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
http://dx.doi.org/10.1023/A:1008192723971

The characterization and phylogenetic relationship of the trichoplusia ni single capsid nuclear polyhedrosis virus polyhedrin gene

Burtram Clinton Fielding and Sean Davison

Abstract. The polyhedrin gene (*polh*) was identified from the *Trichoplusia ni* (*Tni*) single capsid nuclear polyhedrosis virus (SNPV). An EcoRI fragment containing the truncated polyhedrin gene was detected by hybridization with an AcMNPV expression vector probe; the remaining portion of the gene was amplified by reverse PCR. An open reading frame (ORF) of 741 nucleotides (nt), encoding a putative protein of 246 amino acids (a.a) with M_r 28,780 Da was identified. The 5^t-noncoding region contained the putative late (TAAG) transcription initiation motif. The 3^t end, downstream of the translation stop codon, lacked an obvious putative poly (A) signal. Nucleotide and amino acid homology are greater than 80% to that of *Mamestra brassicae* polyhedrin sequences. Results suggest that *T. ni*SNPV is a member of the group II nuclear polyhedrosis viruses.

Introduction

The Family *Baculoviridae* are invertebrate-specific pathogens that contain a single circular supercoiled double stranded DNA molecule that ranges in size from 80-220 kilobases (1). In excess of 600 insect species are susceptible to baculovirus infection, with a single virus isolate normally restricted to one host or a few closely related species.

The nuclear polyhedrosis virus (NPV) genome is complex and is capable of encoding for at least 150 proteins (2,3), of which polyhedrin is hyperexpressed late in the infection cycle (4). Polyhedrin, a protein approximately 29-kD in size, is the primary constituent of the crystalline matrix surrounding enveloped nucleocapsids or occluded viruses (5,6). Since non- occluded viruses have been found to be rapidly inactivated in the soil and on plant tissues as well as in dead larvae (7), it is accepted that polyhedrin is essential for the persistence of the biological activity of the virus in the environment.

Homology studies have identified polyhedrin as the most conserved baculovirus protein so far characterized. Among lepidopteran baculovirus polyhedrin genes more than 80% sequence identity has been reported. Additionally, in excess of 50% identity has been reported between lepidopteran baculovirus granulin and polyhedrin sequences (8). The introduction of foreign genes *in lieu* of the polyhedrin gene to increase the efficacy of the virus, has resulted in varying success (9-12).

Based on phylogenetic studies using polyhedrin sequences, NPVs have been classified into two groups, namely Group I and Group II. This suggests NPV evolution comprises of two distinct branches (13,14). However, conclusive support for this hypo- thesis could only be gained through continued molecular analysis of various other conserved baculovirus genes.

In this study we characterize the polyhedrin gene of a nuclear polyhedrosis virus isolated from a field population of *Trichoplusia ni* as a prelude to the introduction of a foreign gene to increase the efficacy of the virus. The *Eco*RI-V fragment of 1.35 kbs, shown to hybridize to a portion of the AcMNPV polyhedrin gene coding region, was cloned and sequenced. The predicted polyhedrin gene product is compared to 9 other occlusion proteins and a phylogenetic relationship is proposed. The nucleotide and amino acid sequences were most homologous (83% and 94%, respectively) to that of *Mamestra brassicae* MNPV. Sequence comparisons suggests that *Tni*SNPV is a member of the previously defined group II nuclear polyhedrosis viruses.

Methods

Viruses and Insect Hosts

Virus was isolated from diseased *Trichoplusia ni* (Noctuidae: Lepidoptera) larvae collected in the Eastern Cape, South Africa. Propagation of virus was done using second instar larvae reared at 26^{O}C (artificial 12 h day-night cycle; 65% humidity) on an artificial lepidopteran diet. Virus killed larvae were collected and virus capsules isolated and purified as previously described (Crook, personal communication).

DNA Extraction and Digestion

Purified viral capsules were alkali lysed and phenol extracted. Extracted DNA was digested with *Eco*RI, *Pst*II, *Dra*I and *Xba*I according to the manufacturers instructions (Boehringer Mannheim) for 1.5 h at 37^{O}C.

