Whole genome sequence and comparative genomic sequence analysis of *Helicoverpa armigera* nucleopolyhedrovirus (HearNPV-L1) isolated from India

Ashika T. Raghavendra^{1,2} · Sushil K. Jalali¹ · Rakshit Ojha^{1,2} · Timalapur M. Shivalingaswamy¹ · Raj Bhatnagar³

Received: 2 July 2016 / Accepted: 24 November 2016 / Published online: 13 February 2017
© Indian Virological Society 2017

Abstract The whole genome of *Helicoverpa armigera* nucleopolyhedrovirus (HearNPV) from India, HearNPV-L1, was sequenced and analyzed, with a view to look for genes and/or nucleotide sequences that might be involved in the differences and virulence among other HearNPVs sequenced from other countries like SP1A (Spain), NNg1 (Kenya) and G4 (China). The entire nucleotide sequence of the HearNPV-L1 genome was 136,740 bp in length having GC content of 39.19% and contained 113 ORFs that could encode polypeptides with more than 50 amino acids (GenBank accession number KT013224). Two ORFs, viz., ORF 18 (300 bp) and ORF 19 (401 bp) identified were unique in HearNPV-L1 genome. Most of the HearNPV-L1 ORFs showed high similarity to NNg1, SP1A and G4 genomes. HearNPV-L1 genome contains 5 h (*hr1-hr5*), these regions were found 84–100% similar to *hr* region of NNg1, SP1A and G4 genomes. A total of four *bro* genes were observed in HearNPV-L1 genome, of which *bro-a* gene was 12 and 351 bp bigger than SP1A and G4 *bro-a*,

respectively, while *bro-b* was 15 bp bigger SP1A and NNg1 *bro-b*, whereas 593 bp shorter than G4 *bro-b*, while *bro-c* was 12 bp shorter than NNg1, however *bro-c* was absent in G4 genome. HearNPV-L1 *bro-d* was 100% homologous to *bro-d* of SP1A, NNg1 and G4 genomes, respectively. The comparative analysis of HearNPV-L1 genome indicated that there are several other putative genes and nucleotide sequences that may be responsible for insecticidal activity in HearNPV-L1 isolate, however, further functional analysis of the hypothetical (putative) genes may help identifying the genes that are crucial for the virulence and insecticidal activity.

Keywords *Helicoverpa armigera* NPV · Whole genome sequencing · Lepidoptera · Comparative genome analysis

Introduction

Helicoverpa armigera nucleopolyhedrovirus HearNPV (*Baculoviridae*), is a rod shaped, single nucleocapsid alphabaculovirus with circular closed, double stranded DNA genomes that ranges in size from 80 to 180 kb [27]. It has proved to have excellent insecticidal propriety for control of *H. armigera* (Lepidoptera: Noctuidae) larvae on number of crops [14, 25]. The family *Baculoviridae* comprises of four genera, viz., Alphabaculovirus (Lepidopteran specific nucleopolyhedroviruses), Betabaculovirus (Lepidopteran specific granuloviruses), Gammabaculovirus (Hymenopteran specific nucleopolyhedroviruses) and Deltabaculovirus (Dipteran specific nucleopolyhedroviruses). The NPVs produce large, polyhedron-shaped occlusion bodies containing many virions, are further designated as single nucleocapsid or multinucleocapsid. Depending on the number of nucleocapsids packaged in an

Electronic supplementary material The online version of this article (doi:10.1007/s13337-016-0352-6) contains supplementary material, which is available to authorized users.

✉ Sushil K. Jalali
jalalisk1910@yahoo.co.in

¹ Division of Molecular Entomology, ICAR-National Bureau of Agricultural Insect Resources, Post Bag No. 2491, H. A. Farm Post, Bellary Road, Hebbal, Bangalore, Karnataka 560024, India

² Department of Biotechnology, Centre of Post Graduate Studies, Jain University, Jayanagar, Bangalore, Karnataka 560011, India

³ International Centre for Genetic Engineering and Biotechnology, Aruna Asaf Ali Marg, New Delhi 110067, India

envelope of the virion, NPVs are called as SNPV (single) and MNPV (multi). NPVs that are isolated from different insects in the field populations differ from one another with regard to virulence against respective insects. [2, 8, 10, 17, 23]. These viruses are known to play role for gene delivery vectors for human gene therapy [13] and for display system for complex eukaryotic proteins [6, 20].

