Characterization of the *Helicoverpa assulta* nucleopolyhedrovirus genome and sequence analysis of the polyhedrin gene region

Soo-Dong Woo*, Jae Young Choi[a], Yeon Ho Je[a] and Byung Rae Jin[b]

Department of Plant Medicine, College of Agriculture, Life and Environment Sciences, Chungbuk National University, Cheongju 361-763, Korea

[a]*School of Agricultural Biotechnology, College of Agriculture and Life Sciences, Seoul National University, Seoul 151-742, Korea*

[b]*College of Natural Resources and Life Science, Dong-A University, Busan 604-714, Korea*

**Corresponding author (Fax, 82-43-271-441; Email, sdwoo@cbnu.ac.kr)*

A local strain of *Helicoverpa assulta* nucleopolyhedrovirus (HasNPV) was isolated from infected *H. assulta* larvae in Korea. Restriction endonuclease fragment analysis, using 4 restriction enzymes, estimated that the total genome size of HasNPV is about 138 kb. A degenerate polymerase chain reaction (PCR) primer set for the polyhedrin gene successfully amplified the partial polyhedrin gene of HasNPV. The sequencing results showed that the about 430 bp PCR product was a fragment of the corresponding polyhedrin gene. Using HasNPV partial predicted polyhedrin to probe the Southern blots, we identified the location of the polyhedrin gene within the 6 kb *Eco*RI, 15 kb *Nco*I, 20 kb *Xho*I, 17 kb *Bgl*II and 3 kb *Cla*I fragments, respectively. The 3 kb *Cla*I fragment was cloned and the nucleotide sequences of the polyhedrin coding region and its flaking regions were determined. Nucleotide sequence analysis indicated the presence of an open reading frame of 735 nucleotides which could encode 245 amino acids with a predicted molecular mass of 29 kDa. The nucleotide sequences within the coding region of HasNPV polyhedrin shared 73.7% identity with the polyhedrin gene from *Autographa californica* NPV but were most closely related to *Helicoverpa* and *Heliothis* species NPVs with over 99% sequence identity.

[Woo S-D, Choi J Y, Je Y H and Jin B R 2006 Characterization of the *Helicoverpa assulta* nucleopolyhedrovirus genome and sequence analysis of the polyhedrin gene region; *J. Biosci.* **31** 329–338]

1. Introduction

The Baculoviridae are a diverse family of virus pathogens, which are infectious for arthropods, particularly insects of the order Lepidoptera. Baculovirus infections have been reported in over 600 insect species of the orders Hymenoptera, Diptera, Coleoptera, Neuroptera, Trichoptera, and Thysanura, as well as in the Crustaceae order Decapoda (Murphy *et al* 1995). The family is divided into two genera, nucleopolyhedrovirus (NPV) and granulovirus (GV). The NPV has been well investigated because of its potential for use as an insect pest control agent and a vector for the expression of various heterologous genes under the control of polyhedrin promoter (Smith *et al* 1983; Miller 1988; Maeda 1994; Woo *et al* 2000). The large number of review papers and books reflect the broad interest in baculoviruses (Maeda 1989; Adams and Bonami 1991; King and Possee 1992; O'Reilly *et al* 1992). The current interest in molecular biology of these viruses is fostered by their potential as modified virus pesticides with

Keywords. Baculovirus; genome; HasNPV; *Helicoverpa assulta*; polyhedrin gene

Abbreviations used: AcNPV, *Autographa californica* nucleopolyhedrovirus; HasNPV, *Helicoverpa assulta* nucleopolyhedrovirus; ORF, open reading frame; PCR, polymerase chain reaction; REN, restriction endonuclease.

increased toxicity (Inceoglu *et al* 2001) and as gene-therapy vectors in medical sciences (Kost and Condreay 2002; Kost *et al* 2005). The improvement of both applications requires a detailed knowledge of distinct baculovirus features and the extent of their diversity. In order to understand the evolution of baculoviruses and the molecular mechanism behind baculovirus infection and replication, the sequencing of baculovirus genomes has been undertaken by a number of research groups (Zhang *et al* 2005). Despite the large number of baculoviruses in nature, only a small fraction of these viruses has been studied in any detail. To date, the genomes of 26 NPVs have been sequenced completely. Knowledge on the genetic relationship of the majority of baculoviruses is still fragmentary. Reports of numerous insect species infected by baculoviruses were catalogued, but the characterization of *Helicoverpa assulta* NPV (HasNPV) has not been reported yet. The isolation of a NPV infecting *H. assulta* was only first reported by Jin *et al* (1995) in Korea. Since that time, however, there has been no information in molecular level for HasNPV. The oriental tobacco budworm, *H. assulta*, is distributed in Asia, Africa and Australia. The insect uses a number of *Solanacea* species as host plants and the larvae feed inside the fruits of hot pepper, or on tobacco or onion leaves (Hill 1983). Since larvae feed within the fruits of host plants and are increasingly resistant to insecticides (Cork *et al* 1992), the control of this pest is an important challenge. Adverse effects of insecticides and difficulties in *H. assulta* control led the researchers to study alternative method such as integrated pest management (IPM) programs, including the use of pheromone trapping (Cork *et al* 1992) and natural enemies (Choi *et al* 1975). However, because the identification and characterization of HasNPV was not analysed well the viral insecticide using it did not developed yet. To control *H. assulta* effectively as IPM program, the development of viral insecticide is needed and the characterization of HasNPV should be analysed in molecular level.

The baculoviruses produce polyhedrin at very high levels in the late phase of infection. Polyhedrin is the major component of polyhedra and has been well studied. Polyhedrin is protein of about 245 to 250 amino acids and appears to be most highly conserved baculovirus protein. These characteristics lead to the use of polyhedrin sequences as the base of baculovirus phylogenetic studies (Zanotto *et al* 1993). The investigation of polyhedrin gene structure, therefore, is important to identify the NPV. Recently, we developed a technique using the polymerase chain reaction (PCR) for the detection of NPV polyhedrin gene (Woo 2001). In this study, we describe the localization of polyhedrin gene using this technique in the HasNPV isolated from Korea and its characterization by nucleotide sequencing. In addition, HasNPV DNA was analysed by restriction endonuclease (REN) analysis to characterize its genome.