Clone Construction and Location of Polyhedrin Gene

*Eco*RI restriction fragments were cloned into the compatible site of the multiple cloning region of SK $^{+}$ -Bluescript and transformed into supercompetent *E. coli* JM 105 cells. Following blue-white selection, plasmid DNA was extracted by the mini-prep method (15). Positive clones were *Eco*RI- digested, run on 0.8% agarose gels, visualized and transferred to Hybond-N nitrocellulose membranes (*Amersham,* Life Sciences) by capillary Southern Blot techniques; these were routinely screened for the presence of the polyhedrin gene by hybridization to a labeled probe. The probe consisted of an expression vector (pAcRP23), containing a portion of the AcMNPV polyhedrin gene, labeled using the DIG Nick Translation Mix kit (Boehringer Mannheim) according to the manufacturers' specifications. Prehybridizations and hybridizations were performed under stringent conditions, at 68^{O}C for 2 and 18 h respectively. Detection of results were done using the DIG-DNA Labeling and Detection Kit according to the manufacturers instructions (Boehringer Mannheim) and visualized after 18 h.

Sequencing and Homology Searches

Sequencing of the *Eco*RI-V fragment (clone pSKE1V) was done by automated sequencing using a Pharmacia ALF/Express automated sequencer, with universal primers. Amino acid and nucleotide sequences were compared with those in GenBank at the National Centre for Biotechnology by using the Advanced Blast Search Server. The initial results indicated an incomplete polyhedrin ORF at the 3t end of the sequence. PCR primers were designed and the missing sequence amplified by reverse PCR methods (15).

Results and Discussion

The TniSNPV polyhedrin gene was found by using an expression vector containing the AcNPV polyhedrin gene to probe an *Eco*R1 DNA library from TniSNPV by Southern blotting. This identified a positive signal from a 1.35 kb clone. End sequencing of this clone using universal sequencing primers found it to be similar to other baculovirus polyhedrin genes at both ends, however the 3t end was found to be truncated. To find the missing sequence from the truncated end, reverse PCR primers were designed. By using reverse PCR on Hindlll digested genomic DNA, it was possible to amplify the remaining portion of the gene. Sequencing of the positive *Eco*R1 clone, subclones from it and the reverse PCR product was achieved using the sequencing strategy indicated in Fig. 1.

Fig. 2 shows the nucleotide sequence of a 1551-bp region of the TniSNPV DNA that contains the complete coding region for the polyhedrin. The sequences obtained were confirmed by at least two overlapping sequences. Two restriction endonuclease cleavage sites (*Dra*l and *Hind*lll) had been previously located by restriction enzyme analysis and were correctly predicted by the nucleotide sequence. The sequencing data from the 1551 nucleotides sequenced suggests an ORF of 741 nucleotides encoding a putative protein of 247 amino acids with an estimated size of 28.89 kDa (Fig. 2).