The gram pod borer *H. armigera* is an important pest that feeds on a number of agricultural and other economically important plants and is recorded on more than 120 host-plants causing severe crop losses to over 80 cultivated crops throughout tropical and sub-tropical region of our planet [30, 31]. *H. armigera* has developed a resistance to all groups of chemical insecticides [29]. In order to mitigate the resistance developed, several options have been explored for the control of this devastating pest. Among them, NPVs have come as one of the most promising larvicidal agents against such pests [4]. To date, more than 60 baculoviruses have been isolated and sequenced completely, a number of NPVs have also been isolated from *H. armigera* so far and each one has very good biological activity [9, 12].

HearNPV to control *H. armigera* was one of the first commercially available baculoviral pesticide produced in China, and has been successfully used to control the pest in agricultural crops [32]. Whole genome sequencing and comparative studies with HearNPV-NNg1 (Kenya) and HearNPV-G4 (China) exhibit the most and least virulence respectively against *H. armigera* larvae [24]. Recently five naturally occurring HearNPV isolates, viz., SP1A, SP1B, LB1, LB3 and LB6 from Spain have been fully sequenced. The insecticidal properties of these five genotypes showed significant differences among them, the principal variation among them was located in homologous regions (*hrs*) and baculovirus repeat ORFs (*bro*) genes [1]. The current study is the first report on the complete genome of HearNPV of the Indian isolate.

In this study, we sequenced and analyzed HearNPV-L1 isolate from Ludhiana, Punjab, India. The HearNPV-L1 isolate showed high insecticidal activity against *H. armigera* larvae in the laboratory conditions, having lowest LC₅₀ values compared to several isolates, viz., Hingoli, Nanded, Nagpur, Bengaluru, Faridkot and Manasa obtained from different parts of India. We also compared the complete genome of HearNPV-L1 isolate with those of previously reported NNg1 from Kenya, G4 from China and a recently sequenced SP1A isolate from Spain. Our data showed that two ORFs were unique to HearNPV-L1 genome and showed highest similarity to NNg1 isolate of Kenya, apart from these there are marked similarities were also found between HearNPV-L1 and NNg1 in homologous region and *bro* genes.

Materials and methods

Virus and insects

HearNPV-L1 was originally isolated from a naturally infected *H. armigera* from Ludhiana (30.9100°N, 75.8500°E), Punjab, India, in the year 2011. The virus was propagated in laboratory stocks of healthy third instar *H. armigera* larvae. These insects were maintained on artificial diet using BOD (Biological Oxygen Demand) incubator maintained at 27.5 °C and 60% RH.

Virus purification, isolation and production

The original virus isolated was obtained from diseased larvae collected from the pigeonpea (*Cajanus cajan*) crop in Ludhiana, India. Occlusion bodies (polyhedra) from the infected larvae were purified by centrifugation at 2500 revolutions per minute (rpm) for 1 min at 4 °C. The debris was pelleted, whereas the occlusion bodies (OBs) remained in the supernatant. For the isolation of pure virus, the suspension was again subjected to centrifugation at 10,000 rpm for 5 min at 4 °C. The OBs settled as a pellet at the sides of the walls of the centrifuge tube. This process was repeated twice to isolate only the OBs leaving behind the contaminants. The pellet obtained was then suspended in 1 ml of double distilled water and stored at −20 °C for further use. The virus obtained from the above procedure isolated from the field collected diseased larvae was further confirmed by conducting a series of bioassay tests on *H. armigera* in order to determine their specificity. The same stock was used for isolation of Genomic DNA for Next generation sequencing through Illumina NextSeq 500 technology.

Isolation of viral DNA

DNA extractions were performed using 0.1 M sodium carbonate solution, the final suspension included 0.5 mgml^{−1} of proteinase K and 1% SDS and incubated overnight at 37 °C. This was followed by Phenol: Chloroform: Isoamyl Alcohol (25:24:1 v/v) extraction with the final absolute ethyl alcohol precipitation [28]. The DNA was re-suspended in a TrisEDTA buffer (composed of 10 mM Tris-HCl, 1 mM EDTA, pH 8.0). The extracted DNA was visualized in Vilber Lourmat gel documentation system on 1% agarose gel.