2. Materials and methods

2.1 *Virus and DNA isolation*

An isolate of HasNPV was obtained from infected larvae with NPV symptom. The virus multiplied in third instars larvae of *H. assulta* were fed with polyhedra for mass production of HasNPV. Virus DNA was prepared from polyhedra of infected larvae. The polyhedra produced in insect larvae were purified by standard method (O'Reilly *et al* 1992). To extract virus DNA, the purified polyhedra were resuspended in 0.1 M sodium carbonate solution [0.1 M Na_2CO_3, 0.17 M NaCl, 0.01 M EDTA (pH 10.9)], and incubated at 37°C overnight with 0.5 mg/ml of proteinase K (Sigma) and 1% of SDS. A further extraction with phenol and chloroform: isoamylalcohol (24:1) was performed and DNA was ethanol-precipitated. DNA was resuspended in TE buffer (Tris-HCl 10 mM, pH 8: EDTA 1 mM).

2.2 *Restriction endonuclease analysis*

Purified viral DNA was digested with restriction endonucleases (Takara Korea Biomedical) and DNA fragments were separated on a 0.6% agarose gel at 50 V for 12 h. The sizes were estimated according to the migration distances on the gels compared with known sizes of phage lambda *Hind*III fragments, 1 kb plus DNA ladder (Takara Korea Biomedical) and various restriction fragments of *Autographa californica* NPV (AcNPV) used as size standards. For fragments with sizes larger than 15 kb a second enzyme was used to generate smaller fragments, providing a better estimate.

2.3 *PCR of the polyhedrin gene*

To identify the polyhedrin gene in virus DNA, the PCR was performed using previous reported degenerate PCR primer set and condition (Woo 2001). The reaction was carried out using PreMix™-Top (Bioneer), 50 ng of viral DNA and 1μl of each primer in a total volume of 20 μl. Amplification was accomplished with the DNA Thermal Cycler (Perkin Elmer Cetus). Following amplification, the PCR products were analysed by 0.9% agarose gel electrophoresis. The amplified PCR products were cloned into a pGemT-Easy PCR cloning vector (Promega) and sequenced.

2.4 *Southern blot analysis*

The virus DNA was digested with several restriction endonucleases and then electrophoretically separated in 0.7% agarose gels and transferred to Hybond-N-nylon membrane

Table 1. Nucleopolyhedroviruses and sources of sequence information for this study.

Virus	Group	Host	Reference or database accession for polyhedrin gene sequence
AcNPV	I	*Autographa californica*	Hooft van Iddekinge *et al* (1983)
AceNPV	I	*Archips cerasivoranus*	U40834
ArNPV	I	*Attacus ricini*	Hu *et al* (1993a)
BmNPV	I	*Bombyx mori*	Gomi *et al* (1999)
BsNPV	II	*Buzura suppressaria*	Hu *et al* (1993b)
HarNPV	II	*Helicoverpa armigera*	Le *et al* (1997)
HcNPV	I	*Hyphantria cunea*	D14573
HtarNPV	II	*Heliothis armigera*	U97657
HzNPV	II	*Helicoverpa zea*	Cowan *et al* (1994)
LdNPV	II	*Lymantria dispar*	Smith *et al* (1988)
LsNPV	II	*Leucania seperata*	U30302
MbNPV	II	*Mamestra brassicae*	Cameron and Possee (1989)
OpMNPV	I	*Orgyia pseudotsugata*	Leisy *et al* (1986b)
OpSNPV	II	*Orgyia pseudotsugata*	Leisy *et al* (1986a)
PfNPV	II	*Panolis flammea*	Oakey *et al* (1989)
PnNPV	I	*Perina nuda*	Chou *et al* (1996)
SeMNPV	II	*Spodoptera exigua*	van Strien *et al* (1992)
SlNPV	II	*Spodoptera litura*	X94437
SlrNPV	II	*Spodoptera littoralis*	Croizier and Croizier (1994)
TnGV	-	*Trichoplusia ni*	Akiyoshi *et al* (1985)

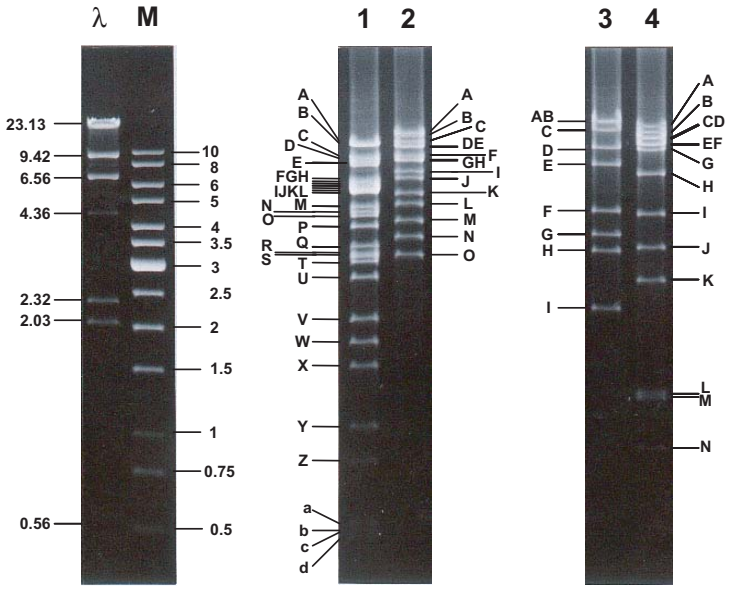

Figure 1. Gel photograph and schematic representation of restriction fragment profiles of HasNPV DNA cleaved with *Eco*RI (lane 1), *Nco*I (lane 2), *Xho*I (lane 3) and *Bgl*II (lane 4) and electrophoreses on a 0.6% agarose gel at 50V for 12 h to separate the fragments. Restriction fragment designations are indicated by letters in the schematic drawings of individual restriction endonuclease profiles. Lambda *Hind*III fragments (*λ*) and 1 kb DNA ladder (M) were used as size standards.

