Kumar, P., et al. (2007). The nonstructural protein 8 (nsp8) of the SARS coronavirus interacts with its ORF6 accessory protein. Virology, 366(2): 293-303
http://dx.doi.org/10.1016/j.virol.2007.04.029

UNIVERSITY of the WESTERN CAPE

The Genetic Organization of a 2,966 basepair DNA Fragment of a Single Capsid Nucleopolyhedrovirus Isolated from Trichoplusia ni*

Burtram Clinton Fielding, Sehaam Khan, Weizhou Wang, Courtney Kruger, Rayaana Abrahams and Sean Davison

Abstract
In order to investigate the genomic organization of the *Trichoplusia ni* Single Capsid Nucleopolyhedrovirus (TnSNPV), a 2,966 basepairs (bp) genomic fragment was sequenced. The fragment was found to contain five open reading frames (ORFs) homologous to baculovirus genes, including p26, fibrillin (p10), AcMNPV ORF-29, late expression factor 6 (*lef*-6) and the C-terminal portion of p*74*, on either strand of DNA. Predicted amino acid sequences for the ORFs were compared and identity values of between 12% and 54% were observed. TnSNPV has previously been tentatively identified as a member of the Group II NPVs. Clustering and arrangement of the TnSNPV genes were similar to the clustering reported for SeMNPV, confirming TnSNPV as a Group II NPV.

Introduction
Two genera make up the family Baculoviridae, the *Nucleopolyhedrovirus* (NPV) and *Granulovirus* (GV) [1], all of which infect arthropods [2]. GVs have been isolated exclusively from lepidopteran larvae [3-5]. On the other hand, although the vast majority of NPVs infect insects, a few infect crustaceans [6].

The cabbage looper (*Trichoplusia ni*) is an agri- cultural pest of many different economically impor- tant crops. Preliminary tests have shown the South African isolate of *Trichoplusia ni* Single Capsid Nucleopolyhedrovirus (TnSNPV) to be highly patho- genic to *Trichoplusia ni*, killing larvae two days after application. TnSNPV, a member of the family *Baculoviridae*, possesses a genome about 160 kb in size. The virus is not well characterized with only the polyhedrin [7], *plO* [8] and *ie*-1 [9] genes sequenced and analysed to date. To gain some insight into the reason why the TnSNPV genome is 30 kb larger than AcMNPV, genome libraries are being sequenced and mapped. In this paper, the sequence of a 2,966 bp fragment of the TnSNPV genome was determined and analysed. Analysis of this fragment was used to define differences between baculoviruses and to provide information for the further investigation of TnSNPV phylogeny.

Materials and Methods

TnSNPV was propagated in field-collected late fourth or early fifth instar *Trichoplusia ni* larvae. Larvae were reared on an artificial lepidoptera diet, with a synthetic 12 : 12 h day-night cycle at 65% humidity. TnSNPV was purified and purified OBs were used to extract genomic DNA, which was digested with various enzymes and analysed by agarose gel electrophoresis [10].

Cloning and Sequencing

Genomic DNA was digested with *Pst*I (Boehringer Mannheim) and shotgun-cloned into the compatible restriction sites of pBluescript SK^+ to construct a partial library [10]. A 2,966 bp *Pst*I-*Hind*III subclone was selected for sequencing. Automated sequencing of the end-termini revealed the presence of the C-terminal portion of the TnSNPV *p74* gene and a portion of a *lef*-6 homologue. Subsequently, subclones were constructed and sequenced using universal primers with a Pharmacia ALF/Express automated sequencer.

