Genomics 112 (2020) 3903–3914

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Characterization of a novel alphabaculovirus isolated from the Southern armyworm, *Spodoptera eridania* (Cramer, 1782) (Lepidoptera: Noctuidae) and the evolution of *odv-e66*, a bacterium-acquired baculoviral chondroitinase gene

Daniela T. Rodrigues[a], Lenen Peterson[a], Lucas Boeni de Oliveira[a], Daniel R. Sosa-Gómez[b], Bergmann Morais Ribeiro[c], Daniel M.P. Ardisson-Araújo[a,*]

[a] Laboratory of Insect Virology, Department of Biochemistry and Molecular Biology, Federal University of Santa Maria, Santa Maria, RS, Brazil
[b] Embrapa Soja, Londrina, PR, Brazil
[c] Laboratory of Baculovirus, Cell Biology Department, University of Brasilia, Brasília, DF, Brazil

ARTICLE INFO

Keywords:
Baculovirus
Spodoptera eridania
Alphabaculovirus
Evolution
odv-e66

ABSTRACT

The Southern armyworm *Spodoptera eridania* (Lepidoptera: Noctuidae) is native to the American tropics and a polyphagous pest of several crops. Here we characterized a novel alphabaculovirus isolated from *S. eridania*, isolate Spodoptera eridania nucleopolyhedrivurus CNPSo-165 (SperNPV-CNPSo-165). SperNPV-CNPSo-165 occlusion bodies were found to be polyhedral and to contain virions with multiple nucleocapsids. The virus was lethal to *S. eridania* and *S. albula* but not to *S. frugiperda*. The SperNPV-CNPSo-165 genome was 137.373 bp in size with a G + C content of 42.8%. We annotated 151 ORFs with 16 ORFs unique among baculoviruses. Phylogenetic inference indicated that this virus was closely related to the most recent common ancestor of other *Spodoptera*-isolated viruses.

1. Introduction

The Southern armyworm *S. eridania* (Cramer, 1782) (Lepidoptera: Noctuidae) is a moth native to the American tropics [1] with larvae that are extensively polyphagous [2]. In Brazil, *S. eridania* has become a pest of expanding importance in crops of soybean, cotton, fruits, and weeds [3–6] due to both tolerance to high density population and a high degree of defoliation caused by feeding larvae [5,7]. The use of broad-spectrum chemical insecticides is the main method to control *S. eridania* [8], which can lead to selection of resistant pests and death of non-target organisms (*e.g.* natural enemies, pollinators, and soil arthropods). Moreover, as xenobiotics, chemical insecticides may cause bioaccumulation and intoxication of human and other vertebrate animals [9].

Methods using biocontrol agents, like insect viruses, are important options to complement or even replace pesticides in an integrated pest management program [10,11]. Among the insect viruses found in nature, members of family *Baculoviridae* have been used as effective biopesticides [10]. For example, isolate 2D of the Anticarsia gemmatalis multiple nucleopolyhedrovirus (AgMNPV-2D) has been applied since early 1980s for biocontrol of the soybean pest *Anticarsia gemmatalis* in Brazil [12]. The success of the program in Brazil allowed the use of AgMNPV-2D in other countries of South America, including Argentina, Bolivia, Mexico, and Paraguay [12,13]. The family *Baculoviridae* contains a diverse group of insect-specific viruses with circular double-strand DNA genome, whose sizes range from 80 to 180 kbp and code for 90–180 genes [14]. The family is currently divided into four genera: *Alphabaculovirus* and *Betabaculovirus* that contain members infectious to larvae of lepidopterans (caterpillars of butterflies and moths), *Gammabaculovirus* that contains members infectious to larvae of hymenopterans (specifically sawflies, wasps with caterpillar-like behavior), and *Deltabaculovirus* that contains members infectious to larvae of dipterans (specifically mosquitoes) [14,15]. A hallmark feature of baculovirus infection is the assembly of virions into occlusion bodies (OBs), which protect the virions from environmental adversities [16]. The viral infection begins when the host feeds on substrates contaminated with OBs. Two viral phenotypes are produced during the complete infection cycle. The occlusion-derived virus (ODV), which is responsible for the primary infection, is released in the insect midgut after dissolution of the OBs and infects the midgut epithelium cells. Then, the infected cell produces the budded virus (BV), which is

* Corresponding author.
E-mail address: daniel.araujo@ufsm.br (D.M.P. Ardisson-Araújo).

https://doi.org/10.1016/j.ygeno.2020.06.047
Received 14 February 2020; Received in revised form 24 May 2020; Accepted 28 June 2020
Available online 03 July 2020

responsible for secondary infection and spreads the viral disease from the midgut throughout the insect body [15,17]. At the end of the infection, the larvae exhibit a weakened, melanized tegument and an internal anatomy, which has been largely liquefied [18].

In a published study [19] a baculovirus isolated from the caterpillar *S. eridania* (Spodoptera eridania nucleopolyhedrovirus isolate 251, or SperNPV-251) was sequenced and characterized, but according to the phylogeny and parameters of species demarcation, this virus was related to Spodoptera litura nucleopolyhedrovirus II (SpltNPV-II), a representative of an unclassified virus lineage. In this work, we characterized a second baculovirus isolated from the Southern armyworm *S. eridania* at the structural, biological, and molecular levels. The virus was found in larval cadavers of *S. eridania* exhibiting symptoms of baculovirus infection, which were obtained from the virus collection of the Brazilian Agricultural Research Corporation (EMBRAPA, 'Empresa Brasileira de Pesquisa Agropecuária') and named Spodoptera eridania nucleopolyhedrovirus isolate CNPSo-165 (SperNPV-CNPSo-165). Ultrastructural and bioassay analysis of the OBs and sequence analysis of the genome were carried out. This virus was found to be distinct from SperNPV-251 and a representative of a novel tentative species inside *Alphabaculovirus*.

2. Materials and methods

2.1. Virus sample and purification of OBs

Carcasses of *S. eridania* larvae with symptoms of baculovirus infection found between the cities of Ibipora and Jataizinho (State of Parana, Brazil) were sent to EMBRAPA in March 2011, and kept at −20 °C until the purification of the OBs [20]. The insect cadavers were homogenized with an equivalent volume of ddH2O (*w/v*), filtered through cotton gauze, and centrifuged at 5000 ×*g* for 10 min. The supernatant was discarded, and the pellet was suspended in the same volume of 0.5% SDS, and centrifuged at 5000 ×*g* for 10 min. The washing step was repeated three more times. The pellet was suspended in 0.5 M NaCl, centrifuged at 5000 ×*g* for 10 min and suspended in 2 ml ddH2O. OBs were loaded onto a sucrose gradient (40–65%), centrifuged at 130,000 ×*g* for 3 h. OBs were collected as a band and diluted five times with ddH2O. The suspension were collected by centrifugation at 7000 ×*g* for 10 min, diluted in ddH_2O (10^6 OBs / ml ddH_2O), and stored at 4 °C [20,21].

2.2. Electron microscopy

For SEM and TEM analysis, 100 μl of the OB-containing suspension at a concentration of 10^9 OBs/mL ddH2O were used for preparation according to previously published protocols [22]. For SEM, OBs (10^9 OBs/ml) were treated with acetone 1 X and then incubated at 25 °C for 1 h. The solution was loaded onto a metallic stub, dried overnight at 37 °C, coated with gold in a Sputter Coater (Balzers) for 3 min, and observed in a SEM Jeol JSM 840A at 10 kV. For TEM, pellets of purified OBs were fixed in Karnovsky fixative (2.5% glutaraldehyde, 2% paraformaldehyde, in 0.1 M, pH 7.2, cacodylate buffer) for 2 h, post-fixed in 1% osmium tetroxide in the same buffer for 1 h and then stained en bloc with 0.5% aqueous uranyl acetate, dehydrated in acetone, and embedded in Spurr's low viscosity embedding medium. The ultrathin sections were stained with uranyl acetate/lead citrate and observed in a TEM Jeol 1011 at 80 kV. Nucleocapsids were counted from five independent fields.

2.3. Viral DNA purification, genome sequence, assembly, and annotation

Viral DNA was purified from 200 μl of the OB-containing suspension (10^{10} OBs/mL of ddH_2O) according to previous protocols [21]. The viral genome was sequenced with the 454 Genome Sequencer (GS-FLX) Titanium. The reads were trimmed and used for the *de novo* assembly method in the software Geneious R9 [23] with a minimum pairwise identity of 98.4%. The open reading frames (ORFs) that started with a methionine codon (ATG) and encoded polypeptides of at least 50 amino acids were annotated using the same software and BLAST-X [24]. In order to identify homologous regions (*hrs*) present in the genome, DOTPLOT and Tandem Repeat Finder searches were performed using the Geneious R9 program to analyze the composition of the repeat region. The genomic DNA sequence was submitted to GenBank under the accession number **MT040195**.

2.4. Baculovirus phylogeny and species demarcation criterion

For phylogenetic analysis of baculoviruses, the MAFFT alignment [25] was carried out upon the nucleotide sequence of the 38 baculoviral core genes obtained from 93 baculovirus genomes (Supplementary Table 1). Afterwards, the alignments were concatenated and used to infer a maximum likelihood tree by using the Fast-tree method [26] and a Shimodaira-Hasegawa-like test [27]. To verify whether this virus corresponds to a new species, the nucleotide distances was estimated using the adjusted Kimura-2 parameter from partial sequences of three conserved baculovirus genes, including *lef-8*, *lef-9* and *polyhedrin* [28].

2.5. Gene content analyses

Each ORF found in the genome sequence was submitted individually to BLASTX [24] in order to find the identity to other baculoviruses. The ORFs with no BLAST matches were submitted to HHpred and SMART [29,30] to search for conserved domains. Moreover, the complete genome of the novel virus was compared to other alphabaculoviruses through the construction of syntenic maps using the progressive Mauve algorithm implemented in the software Geneious R9 [23]. In order to compare the gene content of the SperNPV-CNPSo-165 genome and other related baculoviruses, the genomes of several *Spodoptera*-isolated alphabaculoviruses were re-annotated according to the same criteria used for the novel virus and constructed a Venn Diagram (http://bioinformatics.psb.ugent.be/webtools/Venn/) to represent the number of ORFs shared between SperNPV-CNPSo-165 and the closest relatives.

2.6. In silico characterization of the baculovirus odv-e66 gene

A homolog of the baculoviral chondroitinase gene *odv-e66* was identified in the SperNPV-CNPSo-165 genome (SperNPV-ORF-124). To understand the evolution and acquisition of the *odv-e66* gene, the genome context was evaluated in relation to the genome of closely related species for each homolog found. Phylogenetic analyses based on the predicted amino acid sequence of *odv-e66* were performed using sequences retrieved from the BLASTX. The sequences were aligned by the MAFFT method [25] and the alignment used for phylogenetic inference with the PHYML method [31] under the substitution models LG + G (2.11). The optimal model was predicted by the MEGA7 software [32].