(Amersham Biosciences) by using a vacuum transfer unit (Hoefer TE-80) for 30 min. Southern hybridization was carried out according to the method recommended by the supplier, with the nonradioactive DNA labelling and detection kit (Boehringer Mannheim). The amplified putative HasNPV polyhedrin gene fragment was used as a probe. Cloning of positive restriction enzyme fragments of HasNPV DNA was achieved by isolated from agarose gel and subcloning into the plasmid pBluescriptII KS(+).

2.5 *Construction of the deletion mutant set*

One set of unidirectional deletion plasmids was constructed by using Exonuclease III and Klenow fragment of DNA polymerase I. The 2.3 kb plasmid containing HasNPV polyhedrin gene was digested with *Pst*I and *Eco*RI. After treatment with Exonuclease III by the time interval, the fragment was treated with the Klenow fragment of DNA polymerase I to fill in the ends, and the desired fragment was purified by gel electrophoresis and circularized. The ligated plasmids, which had the insert in each size, were obtained from the transformant *Escherichia coli* XL1-Blue. These plasmids were subjected to DNA sequencing. DNA sequence analysis was performed by the dideoxynucleotide chain termination method at the DNA sequencing facility, Takara Korea Biomedical Institute.

2.6 *Analysis of the polyhedrin gene*

The polyhedrin gene sequence of HasNPV has been deposited in GenBank under accession No. DQ157735. An alignment of the complete DNA sequences of 20 polyhedrin genes and 1 granulin gene (table 1) was generated by CLUSTAL_X (Thompson *et al* 1997), and phylogenetic trees were constructed using the neighbour-joining and maximum parsimony algorithms in PAUP* Version 4.0b10 (Swofford, 2001). The robustness of the tree topologies was estimated by bootstrap analysis with 1000 replicates (PAUP*, Version 4.0b10). The granulin sequences were defined as an out-group.

3. Results

3.1 *Restriction endonuclease fragment analysis*

The DNA of HasNPV was digested with *Eco*RI, *Nco*I, *Xho*I and *Bgl*II with 30, 15, 9, and 14 visible fragments detectable for each digest, respectively, with size ranging from 42 to 0.3 kb (figure 1). The fragments were designated alphabetically starting with A for the largest fragment for each enzyme digest as proposed by Vlak and Smith (1982).

Table 2. Size of restriction endonuclease fragments (kb) of HasNPV.

Fragment	*Eco*RI	*Nco*I	*Xho*I	*Bgl*II
A	14.05	18.34	42.53	24.78
B	13.92	17.53	41.24	19.21
C	9.35	15.24	20.25	16.82
D	8.92	13.12	12.54	16.71
E	8.35	12.79	8.20	14.25
F	6.52	9.32	4.59	13.94
G	6.38	8.95	3.83	12.12
H	6.21	8.71	3.36	6.80
I	5.78	7.15	2.25	4.51
J	5.68	6.50		3.45
K	5.57	5.25		2.71
L	5.42	4.75		1.23
M	4.68	4.05		1.20
N	4.59	3.55		0.85
O	4.43	3.19		
P	4.06			
Q	3.42			
R	3.20			
S	3.17			
T	3.02			
U	2.85			
V	2.05			
W	1.76			
X	1.49			
Y	1.01			
Z	0.78			
a	0.55			
b	0.49			
c	0.45			
d	0.38			
Total	138.53	138.44	138.79	138.58

These numbers represent the minimum number of cleavage sites for each of these four enzymes, since fragments smaller than 0.3 kb are not detected. The size estimates for the REN fragments and the total genome size for HasNPV DNA are given in table 2. These estimates are the means from 8 gels and were determined comparing REN fragment mobility with those of size standards and/or AcNPV DNA. The total genome size of HasNPV was estimated to be about 138 kb.

3.2 *Polyhedrin gene localization, cloning, and sequencing*

The polyhedrin gene of HasNPV was successfully amplified by previous-reported degenerate PCR primer set for the polyhedrin gene. Size of the amplified product was about 430 bp, which was identical with the previous report (Woo 2001). The amplified PCR product was cloned into a pGemT-Easy PCR cloning vector and sequenced. The sequencing results showed that the PCR product was a fragment corresponding to the previous-reported other NPV polyhedrin genes (*H. armigera, H. zea* etc.) (Le *et al* 1997; Cowan *et al* 1994).

To identify the location of the polyhedrin gene in the HasNPV genome, Southern blot analysis was applied using PCR-amplified partial HasNPV polyhedrin gene fragment as a probe. As the result, we identified the location of the polyhedrin gene within the 6 kb *Eco*RI, 15 kb *Nco*I, 20 kb *Xho*I, 17 kb *Bgl*II and 3.0 kb *Cla*I fragments, respectively (data not shown). The 3.0 kb *Cla*I fragment of HasNPV DNA was isolated and subcloned into the same site on plasmid pBluescriptII KS (+) to construct plasmid pHaC. In order to determine the restriction map of the 3.0 kb *Cla*I fragment, the pHaC DNA was digested with *Bam*HI, *Eco*RI, *Hinc*II, *Hind*III, *Pst*I, *Sac*I, *Sal*I, *Xho*I and *Xba*I, respectively, and mapped (figure 2). From repetitive subcloning and Southern hybridization, we were able to localize HasNPV polyhedrin gene to a smaller region of the 2.3 kb *Eco*RI–*Cla*I fragment. This fragment was cloned and the deletion mutant sets were constructed to determine nucleotide sequence. The arrows indicate the direction and location of deletion mutants (figure 2). A total of 2358 nucleotides sequence encompassing the entire coding region and 5`, 3` non-coding flanking sequences were determined (figure 3). Nucleotide sequence analysis indicated the presence of an open reading frame (ORF) of 738 nucleotides, which could encode 246 amino acid residues with a predicted molecular mass of 29 kDa.

3.3 *Characterization of the polyhedrin gene*

Comparison of the nucleotide sequence of the ORF with the polyhedrin coding sequences of AcNPV (Hooft van Iddekinge *et al* 1983), *H. armigera* NPV (HarNPV) (Le *et al* 1997), *Heliothis armigera* NPV (HtarNPV) and *H. zea* NPV (HzNPV) (Cowan *et al* 1994) indicates 73.7%, 99.2%, 99.2% and 99.1% homologies respectively (table 3). The other NPVs showed lower homologies from 70% to 79%. At the amino acid level, 100%, 99% and 100% sequence identity is observed with HarNPV, HtarNPV and HzNPV, respectively. This high degree of sequence identity with polyhedrin gene sequences from this presumably closely related NPV suggests that the ORF sequenced is the HasNPV polyhedrin gene. HasNPV polyhedrin gene region also indicated that there is an ORF immediately downstream from the putative HasNPV polyhedrin gene (figure 3). This ORF has a small region of homology to AcNPV ORF 1629 in complement direction.

