Nucleotide Sequence Analysis of the Polyhedrin Gene of *Antheraea pernyi* Nucleopolyhedrovirus and Construction of a Transfer Vector Plasmid

Jun Kobayashi, Ryo Ando, Xue Ying Wang, Yuan Jiao Huang and Shigetoshi Miyajima

Reprinted from
Int. J. Wild Silkmoth & Silk
Vol. 6, 2001

Int. J. Wild Silkmoth & Silk **6**, 53–58 (2001)

Nucleotide Sequence Analysis of the Polyhedrin Gene of *Antheraea pernyi* Nucleopolyhedrovirus and Construction of a Transfer Vector Plasmid

Jun Kobayashi[1)], Ryo Ando[1)], Xue Ying Wang[2)], Yuan Jiao Huang[1)] and Shigetoshi Miyajima[1)]

[1)] *Faculty of Engineering, Mie University, Tsu, Mie 514-8507, Japan.*
[2)] *Bioscience and Technology College of Shenyang Agricultural University, Shenyang 110161, China.*

Abstract The polyhedrin gene of *Antheraea pernyi* nucleopolyhedrovirus (AnpeNPV) was identified in a 6.6 kbp *Pst*I fragment of the viral DNA genome by Southern blot hybridization using the polyhedrin gene of *Bombyx mori* NPV (BmNPV) as a probe. The nucleotide sequence analysis revealed that the AnpeNPV polyhedrin gene contained a promoter sequence of 51 bp and a coding sequence of 738 bp encoding 245 amino acids. Amino acid sequence comparison was made with both polyhedrins and granulins of other baculoviruses and very high homologies were detected with members of the group I NPV. Among them a NPV of another saturuniid moth, *Attacus ricini* (= *Samia cynthia ricini*) showed the highest homology, indicating that the two NPVs are closely related or variants. For the establishment of a novel baculovirus expression vector system using AnpeNPV and diapausing pupae of *A. pernyi*, we have constructed a transfer vector plasmid pApCH1, which enable the foreign gene expression under the control of the strong AnpeNPV polyhedrin promoter.
Key words: AnpeNPV, polyhedrin, transfer vector

Introduction

The baculovirus expression vector systems using *Autographa californica* nucleopolyhedrovirus (AcNPV) and *Bombyx mori* NPV (BmNPV) are now widely used for the high level expression of foreign genes. Among them, the BmNPV vector system was expanded to in vivo expression using larvae of the mulberry silkworm *B. mori* and enabled higher production levels in larvae than in cultured insect cells (Maeda *et al.*, 1985). Like the BmNPV, *Antheraea pernyi* NPV (AnpeNPV) has a potential to produce large amounts of heterologous proteins in not only larvae but also pupae of the Chinese oak silkworm *A. pernyi*. Pupal diapause of *A. pernyi* is an advantageous characteristic over *B. mori*, because diapausing pupae can be stored for more than six months in the refrigerator until using them for the recombinant virus infection. To establish a novel AnpeNPV vector system, we have already developed in vitro replication and plaque-purification system of AnpeNPV using an *A. pernyi* cell line, NISES-AnPe-428 (AnPe) (Inoue and Hayasaka, 1995; Wang *et al.*, 2000). In this paper, we describe the cloning of the polyhedrin gene of AnpeNPV and construction of the polyhedrin promoter-based transfer vector plasmid, which enable us to insert heterologous genes in the AnpeNPV genome and to express them in the virus-infected *A. pernyi* cells.

Materials and Methods

Virus, Bacterium and Plasmid

DNA genome of AnpeNPV A strain plaque-

purified from the Chinese isolate by using AnPe cells (Wang *et al.*, 2000) was used for the polyhedrin gene analysis and transfer vector construction. Competent *E. coli* strain XL1-Blue cells (Stratagene) were used for plasmid DNA transformations. Plasmids, pUC19 (Takara-Shuzo), pBluescriptII (Stratagene) and pCR 2.1 (Invitrogen), were used for construction of the transfer vector pApCH1 as described later.

DNA manipulations

All plasmid DNA recombination techniques were essentially as described by Sambrook *et al.* (1989). Restriction enzymes and other DNA modifying enzymes were purchased from Takara-Shuzo Co. Ltd.