Next generation sequencing of HearNPV-L1

The quality of DNA extracted from the NPV was checked on 1% agarose gel for the single intact band. 1 µL of each

sample used for determining concentration using Qubit 2.0 fluorometer. The paired end sequencing libraries was prepared using Illumina True seq Nano DNA HT library preparation kit as per the manufacturer's instructions. The amplified libraries were analyzed in bioanalyser 2100 (M/s Agilent Technologies) using High sensitivity HS DNA chip as per manufacturer's instructions. The library was further sequenced on Next Seq 500 Illumina 2 × 150 bp chemistry.

Sequence analysis

The genome analysis was performed by reference genome assembly of the Nucleopolyhedrosis virus from National Centre for Biotechnology Information (NCBI). The analysis revealed that high quality data covered more than 95% of reference genome. All mapped reads were de-novo assembled by SPAdes (3.5.0) software on default parameter. A total of 99 scaffolds were generated. A gene prediction was carried out from generated scaffold, which was annotated by using BlastX against NCBI nr database. ORFs encoding more than 50 amino acids (150 bp) were considered to be protein-encoding and named as putative genes. One hundred and thirteen (113) ORFs in the final dataset were checked for maximum alignment with known baculovirus gene homologues from GenBank. ORFs with an overlaps of *hr* region were not included from the alignment analysis. Homology searches were done through the NCBI using BLAST. Multiple alignment and percentage identities were performed using the Clustal W 2.0.

Results

Analysis of HearNPV-L1 genome

The HearNPV-L1 genome was assembled into a contiguous sequence of 136,740 bp in length (GenBank accession number; KT013224) with GC content of 39.19% and contained 113 ORFs that could encode polypeptides with more than 50 amino acids. The contiguous sequence of the entire HearNPV-L1 genome was analyzed and compared with SP1A [1], NNg1 [21], G4 [34] and Ac [2] genomes. The identified HearNPV-L1 ORFs were plotted on the circular map shown in (Fig. 1). Detailed information on the size, location and percent homology of HearNPV-L1 ORFs is given in electronic supplementary material Table 1 together with those of SP1A, NNg1, G4 and *Autographa californica* multicapsid NPV (AcMNPV).

The 113 ORFs identified in the HearNPV-L1 genome were distributed almost evenly along both strands of genomic DNA, with 55 ORFs facing towards clockwise and 58 ORFs towards anti clockwise. The HearNPV-L1

genome is almost collinear with those of SP1A, NNg1 and G4. In addition, a total of 76 L1 ORFs had homologous ORFs on the AcMNPV genome, sharing low amino acid sequence identities ranging from 7 to 64.6%, except of *Polh*, which shared higher amino acid sequence identities of 85% (Electronic supplementary material Table 1).

Presence and absence of ORFs in genomes with higher virulence

The HearNPV-L1 genome contained two putative ORFs, ORF 18 and ORF 19 of 300 bp and 401 bp respectively unique to HearNPV-L1 genome. These ORFs even not found in AcMNPV also. However, the function of these ORFs is unknown as per comparison of L1 with SP1A, NNg1 and G4. It shows that ORF 62 of NNg1 is absent in L1, SP1A and G4 genome. Out of 113 ORFs only one had early promoter motif. Out of 113 ORFs 20 ORFs were 100% identical in SP1A, 21 in NNg1 and 25 in G4 genomes, respectively.

Comparative analysis of five core genes

Five Core genes were chosen for comparison among L1, SP1A, NNg1 and G4 genomes which play an important role in insecticidal activity. *Polh* (1), DNA binding protein (*dbp*) (1), Inhibitor of apoptosis (*iap*) (2), per os infectivity factor (*pif*) (3) and Late expression factor (*lef*) (9) were studied. Upon their comparison it was found that there was slight nucleotide level difference was present in *lef5* of HearNPV-L1 (849 bp in size) which was 99 bp shorter than the SP1A, NNg1 and G4 genome, whereas *lef2* (696 bp in size) and was 21 bp shorter than *lef2* of SP1A, 15 bp greater than *lef2* of NNg1 and 30 bp shorter than *lef2* of G4 genome. All the core protein genes of HearNPV-L1 were 97–100% identical to SP1A, NNg1 and G4 genome (Table 1).