Computer Analysis

Sequence analysis was performed using the GCG computer program. Conceptual amino acid sequences were compared with sequence homologues at GenBank/EMBL using the Advanced Blast Search Server [11]. Sequence alignments were done using CLUSTAL X [12] and GENEDOC software [13] was used for homology shading of the aligned amino acid sequences. Baculovirus sequences used in the com- parative analysis were (GenBank accession numbers included): HaSNPV, *Heliothis armigera* SNPV (NC003094); AcMNPV, *Autographa californica* MNPV (L22858); BmNPV, *Bombyx mori* NPV (L33 180); SeMNPV, *Spodoptera exigua* MNPV (AF169823); LdMNPV, *Lymantria dispar* MNPV (AF081810); OpMNPV, *Orgyia pseudotsugata* MNPV (T10403), SlituraNPV, *Spodoptera litura* NPV (NC0003102); Slitt, *Spodoptera littoralis* NPV (X99376); CfNPV, *Choristineura fumiferana* MNPV (M97904); XcGV, *Xestia c-nigrum* GV (U70896); PxGV, *Plutella xylostella* GV (NC002593); Buzu- SNPV, *Buzura suppressaria* SNPV (AAC77813); HzSNPV, *Heliothis zea* SNPV (NC002593); CpGV, *Cydia pomonella* GV (NC002816); EpNPV, *Epiphyas postvittana* NPV (NC003083).

Results and Discussion

Sequence Analysis

Five ORFs homologous to baculovirus proteins were identified within the fragment. These included *p26*, AcMNPV ORF29-homologue, *plO*, *lef*-6 and the C-terminal portion of *p74*. The nucleotide and putative amino acid sequences of the ORFs are presented in Fig. 1.

Lef-6

Lef-6 has been shown to be involved in the expression of both the late major capsid protein (*vp39*) and the very late protein, polyhedrin (*Polh*). It is believed that *lef*-6 functions either as a *trans*-acting transcriptional or translational activator, or as a primary or accessory replication factor [14,15].

The TnSNPV *lef*-6 homologue encoded for a potential protein 158 amino acids in length. A potential late transcription initiation motif (ATAAG) was located 13 nt upstream of the putative translation start codon (ATG). No obvious polyadenylation signal was found downstream of the putative translation stop codon (TAA). The 474 nt ORF encodes for a putative protein with M_r = 18.29 kDa. Overall, with the exception of AcMNPV and BmNPV, low amino acid sequence identity was observed between the different baculoviruses compared. The conceptual TnSNPV *lef*-6 amino acid sequence was most homo- logous to HzSNPV and HaSNPV, with 39% sequence identity and 56% similarity. Two conserved amino acid sequence motifs (longer than two amino acids) appear to be present in the traditional Groups I and II NPVs: K(R/K)F and TRK (Fig. 2).

p74

During the baculovirus replication cycle, two phenotypically distinct viruses are produced [16]. One of these phenotypes, Occlusion Derived Virus (ODVs), is responsible for the horizontal transmission of virus between insects. After ingestion, ODVs are dissolved by the alkaline midgut, releasing occlusion bodies (OBs). Polypeptide p74 appears to be important for the successful infection of insects, but not for cell culture [17]. Peptide p74 is not glycosylated and it has been shown that expression of the AcMNPV *p74* gene is essential for the production of infectious OBs [17,18]. The peptide appears to be a structural polypeptide of OBs, most probably associated with the outside surface of the virion envelope [19]. Although it is expressed at low levels late in the AcMNPV infection cycle, no conventional late transcription initiation motif (TAAG) is present [20]. The C-terminal portion of a TnSNPV *p74* homologue was identified and compared to other baculovirus p74 proteins. The partial TnSNPV gene was most homologous to the SeMNPV *p74* gene (54% identity and 73% similarity). As expected, the GV *p74* homologue showed relatively low identity values (23-26%) to the TnSNPV *p74* gene. TnSNPV *p74* is in the opposite orientation to the *plO* homologue identified. A possible poly-A motif was identified 7 nt downstream of the putative stop codon (TAA). Several conserved motifs were observed in both the NPV and GV *p74* homologues (black shaded areas in Fig. 3), the largest being (I/V)NS(N/D)GQ(M/L)(L/I).

p26

Polypeptide p26 is transcribed as early as 18 h post-infection (hpi) in the AcMNPV infection cycle [21]. The transcripts are synthesized by the host RNA polymerase II throughout the infection cycle [22], accumulating in the cytoplasm. The peptide is normally present in the membrane fraction of budded virus [23] and has been shown to be non-essential for *in vitro* replication [24]. AcMNPV *p26* gene produces two transcripts: a minor transcript terminating upstream of the p10 ORF and a major transcript transcribed through and co-terminating with the *plO* gene [25].

