Isolation, identification and molecular phylogenetic analysis of *Hyblaea puera* Nucleopolyhedrovirus

Saranya Vijay Krishnan Corresp., [1] , **Sajeev TV** [1]

[1] Department of Forest Entomology, Kerala Forest Research Institute, Thrissur, Kerala, India

Corresponding Author: Saranya Vijay Krishnan
Email address: saranyasivadas89@gmail.com

Hyblaea puera (Lepidoptera: *Hyblaeidae*), is considered as a serious pest of teak in India and other tropical regions. It causes entire defoliation of teak trees and results in huge timber loss thereby decreasing forest productivity. *Hyblaea puera* Nucleopolyhedrovirus (HpNPV) is a baculovirus that has been employed in various parts of India as a bio-control agent against the pest *H. puera*. An unfeigned nucleopolyhedrovirus was isolated from the larvae of the moth, *H. puera* in Kerala, South India. *Polh, lef-8, pif-2* gene sequences were amplified by PCR with degenerate primers and extracted for phylogenetic analysis. *Hyblaea puera* Nucleopolyhedrovirus appeared to be a distinct species of Group II NPV alphabaculovirus. Polyhedrin coding region was characterized by nucleotide sequence analysis. To date, Polyhedrin is the first isolated and characterized gene of HpNPV. It indicated the presence of ORF comprising 741 nucleotides which encode 246 amino acids with a predicted molecular mass of 28 KDa. Phylogeny based on three conserved baculovirus genes showed the highest homology of HpNPV to *Helicoverpa armigera* NPV. These findings were hardened by restriction endonuclease analysis, even though some differences in restriction pattern were observed. The current study will encourage future efforts to improve the efficacy of HpNPV against its natural host.

Isolation, identification and molecular phylogenetic analysis of *Hyblaea puera* Nucleopolyhedrovirus

Saranya Vijay Krishnan[1] and Sajeev TV[1]

[1]*Department of Forest Entomology, Division of Forest Health, Kerala Forest Research Institute, Peechi, 680653, Thrissur, Kerala, India*

Corresponding author:

Saranya Vijay Krishnan

KFRI, Thrissur, Kerala, 680653, India

Email: saranyasivadas89@gmail.com)

Abstract

Hyblaea puera (Lepidoptera: *Hyblaeidae*), is considered as a serious pest of teak in India and other tropical regions. It causes entire defoliation of teak trees and results in huge timber loss thereby decreasing forest productivity. *Hyblaea puera* Nucleopolyhedrovirus (HpNPV) is a baculovirus that has been employed in various parts of India as a bio-control agent against the pest *H. puera*. An unfeigned nucleopolyhedrovirus was isolated from the larvae of the moth, *H. puera* in Kerala, South India. *Polh, lef-8, pif-2* gene sequences were amplified by PCR with degenerate primers and extracted for phylogenetic analysis. *Hyblaea puera* Nucleopolyhedrovirus appeared to be a distinct species of Group II NPV alphabaculovirus. Polyhedrin coding region was characterized by nucleotide sequence analysis. To date, Polyhedrin is the first isolated and characterized gene of HpNPV. It indicated the presence of ORF comprising 741 nucleotides which encode 246 amino acids with a predicted molecular mass of 28 KDa. Phylogeny based on three conserved baculovirus genes showed the highest homology of HpNPV to *Helicoverpa armigera* NPV. These findings were hardened by restriction endonuclease analysis, even though some differences in restriction pattern were observed. The current study will encourage future efforts to improve the efficacy of HpNPV against its natural host.

Keywords: Nucleopolyhedrovirus, *Hyblaea puera*, Phylogeny, HpNPV, Baculovirus.

Introduction

Baculovirus genomes are delineated by large circular double stranded DNA molecules, straggling from 80-180 kbp length [Miele et al.,2011; Monique & Vlak,2007; Zhu et al.,2014]. The family *baculoviridae* is stratified into four distinct genera based on the phylogeny and the host specificities scilicet *Alphabaculovirus* (lepidopteran specific NPV), *Betabaculovirus* (lepidopteran specific GV), *Gammabaculovirus* (hymenopteran specific NPV), and *Deltabaculovirus* (dipteran specific NPV). Alphabaculovirus can be further indexed into Group I NPV and Group II NPV under the basis of phylogenetic analyses [Bulach et al.,1999; Herniou et al.,2001; Jehle et al.,2006a; Zanotto, kissing&Maruniak,1993]. Till date, sixty-two baculovirus genomes have been radically sequenced [NCBI GenBank May 2018]. Among those sequenced, forty-two discern a kinship with *Alphabaculovirus*, fifteen establish a good rapport with *Betabaculovirus*, three have been accorded to *Gammabaculovirus*, one is lone classified under *Deltabaculovirus* and one belongs to an unclassified virus. A clique of thirty-seven core genes are extant in all baculoviruses sequenced until now [Miele et al.,2011; Zhu et al.,2014].

The combined sequences of conserved genes can be used to study molecular phylogeny which in turn is a powerful tool to identify lepidopteran specific baculoviruses [Herniou et al.,2001; Herniou et al.,2003; Jehle,2004; Jehle et al.,2006a; Jehle et al.,2006b; Lange et al.,2004]. Earlier phylogenetic analyses copiously pivoted on a single gene, but nowadays it is being gentrified by combined gene sets. The verity is that polyhedrin is a chimeric gene that arises from recombination [Jehle,2004]. Amongst the conserved genes-*lef-8, pif-2* and *pol h* used for phylogenetic analysis and species identification, *lef-8* and *pif-2* corroborated to be the bulk steadfast markers [Herniou et al.,2004; Van oers et al.,2004]. *Lef-8* gene encodes the largest subunit of RNA polymerase incumbent for late gene transcription [Acharya & Gopinathan,2002; Van oers et al.,2004]. Oral infectivity of the virus requires *pif-2* which is crucial for instigating an evolutionarily conserved complex on ODV surface alongside other noted proteins [Peng et al.,2012., Pijilman, Pruijssers & Vlak,2003]. *Pol h*, encodes for *polyhedrin,* a key protein of occlusion body [Bideshi, Bigo&Federici,2000 Zanotto, Kissing&Maruniak,1993].