Southern blot hybridization

For Southern blot hybridization, AnpeNPV DNA was digested with *Pst*I and electrophoresed using a 0.8 % agarose gel. Then the *Pst*I fragments in the gel were transferred on to a Hybond-N+ membrane (Amersham) and hybridized using ECL direct nucleic acid labeling and detection system (Amersham) with the *Sal*I-*Hin*dIII fragment of BmNPV containing the entire polyhedrin gene as a probe.

DNA sequencing and sequence analysis

The nucleotide sequence of the AnpeNPV polyhedrin gene was determined by DNA sequencer (Model 373A, PE Applied Biosystems) and Taq dye-deoxy terminator sequencing kit (PE Applied Biosystems) with the following oligonucleotide primers.

M13 forward, 5'-CGTTGTAAAACGACGGCCAG-3'
M13 reverse, 5'-CAGGAAACAGCTATGACCAT-3'
AP1, 5'-AGTTATAGGAAATTTTACTACAAAG-3'
AP2, 5'-GCGCAAGAAGCATTTAGTCG-3'
AP3, 5'-TGGCGAGGCTAATTCTGTAC-3'
AP4, 5'-GAGATTGTGGACCGCATCTA-3'
AP7, 5'-GTTTAGCAAGAGCGTTCAGA-3'

Homology search of predicted amino acid sequence of the AnpeNPV polyhedrin were carried out using the BLAST program supplied in the GenomeNet BLAST Internet Server (Institute for Chemical Research, Kyoto University).

Phylogenetic relationships among baculoviruses were analyzed by comparing amino acid sequences of both polyhedrins and granulins using UPGMA method in GENETYX-MAC programs (Software Development Co. Ltd.).

Results and Discussion

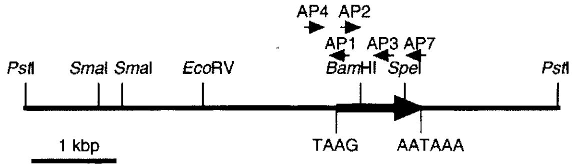
Identification and sequencing of the AnpeNPV polyhedrin gene

AnpeNPV DNA was digested with *Pst*I and the resulting fragments were cloned into pBluescriptII. Among the cloned fragments, the AnpeNPV polyhedrin gene was identified in a 6.6 kbp *Pst*I fragment by Southern blot hybridization using the BmNPV polyhedrin gene probe (data not shown). The fragment was further analyzed by digesting with several restriction endonucleases to make a physical map (Fig. 1a). Then, the nucleotide sequence of the 6.6 kbp *Pst*I fragment was determined. Sequence analysis revealed that the consensus late promoter transcription start signal (TAAG) and translation start codon (ATG) of the polyhedrin gene were located at 192 and 141 nucleotides (nt) from the *Bam*HI site, respectively, and that the translation stop codon (TAG) and putative polyadenylation signal (AATAAA) were located within and at 276 nt from the *Spe*I site, respectively (Fig. 1b). The results indicated that the AnpeNPV polyhedrin gene consisted of a promoter sequence of 51 bp and a coding sequence of 738 bp encoding 245 amino acids (DDBJ/Genbank/EMBLE accession number AB062454), which well agreed with previously reported partial sequences of the AnpeNPV polyhedrin gene region (Yuwen *et al.*, 1987, 1988) except several nucleotide mismatches.

Polyhedrin gene comparison between AnpeNPV and other NPVs

In homology searches for the predicted amino acid sequence of the AnpeNPVpolyhedrin uaing BLAST program, very high identities (more than 95%) were found in those of *Attacus ricini* NPV, *Epiphyas postvittana* NPV, *Choristoneura fumiferana* NPV, *C. rosaceana* NPV, *Rachiplusia ou* MNPV, *Thysanoplusia orichal-*

a.