Analysis of homologous regions (*hrs*)

Most of the baculovirus genome contain homologous repeated region, which also contains a “core” imperfect palindrome and dispersed unevenly among the genome in AT rich intergenic region [15]. The *hr* repeats initiates the DNA replication in transient assays [22]. The baculovirus *hrs* acts as enhancers for RNA polymerase mediated transcription of baculovirus early promoter [7, 26]. It functions as origin of the DNA Replication in transient replication assays [5, 6]. They are also site for recombinant and rearrangement in baculovirus genomes [16]. There are five *hr* regions are identified as *hr1*, *hr2*, *hr3*, *hr4* and *hr5*, whose sizes were 1540, 1395, 133, 1790 and 1011 bp, respectively. A palindrome was found in *hr1* region of

Fig. 1 Circular genome map of HearNPV-L1 isolated from *Helicoverpa armigera* collected from Ludhiana, India. The position, size and direction for the 113 open reading frames (ORFs) listed in Table 1 are presented as

different colour bars, along with ORF numbers and gene names. The Pink coloured bars indicate the location and relative size of the five homologous regions (*hrs*) and a direct repeat sequence (*dr*)

HearNPV-L1 genome, which looks quite similar to those of NNg1, SP1A, and G4 genomes, respectively. The repeat contains an imperfect palindrome sequence of 31 bp, in which two homologous sequences of the IE-1 binding motif (IBM) represented by 5'-ACBYGTAA-3' in AcMNPV *hrs* [18] were located in an inverted manner (Fig. 2 5'-ATATGAAA-3') in HearNPV-L1 genome.

The locations of *hr2* and *hr3* were conserved among SP1A, NNg1 and G4. However, *hr2* and *hr3* in HearNPV-L1, SP1A and NNg1 were separated by three bro genes (ORF43 (*bro-a*), ORF44 (*bro-b*) and ORF45 (*bro-c*) for HearNPV-L1, ORF57 (*bro-a*), ORF58(*bro-b*) and ORF59 (*bro-c*) for SP1A and ORF59 (*bro-a*), ORF60(*bro-b*) and ORF61 (*bro-c*) for NNg1 whereas there was only two *bro*

Table 1 Comparative analysis of five core genes of HearNPV-L1 isolated from *Helicoverpa armigera* collected from Ludhiana, India

Sl no	L1				SPA1				NNg1				G4			
	Genes	ORF	Bp	Aa	ORF	Bp	Aa	ID%	ORF	Bp	aa	ID%	ORF	Bp	aa	ID%
1	Polyhedrin	1	741	246	1	741	246	100	1	741	246	100	1	741	246	100
2	Dbp	16	972	323	23	972	322	100	24	972	323	100	25	972	323	100
3	DNAPol	52	3063	1020	66	3063	1020	99	68	3063	1020	99	67	3063	1020	99
4	Iap2	47	753	250	61	759	252	98	64	753	250	99	62	753	250	99
5	Iap3	83	807	267	102	807	268	97	106	807	267	97	103	807	268	97
6	Pif3	78	600	199	97	600	199	99	101	600	199	99	98	600	199	99
7	Pif1	90	1587	528	110	1587	528	100	114	1587	528	99	111	1587	528	99
8	Pif2	110	1152	383	133	1152	383	99	140	1152	383	99	132	1152	383	100
9	Lef6	15	564	187	22	564	187	100	23	564	187	100	24	564	187	99
10	Lef11	24	384	127	30	384	127	100	32	384	122	100	32	384	127	100
11	Lef8	28	2706	901	36	2706	901	99	38	2706	901	100	38	2706	901	98
12	Lef9	39	1560	519	53	1560	519	99	55	1560	519	99	56	1560	519	99
13	Lef3	50	1140	379	64	1140	379	99	67	1140	379	99	65	1140	379	99
14	Lef4	62	2019	461	77	2019	461	98	81	2019	461	100	78	2019	461	100
15	Lef5	70	849	282	85	948	315	100	89	948	315	100	87	948	315	99
16	Lef2	95	696	231	117	717	238	98	123	681	226	98	117	726	241	93
17	Lef1	102	738	245	125	738	245	99	131	738	245	98	124	738	245	96

(A)