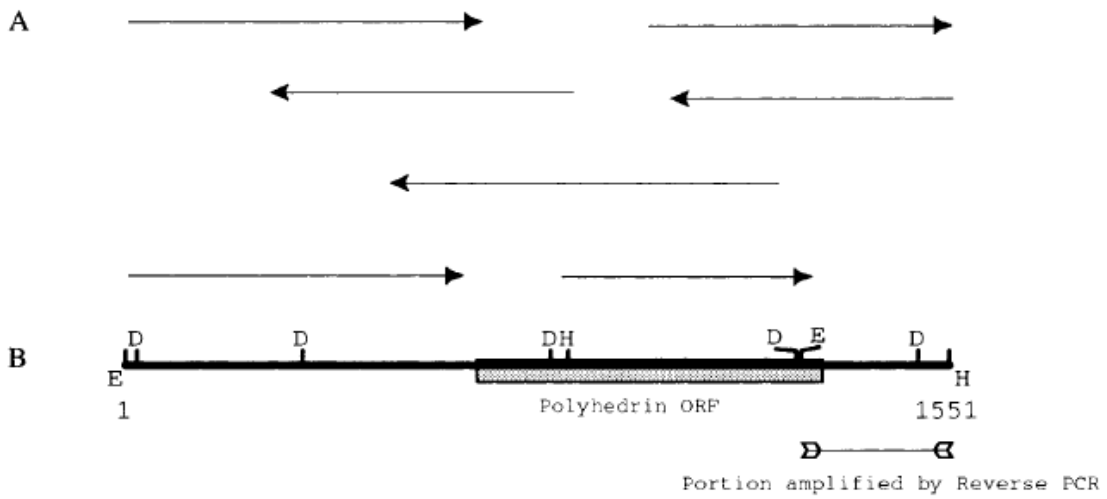

Fig. 1. Enzyme map, sequencing strategy and ORF distribution of the TniSNPV DNA region. A. Sequencing strategy. Deletion clones are represented by arrows starting at the beginning of the insert and pointing in the sequencing direction. B. Map of selected restriction enzyme recognition sites; E = *Eco*RI; H = *Hind*III; D = DraI.

The nontranslated leader sequence of TniSNPV was most similar to that of *Panolis flammae* MNPV (92%) and *Mamestra brassicae* MNPV (89%). Overall, elements in the noncoding leader sequence were found to be very similar to those of other NPV polyhedrin genes (Fig. 3a). Similarities included an adenine at position – 3 and the possible duplication of the TTNGTA motif (16). The non-coding leader sequence containing the baculovirus late transcription initiation motif (TAAG) is AT rich consistent with other baculovirus late expression genes (8). Comparison of the nucleotide sequence of the ORF shows the TniSNPV to be most closely related to the polyhedrin sequences of *Mamestra brassicae* NPV (strain Oxford) (83% identity) and *Panolis flammae* MNPV (80%). Also, homology values for the prototype baculoviruses, *Orgyia pseudotsugata* SNPV (80%) and *Autographa californica* MNPV (84%) were in excess of 80.

At the amino acid level the greatest homologies were found to *Mamestra brassicae* NPV (strain Oxford), *Panolis flammae* MNPV, *Orgyia pseudotsu- gata* SNPV, and *Autographa californica* MNPV, with homologies of 94%, 93%, 93% and 92% respectively (Fig. 3b). Attempts have been made to design a phylogenetic relationship between the different baculoviruses based on polyhedrin and ecdysteroid glucosyltransferase gene sequences (16,17). On this basis it has been proposed that the baculoviruses fit into two phylogenetically diverse groups. Group I is characterised by a N-terminal peptide sequence: MP (D/N) YS, a length of 245 amino acids and an obvious poly (A) terminal motif. Group II is characterised by a N-terminal peptide sequence MYT (R/P) YS, a length of 246 amino acids and some members of this group lack a poly (A) tail. Group II has been further divided into two clades based on sequence similarity (13). Consistent with the characteristics of group II, the TniSNPV was found to have a MYTRYS polypeptide sequence at the start of the ORF and a length of 246 amino acids, also there was no apparent poly (A) tail. Previous reports suggest that the appearance of a poly (A) tail is a defining characteristic of group I viruses (16). Since no poly A tail was found here it is further evidence for TniSNPV being in group II.

A knowledge of the phylogenetic relationships among baculoviruses is highly relevant to understanding the virus host range and the development of biological control agents. Gene sequence data relating to polyhedrin genes is providing an important tool for discerning phylogenetic relationships. Such analysis has clearly indicated that NPV evolution is divided into two distinct branches. Although the evidence suggests that TniSNPV is phylogenetically related to the other group II NPVs, in order to assess phylogenetic relationships more than one gene family needs to be studied (17). It is therefore necessary for more baculovirus genes to be characterised in order to establish accurate phylogenetic relationships.

Acknowledgments

This work was funded by the Foundation for Research Development. We are grateful for sequencing carried out by Di James from the Department of Microbiology at the University of Cape Town.