Hyblaea puera (order: Lepidoptera) is reviewed as a profound pest of teak, which feeds on tender foliage of teak leaves, consequently precipitating conspicuous economic forfeiture on teak productivity [Chandrasekhar et al.,2005; Nair & Sudheendrakumar,1986]. This major pest is prevalent in forests of India, Thailand, Bangladesh, South East Asia, Mexico and Northern Queensland in Australia [Moraes & Maruniak,1997; Nair & Sudheendrakumar,1986; Cibrien et al.,2015]. *Tectona grandis* is considered as an economically important timber indigenous to India, Myanmar, Malaysia, Thailand, Laos and certain parts of Australia. *Hyblaea puera nucleopolyhedrovirus*(HpNPV), an imminent biopesticide maneuvered in resistance to the teak defoliator *H. puera*, is a part of the large family Baculoviridae [Biji, Sudheendrakumar & Sajeev,2006; Sudheendrakumar, Mohammed & Verma,1988]. Molecular characterization of HpNPV is still very much in its infancy. Even, the studies on its genomic sequence analysis is totally lacking. In this paper, we characterized this virus(HpNPV) on a molecular basis and sequenced three conserved baculovirus core genes. Its evolutionary relationship with other NPV's was evaluated using these three core gene sequences, thereby resolving the phylogenetic position of HpNPV.

Materials and methods

Invitro rearing of Hyblaea puera larvae and isolation of the virus

First instar larvae of *H. puera* were collected from infested Palapilly and Nilambur forest areas. The collected larvae were aseptically transferred into sterile glass bottles. In the laboratory, the larval rearing room was facilitated with a temperature of 28±4ºC and relative humidity of 60±10%. HpNPV used in this study was obtained from the stock culture maintained in Entomology laboratory of KFRI, Nilambur, Kerala. The virus was multiplied by infecting the fourth instar *H. puera* larvae. Each larva was individually fed with HpNPV coated leaf discs (0.5 cm diameter) at a dosage of 10^6 POB's per larva. The larvae were then individually reared in rearing tubes(5.5x2.3cm) with a perforated lid on an artificial diet [Bindu, Sajeev&Sudheendrakumar,2014; Mathew et al.,1990; Sudheendrakumar, Sajeev& Biji,2008] at 28±4ºC with 60±10% humidity. [Bindu, Sajeev&Sudheendrakumar,2014; Mathew et al.,1990].

After 96 hours of post-infection, fully infected and dead larvae were retrieved and processed for virus extraction. Extraction and purification of Polyhedral Occlusion Bodies (POB's) were initiated by cutting the abdominal epithelium, filtering and centrifugation at 5000 rpm for 5 minutes. The entire process was repeated thrice. Purified POB's were then enumerated with the aid of a Neubauer's haemocytometer (0.1 mm depth) under a light microscope.

To extract virus DNA, the purified polyhedral bodies were re-suspended in lysis buffer (10Mm Tris-Hcl, 10 Mm EDTA, 1% SDS, 1M Na_2CO_3, p^H 8.3) and incubated at 37ºC overnight with proteinase k (0.5 mg/ml). Further purification of viral DNA was carried out by phenol: chloroform: isoamyl alcohol (25:24:1) extraction and precipitation. Purified viral DNA was digested with Hind III restriction enzyme and DNA fragments were separated on 0.6% agarose gel at 50V overnight.

PCR of Polyhedrin, lef-8 and pif-2 genes

Purified HpNPV DNA was used as a template for PCR reactions. To acquire a partial *pol h* gene for nucleotide sequencing, the conserved domain was amplified with degenerate polh primer sets as described previously by Moraes and Maruniak [Moraes & Maruniak,1997] under standard PCR conditions. The reaction product was cloned into a pGEM-T easy plasmid (Promega). DNA sequence towards the 3'end was extended by three successive steps following linear amplification, homopolymer tailing and PCR amplification. Linear amplification mix contained 2 µl of purified HpNPV DNA as the template, 2 pmol of Polyhedrin F2 primer (Table 1), 200 µM dNTP,1.5 mMMgCl_2 and 2.5 U of Taq DNA polymerase (Roche applied science) in a total volume of 50 µl. Amplification protocol consisted of initial denaturation at 94ºC for 5 min followed by 25 cycles of 94ºC for 30 seconds, 45ºC for 30 sec and 72ºC for 1 min followed by final elongation at 72ºC for 10 min. The DNA amplicon was purified by reaction cleanup kit(Qiagen). Amplicon thus obtained was subjected to homopolymer tailing, containing 20 µl of final reaction mix (5µl amplicon, 400 U terminal transferase (Roche applied science), 1X reaction buffer, 5 mM $COCl_2$, 0.5 mM dCTP). The reaction was incubated at 37ºC for 15 minutes after which 0.2 µM EDTA p^H 8.0 (2µl) was added to stop the reaction. The DNA was further purified by clean up kit (Qiagen) prior to final PCR amplification.

PCR amplification contained a final 20 µl of reaction mix consisting of 50 pmol of each polyhedrin F3 primer and polyG primer (Table 1), 200M dNTP, 1.5 mM $MgCl_2$, 2.5U OF Taq DNA polymerase and 2 µl of tailed fragments. The PCR protocol consisted of initial denaturation at 94ºC for 5 min followed by 35 cycles of 94ºC for 30 secs, 55ºC for 30 secs, 72ºC for 1 min followed by final elongation at 72 ºC for 15 minutes. Finally, the amplification products were cloned into pGEM -T easy vector (Promega) to generate a library of fragments for sequencing [Bioserve Biotechnologies, Hyderabad]. The DNA sequence towards the 5' end of polyhedrin gene was extended by the same method described above except for a change in primer to polyhedrin R2 (Table 1) for linear amplification and polyhedrin R3 (Table 1) in final PCR amplification.