A neighbour-joining tree was generated from the amino acid sequences of 20 NPV polyhedrin genes (figure 4) using the sequence of granulin gene from *Trichoplusia ni* granulovirus (TnGV) as an out-group. This shows that HasNPV is a group II NPV and sits in a group of closely related NPVs including HarNPV.

4. Discussion

The size of the HasNPV genome was estimated at 138 kb by REN fragment analysis (figure 1 and table 2). Although

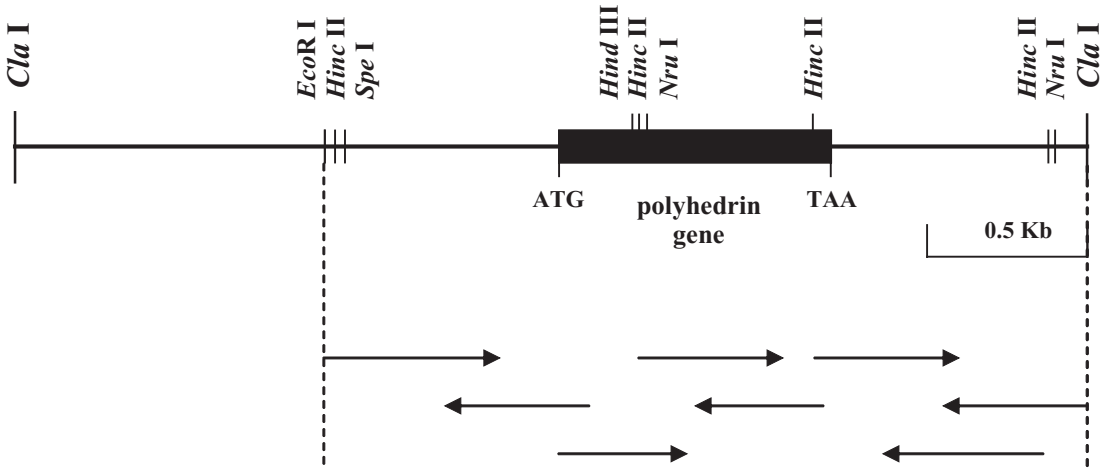

Figure 2. Physical mapping and sequencing strategy of the polyhedrin gene region of HasNPV.

```
GAATTCATTTTAATTACAGTTTTTATAATTGTACAAACAGTCTATAACCAACCATGTGTAACGTGTGGCCAGTGGTTAACCGTGTGCTTT   90
GCAAACTAGTCATGCAAAATTTGTCCAAAATATATGGCAATATACAATTTTTATATTTGGTGGGCAACAGGCCCAAGGAAATTCAAGAGG  180
AACAAGCCAATTTTAACGAATTATATTACAAGTTCAAAGTGTTTAGATCACAATTGCCCGACATGAATTGTGAAACTTTTGCTCATAAAT  270
TGATTGACCAGAAAATATTGTATTGCAGAGAAATTCATAATTTGTATTTAAATTTTTTATATTGCTTTTACAAACAATACTTTGACACGC  360
TGAAGATTGACTGCAATATTTTTAAGGATTTGATAGATGACGATGTACCATTGCAAGATTTTGAAGAGTTAAATGTTGTCCTACTCGACA  450
ATAACATAGCAATGTATACGGCTTTGTGTGATGATGTGTTTGAAAAGAAAACCATTATACAAGATATAGAGTATGTGATGAACAAAATAT  540
GCGTTGAAGGAGCGTACGTGCCATTTCAAGAAGAAATTTTACAATATCAAATCTTTTTGCAAGAATATGAAGATTTCTGTCGTCGTGTTG  630
AAAATTTGTAATAAAACCCAAATAAACCTTTAATATAAATATTAAACATACACTTTTATTTCTAAAATAAGTATTTTTTTCCTATTGTTC  720
```