AAACAAATTACGTCATTCGTTTTAAAAATATTGCATCATCTTTAAT
 TTGAAACTAGCCCGCGCTTTCATATGAACCGTCGGCGAAGATCG
 ATAAATTTGTTCTAGAACATTCGATGGTTTGACCCAAAAAACA
 AATGACGTCATATAGCGTGCCTCCAATC

(B)

5' TAGCCCGCGCTTTCATATGAACCGTCGGCG
 3' ATCGGGCGCGAAAGTATCTTTGGCAGCCGC

Fig. 2 The consensus repeat sequence identified in the genome of HearNPV-L1 h. An imperfect palindrome in repeat region indicated in grey shade (a) and boxed sequence represent putative IE-1 binding motif (AcMNPV and IBM) like sequence

genes were found in G4 genome, viz., ORF59 (*bro-a*) and ORF60 (*bro-b*) between *hr2* and *hr3*. It has previously been suggested that *hr2* and *hr3* comprised a single *hr* before the insertion of *bro* genes occurred between them [3]. Our comparative analysis of *bro* gene sequences indicated that the NNg1 *bro-b* gene, the third *bro* gene present between *hr2* and *hr3*, was generated through duplication of the *bro-c* gene, thus suggesting that this region was a hot spot for recombination. These baculovirus variants contain different numbers of repeats, and this suggests that *hrs* show high structural plasticity. It is also interesting to note that, despite the high degree of structural similarity among *hrs*, the actual sequences are entirely different among baculovirus species. The arrangement of these *hrs* on the genome is almost the same as NNg1 and it is possible that

variability in the *hr* sequences affect not only progeny of virus production, but also the insecticidal activity of the *Helicoverpa* spp. NPVs [17]. The homologous regions are also suggested to be responsible for the difference in the virulence factor of NPVs [11].

Analysis of *bro* genes

Baculovirus-repeated ORF genes (*bro*) were first identified in baculovirus genomes as a repetitive gene and constitute a gene family. Most *bro* genes contain a 41 aa core sequence at the N-terminal half and several different domains throughout the sequence. Based on the similarity of those different domains, *bro* genes were separated into four groups, I–IV [9].

The HearNPV-L1 genome contained four *bro* genes, *bro-a* (ORF44), *bro-b* (ORF45), *bro-c* (ORF46) and *bro-d* (ORF85), encoded 246, 412, 267 and 500 amino acids, respectively (Table 2). The HearNPV-L1 *bro-a*, which shared relatively high amino acid sequence identities of 75% with *bro-a* and 57% with *bro-b* (ORF60) from SP1A and G4 respectively, however, *bro-c* was found absent in G4 genome. The HearNPV-L1 *bro-b* had no homologues in any of NNg1, SPA1 and G4. The HearNPV-L1 *bro-c*, which shared 91% amino acid sequence identity with *bro-b* (ORF61) of SP1A and (ORF61) of NNg1 (ORF60) in G4. The HearNPV-L1 *bro-d*, was the most conserved *bro* gene among the *Helicoverpa* spp. NPVs, sharing amino acid sequence identities of 99% with *bro-c* (ORF 59) from

Table 2 Comparative analysis of bro genes of HearNPV-L1 isolated form *Helicoverpa armigera* collected from Ludhiana, India

Sl no	L1				SP1A				NNg1				G4			
	Genes	ORF	Bp	Aa	ORF	Bp	aa	Identity	ORF	Bp	aa	Identity	ORF	Bp	Aa	Identity
1	Bro-a	44	1086	246	57	1074	357	91	59	1086	361	99	59	735	244	100
2	Bro-b	45	465	412	58	306	101	84	60	306	101	81	60	1058	352	99
3	Bro-c	46	1047	267	59	1047	352	75	61	1059	352	73				
4	Bro-d	85	1506	500	104	1506	501	99	108	1506	501	99	104	1506	501	99

SP1A and ORF 61 of NNg1. NPVs, shared high amino acid sequence identity with other NPV genomes based on the fact that HearNPV-L1 *bro-b* and *bro-c*, along with *bro-a*, were located between *hr2* and *hr3*, which was considered to be a hot spot for baculovirus genome recombination [19], these *hrs* may be responsible for *bro-c* duplication.