b.

```
                                                      AP1
             -51 TAAGTACATTGCTGTTATTGTAGCAACTTTGTAGTAAAATTTCCTATAACT -1

  1 ATG CCA GAT TAC TCA TAC CGG CCG ACC ATT GGT CGC ACC TAT GTG TAC GAC AAC AAG TAT 60
  1 Met Pro Asp Tyr Ser Tyr Arg Pro Thr Ile Gly Arg Thr Tyr Val Tyr Asp Asn Lys Tyr 20

                                                         AP2
 61 TAC AAA AAC TTA GGG TCC GTC ATT AAG AAC GCC AAG CGC AAG AAG CAT TTA GTC GAA CAT 120
 21 Tyr Lys Asn Leu Gly Ser Val Ile Lys Asn Ala Lys Arg Lys Lys His Leu Val Glu His 40

                                BamHI
121 GAA GAG GAA GAA AAG CAT TGG GAT CCT TTA GAC AAT TAC ATG GTC GCG GAA GAC CCT TTC 180
 41 Glu Glu Glu Glu Lys His Trp Asp Pro Leu Asp Asn Tyr Met Val Ala Glu Asp Pro Phe 60

181 CTG GGG CCG GGT AAA AAC CAA AAA CTG ACA CTT TTC AAG GAA ATC CGC AAC GTT AAA CCC 240
 61 Leu Gly Pro Gly Lys Asn Gln Lys Leu Thr Leu Phe Lys Glu Ile Arg Asn Val Lys Pro 80

241 GAC ACA ATG AAA CTT ATT GTC AAC TGG AGC GGT AAA GAA TTT CTG CGC GAA ACT TGG ACC 300
 81 Asp Thr Met Lys Leu Ile Val Asn Trp Ser Gly Lys Glu Phe Leu Arg Glu Thr Trp Thr 100

301 CGC TTT GTT GAG GAT AGC TTT CCG ATT GTA AAC GAC CAA GAG GTC ATG GAT GTG TTC CTC 360
101 Arg Phe Val Glu Asp Ser Phe Pro Ile Val Asn Asp Gln Glu Val Met Asp Val Phe Leu 120

361 GTC ATT AAC CTG CGC CCC ACG CGC CCC AAC AGG TGC TAC AAG TTC TTG GCG CAG CAC GCG 420
121 Val Ile Asn Leu Arg Pro Thr Arg Pro Asn Arg Cys Tyr Lys Phe Leu Ala Gln His Ala 140

421 CTC AGA TGG GAC TGC GAT TAC GTG CCG CAC GAG GTA ATC CGC ATT GTG GAG CCA TCC TAC 480
131 Leu Arg Trp Asp Cys Asp Tyr Val Pro His Glu Val Ile Arg Ile Val Glu Pro Ser Tyr 160

                                        AP3
481 GTG GGC ATG AAC AAC GAG TAC AGA ATT AGC CTC GCC AAG AAA GGC GGC GGC TGC CCC ATC 540
161 Val Gly Met Asn Asn Glu Tyr Arg Ile Ser Leu Ala Lys Lys Gly Gly Gly Cys Pro Ile 180

541 ATG AAC ATT CAC AGC GAG TAC ACC AAC TCG TTT GAA TCG TTT GTA AAC CGC GTA ATC TGG 600
181 Met Asn Ile His Ser Glu Tyr Thr Asn Ser Phe Glu Ser Phe Val Asn Arg Val Ile Trp 200

601 GAG AAC TTT TAC AAG CCC ATT GTG TAC ATC GGC ACG GAC TCG GGT GAG GAG GAG GAA ATT 660
201 Glu Asn Phe Tyr Lys Pro Ile Val Tyr Ile Gly Thr Asp Ser Gly Glu Glu Glu Glu Ile 220

661 CTC ATC GAG GTT TCG CTT GTG TTC AAG GTC AAG GAG TTT GCG CCC GAC GCG CCA CTG TTT 720
221 Leu Ile Glu Val Ser Leu Val Phe Lys Val Lys Glu Phe Ala Pro Asp Ala Pro Leu Phe 240

                        SpeI
721 ACT GGC CCC GCG TAC TAG TTCGCCGACAATTCGTTGGCTACTCTTAACAAGTTTTTAACTTTAAATCGTCGG 793
241 Thr Gly Pro Ala Tyr  *  245

794 GCGCGTCAACCGCCTGCACGAAAGCGCGCACGTCGTCGTACCGGCCGGCAAAAATTAGGTCCTCAATTTGTACATAAAA 872

                                                                              AP7
873 TTGCGGCGCGTTCGCCGCGTATAACTCTACGGGCGACTGGCTGTCCAACGCGTTTTCAGTCAGCTGAACGCTCTTGCTA 951

952 AACAAACGCAAATCCATTTTTTGCCTTTTCCACGTTTTCAACCGTTCCTGGGTCTTTTTCCGCAATAAA 1020
```