2.7. Insects and bioassays

S. eridania, *S. frugiperda*, and *S. albula* caterpillars used in this work were obtained from laboratory colonies established in 2015 with insects collected in the city of Londrina (Parana, Brazil). As previously described, early third-instar caterpillars were fed *ad libitum* by an artificial diet contaminated with the virus [33,34]. The insects were kept at 26 ± 1.5 °C, with relative humidity of 75 ± 10% and photoperiod of 14:10 (L:D). The assays with *S. eridania* were performed in triplicate using six virus concentrations (n = 45 per concentration) 2.0×10^3, 6.0×10^3, 18.00×10^3, 54.00×10^3, 162.00×10^3, and 486.00×10^3 OBs/ml and an untreated group (n = 44) was set up as control. The *S. frugiperda* caterpillar assays were performed using five virus concentrations (n = 30 insects per concentration) 1×10^6, 1×10^5, 1×10^4, 1×10^3 and 1×10^2 OBs/mL and an untreated group (n = 40) was set up as control. The *S. albula* caterpillar assays

were performed using five virus concentrations (n = 20 insects per concentration) 5 × 10^3, 2.5 × 10^3, 1.25 × 10^3, 0.62 × 10^3 and 0.31 × 10^3 OBs/mL and an untreated group (n = 20) established as control. Mortality was determined after 12 days. The results were analyzed by Probit in PoloPlus version 1.0. The LC50s were considered significantly different based on the non-overlap of the 95% confidence limits.

3. Results and discussion

3.1. Virus isolation and OBs ultrastructure

In 2011, dead *S. eridania* larvae were collected in soybean crops with clear symptoms of baculovirus infection, including tree top disease behavior and tegument discoloration and liquefaction (data not shown). The virus was catalogued in the EMBRAPA virus collection and called Spodoptera eridania NPV isolate CNPSo-165 (SperNPV-CNPSo-165). Ultrastructural analysis of purified OBs by SEM showed OBs with a predominantly polyhedral shape (Fig. 1A). The OB size was a mean diameter of 2.7 ± 0.4 μm. TEM analysis showed OBs occluding virions with several nucleocapsids per envelope, with a mean of 5.8 ± 2.6 nucleocapsids/envelope (Fig. 1B). The calyx, an electron-dense structure that surrounds mature polyhedra, was also observed (Fig. 1B, black arrow). All structures observed were similar to those observed previously in other alphabaculovirus OBs [35–37].

3.2. Virus etiology confirmation and bioassays

To confirm the infection etiology, we carried out a dose-mortality response bioassay in a laboratory colony of *S. eridania*. We confirmed that the virus was lethal to *S. eridania* larvae with a LC_{50} of 1.04 × 10^5 OB/mL (Table 1) towards 3rd instar larvae. The infected caterpillars exhibited a yellowish, easily ruptured tegument with melanotic pigment (data not shown) of the sort typically seen with other baculovirus infections [18]. We also tested the ability of SperNPV-CNPSo-165 to kill larvae of other species of the *Spodoptera*-complex, including *S. frugiperda* and *S. albula*. The isolate was not able to orally infect *S. frugiperda*, although was found to be lethal to *S. albula* with a LC_{50} of 735 OBs/mL in a much lower OB concentration than that observed for *S. eridania*. It is not clear why *S. albula* is more susceptible to SperNPV-CNPSo-165 OBs than the own *S. eridania*. In Brazil, population of *S. eridania* had become more common than *S. albula* (personal communication). Until now, we did not observe high prevalence of SperNPV in *S. eridania* population (data not shown). Importantly, both species have a very similar size and based on that we could assume that they use the same food amount. The host range of baculovirus may vary according to the viral species isolate; for instance, some isolate may be infectious to more than 20 hosts (*e.g.* AcMNPV, [38]) whereas others only infect a single host [39]. The ability to kill both *S. eridania* and *S. albula* may reflect the close degree of relatedly between these species [40]. We found statistical difference between the lethal concentrations observed for these two insects.

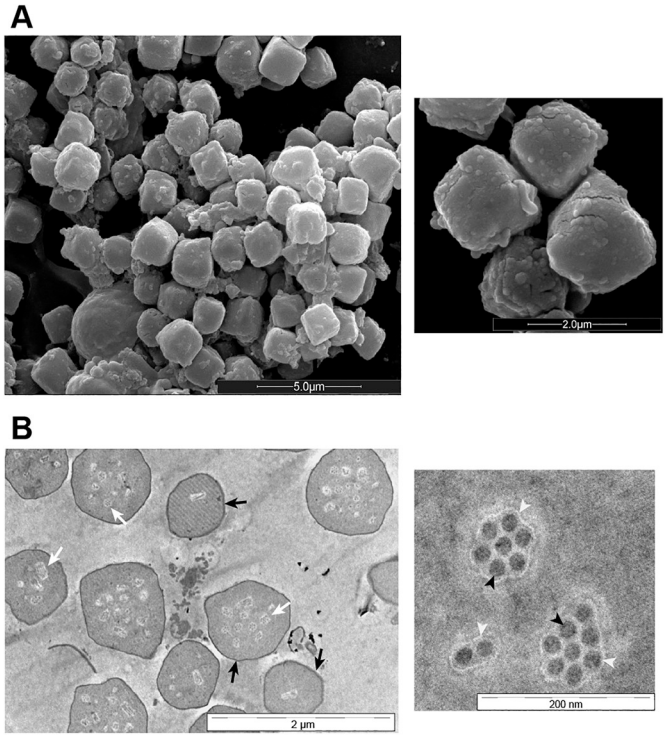

Fig. 1. Characterization of the SperNPV-CNPSo-165 isolate from dead Southern armyworm larvae with symptoms of baculovirus infection. (A) Scanning electron micrograph of SperNPV-CNPSo-165 OBs reveals their polyhedral shape (scale bar = 5.0 μm and 2.0 μm). (B) Transmission electron micrograph of SperNPV-CNPSo-165 OBs showing embedded virions (white arrow) with multiple rod-shaped nucleocapsids (black arrowhead) per ODV envelope (white arrowhead). The calyx is pointed by the black arrow (scale bar = 2.0 μm and 200 nm).

Table 1
Dose-mortality response of third instar larvae of *Spodoptera eridania* infected orally with SperNPV-CNPSo165.

Species	n[a]	Slope	LC_{50}	Fiducial limits
			(OB/ml)	(95%)
S. eridania	315	1.039 ± 0.126	1.04 × 10^5	7.07 × 10^5–16.76 × 10^5
S. albula	200	1.360 ± 0.335	7.35 × 10^2	3.66 × 10^2–11.53 × 10^2

[a] Number of insects tested.

3.3. Features of the SperNPV-CNPSo-165 genome sequence

Sequencing of the SperNPV-CNPSo-165 produced almost 10,000 reads with a mean size of 762.8 ± 214.3 bp and coverage of 40×. The reads were assembled into one single circular genome contig of 137.373 bp in size with a G + C content of 42.8%. The genome is in a range of the genome size and nucleotide distributions reported for other alphabaculoviruses (Supplementary Table 1). 151 ORFs potentially encoding proteins of 50 or more amino acids were identified and annotated (Supplementary Table 2), covering 88.54% of the genome, whereas 11.46% of the genome was found to be intergenic space. Among the annotated genes, we identified the 38 currently defined baculovirus core genes, 26 genes shared between alpha- and betabaculovirus genomes [41], and several auxiliary genes. Interestingly, regarding the intergenic spaces, the genome presented no typical homologous region, besides six direct repeats, four repeat regions, and two short repeats were found. Direct repeats were called dr1–6 and varied from 67 to 148 bp. Most of the drs (dr2–6) consist of only two repeats at the same direction with size varying 32–77 bp (data not shown). Only dr1 consisted of four concatenated repeats of 32 bp long. The repeat regions were called ReapReg1–4 and varied in size from 303 to 1.001 bp (Table 2). While other Spodoptera sp. NPVs possess *hr*s, a small number of other baculoviruses (RanuNPV [16], UrprNPV [21], and also betabaculoviruses like ErelGV [42] have been reported not to contain them. We inspected the percentage of both coding regions and intergenic spaces for each of the SperNPV-CNPSo165-related alphabaculovirus genomes and found that the novel virus is one of the three virus with the highest percentage of coding region (Supplementary Table 3).

3.4. Baculovirus phylogeny

We inferred the phylogenetic relationship of SperNPV-CNPSo-165 to other baculoviruses from core gene nucleotide alignments. The

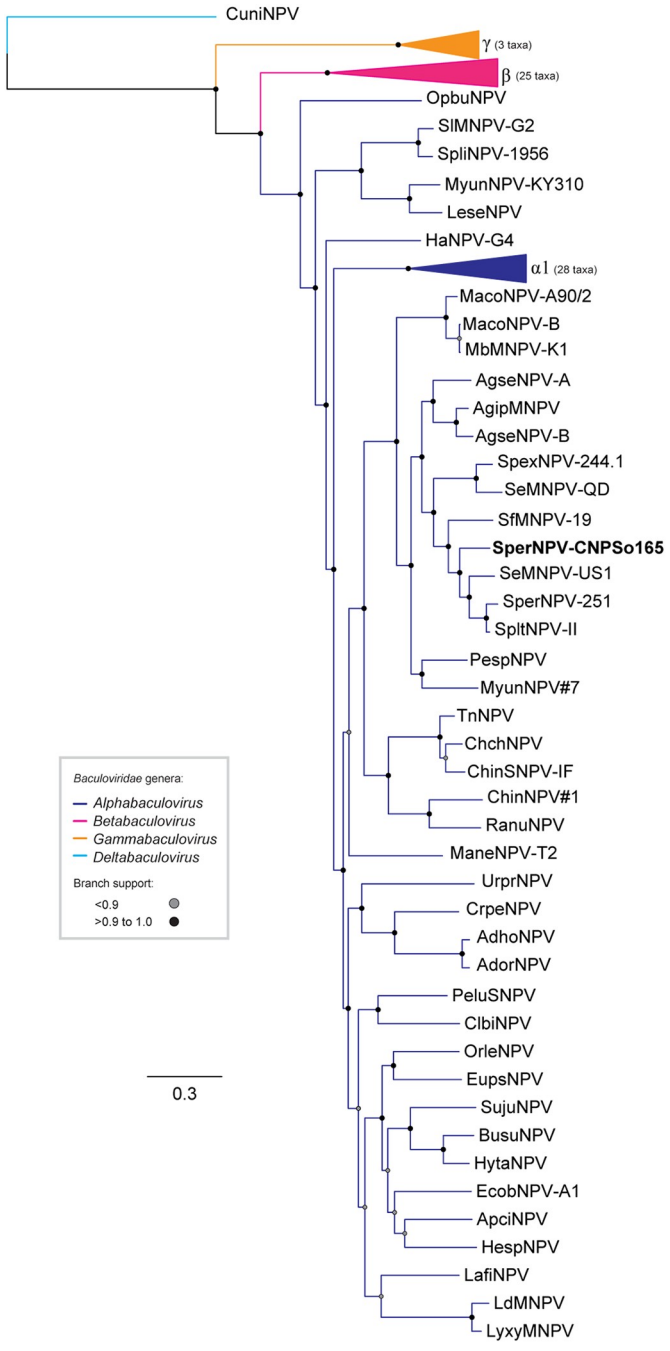

Fig. 2. Baculovirus phylogeny. The phylogeny shows that SperNPV-CNPSo-165 is an alphabaculovirus closely related to other *Spodoptera*-isolated viruses. The novel virus shares a common ancestor with a branch containing SeMNPV, SpltNPV-II, and SperNPV-251. The maximum likelihood tree was inferred based on the concatenated nucleotide sequences of the 38 core genes from several selected baculovirus genomes (Table S1) using the FastTree method. The branch support was determined by the SH-like method (black and grey closed circles). Some branches were collapsed for clarity: alphabaculovirus group 1, betabaculovirus (pink), gammabaculovirus (orange), and deltabaculovirus (CuniNPV, light blue). CuniNPV was used to root the tree. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

