

**UNIVERSIDADE FEDERAL DE SANTA MARIA  
CENTRO DE CIÊNCIAS NATURAIS E EXATAS  
PROGRAMA DE PÓS GRADUAÇÃO EM CIÊNCIAS BIOLÓGICAS:  
BIOQUÍMICA TOXICOLÓGICA**

**Ethiane Rozo dos Santos**

**GENÔMICA DE UM BACULOVÍRUS ISOLADO DA LAGARTA  
PRAGA DE LEGUMINOSAS *Urbanus proteus* (LINNAEUS, 1758)  
(LEPIDOPTERA: HESPERIIDAE)**

Santa Maria, RS  
2018

**Ethiane Rozo dos Santos**

**GENÔMICA DE UM BACULOVÍRUS ISOLADO DA LAGARTA PRAGA DE  
LEGUMINOSAS *Urbanus proteus* (LINNAEUS, 1758) (LEPIDOPTERA:  
HESPERIIDAE)**

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Biológicas: Bioquímica Toxicológica, da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do título de **Mestre em Ciências Biológicas: Bioquímica Toxicológica.**

Orientador: Daniel Mendes Pereira Ardisson-Araújo

Santa Maria, RS  
2018

Santos, Ethiane Rozo dos  
GENÔMICA DE UM BACULOVÍRUS ISOLADO DA LAGARTA PRAGA DE  
LEGUMINOSAS *Urbanus proteus* (LINNAEUS, 1758)  
(LEPIDOPTERA: HESPERIIDAE) / Ethiane Rozo dos Santos.-  
2018.  
81 p.; 30 cm

Orientador: Daniel Mendes Pereira Ardisson-Araújo  
Dissertação (mestrado) - Universidade Federal de Santa  
Maria, Centro de Ciências Naturais e Exatas, Programa de  
Pós-Graduação em Ciências Biológicas: Bioquímica  
Toxicológica, RS, 2018

1. Genômica 2. Baculovírus 3. Alphabaculovirus 4.  
*Urbanus proteus* 5. Evolução I. Mendes Pereira Ardisson  
Araújo, Daniel II. Título.

**Ethiane Rozo dos Santos**

**GENÔMICA DE UM BACULOVÍRUS ISOLADO DA LAGARTA PRAGA DE  
LEGUMINOSAS *Urbanus proteus* (LINNAEUS, 1758) (LEPIDOPTERA:  
HESPERIIDAE)**

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Biológicas: Bioquímica Toxicológica, da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do título de **Mestre em Ciências Biológicas: Bioquímica Toxicológica.**

**Aprovado em 17 de julho de 2018:**

---

**Daniel Mendes Pereira Ardisson-Araújo, Dr. (UFSM)**  
(Presidente/Orientador)

---

**Bergmann Morais Ribeiro, PhD (UnB)**

---

**Eduardo Furtado Flores, Dr. (UFSM)**

Santa Maria, RS  
2018

## DEDICATÓRIA

*A minha família, minha mãe Neli e ao meu pai Clenio. Dedico também à minha avó Amélia Valandro Rozo e ao meu avô Marcelino Rozo, por todo o amor, carinho e educação recebidos, vocês são o exemplo de luta e persistência que me fazem seguir em frente.*

## AGRADECIMENTOS

Primeiramente agradeço a minha mãe Neli, essa mulher que é meu exemplo de energia e perseverança, movendo céus e terra para me criar praticamente sozinha. Sempre me apoiando e me dando forças para seguir em frente. Não consigo descrever o quanto eu te amo e o orgulho que sinto em ser tua filha.

Agradeço ao meu pai Clenio, que mesmo com todas as dificuldades, foi um excelente pai, amigo, as vezes filho. Amo você meu teimoso.

Aos meus avós, Amélia e Marcelino que me ensinaram ser o melhor que eu posso a cada dia. Tudo o que eu faço é por vocês.

Aos meus familiares, principalmente a minha tia Geny, pela confiança e pelos seus conselhos sábios que me faz pôr os pés no chão e pensar melhor, e a minha tia Clenir que mesmo estando distante sempre me apoiou e foi uma mãezona.