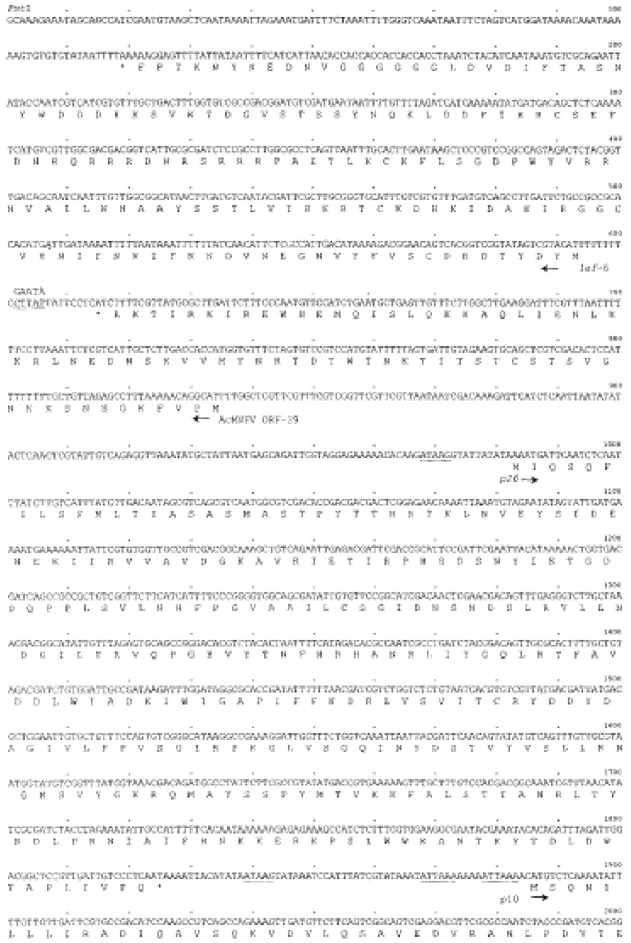

Fig l Nucleotide sequence of the TnSNPV genomic fragment. The predicted amino acid sequence is represented by the one letter code designation below the nucleotide sequence. The ORFs and their direction of transcription are shown. Putative TAAG motifs for late gene transcription and putative poly A motifs are underlined. Unique restriction endonuclease sites are indicated. Signals on the complementary strand of the DNA are underlined and their complementary sequences are shown.

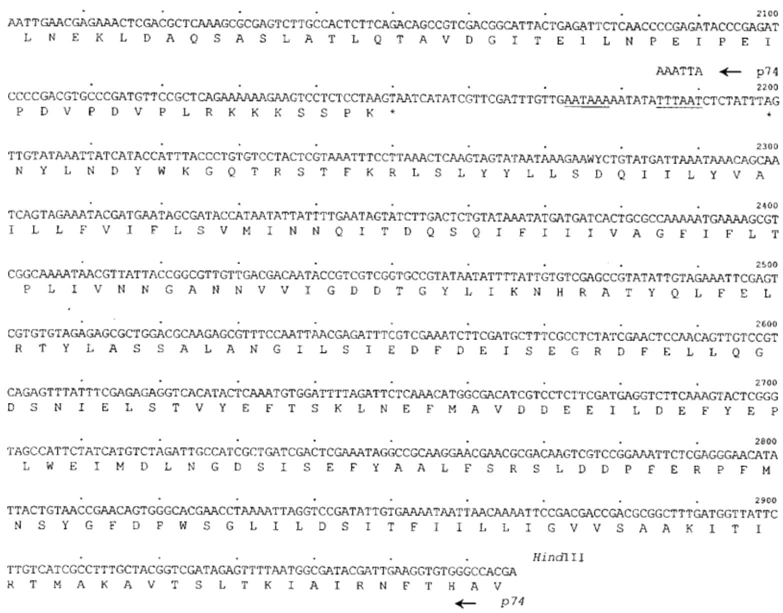

Fig l Continued.