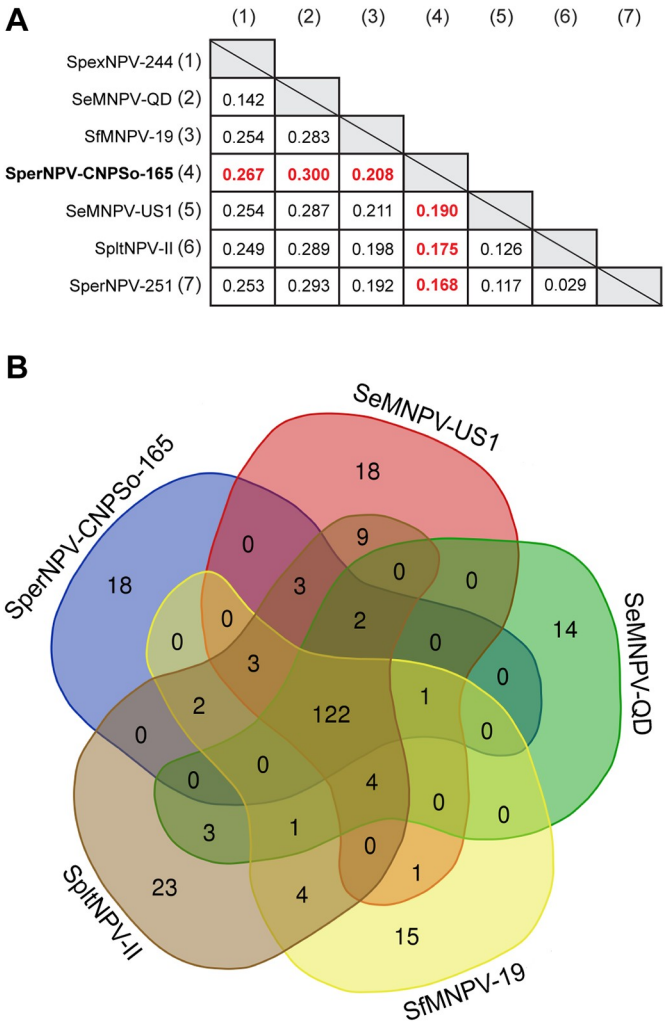

Fig. 3. Species demarcation criterion with the adjusted Kimura-2 parameter (aK2P) and gene content analysis of SperNPV-CNPSo-165 and other closely related viruses. (A) aK2P based on the concatenated fragments of partial polh/lef-8/lef-9 of the SperNPV-CNPSo-165 cluster. The distances were calculated using MEGA (Kimura 2-parameter model) [32], based on the species demarcation criteria [28]. In red, we show SperNPV-CNPSo-165 values that fulfill the criterion to establish a new species (more than 0.072 substitution/site). (B) Venn diagram comparing the gene content among SperNPV-CNPSo-165 and its closest relatives (SeMNPV-US1, SeMNPV-QD, SfMNPV-19, and SpltNPV-II). The gene content was compared by BLASTX to find homologs. A total of 243 genes were found: 122 were shared among all four virus genomes, and 16 were found only in the SperNPV-CNPSo-165 genome. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

phylogenetic tree exhibited a topology with a single branch containing all alphabaculoviruses as that observed for other previously published core gene baculovirus trees. SperNPV-CNPSo-165 clustered in a highly supported clade with SpexNPV-244.1, SeMNPV-QD, SfMNPV-19, SeMNPV, SpltNPV-II, and SperNPV-251. These viruses are part of a group of alphabaculovirus isolated solely from noctuid hosts. The SperNPV-CNPSo-165 isolate is closely related to the most recent common ancestor (m.r.c.a.) of the SeMNPV and SpltNPV-II viruses (Fig. 2). The most closely related alphabaculovirus to SperNPV-CNPSo-165, based on the nucleotide pairwise identity among the core genes, was found to be SeMNPV-US1 with 77.8% of identity, whereas the most distant alphabaculovirus was EppoNPV with 48.1% identity (Supplementary Table 1).

3.5. Species demarcation criteria

We investigated whether SperNPV-CNPSo-165 is a representative member of a new species inside genus *Alphabaculovirus*. Comparative analysis using the adjusted Kimura-2-parameter (aK2P) substitution

Table 2
Gene content and BLAST. Characteristics of the Spodoptera eridania nucleopolyhedrovirus isolate CNPSo165 (SperNPV-CNPSo165) genome: number, position, nucleotide and amino acid size of each ORF and homology search. Predicted ORFs are compared with homolog genes in four related genomes: Spodoptera litura nucleopolyhedrovirus II (SpltNPV-II), Spodoptera exigua multiple nucleopolyhedrovirus (SeMNPV-US1), Spodoptera frugiperda nucleopolyhedrovirus isolate 19 (SfMNPV-19), and Spodoptera exigua multiple nucleopolyhedrovirus isolate QD (SeMNPV-QD) and also Autographa californica multiple nucleopolyhedrovirus clone C6 (AcMNPV-C6).

ORF	Name	Position			Size (nt)	Size (aa)	SeMNPV-US1		SeMNPV-QD		SfMNPV-19		SpltNPV-II		AcMNPV	
							ORF	ID (%)	ORF	ID (%)	ORF	ID (%)	ORF	ID (%)	ORF	ID (%)
1	*polyhedrin*	1	>	741	741	247	1	97.6	1	94.72	1	97.15	1	97.56	8	85.95
2	*pp78/83*	820	<	2215	1396	465	2	64.79	2	61.19	2	65.99	2	78	9	27.93
3	*pk-1*	2271	>	3103	833	278	3	84.39	3	74.55	3	75.94	3	90.23	10	42.54
4	*hoar*	3144	<	5,42	2277	759	4	76.25	4	52.30	4	50.85	4	77.91	–	–
–	*Direct region 1*	–	–	–	135	–	–	–	–	–	–	–	–	–	–	–
5	*ac152-like*	5785	>	6,09	306	102	–	–	–	–	–	–	–	–	–	–
6	*sperNPV-ORF-6*	6221	>	7924	1704	568	–	–	–	–	–	–	–	–	–	–
–	*Direct region 2*	–	–	–	150	–	–	–	–	–	–	–	–	–	–	–
–	*Direct region 3*	–	–	–	126	–	–	–	–	–	–	–	–	–	–	–
7	*odv-e56(pif-5)*	8161	>	9282	1122	374	6	84.70	5	63.17	8	68.28	7	82.53	148	48.69
8	*me53*	9529	>	10,635	1107	369	7	75.72	6	60.12	9	67.43	8	80.06	139	25.22
9	*SperNPV-ORF-9*	11,143	>	11,292	150	50	–	–	–	–	–	–	–	–	–	–
–	*Direct region 4*	–	–	–	75	–	–	–	–	–	–	–	–	–	–	–
–	*Repeat region 1*	–	–	–	303	–	–	–	–	–	–	–	–	–	–	–
10	*SperNPV-ORF-10*	12,151	>	12,303	153	51	–	–	–	–	–	–	–	–	–	–
11	*f protein*	12,498	>	14,54	2043	681	8	87.21	7	63.12	11	73.44	10	83.03	23	24.73
12	*gp16*	14,646	<	14,936	291	97	9	83.52	8	61.29	12	65.26	11	87.50	130	31.46
13	*p24*	14,981	<	15,625	645	215	10	76.99	9	69.81	13	73.58	12	79.26	129	32.54
14	*SperNPV-ORF-14*	15,822	>	16,103	282	94	11	60.75	10	72.22	14	61.76	13	94.12	–	–
15	*lef-2*	16,066	>	16,707	642	214	12	76.06	11	61.21	15	64.93	14	79.62	6	40.09
16	*38.7 kDa*	16,749	<	17,891	1143	381	13	70.36	12	56.85	16	62.40	15	70.65	13	26.86
17	*lef-1*	17,891	<	18,541	651	217	14	79.63	13	64.11	17	64.62	16	78.24	14	43.98
18	*SperNPV-ORF-18*	18,585	>	19,016	432	144	15	73.23	14	48.31	18	57.36	–	–	–	–
19	*cathepsin*	19,004	<	20,011	1008	336	16	91.30	15	88.13	19	89.12	17	90.80	127	56.21
20	*chitinase*	20,061	>	21,773	1713	571	19	86.27	16	84.29	21	82.17	18	89.98	126	64.43
21	*SperNPV-ORF-21*	21,81	<	22,124	315	105	20	42.11	–	–	–	–	137	34.95	–	–
22	*SperNPV-ORF-22*	22,09	>	22,272	183	61	–	–	–	–	–	–	–	–	–	–
23	*SperNPV-ORF-23*	22,342	>	22,797	456	152	21	54.67	–	–	–	–	20	57.72	–	–
24	*nad-glutamate dehydrogenase*	23,205	>	23,495	291	97	–	–	–	–	–	–	–	–	–	–
25	*SperNPV-ORF-25*	23,659	>	25,008	1,35	450	23a	68.97	–	–	–	–	21	73.56	–	–
26	*gp37*	25,121	>	25,915	795	265	25	90.38	17	79.51	23	86.29	22	89.84	64	57.98
27	*ptp-2*	25,912	<	26,404	495	165	26	78.79	18	59.35	24	64.38	23	83.44	–	–
28	*egt*	26,516	>	28,093	1578	526	27	87.67	19	71.60	25	76.18	24	87.24	15	46.46
29	*SperNPV-ORF-29*	28,147	>	28,323	177	59	–	–	–	–	–	–	–	–	–	–
30	*SperNPV-ORF-30*	28,284	>	28,838	555	185	28	76.19	20	46.99	26	64.13	25	73.08	–	–
31	*SperNPV-ORF-31*	28,863	>	29,516	654	218	29	61.06	–	–	27	50.46	29	61.61	–	–
32	*SperNPV-ORF-32*	29,55	<	32,09	2541	847	30	66.32	21	49.76	28	46.11	30	67.87	–	–
33	*SperNPV-ORF-33*	32,257	>	32,877	621	207	31	61.23	22	41.92	29	60.50	31	66.99	–	–
34	*pkip*	32,953	<	33,471	519	173	32	64.85	23	66.86	30	66.86	32	70.76	24	26.28
35	*SperNPV-ORF-35*	33,514	<	33,843	330	110	33	61.61	25	56.76	–	–	33	60.36	–	–
36	*arif-1*	33,845	<	34,75	906	302	34	67.97	26	43.65	32	53.55	35	60.31	21	24.65
37	*pif-2*	34,692	>	35,843	1152	384	35	86.02	27	71.17	33	79.57	36	90	22	60.39
38	*pif-1*	35,961	>	37,475	1515	505	36	71.85	28	50.19	34	65.03	37	74.47	119	47.69
39	*ac120-like*	37,472	>	37,72	249	83	37	68.35	29	55.74	35	61.25	38	69.01	120	31.58
40	*fgf*	37,73	<	38,863	1134	378	38	66.08	30	43.95	36	47.12	39	79.89	32	24.65
41	*SperNPV-ORF-41*	39,092	<	39,214	123	41	–	–	–	–	–	–	–	–	–	–
42	*SperNPV-ORF-42*	39,279	>	39,989	711	237	40	65.13	31	53.65	37	60.34	41	70.89	–	–
43	*alk-exo*	40,04	<	41,281	1242	414	41	66.59	32	53.30	38	61.95	42	65.62	134	38.27
44	*SperNPV-ORF-44*	41,836	>	42,015	180	60	–	–	–	–	–	–	–	–	–	–
45	*SperNPV-ORF-45*	42,064	<	42,219	156	52	–	–	–	–	–	–	–	–	–	–
–	*Repeat region 2*	–	–	–	1002	–	–	–	–	–	–	–	–	–	–	–
46	*SperNPV-ORF-46*	42,882	<	43,217	336	112	42	69.23	33	57.80	39	71.43	43	74.77	–	–
47	*ac18-like*	43,216	>	44,37	1155	385	43	80.26	34	64.06	40	65.49	44	80.16	18	24.48
48	*SperNPV-ORF-48*	44,417	<	44,809	393	131	44	75.70	35	62.50	41	63.20	45	85.38	–	–
49	*rr2*	44,937	>	45,878	942	314	45	85.94	36	74.44	–	–	46	87.86	–	–
50	*calix,pep*	45,942	<	46,925	984	328	46	86.83	37	79.75	44	86.87	47	92.68	131	41.49
51	*ac117-like*	47,062	<	47,358	297	99	47	73.79	38	59.60	45	52.04	48	71.43	117	34.31
52	*sod*	47,485	<	47,94	456	152	48	84.78	40	75.50	47	82.24	51	86.49	31	68.71
53	*SperNPV-ORF-53*	48,131	>	48,49	360	120	49	35.09	–	–	–	–	52	75.38	–	–
54	*pif-3*	48,595	>	49,161	567	189	50	80.95	41	69.68	48	76.02	53	86.91	115	51.88
55	*SperNPV-ORF-55*	49,161	>	49,586	426	142	51	74.47	42	34.92	49	56.93	54	74.13	–	–
56	*SperNPV-ORF-56*	49,698	>	51,2	1503	501	52	74.51	43	54.22	50	62.38	55	71.34	–	–
57	*SperNPV-ORF-57*	51,238	>	51,918	681	227	53	84.65	44	79.04	51	83.33	56	86.46	–	–
58	*nrk1*	51,958	<	53,043	1086	362	54	67.13	45	61.62	52	57.42	57	66.76	–	–
59	*SperNPV-ORF-59*	53,091	<	53,381	291	97	–	–	–	–	–	–	–	–	–	–
–	*Direct repeat 5*	–	–	–	90	–	–	–	–	–	–	–	–	–	–	–
60	*SperNPV-ORF-60*	53,791	>	54,003	213	71	–	–	–	–	–	–	–	–	–	–
61	*dutpase*	54,065	>	54,496	432	144	55	87.41	46	77.46	53	66.67	58	85.31	–	–