To complement the data set previously used for baculovirus phylogenetic analyses, *lef-8* and *pif-2* gene sequences were obtained by degenerate primer sets (Table 1) as described by Herniou et al [Herniou et al.,2004]. PCR reaction mix contained 30 ng of purified viral DNA, 1.5 mM $MgCl_2$, 200 µM dNTP's, 2.5U of Taq DNA Polymerase(Promega), 1X PCR buffer (Promega), 50 pmol of each primer (Table 1) in a final volume of 50 µl under standard PCR conditions. The amplification products were 704 bp and 480 bp respectively.

Phylogenetic analysis

Since the two genes (*lef-8 and pif-2*) are congruent as suggested by previous reports, both can be concatenated for phylogenetic analyses [Herniou et al.,2001; Herniou et al.,2004]. The *pol h* gene was analyzed separately since it is not congruent to other gene sets [Herniou et al.,2003b; Jakubowska, Vlak&Ziemmicka,2005]. *Pol h, lef-8, pif-2* gene sequences from other baculoviruses were obtained from GenBank and used for phylogenetic analysis (Table 2). Multiple sequence alignments were generated using ClustalX. Phylogenetic trees were constructed using the Maximum Likelihood(ML) method in PAUP* version 4.0 [Swofford,2001]. The robustness of tree topologies was evaluated by bootstrap analysis with 1000 replicates. All the three genes were found suitable for phylogenetic analysis because of the sequence availability for most baculoviruses.

Results

Gene sequencing, polyhedrin characterization and phylogenetic analysis

To determine the taxonomic status of HpNPV, three commemorated baculovirus genes *pol h*, *lef-8* and *pif-2* genes were amplified, sequenced and analyzed phylogenetically. Partial sequences of *lef-8* and *pif-2* obtained after degenerate PCR primers (Table 1) were deposited in GenBank under accession numbers MH254887 and MH362814 respectively. BLAST homology searches for *lef-8* and *pif-2* genes revealed the highest (100%) homology with *Helicoverpa armigera* NPV (HearNPV) at the nucleotide level. Translated BLAST provided the result of 100% amino acid sequence identity. The concatenated *lef-8 + pif-2* phylogenetic analysis which contains only fully sequenced lepidopteran NPV's clearly placed HpNPV among group II NPV's (Figure 1A). The combined *lef-8+pif-2* tree provides greater resolution with high bootstrap support (100%), but greater than 98% in most cases for the elementary relationships of NPV's. The combined analysis of *lef-8* and *pif-2* amino acid sequences provides insights into relatedness of baculovirus genomes [George, Martin & David,2015]. From this phylogenetic analysis, it can be concluded that sequences of both these genes from HpNPV showed the closest relationship to *H. armigera* NPV(HearNPV), which is a member of group II lineage supported by high bootstrap scores.

However, since polyhedrin sequences are available for a majority of the NPV's sequenced till date, HpNPV polyhedrin gene (*pol h)* was characterized. Therefore, the sequence of 1398 bp DNA fragment (GenBank accession no:MH719085), flanking the entire coding region was determined. PCR amplification products were obtained with polyhedrin degenerate primers (Table 1) and sequenced (GenBank accession no: MH254886). The sequence was extended on both the ends by DNA walking using HpNPV polyhedrin specific primers (Table 1) thereby generating the library of fragments covering the flanking region.

Nucleotide sequence analysis of polyhedrin

Nucleotide sequence analysis of *pol h* gene indicated the presence of an open reading frame of 741 nucleotides encoding 246 amino acid residues with a predicted molecular mass of 28.9 KDa. The upstream sequence of the translation initiation site contains a putative baculovirus late promoter element ATAAG at position -73 to -69. It was demonstrated that the pentanucleotide

ATAAG is comprehensively located 100 nucleotides upstream to start codon and are highly conserved among insect baculovirus Polyhedrin promoters [Van oers et al.,2004; Parin et al.,2008]. The isolated *pol h* gene also has an AT-rich region downstream to the stop codon TAA in the 3' noncoding region. It is assumed that the T-rich sequence is essential for termination and polyadenylation [Parin et al.,2008]. Furthermore, this AT-rich region is highly conserved among insect baculovirus polyhedrin promoters [Rohrmann,1986]. Moreover, the coding portion of HpNPV is not interrupted by introns. The salient elements of eukaryotic gene promoters analogous to the consensus sequence TATA and CCAAT boxes were also located in the 5' flanking region of the HpNPV *pol h* gene. The TATA box with DNA sequence, GTATAA was found 245 nucleotides upstream from the start codon. In addition, two more TATA sequences, ATAAAG and TAAATA were also observed at position -120 and -14 respectively. An extra set of CCAAT sequence was recognized some 186 nucleotides upstream to the transcription start site. Notwithstanding the role of any of the eukaryotic promoter signals noted above is merely speculative and should be figured through transcriptional studies.

Amino acid sequence analysis

Amino acid sequence analysis revealed that polyhedrin is a 246-amino acid peptide with a molecular weight of 28.9 KDa. It has an aliphatic index of 78 and hydropathicity around -0.5. Since the number of positively and negatively charged residues are roughly equivalent, the protein can be considered as relatively stable. Although several potential motifs for N-glycosylation and O-glycosylation were found in HpNPV polyhedrin, proteins without signal peptide are unlikely to be exposed to those post-translational modifications. (www.cbs.dtu.dk)

It is quite interesting to detect one more prominent open reading frame of 168 nucleotides encoding 55 amino acids with a predicted molecular mass of 6.2 KDa in the complementary strand. BLAST analysis revealed a striking similarity to ChchNPV with 100% homology. We have also noted 96% and 87% homology with PsinNPV and TnNPV respectively. The latter analysis reported it to be a viral capsid protein. Furthermore, this hypothetical protein of 55 amino acids tends to be an intrinsically disordered protein with a higher proportion of hydrophilic, charged residues (http://iupred.elte.hu/). Our predicted hypothetical protein lacks any tryptophan, tyrosine and cysteine residues which render it invisible under UV-

spectrophotometry. Nevertheless, extensive studies are needed to determine whether the observed protein is just a part of the structural gene. Such studies are now in progress.