Unlike other baculoviruses sequenced to date, SeMPNV possesses two *p26* homologues. It would appear as though the two copies were acquired either independently from different sources, or that the one diverged from the other and then rearranged following duplication [26].

Although various *p26* baculovirus genes have been completely sequenced, the exact function has not been determined. The TnSNPV putative *p26* gene appears to be the largest sequenced to date. TnSNPV *p26* has previously been partially characterized [8]. Amino acid sequence data was most homologous (35% identity and 51% similarity) to Se2MNPV p26. The ORF of 840 nt encodes for a possible protein 280 amino acids in length, with a putative mass of 31.71 kDa. Upstream of the p26 ORF a possible late promoter motif (ATAAG) is present at position -17 nt relative to the putative ATG start codon. Further upstream, at position -230 and -251 nt, two characteristic early gene mRNA start sites (CAGT) were identified [27]. However, these putative start sites were not preceded by perfect TATA boxes as described by [28]. Downstream of the p26 ORF, two possible poly-A motifs were present at positions 41 and 51 nt relative to the putative stop codon TAA. The conserved p26 amino acid sequences HQPPGV, GAPI, LVSVVT, SVYG, QLPY are shown in Fig. 4 [27].

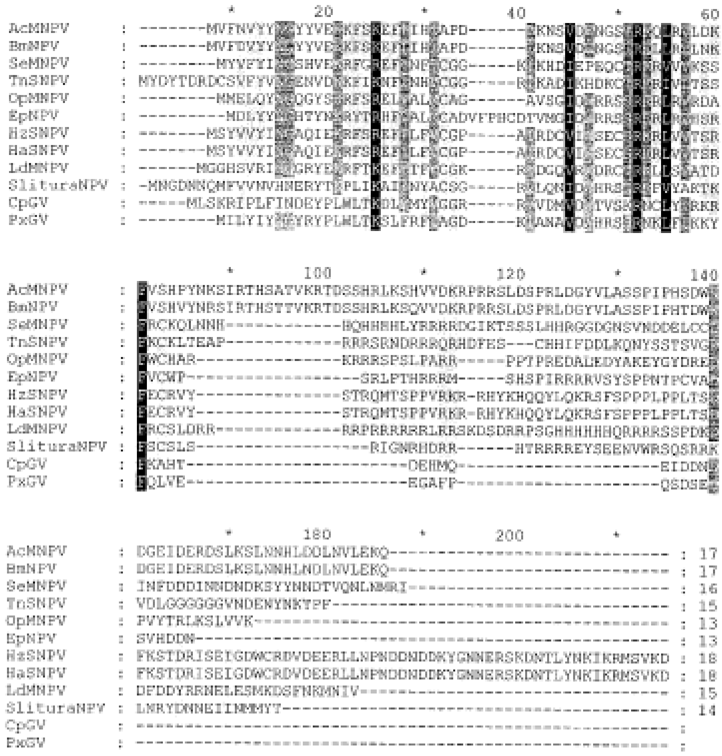

Fig 2 Alignment of the deduced amino acid sequences of putative *lef* 6 genes from 12 baculoviruses was done using CLUSTAL X. Dashes were introduced to align the sequences. Black shaded areas indicate 100% homology; deep grey 80% or more; light grey 60% or more.

However, it appears as though these amino acid sequences are less conserved than previously reported. Only two motifs seem to be highly conserved, i.e. (Y/F)PG and (I/V)S(V/L)(V/I)(T/S) (Fig. 4).