(*continued on next page*)

Table 2 (*continued*)

ORF	Name	Position			Size (nt)	Size (aa)	SeMNPV-US1		SeMNPV-QD		SfMNPV-19		SpltNPV-II		AcMNPV	
							ORF	ID (%)	ORF	ID (%)	ORF	ID (%)	ORF	ID (%)	ORF	ID (%)
62	*SperNPV-ORF-62*	54,674	>	55,327	654	218	–	–	–	–	54	33.97	59	34.74	–	–
63	*p13*	55,39	>	56,238	849	283	56	83.58	47	74.63	55	82.67	60	84.64	–	–
64	*SperNPV-ORF-64*	56,262	<	57,031	771	257	–	–	–	–	–	–	–	–	–	–
65	*ac108-like*	57,129	<	57,464	336	112	58	81.82	49	68.13	57	79.59	62	89.11	108	38.81
66	*odv-ec43*	57,477	<	58,547	1071	357	59	96.63	50	88.76	58	94.66	63	97.75	109	44.05
67	*pif-7*	58,531	<	58,71	180	60	60	94.87	51	74.58	59	83.78	64	83.05	110	44.12
68	*vp80*	58,707	<	60,368	1662	554	61	62.61	52	55.39	60	68.95	65	63.26	104	23.43
69	*p45*	60,404	>	61,522	1119	373	62	90.56	53	82.31	61	86.48	66	92.50	103	52.20
70	*p12*	61,512	>	61,823	312	104	63	74.53	54	76	62	85.53	67	79	102	29.49
71	*p40*	61,852	>	63,009	1158	386	64	86.67	55	77.60	63	82.08	68	87.57	101	41.71
72	*p6.9*	63,08	>	63,319	240	80	65	–	56	–	64	–	69	–	100	–
73	*lef-5*	63,322	<	64,155	834	278	66	88.49	57	78.10	65	85.87	70	91.30	99	49.10
74	*38 k*	64,054	>	64,953	900	300	67	84.67	58	73.49	66	83.89	71	92.33	98	47.56
75	*ac150-like*	65,048	>	65,41	363	121	68	54.10	59	39.36	67	36.29	72a	60.98	150	30.43
76	*bro-a*	65,341	<	66,423	1083	361	–	–	60	33.23	68	56.73	72	70.03	2	41.44
77	*pif-4*	66,559	<	67,068	510	170	69	90.34	61	81.07	70	92.41	74	91.12	96	51.46
78	*dna helicase*	67,037	>	70,714	3678	1226	70	82.90	62	76.55	71	74.38	75	83.58	95	41.89
79	*odv-e25*	70,789	<	71,439	651	217	71	91.67	63	85.19	72	90.74	76	94.44	94	46.01
80	*p18*	71,436	<	71,912	477	159	72	82.35	64	82.91	73	75.16	77	84.42	93	48.70
81	*p33*	71,922	>	72,68	759	253	73	94.84	65	86.51	74	90.87	78	96.43	92	53.28
–	*Short repeat 1*	–	–	–	39	–	–	–	–	–	–	–	–	–	–	–
82	*lef-4*	73,003	<	74,412	1,41	470	74	73.38	66	64.41	75	75.11	79	75.85	90	44.98
83	*vp39*	74,411	>	75,391	981	327	75	96.30	67	71.08	76	84.74	80	97.14	89	43.48
84	*cg30*	75,704	>	76,789	1086	362	76	52.17	68	27.36	77	29.10	81	55.06	88	30.88
–	*Direct repeat 6*	–	–	–	66	–	–	–	–	–	–	–	–	–	–	–
85	*vp91*	76,894	<	79,29	2397	799	77	69.77	70	62.15	78	67.96	82	72.60	83	39.50
86	*ac82-like*	79,337	>	79,96	624	208	78	71.01	71	67.30	79	70	83	71.23	82	36.15
87	*ac81-like*	79,866	>	80,537	672	224	79	91.53	72	75.94	80	84.90	84	95.21	81	61.83
88	*gp41*	80,506	>	81,495	990	330	80	96.80	73	93.92	81	94.26	85	96.62	80	54.46
89	*ac78-like*	81,508	>	81,873	366	122	81	58.46	74	67.50	82	63.64	86	66.94	78	74.07
90	*vlf-1*	81,875	>	82,99	1116	372	82	97.38	75	94.48	83	96.51	87	98.26	77	68.10
91	*SperNPV-ORF-91*	83,484	<	83,831	348	116	–	–	–	–	–	–	–	–	–	–
92	*SperNPV-ORF-92*	84,237	<	84,401	165	55	–	–	–	–	–	–	–	–	–	–
93	*SperNPV-ORF-93*	84,476	>	84,775	300	100	–	–	–	–	–	–	–	–	–	–
94	*p26 a*	84,859	<	85,608	750	250	87	82.40	76	66.27	85	77.63	90	85.02	136	20.71
95	*iap-2*	85,636	<	86,574	939	313	88	63.58	77	59.39	86	60.13	91	67.60	71	29.44
96	*mtase*	86,417	<	87,211	795	265	89	75.37	78	61.17	87	65.15	92	74.18	69	49.34
97	*pif-6*	87,258	<	87,656	399	133	90	85.84	79	76.47	88	84.07	93	82.40	68	47.66
98	*lef-3*	87,709	>	88,848	1,14	380	91	68.06	80	57.92	89	69	94	75.68	67	29.69
99	*desmoplakin*	88,918	<	91,026	2109	703	92	67.04	81	47.38	90	59.93	95	66.48	66	38.46
100	*dna polymerase*	91,07	>	94,234	3165	1055	93	80.61	82	74.12	91	76.54	96	81.19	65	45.53
101	*ac75-like*	94,251	<	94,64	390	130	94	83.72	83	85.27	92	64.04	97	91.47	75	28.89
102	*ac76-like*	94,651	<	94,908	258	86	95	94.12	84	90.59	93	97.65	98	94.12	76	40
103	*SperNPV-ORF-103*	95,02	>	95,343	324	108	96	44.64	84b	47.92	94	50.49	99	52.29	–	–
104	*lef-9*	95,379	<	96,893	1515	505	97	90.80	85	96.97	96	92.06	100	91.24	62	64.61
105	*fp-25 k*	96,956	>	97,546	591	197	98	95.36	86	91.24	97	94.85	101	95.82	61	62.96
106	*p94*	97,697	<	99,814	2118	706	99	45.15	87	36.48	98	33.47	102	41.72	134	21.86
107	*ac60-like*	99,981	>	100,247	267	89	100	86.76	88	73.13	99	86.36	103	88.24	60	46.77
108	*ac58/59-like*	100,252	>	100,827	576	192	101	91.30	89	90	100	91.43	104	91.30	59	49.09
109	*ac57-like*	100,82	<	101,374	555	185	102	69.57	90	62.13	101	69.23	105	72.13	57	37.50
110	*SperNPV-ORF-110*	101,308	>	101,463	156	52	–	–	–	–	–	–	–	–	–	–
111	*ac56-like*	101,525	<	101,806	282	94	103	60.22	91	47.22	102	50	106	61.22	56	34.62
112	*ac55-like*	101,724	>	101,981	258	86	104	57.89	92	40.54	103	42.47	107	65.85	55	33.33
113	*vp1054*	102,087	<	103,115	1029	343	105	81.10	93	72.14	104	71.30	109	85.09	54	42
114	*lef-10*	102,976	<	103,203	228	76	106	81.67	94	68	105	77.33	110	88	53a	46.67
115	*SperNPV-ORF-115*	103,187	>	103,393	207	69	107a	73.53	95	55.88	106	63.64	111	73.53	–	–
116	*SperNPV-ORF-116*	103,412	>	104,422	1011	337	107	71.91	96	49.85	107	55.42	112	68.36	–	–
117	*ac53-like*	104,429	<	104,845	417	139	108	89.05	97	70.59	108	81.75	113	91.97	53	47.79
118	*ac52-like*	104,905	>	105,456	552	184	109	76.32	98	50.56	109	53.76	114	66.12	52	22.73
–	*Repeat region 3*	–	–	–	426	–	–	–	–	–	–	–	–	–	–	–
119	*SperNPV-ORF-119*	106,029	<	106,427	399	133	–	–	–	–	–	–	–	–	–	–
120	*iap-3*	106,221	>	107,066	846	282	110	57.29	99	42.62	110	43.42	115	61.32	27	26.88
121	*bjdp*	107,111	<	108,22	1,11	370	111	64.37	100	43.36	111	61.20	116	64.78	51	23.43
122	*lef-8*	108,328	>	110,985	2658	886	112	87.09	101	79.39	112	86.37	117	89.23	50	61.82
–	*Short repeat 2*	–	–	–	39	–	–	–	–	–	–	–	–	–	–	–
123	*ac43-like*	111,086	<	111,256	171	57	113	80.49	102	50.98	113	62.50	118	86.27	43	42.55
124	*odv-e66*	111,266	<	113,184	1,92	640	114	63.91	103	41.97	114	55.22	119	67.33	46	29.03
125	*p47*	113,223	>	114,431	1209	403	115	84.58	104	75.25	115	83	120	84.83	40	52.61
126	*ac112/113-like*	114,572	>	115,252	681	227	116	47.56	105[a]	22	116	34.57	121	40.20	113	61.52
127	*ac114-like*	115,352	>	115,942	591	197	117[a]	50	105[a]	29	117	35.87	122[a]	55	114	30.43
128	*SperNPV-ORF-128*	116,091	>	116,621	531	177	117[a]	27	105[a]	26	118[a]	35	122[a]	31	–	–
129	*nudix*	116,674	>	117,402	729	243	118	93.84	106	80.99	118[a]	88	123	92.74	38	61
130	*LEF-11*	117,303	>	117,725	423	141	119	72.12	107	67.62	119	72.64	124	74.77	37	34.74