Fig. 1 Physical map of the 6.6 kbp *Pst*I fragment of AnpeNPV genome (a) and nucleotide sequence of the AnpeNPV polyhedrin gene (b). Positions of consensus late gene motif TAAG and polyadenylation signal AATAAA are indicated. Small arrows (a, b) and a large arrow (a) indicate positions and directions of oligonucleotide primers and polyhedrin coding region, respectively. The predicted amino acid sequence of the polyhedrin is given below the nucleotide sequence (b).

cea MNPV, *Anticarcia gemmatalis* MNPV and *Orgyia pseudotsugata* MNPV (OpMNPV) (Fig. 2).

Two evolutionary distinct groups were identified in lepidopteran NPVs (group I and II) based on the homology of the polyhedrin sequence (Zannot *et al.*, 1993; Hyink *et al.*, 1998). All of NPVs whose polyhedrin sequences showed high homologies to the AnpeNPV polyhedrin sequence belong to the group I. Therefore, it is obvious that the AnpeNPV is a

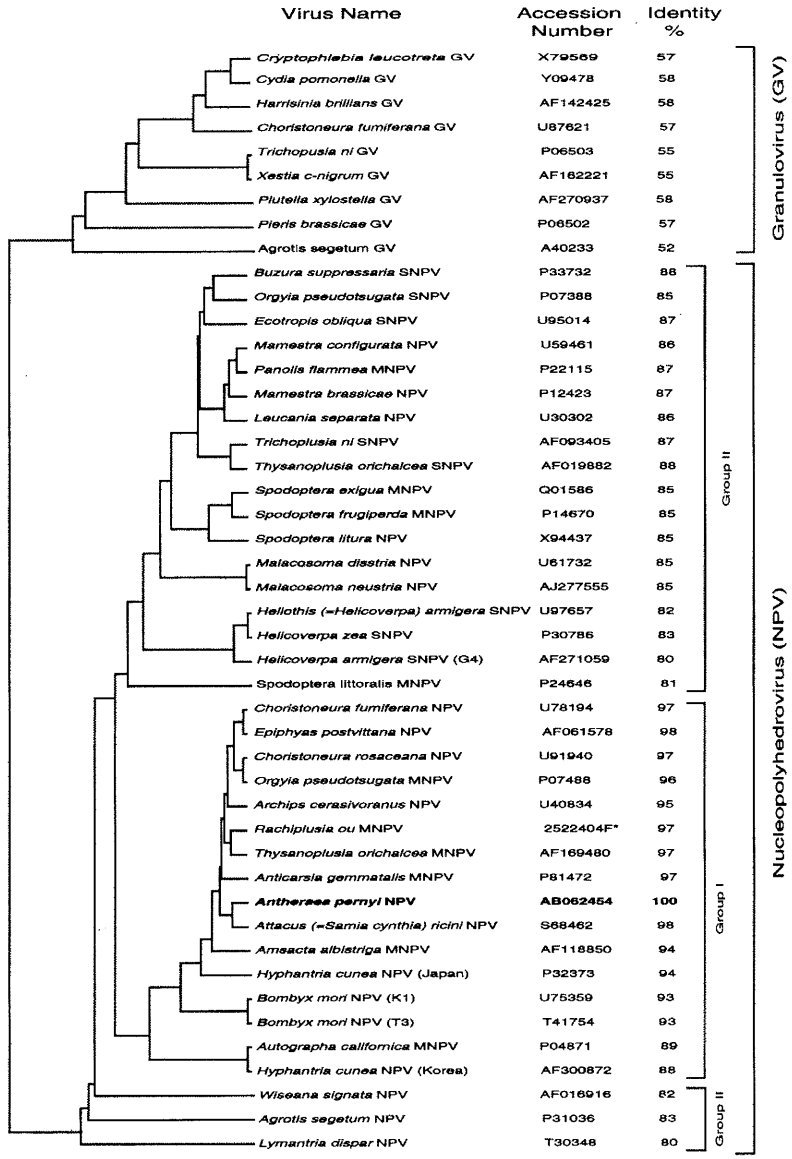


Fig. 2 A phylogenetic tree of baculoviruses based on the amino acid sequences of both polyhedrins and granulins available in the databases. Virus name, accession number in the Genbank database (except *Rachiplusia ou* MNPV, for which code number (*) in PIR