AcMNP6 O7F29-Homologue

The function of the AcMNPV ORF29-homologue is not known. An ORF of 219 nt encoding for a protein of 73 amino acids was identified, encoding for a putative protein of 8.48 kDa. Two conserved motifs were present (shaded in black): (Q/E)HW/FE(R/K)I and (I/L/V)TK (Fig. 5). The TnSNPV ORF was similar in size to the ORFs of AcMNPV and OpMNPV. Comparison of TnSNPV to five homologues showed low identity values. The AcMNPV ORF29-homologue primary amino acid sequence was not well conserved between the various NPVs. The TnSNPV AcMNPV ORF29-homologue was most homologous to SeMNPV with an identity value of 29% and a similarity value of 38%.

p10

This non-structural peptide is expressed very late in the infection cycle. Although the gene is conserved between different baculoviruses, primary amino acid structures differ significantly.

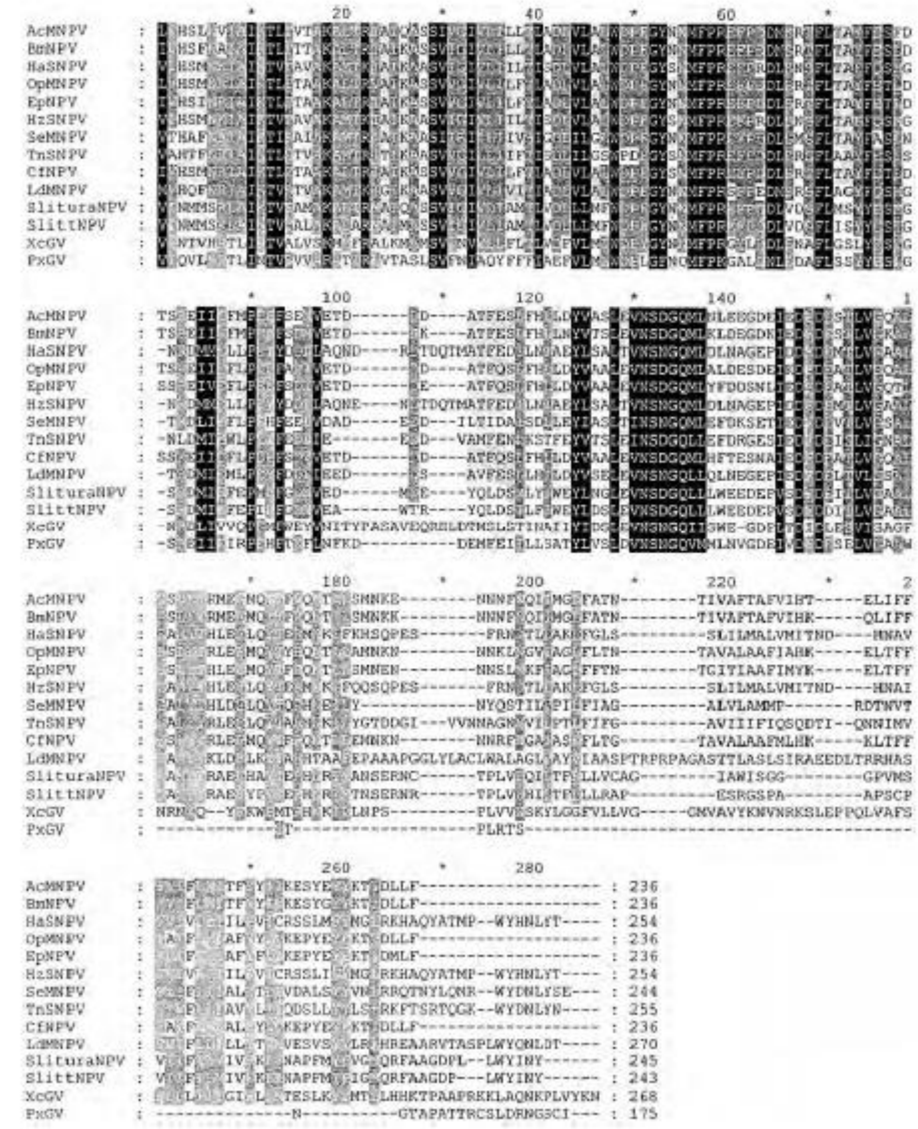

Fig. 3. Alignment of the deduced amino acid sequences of putative baculovirus *p74* genes was done using CLUSTAL X. Dashes were introduced to align the sequences. Black shaded areas indicate 100% homology; deep grey 80% or more; light grey 60% or more.