(*continued on next page*)

Table 2 (*continued*)

ORF	Name	Position			Size (nt)	Size (aa)	SeMNPV-US1		SeMNPV-QD		SfMNPV-19		SpltNPV-II		AcMNPV	
							ORF	ID (%)	ORF	ID (%)	ORF	ID (%)	ORF	ID (%)	ORF	ID (%)
131	*pp31/39 k*	117,724	>	118,638	915	305	120	72.60	108	49.51	120	68.16	125	73.25	36	33.11
132	*SperNPV-ORF-132*	118,706	>	118,987	282	94	121	56.76	109	32.22	121	55.56	126	60.82	–	–
133	*SperNPV-ORF-133*	119,023	<	119,25	228	76	122	72.86	110	52.11	122	59.70	127	65.67	–	–
134	*Ubiquitin*	119,251	<	119,508	258	86	123	95.71	111	86.25	123	96.05	128	96.05	35	77.63
135	*ac34-like*	119,548	>	120,099	552	184	124	77.42	112	67.40	124	73.45	129a	81.41	34	32.98
–	*Repeat region 4*	–	–	–	969	–	–	–	–	–	–	–	–	–	–	–
136	*ac26-like*	121,309	<	121,695	387	129	125	77.27	113	53.21	125	61.19	129	76.92	26	33.33
137	*dbp*	121,787	>	122,749	963	321	126	69.72	114	52.84	126	62.85	130	73.83	25	30.32
138	*lef-6*	122,766	>	123,266	501	167	127	56.08	115	81.08	127	79.45	131	88.31	28	42.65
139	*ac29-like*	123,308	<	123,568	261	87	128	84.42	116	65.52	128	62.79	132	93.02	29	31.43
140	*p26 b*	123,711	>	124,517	807	269	129	73.33	117	59.32	129	67.05	133	76.40	136	32.10
141	*p10*	124,578	>	124,856	279	93	130	86.36	118	75.36	130	78.33	134	86.96	137	32.35
142	*p74 (pif-0)*	124,947	<	126,899	1953	651	131	84.10	119	67.84	132	77.08	135	87.85	138	57.12
143	*SperNPV-ORF-143*	127,002	>	127,268	267	89	–	–	–	–	133	32.86	136	86.36	–	–
144	*ie-1*	127,383	<	129,371	1989	663	132	63.51	120	48.25	134	56.92	139	71.40	141	31.60
145	*ac146-like*	129,401	>	130,045	645	215	133	70.56	121	45.97	135	61.43	140	69.48	146	32.39
146	*ac145-like*	130,075	<	130,353	279	93	134	92.39	122	82.02	136	86.52	141	92.94	145	43.53
147	*odv-ec27*	130,374	<	131,222	849	283	135	92.55	123	86.62	137	90.11	142	93.97	144	54.33
148	*odv-e18*	131,281	<	131,529	249	83	136	91.46	124	80.23	138	78.31	143	96.34	143	83.33
149	*p49*	131,54	<	132,922	1383	461	137	92.19	125	89.57	139	93.48	144	95.87	142	50.85
150	*ie-0*	132,934	<	133,653	720	240	138	84.93	126	66.36	140	72.73	145	84.40	147–0	30.41
151	*rr1*	133,774	<	136,314	2541	847	139	65.58	127	58.21	141	32.78	146	66.79	–	–

[a] Region with identity with SperNPV-CNPSo165-ORF128.

model applied on selected regions of *lef-8*, *lef-9*, and *polh* showed that the SperNPV-CNPSo-165 fulfills the criterion to establish a novel baculovirus species. A virus isolate may represent a new species if the number of substitution per site is higher than 0.072 [28]. In a previous work, the Spodoptera eridania nuclepolyhedrovirus isolate 251 (SperNPV-251) was described as isolated from *S. eridania* larvae [19]. The virus sample was deposited in an insect virus collection at the USDA-ARS in October 1974. Importantly, we found that SperNPV-251 is related to another currently unclassified isolate, SpltNPV-II that may together represent a novel species. On the other hand, SperNPV-CNPSo-165 shows a aK2P-based pairwise distance of 0.168 (Fig. 3A). Even being isolated from the same host, SperNPV-CNPSo-165 was found to present higher global pairwise nucleotide identity with SeMNPV, isolated from *S. exigua* than SperNPV-251 (Supplementary Table 1).

3.6. Genomic structure and gene content analysis

For genomic comparison, we carried out a MAUVE analysis among the genome of all closely related *Spodoptera*-infecting alphabaculoviruses (*i.e.* SpexNPV-244.1, SperNPV-251, SeMNPV-US1, SeMNPV-QD, SfMNPV-19, and SpltNPV-II) and we found strict collinearity with no inversions among the genomes (data not shown). A gene content analysis was also performed based on BlastX results. ORF content is shown in Table 2. A gene content comparison was also carried out among the SperNPV-CNPSo-165 and its closest relatives, including SeMNPV-US1, SeMNPV-QD, SfMNPV-19, and SplitNPV-II and represented by a Venn diagram (Fig. 3B). A total of 243 different ORFs were found. 122 ORFs were shared among the species, which includes the 38 genes found in all baculovirus genomes. Sixteen ORFs were found only in the SperNPV-CNPSo-165 genome. Sixteen ORFs (SperNPV-ORF-6, ORF-9, ORF-10, ORF-22, ORF-24, ORF-29, ORF-41, ORF-44, ORF-45, ORF-59, ORF-60, ORF-91, ORF-92, ORF-93, ORF-110, and ORF-119) were found to be unique to SperNPV- CNPSo-165. Database queries with most of these unique ORFs yielded matches with several non-viral species (Supplementary Table 3) and only three (SperNPV-ORF-22, SperNPV-ORF-44 and, SperNPV-ORF-91) exhibited no significant similarity with any sequence in a database. SperNPV-ORF-6, ORF-9, ORF-22, ORF-44, ORF-92 and ORF119 contained transmembrane domains with no predicted signal peptide (Supplementary Table 3). Moreover, some unique ORFs were located within the repeat regions, including dr2 and dr3 (ORF-6), dr4 (ORF-9), dr5 (ORF-59) and, repeat region 2 (ORF-44 and ORF-45) (Table 2). ORFs occurring inside repeat regions are generally unstable and can differ significantly from isolate to isolate of the same virus; however, when the individual reads were analyzed, most present the coding sequences stable and we have chosen to annotate them.

3.7. The evolution of odv-e66 in baculovirus

Most noctuid-infecting viruses commonly harbor two copies of *odv-e66*, which encodes an occlusion-derived virus envelope protein, ODV-E66. This transmembrane protein is homologous to chondroitinase AC [46] and possesses the ability to degrade non-sulfated chondroitins and chondroitin sulfate C, but not chondroitin sulfate A [42]. The odv-e66 presents an important role in the penetration of the PM during oral infection by degrading chondroitin and may be related to specificity of baculoviruses [14,42,43].

Unlike other *Spodoptera* spp.-infecting NPVs, SperNPV-CNPSo-165 only contains a single copy of *odv-e66*. To evaluate the presence, distribution, and evolution of *odv-e66* genes, a BLASTX query was carried out with the SperNPV-CNPSo-165 sequence. 91 baculovirus sequences were identified, including 71 from alpha- and 20 from betabaculoviruses. An additonal 35 sequence matches were identified with e-value less than 10^{-4}. Twenty-four out of 35 were found in members of the polydnavirus genus *Bracovirus*, while an additional five derived from members of the related large DNA virus families *Nudiviridae* and *Hytrosaviridae*. Other sequences were from several species of bacteria and we selected some with active chondroitin lyases. This may indicate a putative HGT from bacteria to large insect dsDNA viruses, as previously postulated [42]. Many pathogenic bacteria (*e.g.* streptococci) produce extracellular chondroitinase, which are thought to play a role to facilitate the spread of the organism in host tissues [44]. Moreover, several phages are also known to synthesize a bound form of hyaluronidase. It has been suggested that the function of this viral factor allows an easier penetration into the hyaluronan bacterial capsule by the phage and indeed to the cell surface of the host [45]. Interestingly, active site of chondroitin lyases is composed by three conserved residues and are largely described for bacteria [45]. After MAFFT alignment of baculovirus predicted protein sequences and some Bacteria sequences, we found that the active site is maintained in several baculovirus with most of the residues being conserved (Fig. 4, asterisk and