Ao professor Daniel, que me deu um voto de confiança em ser meu orientador, sem ao menos me conhecer pessoalmente. Foi muito mais que isso, foi amigo, pai, aquele que te puxa as orelhas sempre que necessário. O senhor é um exemplo de profissional admirável, sempre disposto a ajudar e que montou praticamente um laboratório sozinho com dinheiro do próprio bolso. Obrigada pelos ensinamentos e pela paciência, sinto-me orgulhosa de ser sua primeira orientada de mestrado.

Aos meus colegas de laboratório Lenen e Lucas, os LaVI raiz, parceiros desde o nosso primeiro laboratório no Chernobyl, guris que me ensinaram a mexer no Geneious, obrigada pelos momentos de descontração, mas também pelos momentos de profissionalismo, vocês que praticamente decoraram a minha apresentação de qualificação e sempre me ajudam sendo críticos e me forçando a estudar mais.

As minhas colegas de laboratório Bruna por me aturar, me ajudar, dar conselhos, mesmo as vezes não fechando nossas ideias e a Luana que tem um coração enorme, a que convive comigo diariamente, sabe quando estou estressada (quase todo dia) e tenta tornar o meu dia mais leve. Vocês, além de ser profissionais, são minhas amigas e me deram forças durante esta jornada.

Ao meu mais novo colega de laboratório Assis, que mesmo com pouco tempo de convivência, é uma pessoa amiga, sempre disposto a ajudar, que ouve minhas queixas e sempre dá bons conselhos.

A professora Ester, pela confiança depositada, para que eu pudesse permanecer em Santa Maria.

A professora Terimar, por ser minha orientadora durante o final da graduação e me possibilitar a oportunidade de conhecer o PPGBTox, ao professor João Batista pelo primeiro projeto de mestrado.

Aos meus professores da graduação, por todos ensinamentos passados.

A Lu e a Thais, minhas colegas da União Estudantil, passamos muitos perrengues juntas no início desta trajetória, mas sempre nos apoiando uma a outra e dando espaço, quando necessário. Amizade que quero levar para sempre.

A Aline Soeiro, amiga de Palmeirinha, mas que Santa Maria me deu, obrigada pelos conselhos, caminhadas na UFSM, dominadas do mundo e por me fazer sair da minha bolha protetora.

As minhas amigas de Palmeira das Missões, Aline Couto, Andressa, Vanessa Padilha, Vanessa Mattana, Iara, Carini, Sandra, Marluci e Éch, mesmo não estando próximas, sempre me apoiaram e me impulsionaram a seguir em frente.

As minhas afilhadas Izabelly e Laura, dedico a vocês todo amor e quero dar-lhes o melhor exemplo sempre.

Aos meus animais de estimação Pink, Ashley, Dóroti, Sophie e Tonzinho, que moram em Palmeira das Missões e ao Preto e o Café que moram comigo em Santa Maria, vocês me ensinaram o verdadeiro sentido das palavras amor e companheirismo.

A Universidade Federal de Santa Maria, que me abriu as portas e pude realizar minha graduação e pós-graduação.

Ao Programa de Pós-Graduação em Ciências Biológicas: Bioquímica Toxicológica pela oportunidade de fazer o mestrado.

A CNPq pelo apoio financeiro recebido.

Ao Laboratório de Virologia de Insetos, por ser esse excelente grupo de pesquisa e me acolher.

A todos, que mesmo não estando citado aqui, contribuíram para minha formação profissional e pessoal por toda a vida acadêmica, muito obrigada.

## RESUMO

### GENÔMICA DE UM BACULOVÍRUS ISOLADO DA LAGARTA PRAGA DE LEGUMINOSAS *Urbanus proteus* (LINNAEUS, 1758) (LEPIDOPTERA: HESPERIIDAE)