Phylogenetic analysis

Phylogenetic analysis of the Polyhedrin sequence placed HpNPV in Group II NPV's. The GC content of *pol h* appeared to be 38%, in contrast to 38.9% for HearNPV. The phylogeny (*pol h tree*) shows a highly supported group (Fig. 1B), comprising HearNPV and BufuNPV. However, only the Polyhedrin sequence is available for the latter virus, further studies with other genes are required to confirm the ancestry of these two baculoviruses.

Restriction analysis

The restriction enzyme analysis with HindIII was performed for HpNPV and *H. armigera* NPV (strain kindly provided by TNAU, Coimbatore). The Hind III restriction pattern (Fig. 2) was used to estimate the genome size (Table 3) which appears to be approximately 138 kbp. The restriction profile of HpNPV was almost identical to *H. armigera* NPV (Fig. 2). This result supports the contention that HpNPV is closely related to *H. armigera* NPV, representing a distinct species of alphabaculovirus.

Discussion

The size of HpNPV DNA was estimated to be 138 kbp by restriction endonuclease analysis (Fig. 2, Table 3). Even though the isolation of HpNPV was previously reported, knowledge of its genomic sequence is embryonic and not proclaimed yet. The restriction profile for a baculovirus is apparently copper-bottomed and hence milked to draw a distinction between closely related species [Woo et al.,2006]. A group of researchers from Kerala Forest Research Institute (KFRI) has patented a coherent composition of biopesticide widely acknowledged as HpNPV. HpNPV is contemplated as one of the swiftest acting insect viruses as it subjugates wholly seventy-two hours to liquidate the larvae [Bindu,Sajeev&Sudheendrakumar,2014;Biji,Susheendrakumar&Sajeev,2006;Sudheendrakumar, Sajeev&Biji,2008;Sudheendrakumar,Mohammed&Verma,1988]The revelation of HpNPV as a potent biopesticide was a quantum leap in the field of *Hyblaea* management research

[Sudheendrakumar,Mohammed&Verma,1988]. Since then the teak plantations in Kerala endured extortionate diminution owing to this pest [Chandrasekhar et al.,2005; Nair & Sudheendrakumar,1986]. It was probed that polyhedrins are perpetuated in all lepidopteran NPV's which shares almost 85 to 90% amino acid homology [Rohrmann,1986; Woo et al.,2006]. HpNPV is a protein of 246 amino acid residues that in fact resemble other NPV's as well. It only differs at position 106 when compared to *H. armigera* NPV polyhedrin which might be the repercussion of point mutation. Evolutionary homology is generally analyzed by comparison of DNA and protein sequences. Besides, amino acid sequence is the key to explore protein structure and function in the cell. Hence amino acid sequence analysis forms a vital part of post- genomic studies. The stability of a protein is greatly contributed by the participating amino acid residues. This accounts for the higher stability of polyhedrin. The greater the disorder promoting amino acids, greater the instability. Such proteins are called as intrinsically disordered proteins(IUP); which may be short or long. Short disordered proteins around 50 amino acids tend to participate in metal ion binding, ion channels, signal transduction and even regulate GTPase functions. However, there are shreds of evidence that they play a significant role in evolution as well [Robin et al.,2014]. The predicted hypothetical protein is one among the short disordered IUP's.

Phylogenetic analysis evinced that HpNPV accords to the group II NPV's, encompassing a pristine and unique species of alphabaculovirus. The proportion of GC content ranges from 33 to 58.9% in group II NPV's whereas in group I NPV it varies from 36 to 55.9% [Miele et al.,2011]. It is not essential that *pol h* tree topologies should invariably replicate with the concatenated sequences topologies that entails complementation from the other gene phylogenies also [Jehle,2004; Jehle et al.,2006b]. Here, we report that HpNPV *pol h* phylogeny locales this virus among group II NPV's, which is staunchly bolstered by combined *lef-8* and *pif-2* sequence analysis. Moreover, the relatedness of baculovirus is materially resolved by colligating *lef-8/pif-2* amino acid sequences which are conserved in all baculoviruses sequenced till date [George, Martin & David,2015]. As already mentioned, HpNPV is a part of the large family baculoviridae. Core knowledge of nucleotide sequence helps in the enhancement of insecticidal activity of HpNPV. By unravelling the sequence information, phylogenetic analysis can be performed which sheds light into baculovirus evolution that infers in host-pathogen interactions [

Jehle et al.,2006b; Rohrmann,2011]. All these factors, undisputedly bestow to the taxonomic classification of the virus.

Conclusion

In this study, we have shown that based on the three conserved gene sets, *polh, lef-8* and *pif-2,* HpNPV can be clearly allocated to group II lineage of alphabaculovirus, with *H. armigera NPV* as the close relative. These data sets were solidly underpinned by restriction enzyme analysis which was nearly cognate for HpNPV and *H armigera* NPV. Knowledge of the taxonomic position of HpNPV is a crucial element to promote its use in Integrated Pest Management (IPM). This study provided information for investigators focusing to enhance the potency of HpNPV against its natural host *H. puera* or expand viral host range of susceptible pests.

Acknowledgements

All staff members of the Entomology Department in KFRI are acknowledged for their help and support throughout the project period. We especially thank Seema P Nair for critically reading the manuscript and Keith McIntosh for improving the use of English in the manuscript

Competing interests

The authors declare that they have no conflict of interest.

Author contributions

Saranya Vijay Krishnan Conceived ad designed the experiments, performed the experiments, analyzed the data, wrote the paper, prepared tables and /or figures, Read and approved the manuscript.