Discussion

Development of an effective biopesticide against *H. armigera* is the current need in India and rest of the world. In this direction an effort was made to identify the most virulent NPV. HearNPV-L1 isolate was identified after a series of bioassay experiments and its insecticidal activity was determined. The insecticidal activity of HearNPV-L1 isolate expressed in terms of 50% lethal concentration was 1.92×10^7 in third instar of *H. armigera* larvae compared to the isolates obtained from Hingoli (8.2×10^7), Nanded (6.8×10^7), Nagpur (4.7×10^7), Bengaluru (4.7×10^7), Faridkot (10.6×10^7) and Manasa (7.5×10^7) collected from different parts of India. In the current study we attempted whole genome sequencing through next generation sequencing (Illumina NextSeq 500) technology of HearNPV-L1 isolate.

We compared L1 genome with those of SP1A, NNg1 and G4 genomes and found that HearNPV-L1 showed 97–99% homology to these genomes. Majority of putative genes identified in HearNPV-L1 genome shared a high degree of similarity and amino acid sequence identity with NNg1 genome than SP1A and G4 genomes. Genes involved in insecticidal activity crucial for infectivity and productivity such as *p74* and *per os infectivity factor* (promote preoral infection of insect larvae) were analyzed at nucleotide level to find observed differences in insecticidal activity between *Helicoverpa* spp NPVs, HearNPV-L1 shared 97–99% amino acid sequence identity to those of SP1A, NNg1 and G4, suggesting that *p74* and *pifs* are not responsible for the insecticidal activity in L1 isolate. Further comparative analysis of five core genes, viz., *Polyhedrin (pol-h)*, *DNA binding protein (dbp)*, *DNA polymerase (DNApol)*, *Inhibitor of apoptosis (iap)* and *Late*

expression factor (lef) resulted'' in 97–100% identity among *Helicoverpa* spp. genomes (SP1A, NNg1 and G4), indicating the role of these genes is not significant in insecticidal activity.

Detailed comparative study of HearNPV-L1 genome identified several putative genes whose functions were unknown, however, might be responsible for insecticidal activity of HearNPV-L1. Two putative genes viz; ORF 18 and ORF 19 of 299 and 399 bp respectively found unique to HearNPV-L1 genome. These two genes had no homologue in any of the genome (SP1A, NNg1 and G4) compared in this study. Functional characterization of these putative ORFs of HearNPV-L1 genome may help to elucidate whether these genes are associated with the differences in insecticidal activity among *Helicoverpa* spp. NPVs.

HearNPV-L1 genome contains 5 h (*hr1-hr5*), these regions were found 84–100% similar to *hr* region of NNg1, SP1A and G4 genomes. However, size of *hr* significantly varied among HearNPV-L1, NNg1, SP1A and G4 genomes. Among all the 5 h present in L1 genome *hr2* was found to be most identical to those of SP1A, NNg1 and G4 genomes. It may be possible that observed differences among *hrs* or variations in overall *hr* sequence length affects not only insecticidal activity but progeny virus production also in *Helicoverpa* spp. NPVs [24]. A direct repeat (*dr*) was found of 152 bp in HearNPV-L1 genome which was 99% identical to the direct repeat of NNg1 (162 bp). The presence of a homologue of HearNPV-L1 direct repeat in NNg1 shows sharing of unique traits among strains of different geographical regions.

A total of four bro genes were observed in L1 genome, viz., *bro-a*, *bro-b*, *bro-c* and *bro-d*, differences were observed between HearNPV-L1 and other three *Helicoverpa* spp. NPVs in *bro* gene diversity. The early promoter present in *bro-b* gene of L1 has a homologue in SP1A and NNg1 genome but lacked in G4 genome [33]. *bro-c* and *bro-d* of L1 genome was homologous to *bro-c* and *bro-d* of SP1A and NNg1 genomes respectively. Based on this comparative analysis of *bro* genes, it is possible that observed *bro* gene diversity may be responsible for observed insecticidal activity between HearNPV-L1, SP1A, NNg1 and G4 genomes.