AUTORA: Ethiane Rozo dos Santos

ORIENTADOR: Daniel Mendes Pereira Ardisson-Araújo

Os baculovírus são vírus de insetos amplamente utilizados como vetores de expressão gênica e biopesticidas. Esses vírus podem infectar eficientemente larvas de várias pragas agrícolas em todo o mundo, causando uma doença letal. Neste trabalho, foi encontrado um novo baculovírus isolado do estágio larval de *Urbanus proteus* (L.), a lagarta desfolhadora do feijão nos anos de 1980 e foi caracterizado o seu genoma completo. Esta foi uma importante praga de várias leguminosas no Brasil e pertence à família das borboletas HesperIIDae, de onde não foi descrita nenhuma sequência genômica de baculovírus. Este novo vírus mostrou ter o menor genoma entre todos os alphabaculovírus sequenciados até o momento, com 105.555 pb e 119 ORFs. Foram encontrados dez genes únicos, sete genes *bro* e os 38 genes compartilhados entre todos os baculovírus. Verificou-se que UrprNPV estava relacionado com os baculovírus que infectam o gênero *Adoxophyes*, AdorNPV e AdhoNPV, com alta distância genética e comprimento longo do ramo. Curiosamente, poucas filogenias baseadas nos genes compartilhados individuais apoiaram a relação de UrprNPV com ambos, AdorNPV e AdhoNPV. É importante ressaltar que o aumento no número de baculovírus completamente sequenciados aponta para uma forma interessante de entender os baculovírus e a sua evolução, podendo ajudar potencialmente o uso de baculovírus tanto como biopesticida e vetores de expressão.

**Palavras-chave:** Genômica. Baculovírus. *Alphabaculovirus*. *Urbanus proteus*. Evolução.



## ABSTRACT

### GENOMICS OF A BACULOVIRUS ISOLATED FROM A LEGUMINOUS PEST, *Urbanus proteus* (LINNAEUS, 1758) (LEPIDOPTERA: HESPERIIDAE)

AUTHOR: ETHIANE ROZO DOS SANTOS

ADVISOR: DANIEL MENDES PEREIRA ARDISSON-ARAÚJO

Baculoviruses are insect viruses largely used as vectors for gene expression and biopesticides. These viruses can efficiently infect larvae of several agricultural pests worldwide, causing a lethal disease. In this work, a novel baculovirus isolated from the larval stage of *Urbanus proteus* (L.), the bean leafhopper caterpillar was found in the 1980s and the complete genome was characterized. This was an important pest of several legumes in Brazil and belongs to the butterfly family HesperIIDae, from where no baculovirus genome sequence has been described. This new virus was shown to have the smallest genome among all alphabaculoviruses sequenced to date, with 105,555 bp and 119 putative ORFs. Ten unique genes, seven *bro* genes and the 38 baculovirus core genes were found. UrprNPV was found to be related to the Adoxophyes-infecting baculoviruses AdorNPV and AdhoNPV, with high genetic distance and long branch length. Interestingly, few individual core gene-based phylogenies were found to support the relationship of UrprNPV to both AdorNPV and AdhoNPV. Importantly to note that the increase in the number of completely sequenced baculoviruses points to an interesting way of understanding baculoviruses and their evolution, potentially helping the use of baculovirus, both as biopesticide and expression vectors.

**Keywords:** Genomic. Baculovirus. Alphabaculovirus. *Urbanus proteus*. Evolution.

## LISTA DE ILUSTRAÇÕES

### INTRODUÇÃO

Figura 1 - Diagrama esquemático dos dois tipos de fenótipos infectivos de baculovírus.....	14
Figura 2 - Diagrama esquemático da estrutura dos corpos de oclusão (OB) e vírions derivados de oclusão (ODV) de baculovírus.....	15
Figura 3 – Esquema do ciclo de infecção baculoviral completo no inseto hospedeiro.....	17
Figura 4 - Ciclo de infecção com baculovírus em uma célula de inseto.....	18
Figura 5 – Diagrama esquemático da família <i>Baculoviridae</i> .....	20
Figura 6 - <i>Urbanus proteus</i> em sua fase larval e adulta.....	24

### ARTIGO 1

Fig. 1. UrprNPV is an alphabaculovirus. Maximum likelihood inference 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 a 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. The genus Alphabaculovirus contained UrprNPV (boldface) that clustered together with both AdhoNPV and AdorNPV (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).....45