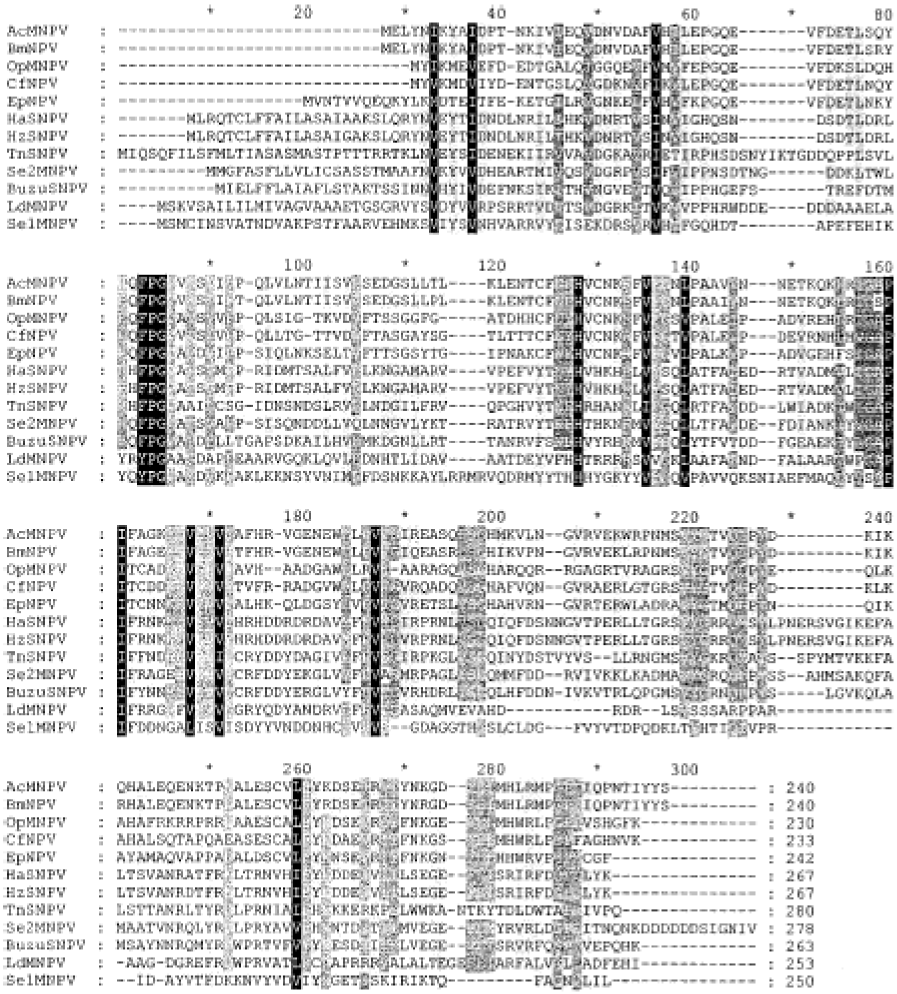

Fig. 4. Alignment of the deduced amino acid sequences of putative *p26* genes from 12 baculoviruses was done using CLUSTAL X. Dashes were introduced to align the sequences. Black shaded areas indicate 100% homology; deep grey 80% or more; light grey 60% or more. Se1MNPV (ORF81) and Se2MNPV (ORF128) were identified from one virus.

Particular segments of the gene confer unique functions on the virus [29]. This suggests that the gene evolved from an ancestor gene faster than other baculovirus genes. The TnSNPV *plO* gene has previously been identified on the *Eco*RI-V fragment and was subsequently characterized [8].