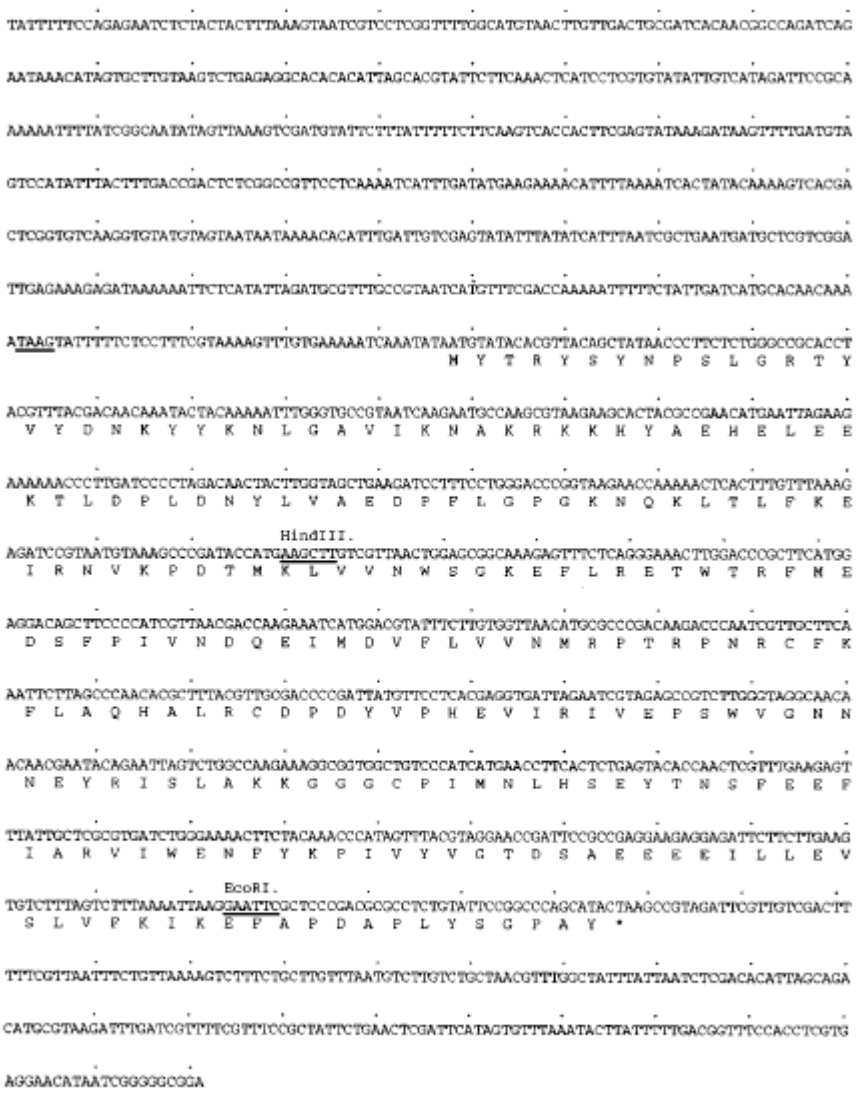

Fig. 2. Nucleotide and predicted amino acid sequences of the *Tni*SNPV polyhedrin gene and flanking regions. Restriction recognition sites for *Hind*III and *Eco*RI are indicated. The putative late translation initiation motif-core (TAAG) is underlined. No obvious poly (A) motif is present.