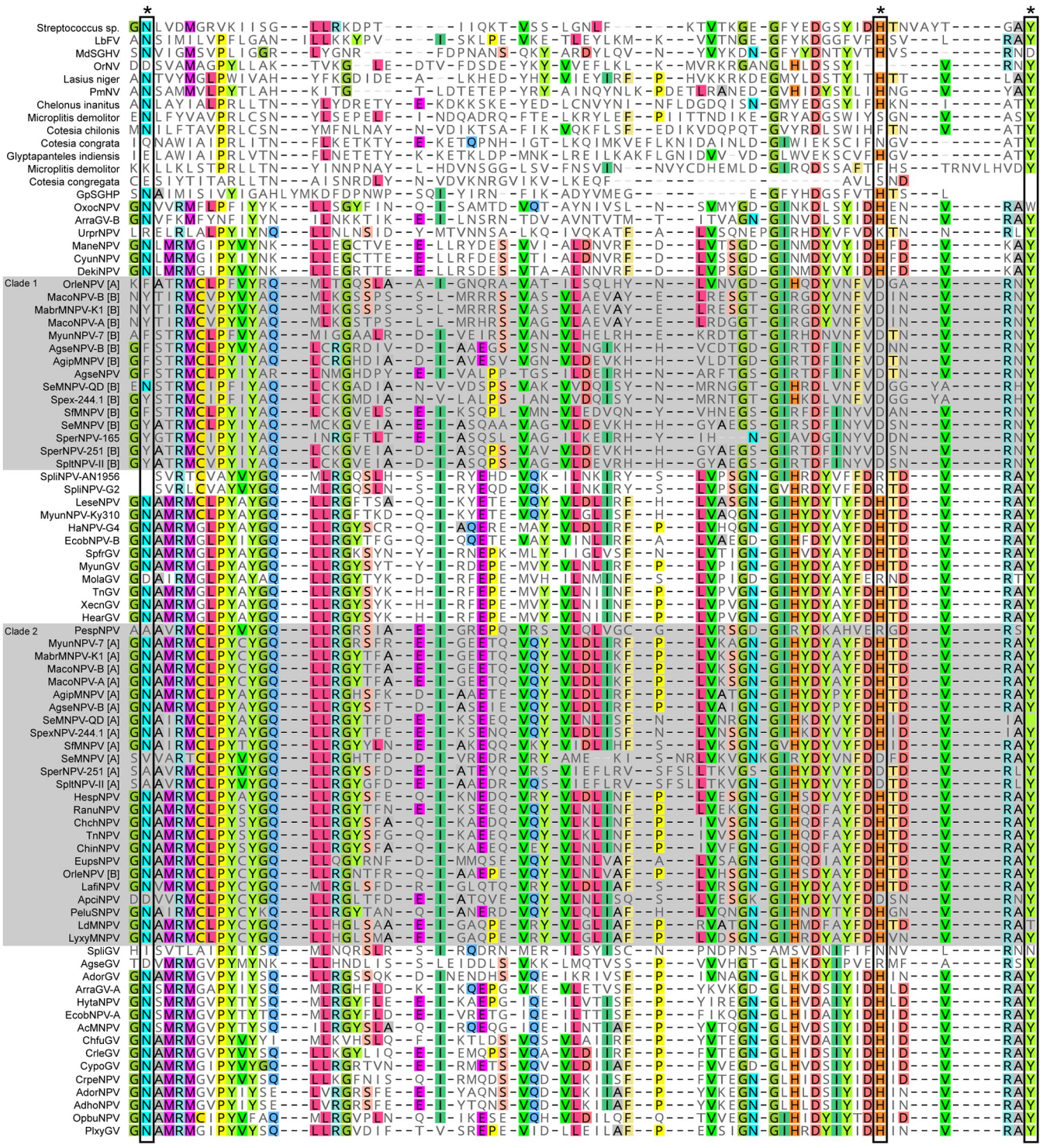

Fig. 4. MAFFT alignment of the predicted amino acid sequence of odv-e66 focusing on its active site of several baculovirus and other organisms, including bacterium, polydnavirus, nudivirus, and hytrosiviruses. The active site is composed by three conserved residues N, H, and Y based on bacterium characterization, which are highlighted by asterisk and black box. Clade 1 and Clade 2 belongs to a baculovirus clade that likely underwent an independent duplication during genome evolution. Clade 1 presented substitutions of N to Y/F and H to D.

black box highlights the the conserved residues, N, H, and Y). Interestingly, when the gene is duplicated in the baculovirus genome, the second copy lose its active site in two residues, N to Y/F and H to D (Clade 1). This finding reinforces that the predicted protein sequence of *odv-e66* is likely an active enzyme in most of the analyzed baculoviruses [42].

We carried out a phylogenetic inference using sequences from baculovirus, other large dsDNA insect viruses, and bacteria (Fig. 5). Large dsDNA viruses formed a monophyletic clade, which could likely depict one putative entrance from bacteria to this viral group (Fig. 5). It is not clear which events took place in the dipteran and hymenopteran baculovirus genomes as both groups lack *odv-e66* homologs in their genomes; therefore, we focused solely on the evolutionary history of lepidopteran baculoviruses. Alphabaculovirus and betabaculovirus

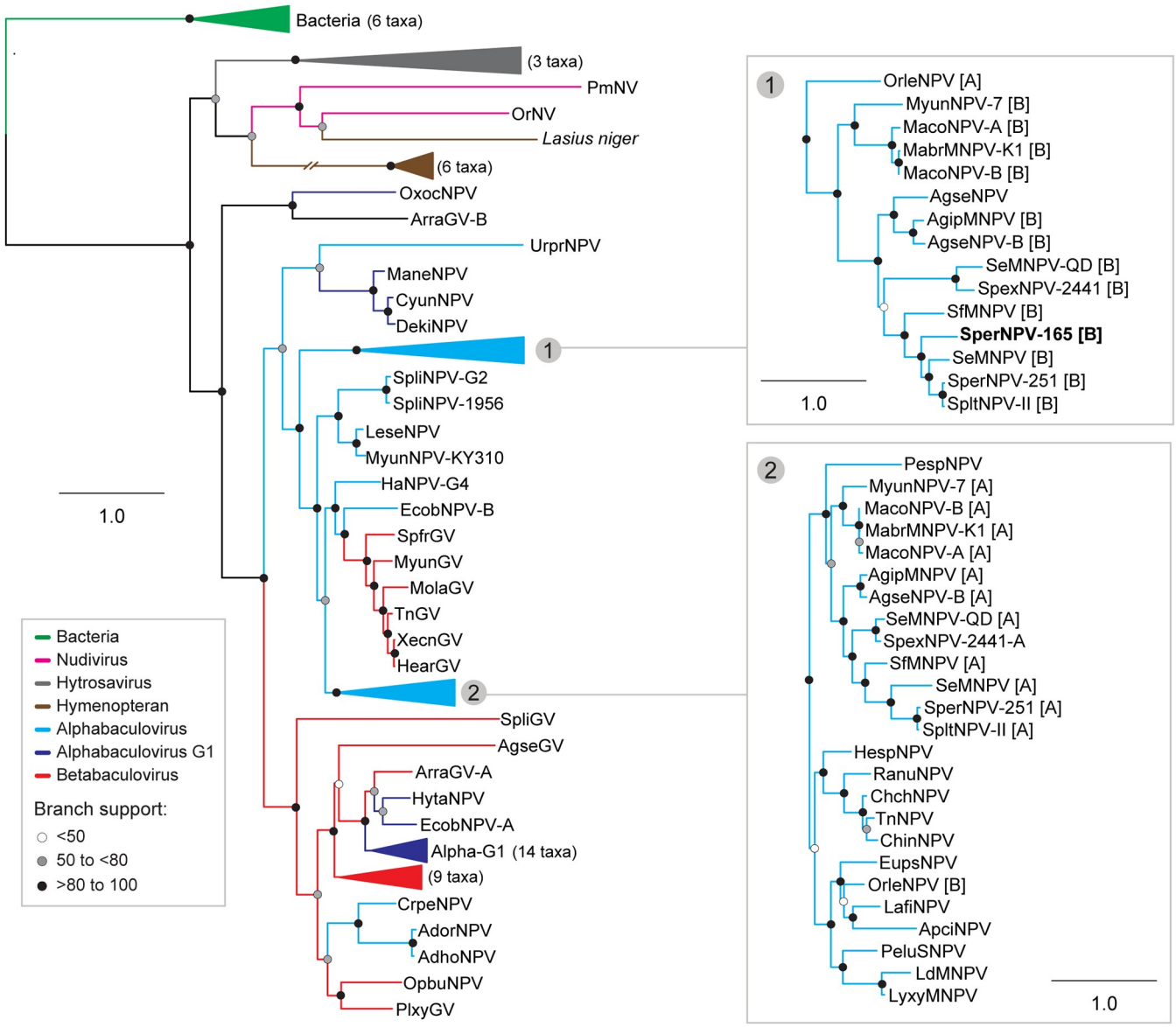

Fig. 5. Phylogenetic inference of *odv-e66*. Phylogenetic inference with the sequences of nudivirus (pink), alphabaculovirus (light blue), and betabaculovirus (red), large dsDNA viruses form a monophyletic clade, as well as branches of baculovirus, representing an entry of the bacteria to this group. Some branches were collapsed for clarity: Bacteria (green), hytrosavirus (grey), hymenopteran (brown), alphabaculovirus group 1(dark blue), betabaculovirus (red). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

branches formed a well-supported monophyletic clade (Fig. 5). However, the presence of mixed taxa along the tree branches reinforces several events of gains and losses in the evolution of *odv-e66* inside baculovirus.

The most parsimonious history of *odv-e66* in lepidopteran baculoviruses portrait 30 steps based on gene phylogeny (Fig. 5), gene *loci* within the genomes (Fig. 6A), and gene distribution among baculoviruses (Table 3 and Fig. 6B): 13 deletions, 16 acquisitions, and one duplication were found. We numbered the steps in Fig. 6B from 1 to 30 and summarized it in Table 3. We found one transfer (which could be plesiomorphic or autapomorphic) in the ancestor of alpha and betabaculoviruses (Fig. 6B, step 1, black square). Only homologs of ArraGV and OxocNPV were placed outside of the baculovirus clade, which could indicate independent acquisitions during evolution or a high genetic divergence. The gene locus reinforces the first hypothesis of being acquired independently (Fig. 6B). The *odv-e66* phylogeny reconstructed partially the evolution of baculovirus, which portrays an intensive gene flow and events of HGT among members of *Baculoviridae*. The duplication of *odv-e66* took place in a single clade of noctuid-infecting baculoviruses. Interestingly, the Independent losses in this clades took place in the lineages of SperNPV-CNPSo-165, AgseNPV-A, and PespNPV (Fig. 6B, steps 28 and 29). Both lineages lost the ancestral gene and retained the duplication (as confirmed by the genomic context, Fig. 6A). PespNPV lost the duplicated copy of *odv-e66* and retained the ancestral gene (Fig. 6B, step 28). A very interesting event took place in the m.r.c.a. of the noctuid-isolated granulovirus with big genomes [47], including MyunGV, SpfrGV-008, MolaGV, TnGV, HaGV, and XecnGV (Fig. 6B, steps 3 and 4). The gene *loci* reinforce the m.r.c.a. acquisition (data not shown). The m.r.c.a lost the ancestral *odv-e66* gene (Fig. 6B, black square) and reacquired it from alphabaculovirus (Fig. 6B, green square). This is reinforced by the gene phylogeny (Fig. 5). ArraGV was the unique lineage of betabaculoviruses with two *odv-e66* copies. From the gene phylogeny, the ArraGV copies seemed to be acquired independently, one from an unknown source and the other one from Alphabaculovirus group I. The genomic context of the gene reinforces the hypothesis of independent acquisition (data not shown).