Overall comparative analysis of HearNPV-L1 genome with three other *Helicoverpa* spp. NPVs SP1A, NNg1 and G4 genomes indicates that apart from *hr* region and *bro* genes there are several other putative genes and nucleotide sequences may be responsible for insecticidal activity in HearNPV-L1 isolate of India and other virulent genomes, viz., SP1A from Spain and NNg1 isolate of Kenya. However, further functional analysis of the hypothetical (putative) genes of HearNPV-L1 isolate may identify the genes which are crucial for the virulence and insecticidal activity.

Acknowledgements This study was funded by “National Agricultural Science Fund” (NASF) of Indian Council of Agricultural Research, New Delhi, India. Authors are thankful to the Director, Dr. Abraham Verghese, ICAR-National Bureau of Agricultural Insect Resources (NBARI), Hebbal, Bengaluru, for his support to conduct the work. First author is thankful to Jain University, Bengaluru, as this work is part of her Ph.D. Thesis.

References

1. Arrizubieta M, Simon O, Williams T, Caballero P. Genomic sequences of five *Helicoverpa armigera* nucleopolyhedrovirus genotypes from Spain that differ in their insecticidal properties. *GenomeA*. 2015;3:e00548–615.
2. Brown SE, Maruniak JE, Knudson DL. Baculovirus (MNPV) genomic variants: characterization of *Spodoptera exempta* MNPV DNAs and comparison with other *Autographa californica* MNPV DNAs. *J Gen Virol*. 1985;66:2431–41.
3. Chen X, Zhang WJ, Wong J, Chun G, Lu A, McCutchen BF, Presnail JK, Herrmann R, Dolan M, Tingey S, Hu ZH, Vlak JM. Comparative analysis of the complete genome sequences of *Helicoverpa zea* and *Helicoverpa armigera* single-nucleocapsid nucleopolyhedroviruses. *J Gen Virol*. 2002;83:673–84.
4. Felipe Alves CA, Ishikawa H, Ikeda M, Kobayashi M. A large homologous region of the *Hyphantria cunea* nucleopolyhedrovirus genome, as a powerful and versatile enhancer in insect expression systems. *Virus Genes*. 2002;39:403–8.
5. Gettig RR, McCarthy WJ. Genotypic variation among wild isolates of *Heliothis* spp nuclear polyhedrosis viruses from different geographical regions. *Virology*. 1982;117:245–52.
6. Grabherr R, Ernst W, Oker-Blom C, Jones I. Developments in the use of baculoviruses for the surface display of complex eukaryotic proteins. *Trends Biotechnol*. 2001;19:231–6.
7. Guarino LA, Summers MD. Functional mapping of a trans-activating gene required for expression of a baculovirus delayed-early gene. *J Virol*. 1986;57:563–71.
8. Guarino LA, Gonzalez MA, Summers MD. Complete sequence and enhancer function of the homologous DNA regions of *Autographa californica* nuclear polyhedrosis virus. *J Virol*. 1986;60:224–9.
9. Hamm JJ. Relative susceptibility of several noctuid species to a nuclear polyhedrosis virus from *Heliothis armiger*. *J Invertebr Pathol*. 1982;39:255–6.
10. Harrison RL, Puttler B, Popham HJR. Genomic diversity in European *Spodoptera exigua* multiple nucleopolyhedrovirus isolates. *J Gen Virol*. 2008;89:775–90.
11. Hayakawa T, Okano RK, Seong SL, Goto C, Meda S. Sequence analysis of the *Xestia c-nigrum* granulovirus genome. *Virology*. 1999;262:277–97.
12. Herz A, Kleespies RG, Huber J, Chen X, Vlak JM. Comparative pathogenesis of the *Helicoverpa armigera* single-nucleocapsid nucleopolyhedrovirus in noctuid hosts of different susceptibility. *J Invertebr Pathol*. 2003;83:31–6.
13. Hu Y-C. Baculovirus vectors for gene therapy. *Adv Virus Res*. 2006;68:287–320.
14. Jones KA, Zelazny B, Ketunuti U, Cherry AC, Grzywacz D. South-east Asia and the Western Pacific. In: Hunter-Fujita FR, Entwistle PF, Evans HF, Crook NE, editors. *Insect viruses and pest management*. Chichester: Wiley; 1998. p. 244–58.