TV Sajeev Conceived and designed the experiments, analyzed the data, Read and approved the manuscript.

DNA Deposition

The following information was supplied regarding data availability:

HpNPV Sequences were deposited in GenBank and available under the following accession numbers: MH719085, MH2254886, MH254887 and MH362814.

References

1. Acharya A., Gopinathan K. 2002. Transcriptional analysis and preliminary characterization of ORF Bm42 from *Bombyx mori* nucleopolyhedrovirus. Virology 299:213-224

2. Bideshi D. K., Bigot Y., Federici B. A.2000. Molecular characterization and phylogenetic analysis of the *Harrisina brillians* granulovirus granulin gene. Archives of Virology. 145:1933–1945.

3. Bindu K. Jose, Sajeev T.V., Sudheendrakumar V.V. 2014. Role of protein and lipids in artificial diets of teak defoliator moth: Hyblaea puera. Journal of Entomology and Zoology Studies 2(2): 97-100

4. Bulach D.M., Kumar C.A., Zaia A., Liang B., Tribe D.E. 1999. Group II nucleopolyhedrovirus subgroups revealed by phylogenetic analysis of polyhedrin and DNA polymerase gene sequences. Journal of Invertebrate Pathology 73: 59-73

5. C.P.Biji, V.V.Sudheendrakumar, T.V.Sajeev. 2006. Quantitative estimation of *Hyblaea puera* NPV production in three larval stages of the teak defoliator, *Hyblaea puera*(Cramer). Journal of Virological Methods 136(1-2):78-82 https://doi.org/10.1016/j.jviromet.2006.04.001

6.David K. Thumbi, Robert J. M. Eveleigh, Christopher J. Lucarotti, Renée Lapointe, Robert I. Graham, Lillian Pavlik, Hilary A. M. Lauzon and Basil M. Arif. 2011. Analysis and Organization of the Orgyia leucostigma Nucleopolyhedrovirus Genome. Viruses 3: 2301-2327 doi: 10.3390/v3112301 ISSN 1999-4915

7. Garavaglia M.J., Miele S.A., Iserte J.A., Belaich M.N., Ghiringhelli P.D.2012.The ac53, ac78, ac101, and ac103 genes are newly discovered core genes in the family Baculoviridae. Journal of Virology 86: 12069–12079.

8. George F rohrmann, Martin A Erlandson, David A Theilmann .2015. A Distinct Group II Alphabaculovirus isolated from a peridroma species. Genome Announcement 3(2) e 00185-15 doi: 101128/ genome A.00185-15

9. Herniou E. A. 2003 Use of comparative genomics and phylogenetics to study the evolution of the Baculoviridae (PhD thesis). Centre for Ecology & Hydrology Imperial college London UK

338 10. Herniou E. A., Luque T., Chen X., Vlak J.M., Winstanley D., Cory J.S., O'Reilly D.R .2001.
339 Use of whole genome sequence data to infer baculovirus phylogeny. Journal of Virology 75:
340 8117-8126
341 11.Herniou EA, Olszewski JA, Cory JS, O'Reilly DR.2003. The genome sequence and evolution
342 of baculoviruses. Annual Review of Entomology 48:211–234
343 12. Herniou E.A, Olszewski J.A, O'Reilly D.R, Cory J.S .2004. Ancient coevolution of
344 baculoviruses and their insect hosts. Journal of Virology 78:3244-3251
345 13. Jakubowska A, Vlak J.M, Ziemnicka J. 2005. Characterization of a nucleopolyhedrovirus
346 isolated from the laboratory rearing of the beet armyworm spodoptera exigua in Poland. Journal
347 of Plant Protection Research 45: 4
348 14. Jehle J. A. 2004. The mosaic structure of the polyhedrin gene of the Autographa californica
349 nucleopolyhedrovirus (AcMNPV). Virus Genes 29: 5-8
350 15. Jehle J. A., Blissard G.W., Bonning. C., Cory J.S., Herniou E.A., Rohrmann G.F., Theilmann
351 D.A., Theim S.M., Vlak J.M. 2006. On the classification and nomenclature of baculoviruses: a
352 proposal for revision. Archives of Virology 151: 1257-1266
353 16. Jehle J.A., Lange M., Wang H., Hu Z., Wang Y., Hauschild R. 2006. Molecular
354 identification and phylogenetic analysis of baculoviruses from Lepidoptera. Virology 346:180-
355 193.
356 17. Lange M., Wang H., Zhihong H., Jehle J.A. 2004. Towards a molecular identification and
357 classification system of lepidopteran specific baculoviruses. Virology 325: 36-47
358 18. Mathew G, Sudheendrakumar V.V., Mohandas K and Nair KSS.1990. An artificial diet for
359 the teak defoliator Hyblaea puera Cramer (Lepidoptera: Hyblaeidae). Entomology 15(4): 159-
360 163
361 19. Miele S.A., Garavaglia M.J., Belaich M.N., Ghiringhelli P.D.2011.Baculovirus: molecular
362 insights on their diversity and conservation. International Journal of Evolutionary Biology 2011:
363 379424. 6
364 20. M.M. Van Oers, Elisabeth A. Herniou., Magda Usmany, Gerben J. Messelink, Just M. Vlak.
365 2004.Identification and characterization of a DNA photolyase- containing baculovirus from
366 chrysodeixis chalcites. Virology 330: 460-470
367 21. Monique M. van Oers, Just M. Vlak .2007. Baculovirus Genomics. Current Drug Targets Oct
368 8(10): 1051-1068

22. Moraes R. R., Maruniak J. E .1997. Detection and identification of multiple baculoviruses using the polymerase chain reaction (PCR) and restriction endonuclease analysis. Journal of Virological Methods 63: 209-217