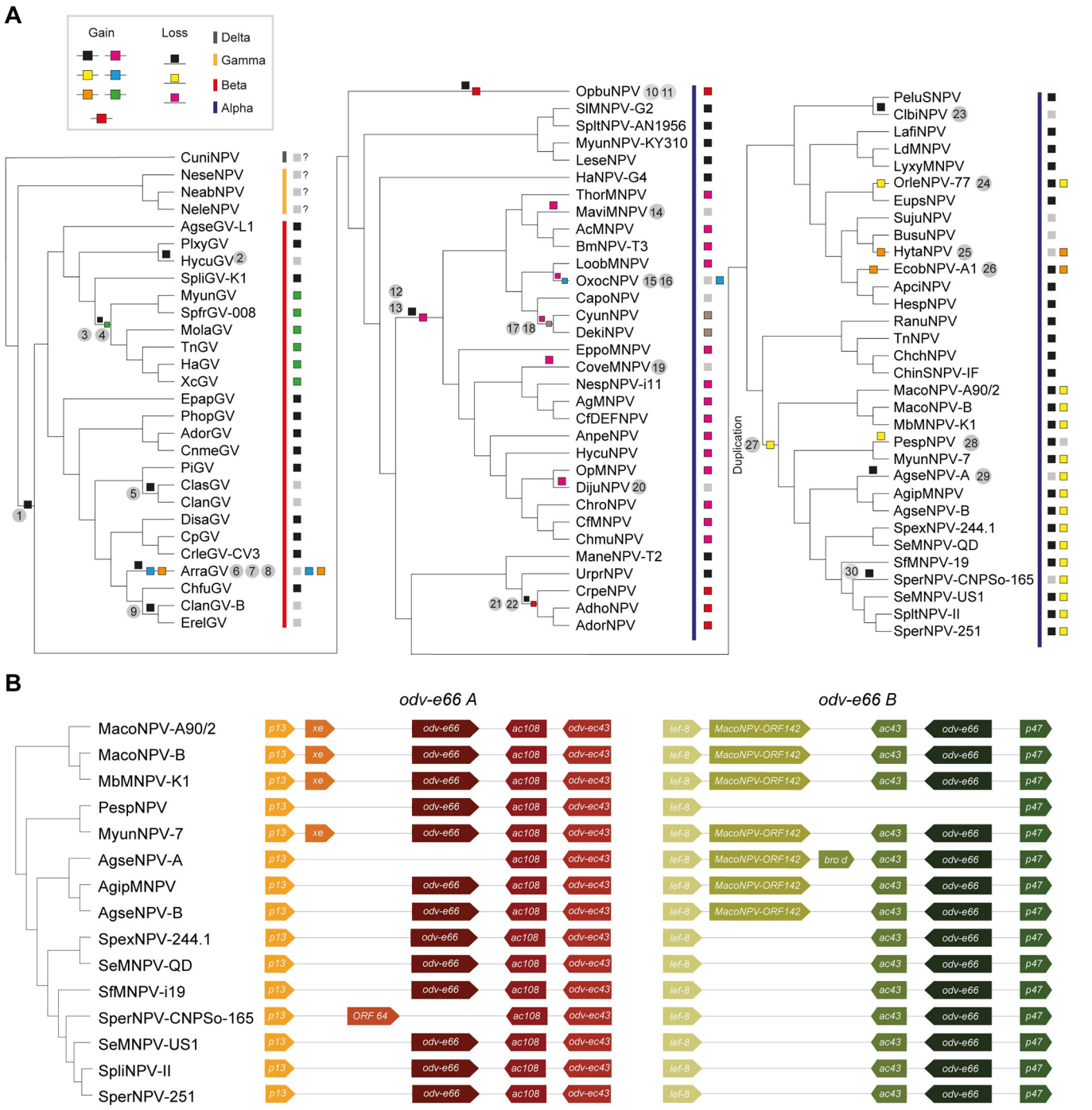

Fig. 6. Genomic context of *odv-e66* and evolutionary analyses of *odv-e66*, a bacterial-related chondroitinase gene homolog. (A) Evolutionary steps with gain, loss, and duplication events for odv-e66 inside the family *Baculoviridae*. Based on the hypothetical phylogeny trees and genomic context, the history of odv-e66 and presence in members of lepidopteran-infecting baculovirus are presented. Table 3 summarizes the evolutionary hypothetical events. Gain is depicted by the square on the line and loss the square above the line. (B) Genomic context of the two copies of *odv-e66* (*odv-e66 A* and *odv-e66 B*) according to evolutionary gene analysis and duplication event in closely related species of SperNPV-CNPSo-165. The arrowheads represent the direction of the genes in the genome. Arrows with similar colors describe gene orthology.

The inability of SperNPV-CNPSo-165 to infect orally *S. frugiperda* could be in part explained by the fact that the SperNPV-CNPSo-165 genome lost one of the copies of *odv-e66*, a gene whose protein product is believed to be implicated in virus host specificity. Moreover, the virus protein is related to a clade in which the two out of three residues in the active site are changed, N to Y and H to D. Therefore, the retained copy of SperNPV-CNPSo-165 is not clear and must be investigated in further studies.

4. Conclusion

In this work, we characterized a novel baculovirus isolated from the Southern armyworm *S. eridania*, Spodoptera eridania nucleopolyhedrovirus CNPSo-165 (SperNPV-CNPSo-165). The virus was found to be more lethal to larvae of *S. albula* than *S. eridania* and not able to kill *S. frugiperda*, all three important agricultural Brazilian pests inside the *Spodoptera*-complex. The isolate seemed to be a member of a new

Table 3
Description for *odv-e66* events that likely took place in lepidopteran-isolated baculovirus genomes presented in Fig. 6.

Step	Lineage/hypothetical m.r.c.a.[a]	Event in *odv-e66* gene evolution	Square colour
1	m.r.c.a of alpha and betabaculovirus	Acquisition from an undisclosed source	Black
2	HycuGV	Independent loss	Black
3	m.r.c.a. of MyunGV, SpfrGV-008, MolaGV, TnGV, HaGV, and	Loss of the ancestral gene	Black
4	XecnGV	Acquisition of an alphabaculovirus-related gene	Green
5	m.r.c.a. of ClasGV and ClanGV	Independent loss	Black
6	ArraGV	Independent loss	Black
7		Independent gain from alphabaculovirus	Blue
8		Independent gain from undisclosed source	Orange
9	m.r.c.a. of ClanGV-B and ErelGV	Independent loss	Black
10	OpbuNPV	Independent loss	Black
11		Independent gain from betabaculovirus	Red
12	m.r.c.a. of group 1 alphabaculovirus	Loss of the ancestral gene	Black
13		Acquisition of a betabaculovirus-related gene	Pink
14	MaviMNPV	Independent loss	Pink
15	OxocNPV	Independent loss	Pink
16		Independent gain from undisclosed source (besides being relalted to one of the copies from ArraGV)	Blue
17	m.r.c.a. of CyunNPV and DekiNPV	Loss of the ancestral gene acquired by group 1 alphabaculovirus	Pink
18		Reacquisition of an alphabaculovirus-related *odv-e66*	Beige
19	CoveMNPV	Independent loss	Pink
20	DijuNPV	Independent loss	Pink
21	m.r.c.a. of CrpeNPV, AdorNPV, and AdhoNPV	Loss of the ancestral gene	Black
22		Acquisition of a betabaculovirus-related gene	Red
23	ClbiNPV	Independent loss	Black
24	OrleNPV-77	Independent gain from the alphabaculovirus duplication gene	Yellow
25	HytaNPV	Independent gain from a betabaculovirus source	Orange
26	EcobNPV-A1	Independent gain from a betabaculovirus source	Orange
27	m.r.c.a. of several noctuid-infecting alphabaculoviruses	Independent duplication	Yellow
28	PespNPV	Independent loss of the duplicated gene version	Yellow
29	AgseNPV-A	Independent loss of the ancestral gene version	Black
30	SperNPV-CNPSo-165	Independent loss of the ancestral gene version	Black

[a] Hypothetical most recent common ancestor.

tentative species inside *Alphabaculovirus*, closely related to the m.r.c.a. of SeMNPV-US1, SpltNPV-II, and SperNPV-251 viruses with a genome of 137.373 bp in size, G + C content of 42.8% and 151 annotated ORFs. SperNPV-CNPSo-165 genome harbored only one copy of *odv-e66*, whereas its closely related viruses present two copies. The evolution of *odv-e66* in baculovirus presented several events, including gene loss, gain, and duplication. Overall, the study of baculovirus allows a better understanding of the virus family evolution, providing important information for the development and improvement of tools for biological control and biotechnology.

Funding statements

This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, grant number 428799/2018-3 and 305468/2019-7), Coordenação de Aperfeiçoamento de Pessoal de nível Superior (CAPES/PROEX, grant number 23038.004173/2019-93) and Fundação de Apoio à Pesquisa do Distrito Federal (FAPDF, grant number 193.001532/2016).

Declaration of Competing Interest

The authors declare that they have no competing interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ygeno.2020.06.047.

References

[1] J.L. Capinera, Southern Armyworm, *Spodoptera eridania* (Cramer) (Insecta: Lepidoptera: Noctuidae), University of Florida Cooperative Extension Service, Institute of Food and Agricultural Sciences, EDIS, 1999 Available at: https://edis.ifas.ufl.edu/pdffiles/IN/IN26300.pdf (Accessed on: March 5, 2019).

[2] C.F.S. Hoo, G. Fraenkel, The selection of food plants in a polyphagous insect, Prodenia eridania (Cramer), J. Insect Physiol. 12 (6) (1966) 693–709, https://doi.org/10.1016/0022-1910(66)90115-6.

[3] A.C. Formentini, D.R. Sosa-Gómez, S.V. de Paula-Moraes, N.M. de Barros, A. Specht, Lepidoptera (Insecta) associated with soybean in Argentina, Brazil, Chile and Uruguay, Ciên. Rural. 45 (12) (2015) 2113–2120, https://doi.org/10.1590/0103-8478cr20141258.

[4] K.B. Santos, A.M. Meneguim, W.J. Santos, P.M. Neves, R.B. Santos, Caracterização dos danos de Spodoptera eridania (Cramer) e Spodoptera cosmioides (Walker) (Lepidoptera: Noctuidae) a estruturas de algodoeiro, Neotrop. Entomol. 39 (4) (2010) 626–631, https://doi.org/10.1590/S1519-566X2010000400025.

[5] L.C. Bortoli, A. Bertin, C.F.S. Efrom, M. Botton, Biologia e tabela de vida de fertilidade de Spodoptera eridania (Cramer) (Lepidoptera: Noctuidae) em morangueiro e videira, Rev. Bras. Frutic. 34 (4) (2012) 1068–1073, https://doi.org/10.1590/S0100-29452012000400013.

[6] K.B.D. Santos, A.M. Meneguim, P.M. Neves, Biologia de Spodoptera eridania (Cramer) (Lepidoptera: Noctuidae) em diferentes hospedeiros, Neotrop. Entomol. 34 (6) (2005) 903–910, https://doi.org/10.1590/S1519-566X2005000600005.

[7] B.H.S. Souza, A.L. Boiça, J.C. Janini, A.G. Silva, N.E.L. Rodrigues, Feeding of Spodoptera eridania (Lepidoptera: Noctuidae) on soybean genotypes/Alimentación de Spodoptera eridania (Lepidoptera: Noctuidae) en genotipos de soya, Rev. Colomb. Entomol. 38 (2) (2012) 215 http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0120-04882012000200009&lng=en (Accessed on: March 18, 2019).

[8] D.F.M. Fragoso, J.R. Carvalho, A.P. Barros, T. Coffler, J.J.P. Marchiori, Lagarta-das-folhas (*Spodoptera eridania*), in: Holtz, Anderson Mathias (Eds.), Pragas Básicas, 9 Ifes, Colatina, 2015, pp. 192–209. Available at: researchgate.net/publication/306097692_LAGARTA-DAS-FOLHAS_Spodoptera_eridania (Accessed on: July 9, 2019).

[9] P.J. Gullan, P.S. Cranston, The Insects: An Outline of Entomology, John Wiley and Sons, Ltd., Chichester West Sussex, 2014, https://doi.org/10.1007/s10841-010-9351-x.

[10] D.R. Sosa-Gomez, F.S. Morgado, R.F.T. Correa, L.A. Silva, D.M.P. Ardisson-Araújo, B.M.P. Rodrigues, E.E. Oliveira, R.W.S. Aguiar, B.M. Ribeiro, Entomopathogenic viruses in the neotropics: current status and recently discovered species, Neotrop. Entomol. (2020), https://doi.org/10.1007/s13744-020-00770-1.

