPV): F proteins from group II NPVs are functionally analogous to AcMNPV GP64. J. Virol. 76 (11), 5729–5736.

Lung, O.Y., Cruz-Alvarez, M., Blissard, G.W., 2003. Ac23, an envelope fusion protein homolog in the baculovirus *Autographa*

californica multicapsid nucleopolyhedrovirus, is a viral pathogenicity factor. J. Virol. 77 (1), 328–339.

Marchler-Bauer, A., Anderson, J.B., DeWeese-Scott, C., Fedorova, N.D., Geer, L.Y., He, S., Hurwitz, D.I., Jackson, J.D., Jacobs, A.R., Lanczycki, C.J., Liebert, C.A., Liu, C., Madej, T., Marchler, G.H., Mazumder, R., Nikolskaya, A.N., Panchenko, A.R., Rao, B.S., Shoemaker, B.A., Simonyan, V., Song, J.S., Thiessen, P.A., Vasudevan, S., Wang, Y., Yamashita, R.A., Yin, J.J., Bryant, S.H., 2003. CDD: a curated Entrez database of conserved domain alignments. Nucleic Acids Res. 31 (1), 383–387.

Milks, M.L., Leptich, M.K., Theilmann, D.A., 2001. Recombinant and wild-type nucleopolyhedroviruses are equally fit in mixed infections. Environ. Entomol. 30 (5), 972–980.

Monsma, S.A., Oomens, A.G., Blissard, G.W., 1996. The GP64 envelope fusion protein is an essential baculovirus protein required for cell-to-cell transmission of infection. J. Virol. 70 (7), 4607–4616.

Nielsen, H., Engelbrecht, J., Brunak, S., von Heijne, G., 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng. 10 (1), 1–6.

O'Reilly, D.R., Crawford, A.M., Miller, L.K., 1989. Viral proliferating cell nuclear antigen. Nature 337 (6208), 606.

Ozers, M.S., Friesen, P.D., 1996. The Env-like open reading frame of the baculovirus-integrated retrotransposon TED encodes a retrovirus-like envelope protein. Virology 226 (2), 252–259.

Pearson, M.N., Rohrmann, G.F., 1995. *Lymantria dispar* nuclear polyhedrosis virus homologous regions: characterization of their ability to function as replication origins. J. Virol. 69 (1), 213–221.

Pearson, M.N., Rohrmann, G.F., 2002. Transfer, incorporation, and substitution of envelope fusion proteins among members of the Baculoviridae, Orthomyxoviridae, and Metaviridae (insect retrovirus) families. J. Virol. 76 (11), 5301–5304.

Pearson, M., Bjornson, R., Pearson, G., Rohrmann, G., 1992. The *Autographa californica* baculovirus genome: evidence for multiple replication origins. Science 257 (5075), 1382–1384.

Pearson, M.N., Groten, C., Rohrmann, G.F., 2000. Identification of the *Lymantria dispar* nucleopolyhedrovirus envelope fusion protein provides evidence for a phylogenetic division of the Baculoviridae. J. Virol. 74 (13), 6126–6131.

Petrik, D.T., Iseli, A., Montelone, B.A., Van Etten, J.L., Clem, R.J., 2003. Improving baculovirus resistance to UV inactivation: increased virulence resulting from expression of a DNA repair enzyme. J. Invertebr. Pathol. 82 (1), 50–56.

Pfeifer, T.A., Hegedus, D.D., Grigliatti, T.A., Theilmann, D.A., 1997. Baculovirus immediate-early promoter-mediated expression of the Zeocin resistance gene for use as a dominant selectable marker in dipteran and lepidopteran insect cell lines. Gene 188 (2), 183–190.

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4 (4), 406–425.

Sancar, A., 2003. Structure and function of DNA photolyase and cryptochrome blue-light photoreceptors. Chem. Rev. 103 (6), 2203–2237.

Saville, G.P., Thomas, C.J., Possee, R.D., King, L.A., 2002. Partial redistribution of the *Autographa californica* nucleopolyhedrovirus chitinase in virus-infected cells accompanies mutation of the carboxy-terminal KDEL ER-retention motif. J. Gen. Virol. 83 (3), 685–694.

Schultz, J., Milpetz, F., Bork, P., Ponting, C.P., 1998. SMART, a simple modular architecture research tool: identification of signaling domains. Proc. Natl. Acad. Sci. U.S.A. 95 (11), 5857–5864.

Seshagiri, S., Miller, L.K., 1997. Baculovirus inhibitors of apoptosis (IAPs) block activation of Sf-caspase-1. Proc. Natl. Acad. Sci. U.S.A. 94 (25), 13606–13611.

Skehel, J.J., Wiley, D.C., 2000. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. Annu. Rev. Biochem. 69, 531–569.

Srinivasan, V., Schnitzlein, W.M., Tripathy, D.N., 2001. Fowlpox virus encodes a novel DNA repair enzyme, CPD-photolyase, that restores infectivity of UV light-damaged virus. J. Virol. 75 (4), 1681–1688.

Theilmann, D.A., Stewart, S., 1992. Tandemly repeated sequence at the 3′ end of the IE-2 gene of the baculovirus *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus is an enhancer element. Virology 187 (1), 97–106.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22 (22), 4673–4680.

Tomalski, M.D., Eldridge, R., Miller, L.K., 1991. A baculovirus homolog of a Cu/Zn superoxide dismutase gene. Virology 184 (1), 149–161.

van Oers, M.M., Herniou, E.A., Usmany, M., Messelink, G.J., Vlak, J.M., 2004. Identification and characterization of a DNA photolyase-containing baculovirus from *Chrysodeixis chalcites*. Virology 330 (2), 460–470.

Washburn, J.O., Trudeau, D., Wong, J.F., Volkman, L.E., 2003. Early pathogenesis of *Autographa californica* multiple nucleopolyhedrovirus and *Helicoverpa zea* single nucleopolyhedrovirus in *Heliothis virescens*: a comparison of the 'M' and 'S' strategies for establishing fatal infection. J. Gen. Virol. 84 (2), 343–351.

Westenberg, M., Veenman, F., Roode, E.C., Goldbach, R.W., Vlak, J.M., Zuidema, D., 2004. Functional analysis of the putative fusion domain of the baculovirus envelope fusion protein F. J. Virol. 78 (13), 6946–6954.

Yasui, A., Eker, A.P., Yasuhira, S., Yajima, H., Kobayashi, T., Takao, M., Oikawa, A., 1994. A new class of DNA photolyases present in various organisms including aplacental mammals. EMBO J. 13 (24), 6143–6151.

Zanotto, P.M., Kessing, B.D., Maruniak, J.E., 1993. Phylogenetic interrelationships among baculoviruses: evolutionary rates and host associations. J. Invertebr. Pathol. 62 (2), 147–164.