23. N Chandrasekhar, TV Sajeev, VV Sudheendrakumar, Moinak Banerjee .2005. Population dynamics of the Teak defoliator (Hyblaea puera Cramer) in Nilambur teak plantations using Randomly Amplified Gene Encoding Primers (RAGEP). BMC Ecology. 5:1 https://doi.org/10.1186/1472-6785-5-1

24. Nair KSS, Sudheendrakumar V.V .1986. The teak defoliator, *Hyblaea puera*: Defoliation dynamics and evidence for short-range migration of moths. Proceedings of Indian National Science Academy Part B Biological Science 95: 7-21

25.Parin Chaivisuthangkura, Chutithorn Tawilert, Thanawan Tejangkura, Sombat Rukpratanporn, Siwaporn Longyant, Weerawan Sithigorngul, Paisarn Sithigorngul.2008. Molecular isolation and characterization of a novel occlusion body protein gene from *Penaeus monodon*nucleopolyhedrovirus. Virology. 381(2): 261-267 https://doi.org/10.1016/j.virol.2008.08.036

26. Peng K, van Lent J.W.M., Boeren S., Fang M., Theilmann D.A., Erlandson M.A., Vlak J.M., Van Oers M.M. 2012.Characterization of Novel Components of the Baculovirus *Per Os* Infectivity Factor Complex. Journal of Virology 86(9):4981-4988. doi:10.1128/JVI.06801-11.

27. Pijilman G.P., Pruijssers A.J., Vlak J.M. 2003. Identification of pif-2, a third conserved baculovirus gene required for per os infection of insects. Journal of General Virology 84: 2041-2049

28. Robin Van der Lee, Benjamin Lang, Kai Kruse, Jorg Gsponer, Natalia Sánchez de Groot, Martijn.A. Huynen, Andreas Matouschek, Monika Fuxreiter, Madan.M. Babu. 2014. Intrinsically Disordered Segments Affect Protein Half-Life in the Cell and during Evolution. Cell Reports, 8(6), 1832–1844. http://doi.org/10.1016/j.celrep.2014.07.055

29. Rohrmann G. F. 1986. Polyhedrin structure. Journal of General Virology 67: 1499-1513

30. Rohrmann G.F .2011. Baculovirus molecular biology, 2nd ed. NCBI, Bethesda, MD.

31. Sudheendrakumar V., Sajeev, T.V., and Biji C. 2008. A New Insect Rearing Container for *in vivo* Mass Multiplication of NPV of *Hyblaea puera*. *Journal of Biological Control, 22*(1): 217-219. doi:10.18641/jbc/22/1/40281

400 32. Sudheendrakumar V. V., Mohamed Ali M.I., Verma R.V. 1988.Nuclear polyhedrosis virus of
401 the teak defoliator Hyblaea puera. Journal of Invertebrate Pathology 51: 307-308.
402 33. Swofford D L.2001.PAUP*(Phylogenetic Analysis using Parsimony and other methods), 4th
403 ed. Sinauer Associates, Sunderland, MA.
404 34. V. C. Cibrián-Llanderal, H. González-Hernandez, D. Cibrián-Tovar, M. Campos-Figueroa,
405 H. de los Santos-Posadas, J. C. Rodríguez-Maciel and A. Aldrete .2015. Incidence of Hyblaea
406 puera (Lepidoptera: Hyblaeidae) in Mexico.Southwestern Entomologist.40(2):441
407 doi:10.3958/059.040.0220
408 35. Woo S-D, Choi J.Y., Je Y. H and Jin B. R .2006. Characterization of the Helicoverpa assulta
409 nucleopolyhedrovirus genome and sequence analysis of the polyhedrin gene region. Journal of
410 Biosciences 31: 329-338
411 36. Zanotto P. M., Kessing B. D., Maruniak J. E**.** .1993. Phylogenetic interrelationships among
412 baculoviruses: evolutionary rates and host associations. Journal of Invertebrate Pathology
413 62:147–164.
414 37. Zhu Z., Yin F., Liu X., Hou D., Wang J., Zhang L., Arif B., Wanh H., Deng F., Zhu Z. 2014.
415 Genome Sequence and Analysis of Buzura suppressaria Nucleopolyhedrovirus: A Group II
416 Alphabaculovirus. PLoS ONE 9(1): e86450.
417 doi: 10.1371/journal.pone.0086450

Table 1(on next page)

PCR primer sequences used in this study.

1 **Table 1: PCR primer sequences**

Name	Oligonucleotide Sequence	Product size	Reference
Polyhedrin F1	TAYGTGTAYGAYAACAAG	645 bp	Moraes & Maruniak et al
Polyhedrin R1	TTGTARAAGTTYTTCCAG		
Polyhedrin F2	CAAGAATTCCATAATGTATACTCG	-	-
Polyhedrin R2	TGTCTGCAGTAGGAACCAAACCG		
Polyhedrin F3	TACTAAGTGGTGATTCGCGA	-	-
Polyhedrin R3	GTTGTCTAGGGGATCAGGA		
Lef-8 F1	TTYTTYCAYGGNGARATGAC	704 bp	Herniou et al
Lef-8 R1	GGNAYRTANGGRTCYTCNGC		
Pif-2 F1	GGWNNTGYATNSGNGARGAYCC	480 bp	Herniou et al
Pif-2 R1	RTYNCCRCANTCRCANRMNCC		
Poly G primer	TATAGGGGGGGGGGGGGG	-	-

2

Table 2(on next page)

Outline of Baculovirus sequences used for phylogenetic analysis in this study

1 **Table 2: Outline of Baculovirus sequences used for phylogenetic analyses in this study**