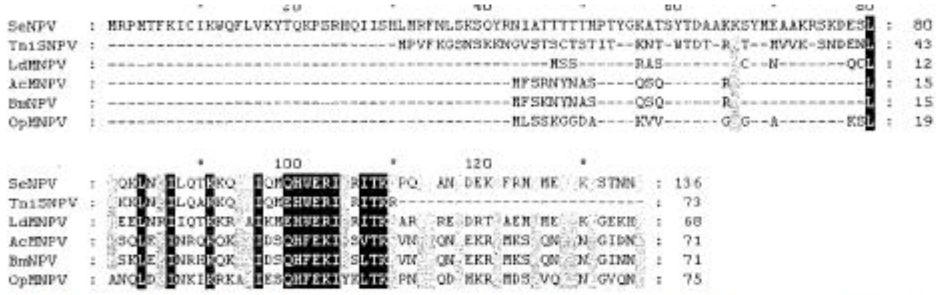

Fig. 5. Alignment of the deduced amino acid sequences of AcMNPV ORF 29 like homologues using CLUSTAL X. Dashes were introduced to align the sequences. Black shaded areas indicate 100% homology; deep grey 80% or more; light grey 60% or more.

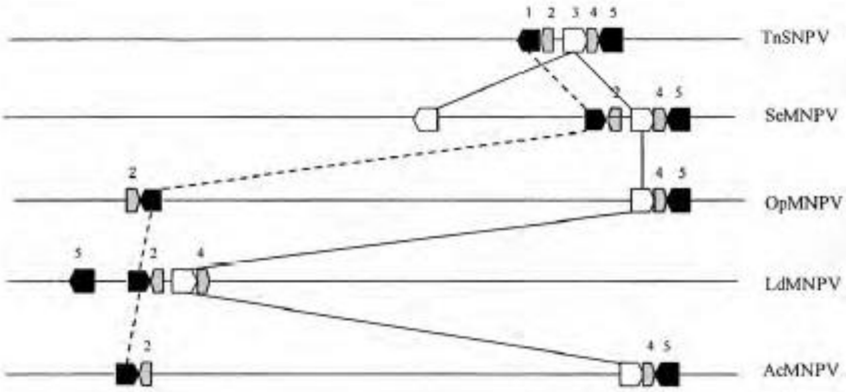

Fig. 6. Comparison of the genomes of TnSNPV, AcMNPV, LdNPV, OpMNPV and SeNPV. Arrows represent the different ORFs and their respective directions of transcription (not drawn to scale). The *lef 6* (1) homologues are linked by dashed lines and the *p26* (3) homologues by solid lines. (2) AcMNPV ORF 29, (4) *p10*, (5) *p74*.

Arrangement of Genes in the Genome

Mechanisms involved in gene re-arrangement are not fully understood. However, it is known that the arrangement of genes of Group I NPVs is fairly conserved. On the other hand, the gene arrangement of Group II NPVs is less conserved, indicating a more diverse group [30]. The cluster of genes identified (*lef*-6-AcMNPV ORF29- homologue-*p26*-*plO*-*p74*) was fairly well conserved between the different baculovirus genomes. With the exception of the transcriptional direction of the TnSNPV *lef*-6, the arrangement and direction of transcription of these genes in SeMNPV and TnSNPV were identical, indicating a possible common baculoviral ancestor. The entire LdMNPV gene cluster showed re-arrangement and a shift to the left of the genome. Gene arrangement and clustering of AcMNPV and OpMNPV are identical, with the gene duplex *lef*-6-AcMNPV ORF29-homologue shifted to the left (Fig. 6). Gene arrangement and gene homology comparisons con- firmed the earlier speculation that TnSNPV was a Group II NPV [7].

Acknowledgements

This research was funded by the National Research
Foundation (NRF, South Africa) and the South African Protea Producers and Exporters Association (SAPPEX).