polyhedrin

```
AAGATTGTGAAAAATCAAATATCCCATAATGTATACTCGTTACAGTTACAGCCCTACTTTGGGCAAAACCTATGTGTACGACAACAAATA  810
                            M  Y  T  R  Y  S  Y  S  P  T  L  G  K  T  Y  V  Y  D  N  K  Y   21
CTTTAAGAATTTAGGTGCTGTTATTAAAAATGCCAAACGCAAGAAGCATTTAGAGGAGCACGAACATGAAGAACGCAACTTGGATTCGCT  900
  F  K  N  L  G  A  V  I  K  N  A  K  R  K  K  H  L  E  E  H  E  H  E  E  R  N  L  D  S  L   51
CGACAAATACTTGGTGGCGGAAGATCCTTTTTTGGGACCCGGCAAAAATCAAAAACTAACTTTGTTTAAAGAGATTCGCAGCGTTAAGCC  990
  D  K  Y  L  V  A  E  D  P  F  L  G  P  G  K  N  Q  K  L  T  L  F  K  E  I  R  S  V  K  P   81
CGACACAATGAAGCTTGTAGTTAACTGGAGCGGTCGCGAATTTCTTCGCGAAACTTGGACGCGTTTCATGGAAGACAGTTTTCCCATTGT 1080
  D  T  M  K  L  V  V  N  W  S  G  R  E  F  L  R  E  T  W  T  R  F  M  E  D  S  F  P  I  V  111
AAACGACCAAGAAATTATGGACGTGTTTCTGTCTGTTAATATGCGACCAACCAAACCGAACCGTTGTTACCGATTCTTAGCGCAACACGC 1170
  N  D  Q  E  I  M  D  V  F  L  S  V  N  M  R  P  T  K  P  N  R  C  Y  R  F  L  A  Q  H  A  141
TCTGCGTTGTGACCCCGACTATATTCCTCACGAAGTCATTCGTATTGTAGAACCTTCCTATGTAGGCAGTAACAACGAGTACAGAATTAG 1260
  L  R  C  D  P  D  Y  I  P  H  E  V  I  R  I  V  E  P  S  Y  V  G  S  N  N  E  Y  R  I  S  171
TTTAGCCAAAAAATACGGCGGTTGTCCCGTTATGAACTTGCACGCTGAATACACTAATTCCTTTGAAGATTTCATTACCAACGTAATTTG 1350
  L  A  K  K  Y  G  G  C  P  V  M  N  L  H  A  E  Y  T  N  S  F  E  D  F  I  T  N  V  I  W  201
GGAGAATTTCTACAAACCAATTGTTTACGTAGGCACTGATTCTGCCGAAGAAGAGGAAATACTCCTAGAGGTTTCTTTGATATTTAAGAT 1440
  E  N  F  Y  K  P  I  V  Y  V  G  T  D  S  A  E  E  E  E  I  L  L  E  V  S  L  I  F  K  I  231

CAAAGAATTTGCACCTGACGCTCCGCTATACACTGGTCCTGCATATTAAACTTGCGATTCAGTTGACATCGTCAATTTGTAACTCATAAT 1530
  K  E  F  A  P  D  A  P  L  Y  T  G  P  A  Y  *                                           246

TTTATCTAAATTCGATCGCAATTCTTGTAATTTTTGATTGGTCGGTTTGGTTCCTAATGCCGACACCACATTAGCTAACGCTTTATCGTA 1620
CTGTTTTTTGAATGTCAAATTTTCCACCGCCATAATGAATTGTTGTAAATTTTTATCGGACAATTGAAGTTCGACATCATCGGATTTGTC 1710
CAAAGGATTATCATACGTTTGTTGTATCAAGTTATCTTCAATAAATATTTGTAGTTTAGCAGAAACTTGTTGTGTTTGTGCATTCGAAAG 1800
CCGTTGATTTAATTGATTTTTTATTGATATTAATGTGTCTTGTGCTTCTGTAGACAAAGGATAATTTTTTATCCATGAACTGTCCAATGT 1890
TATATTGTACAAAGAACGTACATATTGTTTTAATTCGTTGCTGGCTCGCTGCTGTTGTTCGTCGTCGGTCCACCCGTTTTCCGATTCTGA 1980
CGAAACTACAGGACTCGGTTGAACGGCTATGCGTCGTTGCAAAATCTTGGCAGTAGGACTGGCGGCGGCGGTAACGGTATTTACTATCGA 2070
ACCGCCATCGGCGGGTTTTGATACTTTTTTTAATTTAATTCCTTTCTGTATTTGTTCCATCAATTCGGTACGTGGATCTTTTAAAACTTG 2160
CCGAGTCGACGTTGTAAAATCGCGATCTTTACTGGATGGTATTACTATATTTTCTATTAATGGTAATGACGGTGGCGGAGGAGGCGGCGG 2250
CGGAGGAGGTATCGTCGAAGATAAGTTTGTTTGAGGAGGTGGCGGCGGTGGCGGCGGTGGTATTGGTGGCGGCATATGTGTTTGCGGCGA 2340
GGAAACTTCAGAATCGAT                                                                         2358
```

Figure 3. Nucleotide sequence of the HasNPV polyhedrin gene and its flanking regions. The deduced amino acids are indicated with one-letter code designation for polyhedrin.

Table 3. Pairwise comparisons of the nucleotide and deduced amino acid sequences identity (%) of 20 NPVs polyhedrin genes[a].

	HasNPV	AcNPV	AceNPV	ArNPV	BmNPV	BsNPV	HarNPV	HcNPV	HtarNPV	HzNPV	LdNPV	LsNPV	MbNPV	OpMNPV	OpSNPV	PfNPV	PnNPV	SeNPV	SlNPV	SlrNPV
HasNPV	-	73.7	71.0	72.2	72.7	79.0	99.2	73.6	99.2	99.1	70.6	75.4	75.7	74.0	78.3	75.6	74.5	75.6	73.8	73.8
AcNPV	86	-	78.8	77.1	77.7	76.2	73.6	77.8	73.5	73.5	75.5	80.2	80.8	79.5	76.9	82.3	78.7	77.5	78.6	76.6
AceNPV	81	85	-	82.4	79.9	72.4	70.8	81.9	70.8	70.7	74.1	76.2	76.0	86.4	73.6	76.9	86.2	75.0	76.8	73.4
ArNPV	80	88	93	-	81.7	72.9	72.2	82.7	72.0	72.0	75.0	76.5	76.1	84.7	74.0	76.3	83.8	74.6	73.7	72.7
BmNPV	82	86	89	91	-	72.9	72.4	80.7	72.3	72.5	74.7	75.8	75.5	81.4	76.8	76.6	81.2	75.6	74.0	71.8
BsNPV	90	90	84	85	85	-	79.1	74.9	79.0	79.0	70.8	80.2	80.6	74.2	81.0	80.1	74.2	79.3	77.5	75.2
HarNPV	100	86	80	80	82	90	-	73.5	99.2	99.6	70.1	76.0	76.3	73.7	78.6	75.9	74.2	75.9	74.0	73.7
HcNPV	80	87	92	92	89	86	85	-	73.6	73.3	73.3	76.2	76.2	86.1	76.6	77.0	85.4	75.9	76.1	73.9
HtarNPV	99	85	80	79	82	90	99	85	-	99.3	70.0	75.3	75.6	73.7	78.6	75.5	74.5	75.9	73.8	73.8
HzNPV	100	86	80	80	81	91	100	80	99	-	70.3	75.8	76.2	73.6	78.5	75.7	74.1	75.7	74.1	73.6
LdNPV	80	81	79	79	80	83	81	80	80	81	-	74.5	74.4	75.9	73.3	75.8	75.1	73.3	75.2	75.5
LsNPV	86	88	84	84	84	93	86	84	86	87	82	-	97.2	77.1	78.8	90.6	76.1	81.3	81.7	76.4
MbNPV	86	88	84	84	85	94	87	85	86	87	83	98	-	77.4	79.3	90.9	76.0	82.1	81.3	77.5
OpMNPV	84	89	96	95	91	87	83	93	83	84	80	85	86	-	76.1	78.5	96.0	77.8	77.3	77.1
OpSNPV	89	88	83	83	84	95	89	88	89	88	82	94	94	85	-	80.9	76.4	79.0	76.4	75.1
PfNPV	87	89	84	84	85	95	87	85	87	88	84	97	98	86	95	-	77.7	81.4	81.5	76.1
PnNPV	84	89	96	96	91	87	83	94	83	84	80	85	86	99	85	86	-	77.1	76.2	76.1
SeNPV	86	85	82	83	82	91	86	83	86	87	82	91	92	84	91	91	85	-	86.3	78.7
SlNPV	86	84	81	79	80	87	86	80	85	87	81	85	86	84	87	86	84	85	-	77.2
SlrNPV	85	83	81	79	80	86	86	85	85	84	82	85	85	84	87	85	84	84	99	-