Name of Virus	Abbreviation	*polh*	*lef-8*	*pif-2*
Adoxophyes honmai NPV	AdNPV	NP818648	NP818698	NP818742
Adoxophyes orana NPV	AdorNPV	YP002300519	YP002300564	YP002300608
Agrotis ipsilon NPV	AgipNPV	YP002268031	YP0022680160	YP002268072
Agrotis segetum NPVA	AgseNPVA	YP529671	YP529791	YP529706
Agrotis segetum NPVB	AgseNPVB	YP009112562	YP009112680	YP009112597
Apocheima cinerarium NPV	ApciNPV	YP006607771	YP006607790	YP006607853
Autographa californica MNPV	AcMNPV	NP054037	NP054079	NP054051
Bombyx mori NPV	BmNPV	AFJ06797	AFN08967	AIS92745
Buzura suppressaria NPV	BusuNPV	YP009001778	AIW63034	AKN91074
Busseola fusca NPV	BufuNPV	AAT10236	-	-
Chrysodeixis chalcites NPV	ChchNPV	YP249605	YP249641	YP249752
Clanis bilineata NPV	ClbiNPV	YP717539	YP717570	YP717645
Ectropis obliqua NPV	EcobNPV	YP874194	YP874225	YP874299
Euproctis pseudoconspersa NPV	EupsNPV	YP002854611	YP002854631	YP002854731
Helicoverpa armigera NPV	HearNPV	AC105102	AEY77857	AIG63176
Hemileuca species NPV	HeNPV	YP008378219	YP009165657	YP008378325
Hyphantria cunea NPV	HycuNPV	YP473189	-	-
Lambdina fiscellaria NPV	LafiNPV	YP009134716	YP009133306	YP009133237
Leucania separata NPV	LeseNPV	AAA99736	YP758340	YP758460
Lymantria dispar MNPV	LdMNPV	-	AIX47889	AIX47957
Malacosoma sp NPV	MaspNPV	-	ANW12301	ANW12330
Malacosoma disstria NPV	MadiNPV	AAD00095	-	-
Malacosoma neurista MNPV	ManeNPV	AAB31529	-	-
Mamestra brassicae MNPV	MabrNPV	-	AFP95852	YP009011107
Mamestra configurata NPVA	MacoNPVA	NP613084	NP613224	NP613131
Mamestra configurata NPVB	MacoNPVB	NP689176	NP689314	NP689218
Maruca vitrata NPV	MaviNPV	-	YP950765	YP950743
Mythima unipuncta NPV	MyunNPV	AUV65260	-	-
Operophtera brumata NPV	OpbrNPV	AUA60232	AUA60270	AUA60357
Orgyia leucostigma NPV	OrleNPV	YP001650911	YP001650955	YP001651022
Peridroma NPV	PeNPV	YP009049827	YP009049865	YP009049856
Perigonia lusca NPV	PeluNPV	YP009165601	YP009165657	YP009165715

Plutella xylostella NPV	PlxyNPV	ABE68393	-	-
Pseudoplusia includens SNPV	PsinNPV	YP009116914	-	-
Rachiplusia ou MNPV	RoMNPV	-	AAN28015	AAN28046
Spodoptera exigua MNPV	SpexNPV	NP037761	CDG72453	CDG72376
Spodoptera frugiperda MNPV	SpfrNPV	YP001036294	ACA02670	YP001036326
Spodoptera litura NPVII	SpliNPV	YP002332699	YP002332815	AAF72593
Sucra jujuba NPV	SujuNPV	YP009186692	YP009186724	YP009186804
Thysanoplusia orichalcea NPV	ThorNPV	-	YP007250419	YP007250432
Trichoplusia ni SNPV	TnSNPV	AAC64160	YP308923	YP309030
Urbanus proteus NPV	UrprNPV	YP009249983	-	-

2

3

Table 3(on next page)

Size of restriction fragments(kb) digested with Hind III.

1 **Table 3: Size of restriction fragments(kb) digested with Hind III**

Fragment	HpNPV	HearNPV
A	23.1	23.1
B	22.0	22.0
C	14.2	15.2
D	11.3	12.3
E	9.4	9.4
F	8.2	8.5
G	8.0	8.0
H	7.5	7.5
I	7.2	7.2
J	7.0	7.0
K	6.8	6.8
L	6.5	6.5
M	4.3	4.3
N	2.0	2.0
O	0.5	-
Total	**138**	**139.8**

2

3

Figure 1(on next page)

Concatenated *lef-8* + *pif-2* tree.

The phylogenetic trees were based on the amino acid sequences obtained by Maximum Likelihood (ML) analysis for 1000 replicates. Numbers indicate bootstrap scores. Further, bootstrap scores lower than 50% are collapsed. GenBank accession numbers of virus sequences used are listed in the Table 2.