database) and amino acid sequence identity (%) to the polyhedrin of the AnpeNPV (in bold) revealed by BLAST search are indicated. The division of baculoviruses, granulovirus and nucleopolyhedrovirus (group I and II), is also indicated.

member of the group I NPV.

In addition, *A. ricini* NPV was most closely related to the AnpeNPV in the phylogenetic tree. Because both *A. pernyi* and *A. ricini* (=*Samia cynthia ricini*) belong to the family Satruniidae, AnpeNPV and ArNPV may be evolutionary isolated when the two insect species were diverged, or may still be variants of the same virus.

Construction of transfer vector pApCH1

A transfer vector for generating recombinant AnpeNPVs expressing heterologous genes under the control of the polyhedrin promoter was constructed as follows (Fig. 3). The pBluescriptII containing the 6.6 kbp *Pst*I fragment (pApPst6.6) was digested with *Eco*RV and self-ligated. Using the resulting plasmid as a tempelate, a 1.8 kbp fragment containing 5' upstream region of the polyhedrin gene was

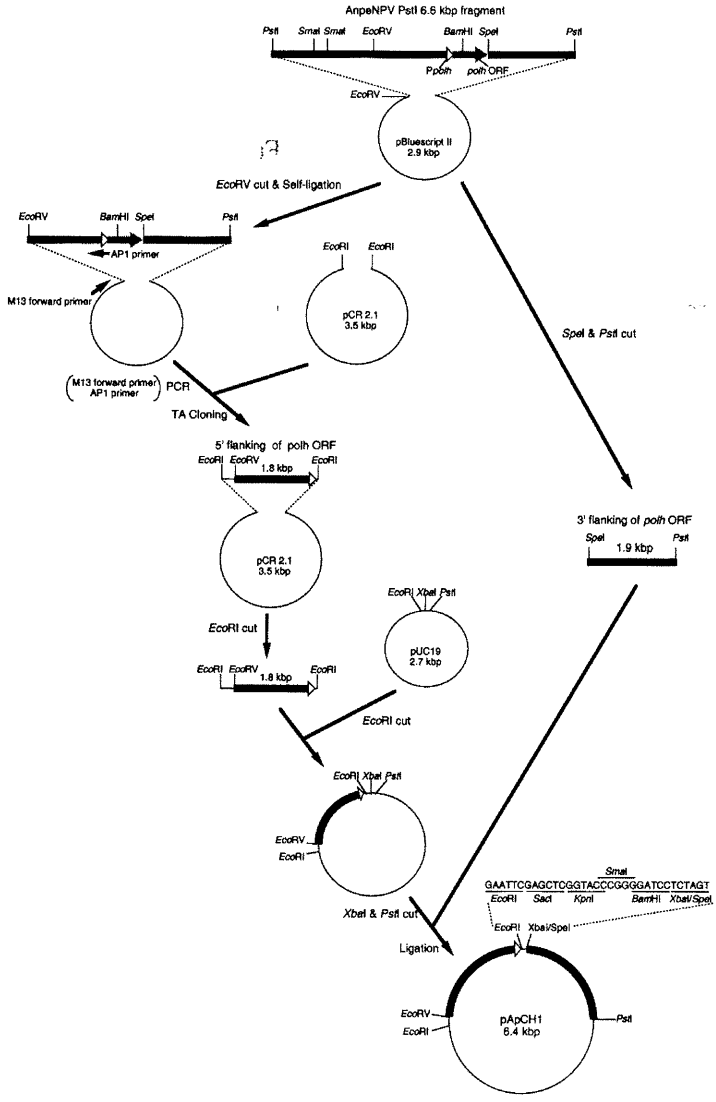


Fig. 3 Schematic diagram of procedure used for the construction of transfer vector pApCH1. The details are described in Results and Discussion.