Fig. 2. The gene content among UrprNPV and its relatives and the genomic context of the *iap-3* of UrprNPV and OpbuNPV. (A) The whole gene content of four alphabaculoviruses (UrprNPV, AdhoNPV, AdorNPV and OpbuNPV) were compared by BLASTX to find homologs. A total of 200 genes were found: 79 were shared among all the four viruses, 16 among UrprNPV, AdhoNPV, and AdorNPV; one between UrprNPV and OpbuNPV (*iap-3*); and 21 were unique to the UrprNPV genome. (B) Comparison of the predicted amino acid sequence of the *iap-3* genes of UrprNPV and its relatives. We show the predicted domains found using HHpred and the encoded amino acid sequences for each gene and found both the BIR and the RING domains, characteristics of baculoviral IAP proteins. OpbuNPV-ORF-61 and UrprNPV-IAP-3 have two BIR domains and one RING domain. In the OpbuNPV-ORF-26, it appears that the C-terminal portion corresponding to the RING domain was lost.....46

Fig. 3. The UrprNPV *bro* gene analyses. (A) Phylogenetic analysis of the UrprNPV *bro-a* gene and some homologs reveals an association with ascovirus genes. The tree was inferred by the FastTree and the branch support was determined by a SH-like method. (B) UrprNPV *bro-a* to *bro-d* locus is compared to other closely related organisms (UrprNPV, HaNPV-G4, AdorNPV, and AdhoNPV). An ancient HaNPV-G4-*bro-b*-related gene seemed to undergo several splits in the genome of UrprNPV, AdorNPV, and AdhoNPV. In the AdorNPV and AdhoNPV genomes, a small region at the amino-terminal portion is missed that corresponded to the *bro-c* and *bro-d* in the UrprNPV. The genome present three other *bro* genes (e, f, and g) in different loci.....47

Fig. 4. Gene parity plots and locally collinear blocks of the UrprNPV genome against AdorNPV and AcMNPV-C6. Plots show the gene content and position in the UrprNPV genome with those from (A) the AdorNPV and (B) the AcMNPV-C6 genomes. Each dot represents a homolog.

The pink colored dots represent the 38 baculovirus core genes. (C) Genomic comparison of the UrprNPV genome and the AdorNPV. Twelve Locally Collinear Blocks (LCB) numbered from 1 to 12 were found. Same colors depict same LCBs across the genomes. Rearrangement can be seen between the analyzed genomes.....48

Fig. 5. Analysis of the baculovirus genome sizes across alphabaculoviruses and other features. (A) Comparison of the size of intergenic spaces and the total genomic size for all the reannotated alphabaculovirus genomes. The six outliers are indicated (names above blue circles). (B) Comparison of the total number of ORFs and the total genomic size for all the reannotated alphabaculovirus genomes. There is a positive correlation among the genomes. UrprNPV is represented by the red closed circle. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).....49

Fig. 6. UrprNPV varies its relationship to other baculovirus based on the individual core gene analyses.(A) Most of the individual the FastTree-based phylogenies of each core gene did not reconstruct the UrprNPV phylogeny observed for the inference based on the concatenated dataset. Thirty eight nodes were found and scored from 0 to 38 depicting all the currently defined baculovirus core genes. Some branches were collapsed for clarity. (B) Correlation of the number of node reconstruction for each core gene and the core gene sizes for several selected baculovirus genomes. We highlighted some outliers and the error bars represent the standard deviations for the gene size.....50