References

1. Hu Z.H., Broer R., Westerlaken J., Martens J.W., Jin F., Jehle J.A., Wang L.M., and Vlak J.M., Virus Res 47, 91-97, 1997.
2. Volkman L.E., Blissard G.W., Friesen P., Keddie B.A., Possee R., and Theilmann D.A., Virus Taxonomy, Sixth Report of the International Committee on Taxonomy of Viruses. Springer Verlag, Vienna, New York, 1995.
3. Blissard G.W. and Rohrmann G.F., An Rev Ento 35, 127-155, 1990.
4. Tweeten K.A., Bulla L.A. Jr and Consigli R.A., Micro Rev 45, 379-408, 1981.
5. Stairs G.R., Parrish W.B., Briggs T.D., and Allietta M., Virology 3O, 583-584, 1966.
6. Washburn J.O., Lyons E.H., Haas Stapleton E.J., and Volkman L.E., J Virol 73(l), 411-416, 1999.
7. Fielding B.C. and Davison P.S., Virus Genes l9(l), 67-72, 1999.
8. Fielding B.C. and Davison P.S., Virus Genes 2O(2), 189-192, 2000.
9. Wang W., Fielding B.C., Leat N., and Davison S., Virus Genes 23(l), 49-58, 2001.
10. Sambrook J., Fritsch E.F. and Maniatis T., Molecular cloning: A laboratory manual. Cold Spring Harbour Press, 1989.
11. Altschul S.F., Madden T.L., Schaffer A.A., Zhang J., Zhang Z., Miller W., and Lipman D.J., Nucl Acid Res 25, 3389-3402, 1997.
12. Thompson J.D., Gibson T.J., Plewniak F., Jeanmougin F., and Higgins D.G., Nucl Acids Res 24, 4876-4882, 1997.
13. Nicholas K.B. and Nicholas H.B. Jr, GENEDOC Distributed by the author 1997.
14. Lu A. and Miller L.K., J Virol 69(2), 975-982, 1995.
15. Pari G.S., Kacica M.A., and Anders D.A., J Virol 67, 25752582, 1993.
16. Volkman L.E., Summers M.D., and Ching Hsuiu H., J Virol l9, 820-832, 1976.
17. Kuzio J., Jaques R., and Faulkner P., Virology l73, 759-763, 1989.
18. Hill J.E., Kuzio J., Wilson J.A., MacKinnon E.A. and Faulner P., Biochim Biophys Acta l772, 187-189, 1993.
19. Faulkner P., Kuzio J., Williams G.V. and Wilson J.A., J Gen Virol 78(l2), 3091-3100, 1997.
20. Rohrmann G.F., J Gen Virol 73, 749-761, 1992.
21. Bicknell J.N., Leisy D.J., Rohrmann G.F., and Beaudreau G.S., Virology l6l, 589-592, 1987.
22. Huh N.E. and Weaver R.F., J Gen Virol 7l, 195-202, 1990.
23. Goenka S. and Weaver R.F., Abstract of the 15th Annual Meeting of the American Society of Virology, 1996, p. 154.
24. Rodems S.M. and Friesen P.D., J Virol 67, 5776-5785, 1993.
25. Rankin C., Ladin B.F., and Weaver R.F., J Virol 57, 18-27, 1986.
26. Ijkel W.F.J., van Strien E.A., Heldens J.G.M., Broer R., Zuidema D., Goldbach R.W., and Vlak J.M., J Gen Virol 8O, 3289-3304, 1999.
27. Van Oers M.M., Hu Z., Arif B.M., van Strien E.A., van Lent J.W.M., and Vlak J.M., J Gen Virol 79, 15531562, 1998.
28. van Strien E.A., Faktor O., Hu Z.H., Zuidema D., Goldbach R.W., and Vlak J.M., J Gen Virol 78(Pt 9), 2365-2377, 1997.
29. Wilson A.W., Hill J.E., Kuzio J., and Faulkner P., J Gen Virol 76, 2923-2932, 1995.
30. Hu Z.H., Arif B.M., Jin F., Martins J.W.M., Chen X.W., Sun J.S., Zuidema D., Goldbach R.W., and Vlak J.M., Virus Res 55, 71-82, 1998.