[a]Dashes indicate identical sequences along the x- and y-axis. The numbers above the diagonal line defined by identical sequences indicate the percent identity of the nucleotide sequences while the numbers below the diagonal line indicate the percent amino acid identity.

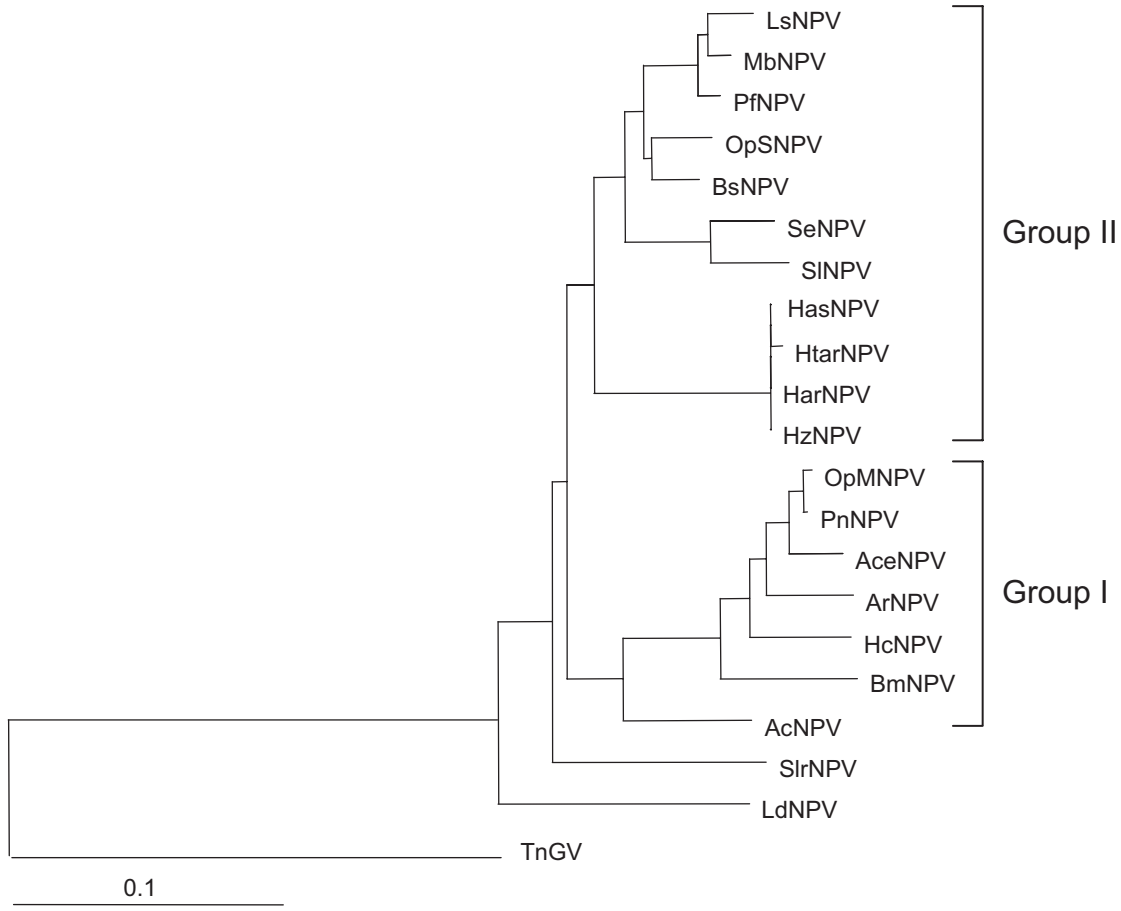

Figure 4. Phylogenic tree of NPV polyhedrin sequences. A phylogenic tree based on the amino acid sequences of the polyhedrin genes available in the GenBank database. The division of group I and group II NPVs is indicated and the position of HasNPV is highlighted. In this study, the granulin gene was included to provide an out-group. The virus used are listed in table 1.

the isolation of HasNPV was previously reported, the characterization of its genome was not reported yet. Genome size in baculoviruses can range from 81 to 160 kb (Mathews 1982; Blissard and Rohrmann 1990; Lauzon *et al* 2004), and is relatively stable and a reliable characteristic for a given virus. The REN profile for a baculovirus is also relatively stable and is often used as an effective tool to differentiate closely related viruses. *H. armigera* and *H. assulta* are sympatric, closely related species in and around Korea. The subfamily *Heliothinae* includes several genera, among them *Helicoverpa* and *Heliothis*. Studies of the phylogeny within species of *Helicoverpa*, using classical morphological characters (Mitter *et al* 1993) as well as genetic markers (Fang *et al* 1997; Cho *et al* 1995) have shown that the oligophagous *H. assulta* is relatively closely related to the polyphagous *H. armigera* and *H. zea*. The genome structure of HarNPV and HzNPV was already reported. The size was estimated at about 130.7 kb and 87 kb for HarNPV and HzNPV, respectively. The genome structure of HtarNPV is not reported exactly yet. Therefore HasNPV has different genome size with these viruses. In addition, REN analysis of the viral DNA showed it to be distinct from published profiles of HarNPV and HzNPV DNAs.