				+1
MbMNPV	AATG	**TAAG**	TAA	TTTTCTCCTTTCGTAGAAGATTGTGAAAAATAAAATATAATG
TniSNPV	AAAA	**TAAG**	TAT	TTTTCTCCTTTCGTAAAAGTTTGTGAAAAATCAAATATAATG
PfMNPV	AATG	**TAAG**	TAA	TTTTCTCCTTTCGTAGAAGATTGTGAAAAATAAAATATAATG
SeMNPV	ATTG	**TAAG**	TAA	TTTTTTCCTTTCGTAAAACATTGTGAAAAAATAAATATAATG
SfMNPV	ATTG	**TAAG**	TAA	TTTTTTCCTTTCGTAAAACATTGTGAAAAAATAAATATAATG
AcMNPV	TAAA	**TAAG**	TAT	TTTACTGTTTTCGTAACAGTTTTGTAATAAAAAAACCTATAAATATG
BmMNPV	TAAA	**TAAG**	TAT	TTTACTGTTTTCGTAACAGTTTTGTAATAAAAAAACCTATAAATATG
AgMNPV	TTAA	**TAAG**	TAT	TTTGCTGTTATTGTAGCAACTTTGTAGTAAAATTTGCTATAACTATG
OpMNPV	TTAA	**TAAG**	TAA	TTTCCTGTTATTGTAACAATTTTGTAAAAAAATTTCCTATAACCATG

Fig. 3a. Alignment of the noncoding leader sequences from selected baculoviruses (16). The adenosine of the translation initiation codon (ATG) is considered as +1. Only numbering for MbNPV is shown. The conserved transcription initiation motif is in bold and the possible duplicated element underlined.