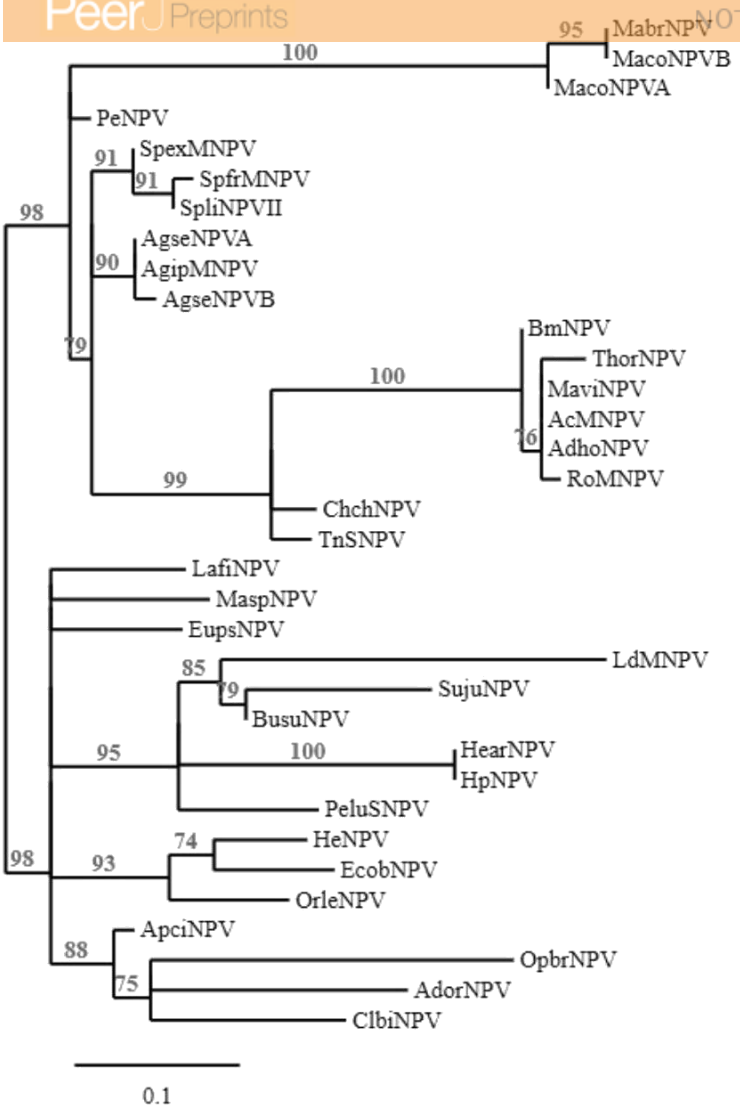
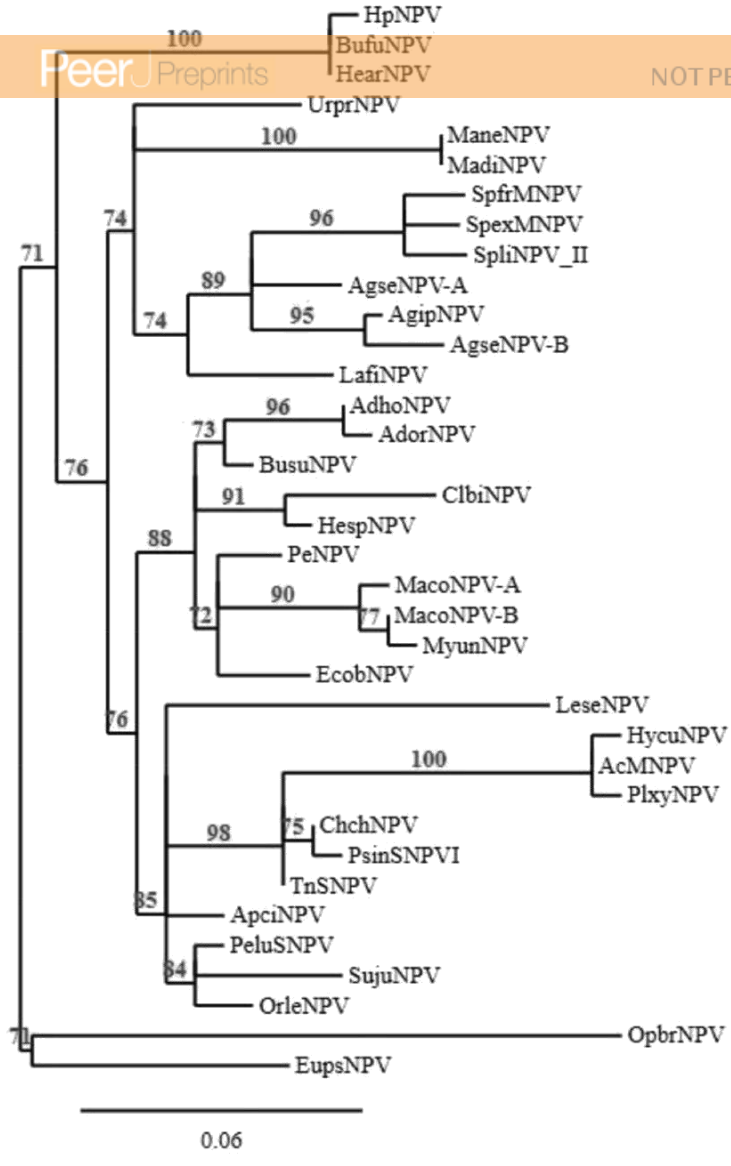


Figure 2(on next page)

Polyhedrin (*Pol h*) tree.

Phylogenetic analysis based on the amino acid sequences obtained by ML method for 1000 replicates. Numbers indicate bootstrap scores. Bootstrap values less than 50% are collapsed. GenBank accession numbers of virus sequences used are listed in Table 2.

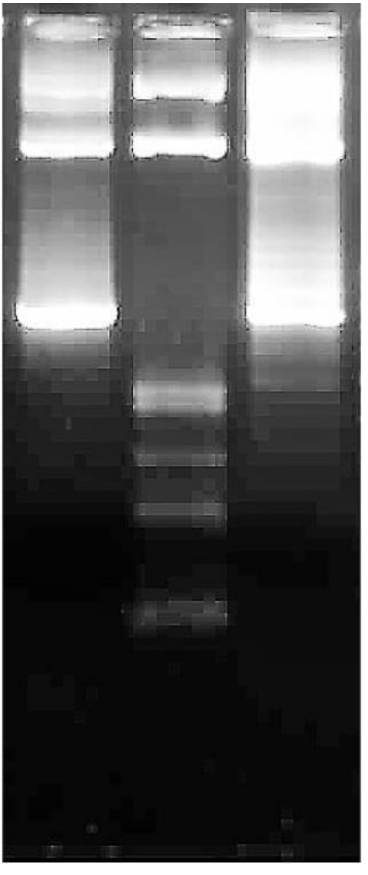


PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27290v1 | CC BY 4.0 Open Access | rec: 19 Oct 2018, publ: 19 Oct 2018

Figure 3(on next page)

Restriction Digestion profile of HpNPV

Gel photograph showing Hind III digestion profiles of HpNPV and HearNPV; 1- HpNPV, 2-HearNPV; M- Marker λ DNA digested with Hind III digest. Electrophoresis carried out on a 0.6% agarose gel at 50V overnight to separate fragments.



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27290v1 | CC BY 4.0 Open Access | rec: 19 Oct 2018, publ: 19 Oct 2018