Analysis of the polyhedrin gene revealed high homology to the HarNPV, HtarNPV and HzNPV polyhedrins in DNA and amino acid sequences. This implies that HasNPV is closely related to these NPVs. In addition, the amino acid sequences of HasNPV, HarNPV and HzNPV polyhedrins are identical, but the genome size and location of polyhedrin gene of HasNPV were different with those of HarNPV and HzNPV. It has been reported that polyhedrins from lepidopteran NPVs are closely related to one another and have 85 to 90% amino acid identity (Rohrmann 1986). HasNPV polyhedrin, a protein of 245 amino acid residues, does indeed resemble other NPVs.

After first report about the localization of polyhedrin gene in AcNPV (Vlak and Smith 1982), Hofft van Iddekinge *et al* (1983) determined its nucleotide sequences. At last count, there are published data on the polyhedrin sequence of some 30 or more baculoviruses (Bulach *et al* 1999).

The polyhedrin gene is highly conserved between NPVs and has been characterized from many NPVs, making it the preferred choice for phylogeny studies. Phylogenetic analysis of polyhedrin from different NPVs separates NPVs into two distinct groups (Zanotto *et al* 1993). Bootstrap analysis using 1000 replicates and the protdist function of PHYLIP predicts that HasNPV is a group II NPV clustered in a closely related group of viruses also including HarNPV, HzNPV, and HtarNPV. Group II NPV have polyhedrins with 246 amino acids instead of the 245 found in Group I. The strong similarity of different NPVs is not necessarily reflected by a close relatedness of their respective host species.

Several studies have compared baculovirus isolates obtained from geographically separate populations of a given host species (Allaway and Payne 1983; Cherry and Summers 1985; Gettig and McCarthy 1982; Hong *et al* 2000; Kislev and Edelman 1982; Maeda *et al* 1990). Investigations of this geographic variability and the role of these genotypic differences in the biology of baculoviruses are an important area of current research. Such studies may provide insight into the evolution of baculoviruses and their hosts and may also aid in the development of more effective virus strains for biological control of insects. The improvement of HasNPV for a successful introduction into biological control for common cutworm requires detailed knowledge of the molecular biology of this virus. This study sets the foundation for this and will serve in genetic engineering of the virus to enhance its potential as a biological control agent.

Acknowledgements

This work was supported by Chungbuk National University Grant in 2005.

References

Adams J R and Bonami J R 1991 *Atlas of invertebrate pathology* (Boca Raton: CRC Press)

Akiyoshi D, Chakerian R, Rohrmann G F, Nesson M H and Beaudreau G S 1985 Cloning and sequencing of the granulin gene from the *Trichoplusia ni* granulosis virus; *Virology* **141** 328–332

Allaway G P and Payne C C 1983 A biochemical and biological comparison of three European isolates of nuclear polyhedrosis viruses from *Agrotis segetum*; *Arch. Virol.* **75** 43–54

Blissard G W and Rohrmann G F 1990 Baculovirus diversity and molecular biology; *Annu. Rev. Entomol.* **35** 127–155

Bulach D M, Kumar C A, Zaia A, Liang B and Tribe D E 1999 Group II nucleopolyhedrovirus subgroups revealed by phylogenetic analysis of polyhedrin and DNA polymerase gene sequences; *J. Invertebr. Pathol.* **73** 59–73

Cameron I R and Possee R D 1989 Conservation of polyhedrin gene promoter function between *Autographa californica* and *Mamestra brassicae* nuclear polyhedrosis viruses; *Virus Res*. **12** 183–199

Cherry C L and Summers M D 1985 Genotypic variation among wild isolates of two nuclear polyhedrosis viruses isolated from *Spodoptera littoralis*; *J. Invertebr. Pathol*. **46** 289–295

Cho S, Mitchell A, Regier J C, Mitter C, Poole R W, Friedlander T P and Zhao S 1995 A highly conserved nuclear gene for low-level phylogenetics: elongation factor-1 recovers morphology-based tree for heliothine moths; *Mol. Biol. Evol.* **12** 650–656

Choi K M, Cho E H, So J S and Hwang C Y 1975 Studies on the seasonal occurrences of the tobacco budworm, *Heliothis assulta* H. (Lepidoptera: Noctuidae), and the parasitism ratio of *Trichogramma* spp. on the eggs; *Korean J. Plant Prot.* **14** 137–140 (in Korean)

Chou C M, Huang C J, Lo C F, Kou G H and Wang C H 1996 Characterization of *Perina nuda* nucleopolyhedrovirus (PenuNPV) polyhedrin gene; *J. Invertebr. Pathol*. **67** 259–266

Cork A, Boo K S, Dunkelblum E, Hall D R, Jee-Rajunga K, Kehat M, Kong Jie E, Park K C, Tepgidagarn P and Xun L 1992 Female sex pheromone of oriental tobacco budworm, *Helicoverpa assulta* (Guenee) (Lepidoptera: Noctuidae): Identification and field testing; *J. Chem. Ecol.* **18** 403–418

Cowan P, Bulach D, Goodge K, Robertson A and Tribe D E 1994 Nucleotide sequence of the polyhedrin gene region of *Helicoverpa zea* single nucleocapsid nuclear polyhedrosis virus: placement of the virus in lepidopteran nuclear polyhedrosis virus group II; *J. Gen. Virol.* **75** 3211–3218

Croizier L and Croizier G 1994 Nucleotide sequence of the polyhedrin gene of *Spodoptera littoralis* multiple nucleocapsid nuclear polyhedrosis virus; *Biochim. Biophys. Acta* **1218** 457–459

Fang Q Q, Cho S, Regier J R, Mitter C, Mattews M, Poole R W, Friedlander T P and Zhao S 1997 A new nuclear gene for insect phylogenetics: dopa decarboxylase is informative of relationships within heliothinae (Lepidoptera: Noctuidae); *Syst. Biol.* **46** 269–283

Gettig R G and McCarthy W J 1982 Genotypic variation among wild isolates of *Heliothis* spp. nuclear polyhedrosis viruses from different geographic regions; *J. Virol*. **117** 245–252

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

Hill D S 1983 *Agricultural insect pests of the tropics and their control* 2nd edition (Cambridge: Cambridge Press)

Hong H K, Woo S D, Choi J Y, Lee H K, Kim M H, Je Y H and Kang S K 2000 Characterization of four isolates of *Bombyx mori* nucleopolyhedrovirus; *Arch. Virol*. **145** 2351–2361