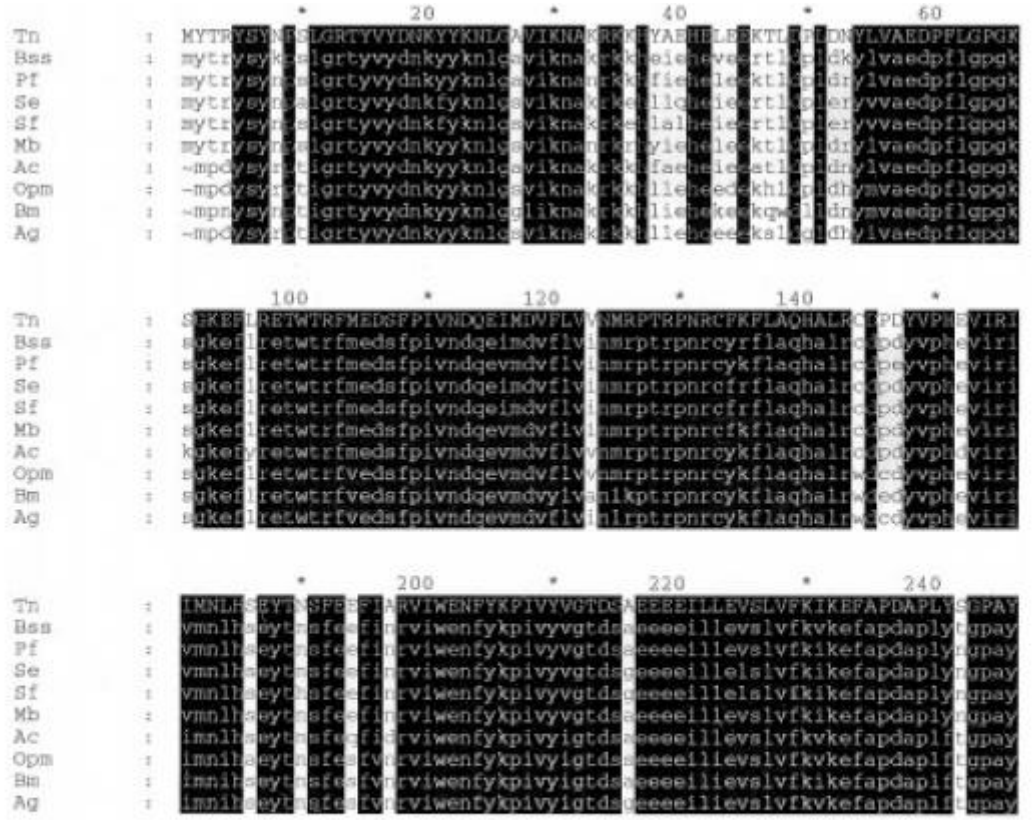

Fig. 3b. Alignment of amino acid sequences of TniSNPV ploy hedrin, to five Group II and four Group I NPVs. Dashes were inserted to align the sequences. The shaded areas indicate homologous and functionally substitutable amino acids between all NPVs. The sequences aligned were as follows: Tn, *Trichoplusia* ni SNPV; Bss, *Buzura superessaria* SNPV (18); Pf, *Panolis flammea* MNPV(19); Ag, *Anticarsia gemmetalis* MNPV (14); Se, *Spodoptera exigua* MNPV (20), Sf, *Spodoptera frugiperda* MNPV (21); Mb, *Mamnestra brassicae* MNPV (22); Ac, *Autographa californica* MNPV (23); Opm, *Orgyia pseudotsugata* MNPV (24) and *Bm*, *Bombyx* NPV (25).

References

1. Miller L.K., Ann Rev Micro 42, 177-199, 1988.
2. Kamita S.G., Majima K., and Maeda S., J Virol 67(1), 445-463, 1993.
3. Fredirici B.A. and Maddox J.V., BioScience 46(6), 410-422, 1996.
4. Vlak J.M., Klinkenberg F.A., Zaal K.J.M., Usmany M., KlingeRoode E.C., Geervliet J.B.F., Roosien J., and Van Lent J.W.M., J Gen Virol 69, 765-776, 1988.
5. Gombart A.F., Pearson M.N., Rorhmann G.F., and Beaudreau G.S, Virol 169, 182-193, 1989.
6. Weyer U. and Possee R.D., J Gen Virol 70, 203±208, 1989.
7. Blissard G.W. and Rorhman G.F., Ann Rev Entomol 35, 127-155, 1990.
8. Rohrmann G.F., J Gen Virol 73, 749-761, 1992.
9. Merryweather A.T., Weyer U., Harris M.P.G., Hirst M., Booth T., and Possee R.D., J Virol 71, 1535-1544, 1990.
10. Tomalski D.T. and Miller L.K., Nature 352, 82-85, 1991.
11. Carbonell L.F., Hodge M.R., Tomalski M.D., and Miller L.K., Gene 73, 409-418, 1988.
12. Hammock B.D., Bonning B.C., Possee R.D., Hanzlik T.N., and Maeda S., Nature 344, 458-461, 1990.
13. Cowan P., Bulach D., Goodge K., Robertson A., and Tribe D.E., J Gen Virol 75, 3211-3218, 1994.
14. Zanotto P.M. de A., Kessing B.D., and Maruniak J.E., J Insect Pathol 62, 147-164, 1993.
15. Sambrook J., Fritsch E.F., and Maniatis T., Molecular Cloning, A Laboratory Manual. Second Edition, Cold Spring Harbor Laboratory Press, 1987.
16. Zanotto P.M. de A., Sampaio M.J.A., Johnson D.W., Rocha T.L., and Maruniak J.E., J Gen Virol 73, 1049-1056, 1992.
17. Clarke E.E., Tristem M., Cory J.S., and O'Reilly D.R., J Gen Virol 77, 2865-2871, 1996.
18. Hu Z.H., Liu M.F., Jin F., Wang Z.X., Liu X.Y., Li M.J., Liang B.F., and Xie T.E., J Gen Virol 74, 1617-1620, 1983.
19. Oakey R., Cameron I.R., Davis B., Davis E., and Possee R.D., J Gen Virol 70, 769-775, 1989.
20. Van Strien E.A., Zuidema D., Goldbach R.W., and Vlak J.M., J Gen Virol 73, 2813-2821, 1992.
21. Gonzalez M.A., Smith D.E., and Summers M.D., J Virol 170, 160-175, 1989.
22. Cameron I.R. and Possee R.D., Virus Res 12, 183-199, 1989.
23. Hooft Van Iddekinge B.J.L., Smith G.E., and Summers M.D., Virology 131, 561-565, 1983.
24. Leisy D., Nesson M., Pearson M., Rohrmann G., and Beaudreau G.S., Virology 153, 157-167, 1986.
25. Iatrou K., Ito K., and Witkiewicz H.J., J Virol 54, 436-445, 1985.