amplified by PCR with M13 forward and AP1 primers and cloned into pCR2.1 by Original TA Cloning Kit (Invitrogen). The 1.8 kbp fragment excised from pCR2.1 by *Eco*RI digestion was subcloned into pUC19 (pAp1.8). The pAp1.8 was digested with *Xba*I and *Pst*I and then ligated with a 1.9 kbp *Spe*I-*Pst*I fragment containing the 3' downstream region of the polyhedrin gene. Finally, the transfer vector pApCH1, which containing both 5' and 3' flanking regions of the polyhedrin gene and between them 3 unique restriction sites (*Sac*I, *Sma*I and *Bam*HI) for insertion of heterologous genes just downstream of the polyhedrin promoter, was obtained (Fig. 3).

We previously reported that the AnpeNPV replication and plaque-purification had become possible by using the *A. pernyi* cell line, NISES-Anpe-428 (Wang *et al.*, 2000). As described in this paper, we have obtained the transfer vector pApCH1, an important tool for the gene transfer to the AnpeNPV genome. Thus, the preparation of two critical components, cell line and transfer vector, necessary for the establishment of the AnpeNPV expression vector system have been completed. To evaluate the recombinant protein production performance of this newly developed AnpeNPV vector system, especially in diapausing pupae, recombinant AnpeNPVs expuessing a reporter gene have been constructed (Huang *et al.*, 2001).

Acknowledgments

This work was partly supported by Enhancement of Center of Excellence, Special Coordination Funds for Promoting Science and Technology, Science Technology Agency, Japan.

References

Huang, Y. J., X. Y. Wang, S. Miyajima, T. Yoshimura and J. Kobayashi, 2001. Efficiency of *Antheraea pernyi* nucleopolyhedro virus-mediated protein production in both an established cell line and diapausing pupae of *A. pernyi*. Int. J. Wild Silkmoth & Silk, **6**: in press.

Hyink, O., S. Graves, F. M. Fairbairn and V. K. Ward, 1998. Mapping and polyhedrin gene analysis of the *Epiphyas postvittana* nucleopolyhedrovirus genome. J. Gen. Virol., **79**: 2853-2862.

Inoue, H. and S. Hayasaka, 1995. A new cell line separated from the contractile muscle cell line of Chinese oak silkworm, *Antheraea pernyi*. J. Seric. Sci, Jpn., **64**: 79-81.

Maeda, S., T. Kawai, M. Obinata, H. Fujiwara, T. Horiuchi, Y. Saeki, Y. Sato and M. Furusawa, 1985. Production of human α-interferon in silkworm using a baculovirus vector. Nature, **315**: 592-594.

Sambrook, J., E. F. Fritsch and T. Maniatis, 1989. Molecular Cloning: A Laboratory Manual. 2nd ed. Cold Spring Harbor Laboratory Press, New York, NY.

Wang, X. Y., J. Kobayashi, S. Miyajima, H. J. Xie, S. L. Ghi, W. Y. Zheng and R. Q. Ji, 2000. Cloning and property analysis of *Antheraea pernyi* nucleopolyhedrovirus (AnpeNPV). Int. J. Wild Silkmoth and Silk, **5**: 51-56.

Yuwen, H., F. Qi, Z. Chunfa, H. Yunte, L. Guangze, L. Shushan, H. Long and J. Qi, 1987. Location and cloning of polyhedrin gene of *Antheraea pernyi* nuclear polyhedrosis virus. Chinese J. Virol., **3**: 157-162.

Yuwen, H., F. Qi, Z. Chunfa, H. Yunde and L. Guangze, 1988. Sequence of noncoding region at 5' end of polyhedrin gene from *Antheraea pernyi* nuclear polyhedrosis virus (ApNPV). Chinese J. Virol., **4**: 88-91.

Zanotto, P. M., B. D. Kessing and J. E. Maruniak, 1993, Phylogenetic interrelationships amomg baculoviruses: evolutionary rates and host associations. J. Inbertebr. Pathol., **62**: 147-164.