Hooft van Iddekinge B J K, Smith G E and Summers M D 1983 Nucleotide sequence of the polyhedrin gene of *Autographa californica* nuclear polyhedrosis virus; *Virology* **131** 5561–5565

Hu J, Ding H and Wu X 1993a Cloning and sequencing of *Attacus ricini* nuclear polyhedrosis virus polyhedrin gene; *Yi Chuan Xue Bao.* **20** 300–304

Hu Z H, Liu M F, Jin F, Wang Z X, Liu X Y, Li M J, Liang B F and Xie T E 1993b Nucleotide sequence of *Buzura suppressaria* single nucleocapsid nuclear polyhedrosis virus polyhedrin gene; *J. Gen. Virol.* **74** 1617–1620

Inceoglu A B, Kamita S G, Hinton A C, Huang Q, Severson T F, Kang K and Hammock B D 2001 Recombinant baculoviruses for insect control; *Pest Manag. Sci.* **57** 981–987

Jin B R, Park H W, Woo S D, Kim W J, Park B S and Kang S K 1995 Morphology and biochemical characteristics of a nuclear polyhedrosis virus isolated from the oriental tobacco budworm, *Helicoverpa assulta* (Guenee); *Korean J. Appl. Entomol.* **34** 218–223 (in Korean)

King L A and Possee R D 1992 *The baculovirus expression system: a laboratory guide* (London: Chapman and Hall)

Kislev N and Edelman M 1982 DNA restriction pattern differences from geographic isolates of *Spodoptera littoralis* nuclear polyhedrosis virus; *Virology* **119** 219–222

Kost T A and Condreay J P 2002 Recombinant baculoviruses as mammalian cell gene-delivery vectors; *Trends Biotechnol.* **20** 173–180

Kost T A, Condreay J P and Jarvis D L 2005 Baculovirus as versatile vectors for protein expression in insect and mammalian cells; *Nat. Biotechnol.* **23** 567–575

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

Le T H, Wu T, Robertson A, Bulach D, Cowan P, Goodge K and Tribe D 1997 Genetically variable triplet repeat in a RING-finger ORF of *Helicoverpa* species baculoviruses; *Virus Res.* **49** 67–77

Leisy D, Nesson M, Pearson M, Rohrmann G and Beaudreau G 1986a Location and nucleotide sequence of the *Orgyia pseudotsugata* single nucleocapsid nuclear polyhedrosis virus polyhedrin gene; *J. Gen. Virol.* **67** 1073–1079

Leisy D, Rohrmann G and Beaudreau G 1986b The nucleotide sequence of the polyhedrin gene region from the multicapsid baculovirus of *Orgyia pseudotsugata*; *Virology* **153** 280–288

Maeda S 1989 Expression of foreign genes in insects using baculovirus vectors; *Annu. Rev. Entomol.* **34** 351–372

Maeda S 1994 Expression of foreign gene in insect cells using baculovirus vectors; in *Insect cell biotechnology* (eds) K Maramorosch and A McIntosh (Boca Raton: CRC Press) pp 1–31

Maeda S, Mukohara Y and Kondo A 1990 Characteristically distinct isolates of the nuclear polyhedrosis virus from *Spodoptera litura*; *J. Gen. Virol.* **71** 2631–2639

Mathews R E F 1982 Classification and nomenclature of viruses; *Intervirology* **17** 1–199

Miller L K 1988 Baculoviruses as gene expression vectors; *Annu. Rev. Microbol*. **42** 177–199

Mitter C, Poole R W and Mattews M 1993 Biosystematics of the Heliothinae (Lepidoptera: Noctuidae); *Annu. Rev. Entomol*. **38** 207–225

Murphy F A, Fauquet C M, Bishop D H L, Ghabrial S A, Jarvis A W, Martinelli G P, Mayo M A and Summers M D 1995 Virus Taxonomy. Sixth report of the international committee on taxonomy of viruses; *Arch. Virol. Suppl*. **10** 85–86

Oakey R, Cameron I R, Davis B, Davis E and Possee R D 1989 Analysis of transcription initiation in the *Panolis flammea* nuclear polyhedrosis virus polyhedrin gene; *J. Gen. Virol.* **70** 769–775

O'Reilly D R, Miller L K and Luckow V A 1992 *Baculovirus expression vectors: a laboratory manual* (New York: Freeman)

Rohrmann G F 1986 Polyhedrin structure; *J. Gen. Virol.* **67** 1499–1513

Smith G E, Summers M D and Fraser M J 1983 Production of human β-interferon in insect cells infected with a baculovirus expression vector; *Mol. Cell. Biol*. **3** 2156–2165

Smith I R, van Beek N A, Podgwaite J D and Wood H A 1988 Physical map and polyhedrin gene sequence of *Lymantria dispar* nuclear polyhedrosis virus; *Gene* **71** 97–105

Swofford D L 2001 *PAUP*. Phylogenetic Analysis Using Parsimony (* and Other Methods*) Version 4. Sinauer Associates, Sunderland, MA, USA

Thompson J D, Gibson T J, Plewniak F, Jeanmougin F and Higgins D G 1997 The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools; *Nucleic Acids Res.* **25** 4876–4882

van Strien E A, Zuidema D, Goldbach R W and Vlak J M 1992 Nucleotide sequence and transcriptional analysis of the polyhedrin gene of *Spodoptera exigua* nuclear polyhedrosis virus; *J. Gen. Virol.* **73** 2813–2821

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

Woo S D 2001 Rapid detection of multiple nucleopolyhedroviruses using polymerase chain reaction; *Mol. Cells* **11** 334–340

Woo S D, Cho K H, Jin B R, Boo K S and Kang S K 2000 Expression of the gene encoding firefly luciferase using *Bombyx mori* nucleopolyhedrovirus vector; *Int. J. Ind. Entomol*. **1** 53–58

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

Zhang C X, Ma X C and 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

Zhang C X and Wu J C 2001 Genome Structure and the p10 Gene of the *Helicoverpa armigera* Nucleopolyhedrovirus; *Sheng Wu Hua Xue Yu Sheng Wu Wu Li Xue Bao (Shanghai)* **33** 179–184

MS received 27 February 2006; accepted 25 July 2006

*e*Publication: 8 August 2006

Corresponding editor: Shahid Jameel