[11] B. Szewczyk, L. Hoyos-Carvajal, M. Paluszek, I. Skrzecz, M.L. De Souza, Baculoviruses—re-emerging biopesticides, Biotechnol. Adv. 24 (2) (2006) 143–160, https://doi.org/10.1016/j.biotechadv.2005.09.001.

[12] D.R. Sosa-Gómez, Microbial control of soybean pest insects and mites, Microbial Control of Insect and Mite Pests, Academic Press, 2017, pp. 199–208 , https://doi.org/10.1016/B978-0-12-803527-6.00013-5.

[13] D.R. Sosa-Gómez, F. Moscardi, B. Santos, L.F.A. Alves, S.B. Alves, Produção e uso de

vírus para o controle de pragas na América Latina. Controle Microbiano de Pragas na America Latina, 3 FEALQ, Piracicaba, Brazil, 2008, pp. 69–110. Available at: https://www.researchgate.net/publication/291300500_Fungos_entomopatogenicos_usados_no_controle_de_pragas_na_America_Latina (Accessed on: January 2, 2019).
[14] G.F. Rohrmann, Baculovirus Molecular Biology, National Center for Biotechnology Information (US), Bethesda (MD), 2013 Available at: http://www.ncbi.nlm.nih.gov/books/NBK49500/ (Accessed on: February 15, 2019).
[15] J.A. Jehle, G.W. Blissard, B.C. Bonning, J.S. Cory, E.A. Herniou, G.F. Rohrmann, D.A. Theilmann, S.M. Thiem, J.M. Vlak, On the classification and nomenclature of baculoviruses: a proposal for revision, Arch. Virol. 151 (7) (2006) 1257–1266, https://doi.org/10.1007/s00705-006-0763-6.
[16] L.B. Trentin, E.R. Santos, A.G.O. Junior, D.R. Sosa-Gómez, B.M. Ribeiro, D.M. Ardisson-Araújo, The complete genome of Rachiplusia nu nucleopolyhedrovirus (RanuNPV) and the identification of a baculoviral CPD-photolyase homolog, Virology 534 (2019) 64–71, https://doi.org/10.1016/j.virol.2019.05.019.
[17] J. Slack, B.M. Arif, The baculoviruses occlusion-derived virus: virion structure and function, Adv. Virus Res. 69 (2006) 99–165, https://doi.org/10.1016/S0065-3527(06)69003-9.
[18] B.A. Federici, Baculovirus pathogenesis, The Baculoviruses, Springer, Boston, MA, 1997, pp. 33–59 , https://doi.org/10.1007/978-1-4899-1834-5_3.
[19] R.L. Harrison, D.L. Rowley, Complete genome sequence of an Alphabaculovirus from the southern armyworm, Spodoptera eridania, Microbiol. Resour. Announc. 8 (3) (2019) e01277-18, https://doi.org/10.1128/MRA.01277-18.
[20] D.R. O'Reilly, L.K. Miller, V.A. Luckow, Baculovirus Expression Vectors: A Laboratory Manual, Oxford University Press on Demand, 1994, https://books.google.com.br/books?hl=pt-BR&lr=&id=lP8VFRX8zHMC&oi=fnd&pg=PA11&ots=Cexne.OzlT&sig=zsizDzuaXuRPGc_SkUy931QT5E0&redir_esc=y#v=onepage&q&f=false (Accessed on: May 8, 2019).
[21] E.R. Santos, L.B. Oliveira, L. Peterson, D.R. Sosa-Gómez, B.M. Ribeiro, D.M.P. Ardisson-Araújo, The complete genome sequence of the first hesperiid-infecting alphabaculovirus isolated from the leguminous pest Urbanus proteus (Lepidoptera: Hesperiidae), Virus Res. 249 (2018) 76–84, https://doi.org/10.1016/j.virusres.2018.03.009.
[22] D.M.P. Ardisson-Araújo, F.L. Melo, M.S. Andrade, R.M.C. Brancalhão, S.N. Báo, B.M. Ribeiro, Complete genome sequence of the first non-Asian isolate of *Bombyx mori* nucleopolyhedrovirus, Virus Genes. 49 (3) (2014) 477–484, https://doi.org/10.1007/s11262-014-1112-6.
[23] M. Kearse, R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data, Bioinformatics 28 (12) (2012) 1647–1649, https://doi.org/10.1093/bioinformatics/bts199.
[24] S.F. Altschul, T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (17) (1997) 3389–3402.
[25] K. Katoh, K. Misawa, K. Kuma, T. Miyata, MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform, Nucleic Acids Res. 30 (14) (2002) 3059–3066, https://doi.org/10.1093/nar/gkf436.
[26] A. Stamatakis, P. Hoover, Jacques Rougemont, A rapid bootstrap algorithm for the RAxML web servers, Syst. Biol. 57 (5) (2008) 758–771, https://doi.org/10.1080/10635150802429642.
[27] M. Anisimova, M. Gil, J.F. Dufayard, C. Dessimoz, O. Gascuel, Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes, Syst. Biol. 60 (5) (2011) 685–699, https://doi.org/10.1093/sysbio/syr041.
[28] J.T. Wennmann, J. Keilwagen, J.A. Jehle, Baculovirus Kimura two-parameter species demarcation criterion is confirmed by the distances of 38 core gene nucleotide sequences, J. Gen. Virol. 99 (9) (2018) 1307–1320, https://doi.org/10.1099/jgv.0.001100.
[29] J. Schultz, F. Milpetz, P. Bork, C.P. Ponting, SMART, a simple modular architecture research tool: identification of signaling domains, Proc. Natl. Acad. Sci. 95 (11) (1998) 5857–5864, https://doi.org/10.1073/pnas.95.11.5857.
[30] J. Söding, A. Biegert, A.N. Lupas, The HHpred interactive server for protein homology detection and structure prediction, Nucleic Acids Res. 33 (2) (2005) W244–W248, https://doi.org/10.1093/nar/gki408.
[31] S. Guindon, J.F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, O. Gascuel, New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0, Syst. Biol. 59 (3) (2010) 307–321, https://doi.org/10.1093/sysbio/syq010.
[32] S. Kumar, G. Stecher, K. Tamura, MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets, Mol. Biol. Evol. 33 (7) (2016) 1870–1874, https://doi.org/10.1093/molbev/msw054.
[33] A.A. Lima, C.W.S. Aragao, M.E.B. Castro, J.V.C. Oliveira, D.R. Sosa-Gómez, B.M. Ribeiro, A recombinant Anticarsia gemmatalis MNPV harboring chiA and v-cath genes from Choristoneura fumiferana defective NPV induce host liquefaction and increased insecticidal activity, PLoS One 8 (9) (2013) e74592, https://doi.org/10.1371/journal.pone.0074592.
[34] G.L. Greene, N.C. Leppla, W.A. Dickerson, Velvetbean caterpillar: a rearing procedure and artificial medium, J. Econ. Entomol. 69 (4) (1976) 487–488, https://doi.org/10.1093/jee/69.4.487.
[35] T. Martins, R. Montiel, J. Medeiros, L. Oliveira, N. Simões, Occurrence and characterization of a nucleopolyhedrovirus from Spodoptera littoralis (Lepidoptera: Noctuidae) isolated in the Azores, J. Invertebr. Pathol. 89 (3) (2005) 185–192, https://doi.org/10.1016/j.jip.2005.06.012.
[36] M. Ordóñez-García, C. Rios-Velasco, J.J. Ornelas-Paz, J.C. Bustillos-Rodríguez, C.H. Acosta-Muñiz, D.I. Berlanga-Reyes, M.A. Salas-Marina, O.J. Cambero-Campos, G. Gallegos-Morales, Molecular and morphological characterization of multiple nucleopolyhedrovirus from Mexico and their insecticidal activity against Spodoptera frugiperda (Lepidoptera: Noctuidae), J. Appl. Entomol. 144 (1–2) (2020) 123–132, https://doi.org/10.1111/jen.12715.
[37] S.R. Escasa, R.L. Harrison, J.D. Mowery, G.R. Bauchan, J.S. Cory, The complete genome sequence of an alphabaculovirus from Spodoptera exempta, an agricultural pest of major economic significance in Africa, PLoS One 14 (2) (2019), https://doi.org/10.1371/journal.pone.0209937.
[38] T.Z. Salem, X.H. Cheng, X.W. Cheng, AcMNPV enhances infection by ThorNPV in Sf21 cells and SeMNPV in Hi5 cells, Arch. Virol. 157 (10) (2012) 1875–1885, https://doi.org/10.1007/s00705-012-1347-2.
[39] L. Wang, T.Z. Salem, D.E. Lynn, X.W. Cheng, Slow cell infection, inefficient primary infection and inability to replicate in the fat body determine the host range of Thysanoplusia orichalcea nucleopolyhedrovirus, J. Gen. Virol. 89 (6) (2008) 1402–1410, https://doi.org/10.1099/vir.0.2008/000695-0.
[40] B. Le Ru, J. Barbut, C. Capdevielle-Dulac, M. Goftishu, G.J. Kergoat, Re-establishment of Spodoptera teferii Laporte in Rougeot (Lepidoptera: Noctuidae, Noctuinae), with an updated molecular phylogeny for the genus Spodoptera Guenée, Ann. Soc. Entomol. France (N.S.) 54 (6) (2018) 497–510, https://doi.org/10.1080/00379271.2018.1528886.
[41] M.J. Garavaglia, S.A.B. Miele, J.A. Iserte, M.N. Belaich, P.D. Ghiringhelli, The ac53, ac78, ac101, and ac103 genes are newly discovered core genes in the family Baculoviridae, J. Virol. 86 (22) (2012) 12069–12079, https://doi.org/10.1128/JVI.01873-12.
[42] Y. Kawaguchi, N. Sugiura, K. Kimata, M. Kimura, Y. Kakuta, The crystal structure of novel chondroitin lyase ODV-E66, a baculovirus envelope protein, FEBS Lett. 587 (24) (2013) 3943–3948, https://doi.org/10.1016/j.febslet.2013.10.021.
[43] D. Hou, W. Kuang, S. Luo, et al., Baculovirus ODV-E66 degrades larval peritrophic membrane to facilitate baculovirus oral infection, Virology 537 (2019) 157–164, https://doi.org/10.1016/j.virol.2019.08.027.
[44] J.R. Baker, S. Dong, D.G. Pritchard, The hyaluronan lyase of *Streptococcus pyogenes* bacteriophage H4489A, Biochem. J. 365 (1) (2002) 317–322, https://doi.org/10.1042/BJ20020149.
[45] W.L. Hynes, L. Hancock, J.J. Ferretti, Analysis of a second bacteriophage hyaluronidase gene from *Streptococcus pyogenes*: evidence for a third hyaluronidase involved in extracellular enzymatic activity, Infect. Immun. 63 (8) (1995) 3015–3020, https://doi.org/10.1128/IAI.63.8.3015-3020.1995.
[46] J. Féthière, B. Eggimann, M. Cygler, Crystal structure of chondroitin AC lyase, a representative of a family of glycosaminoglycan degrading enzymes, J. Mol. Biol. 288 (4) (1999) 635–647, https://doi.org/10.1006/jmbi.1999.2698.
[47] D.M.P. Ardisson-Araújo, A. M. R. Silva, F.L. Melo, E.R. Santos, D. R. Sosa-Gómez, B. M. Ribeiro, A novel betabaculovirus isolated from the monocot pest mocis latipes (Lepidoptera: Noctuidae) and the evolution of multiple-copy genes, Viruses 10(3).