15. Kool M, Ahrens CH, Vlak JM, Rohrmann GF. Replication of baculovirus DNA. *J Gen Virol*. 1995;76:2103–18.
16. Kuzio J, Pearson MN, Harwood SH, Funk CJ, Evans JT, Slavicek J, Rohrmann GF. Sequence and analysis of the genome of a baculovirus pathogenic for *Lymantria dispar*. *Virology*. 1999;253:17–34.
17. Lee HH, Miller LK. Isolation of genotypic variants of *Autographa californica* nuclear polyhedrosis virus. *J Virol*. 1978;27:754–67.
18. Leisy DJ, Rasmussen C, Owusu EO, Rohrmann RF. A mechanism for negative gene regulation in *Autographa californica* multinucleocapsid nuclear polyhedrosis virus. *J Virol*. 1997;71:5088–94.
19. Li L, Li Q, Willis LG, Erlandson M, Theilmann DA, Donly C. Complete comparative genomic analysis of two field isolates of *Mamestra configurata* nucleopolyhedrovirus—a. *J Gen Virol*. 2005;86:91–105.
20. Mäkelä AR, Oker-Blom C. Baculovirus display: a multifunctional technology for gene delivery and eukaryotic library development. *Adv Virus Res*. 2006;68:91–112.
21. Ogembo JG, Caoili BL, Shikata M, Chaeychomsri S, Kobayashi M, Ikeda M. Comparative genomic sequence analysis of novel *Helicoverpa armigera* nucleopolyhedrovirus (NPV) isolated from Kenya and three other previously sequenced *Helicoverpa* spp NPVs. *Virus Genes*. 2009;9:261–72.
22. Possee RD, Rohrmann GF. Baculovirus genome organization and evolution. In: Miller LK, editor. *The baculoviruses*. New York: Plenum Press; 1997. p. 109–40.
23. Smith GE, Summers MD. Analysis of baculovirus genomes with restriction endonucleases. *Virology*. 1978;89:517–27.
24. Sun X. History and current status of development and use of viral insecticides in China. *Viruses*. 2015;7:306–19.
25. Tang P, Zhang H, Li YU, Han B, Wang GZ, Qin QL, Zhang Z. Genomic sequencing and analyses of HearMNPV-a new multi-nucleocapsid nucleopolyhedrovirus isolated from *Helicoverpa armigera*. *Virol J*. 2012;9:168.
26. Theilmann DA, Stewart S. Molecular analysis of the trans-activating IE-2 gene of *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus. *Virology*. 1992;187:84–96.
27. Theilmann DA, Blissard GW, Bonning B, Jehle JA, O'Reilly DR, Rohrmann GF, Thiem S, Vlak JM, Baculoviridae. In: Fauquet CM, Mayo MA, Maniloff J, Desselberger U, Ball LA (eds). *Virus taxonomy*. Eighth report of the international committee on taxonomy of viruses. 2005:177–85.
28. Woo SD. Rapid detection of multiple nucleopolyhedroviruses using polymerase chain reaction. *Mol Cells*. 2001;11:334–40.
29. Yang YL, Li WuY. Current status of insecticide resistance in *Helicoverpa armigera* after 15 years of Bt cotton planting in China. *J Econ Entomol*. 2013;106:375–81.
30. Zalucki MP, Daglish G, Firempong S, Twine PH. The biology and ecology of *Heliothis armigera* (Hubner) and *Heliothis punctigera* Wallengren (Lepidoptera, Noctuidae) in Australia-what do we know. *Aust J Zool*. 1986;34:779–814.
31. Zalucki MP, Murray DHA, Gregg PC, Fitt GF, Twine PH, Jones C. Ecology of *Helicoverpa armigera* (Hubner) and *Heliothis punctigera* (Wallengren) in the Inland of Australia-larval

- sampling and host-plant relationships during winter and spring. Aust J Zool. 1994;42:329–46.
32. Zhang G. Research, development and application of *Heliothis* viral pesticide in China. Resour Environ Yangtze Val. 1994;3:36–41.
33. Zhang CX, Ma XC, Guo ZJ. Comparison of the complete genome sequence between C1 and G4 isolates of the *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus. Virology. 2004;333:190–9.
34. Zhong JG, Jun QG, Wang D, Shao YM, Tang QY, Zhang CX. Biological comparison of two genotypes of *Helicoverpa armigera* single-nucleocapsid nucleopolyhedrovirus. Biocontrol. 2006;51:809–20.