## LISTA DE ABREVIATURAS E SIGLAS

AcMNPV	<i>Autographa californica multiple nucleopolyhedrovirus</i>
AdhoNPV	<i>Adoxophyes honmai nucleopolyhedrovirus</i>
AdorNPV	<i>Adoxophyes orana nucleopolyhedrovirus</i>
AgMNPV	<i>Anticarsia gemmatalis multiple nucleopolyhedrovirus</i>
AgMNPV-2D	<i>Anticarsia gemmatalis multiple nucleopolyhedrovirus-isolado2D</i>
BmNPV	<i>Bombyx mori nucleopolyhedrovirus</i>
<i>bro</i>	<i>baculovirus repeat ORF</i>
BV	do inglês ‘budded virus’
<i>coi</i>	<i>citocromo c oxidase</i>
CONAB	Companhia Nacional de Abastecimento
CoveNPV	<i>Condylorrhiza vestigialis nucleopolyhedrovirus</i>
DisaGV	<i>Diatraea saccharalis granulovirus</i>
EGT	ecdisteróide UDP-glicosiltransferase
EMBRAPA	Empresa Brasileira de Pesquisa Agropecuária
ErelGV	<i>Erinnyis ello granulovirus</i>
ErelGV-1986	<i>Erinnyis ello granulovirus-isolado1986</i>
<i>gran</i>	<i>granulina</i>
GV	<i>Granulovirus</i>
h p.i.	horas pós-infecção
ICTV	International Committee on Taxonomy of Viruses
HaSNPV	<i>Helicoverpa armigera single nucleopolyhedrovirus</i>
HZSNPV	<i>Helicoverpa zea single nucleopolyhedrovirus</i>
<i>iap-3</i>	<i>Inhibitor-of-Apoptosis-3</i>
<i>lef</i>	<i>late expression factor</i>
LoobMNPV-2000	<i>Lonomia obliqua multiple nucleopolyhedrovirus-2000</i>
MAPA	Ministério da Agricultura, Pecuária e Abastecimento
MolaGV	<i>Mocis latipes granulovirus</i>
NGS	do inglês ‘Next Generation Sequencing’
MNPV	<i>Multiple Nucleopolyhedrovirus</i>
NPV	<i>Nucleopolyhedrovirus</i>
OB	do inglês ‘occlusion body’
ODV	do inglês ‘occlusion derived virus’
<i>odv-ec43</i>	<i>occlusion-derived virion envelope capsid -43</i>
OpbuNPV	<i>Operophtera brumata nucleopolyhedrovirus</i>
ORFs	do inglês ‘open read frames’
PeluSNPV	<i>Perigonia lusca single nucleopolyhedrovirus</i>
PIF	do inglês ‘per os infectivity factors’
<i>polh</i>	<i>poliedrina</i>
PsinSNPV-IE	<i>Pseudoplusia includens single nucleopolyhedrovirus-IE</i>
SfMNPV	<i>Spodoptera frugiperda multiple nucleopolyhedrovirus</i>
SNPV	<i>Single Nucleopolyhedrovirus</i>
<i>vlf-1</i>	<i>very late factor-1</i>

## SUMÁRIO

<b>1</b>	<b>APRESENTAÇÃO</b> .....	12
<b>2</b>	<b>INTRODUÇÃO</b> .....	13
<b>2.1</b>	<b>Homeostase celular e o vírus</b> .....	13
<b>2.2</b>	<b>Características gerais de baculovírus</b> .....	13
<b>2.3</b>	<b>Estrutura dos corpos de oclusão e fenótipos virais</b> .....	13
<b>2.4</b>	<b>Ciclo de infecção</b> .....	16
<b>2.5</b>	<b>Família <i>Baculoviridae</i> e os genomas sequenciados</b> .....	19
<b>2.6</b>	<b>Baculovírus como controladores naturais de insetos praga</b> .....	21
<b>2.7</b>	<b><i>Urbanus proteus</i> inseto praga de leguminosa</b> .....	22
<b>2.8</b>	<b><i>Urbanus proteus</i> nucleopolyhedrovírus e a Embrapa-Soja</b> .....	24
<b>3</b>	<b>OBJETIVOS</b> .....	25
<b>3.1</b>	<b>Objetivo geral</b> .....	25
<b>3.2</b>	<b>Objetivos específicos</b> .....	25
<b>4</b>	<b>RESULTADOS</b> .....	26
	The complete genome sequence of the first hesperiid-infecting alphabaculovirus isolated from the leguminous pest <i>Urbanus proteus</i> (Lepidoptera: Hesperidae).....	26
<b>5</b>	<b>DISCUSSÃO</b> .....	52
<b>6</b>	<b>CONCLUSÃO</b> .....	56
	<b>REFERÊNCIAS BIBLIOGRÁFICAS</b> .....	57
	<b>APÊNDICE A – FIGURA SUPLEMENTAR 1</b> .....	63
	<b>APÊNDICE B – FIGURA SUPLEMENTAR 2</b> .....	64
	<b>APÊNDICE C – FIGURA SUPLEMENTAR 3</b> .....	65
	<b>APÊNDICE D – TABELA SUPLEMENTAR 1</b> .....	67
	<b>APÊNDICE E – TABELA SUPLEMENTAR 2</b> .....	72
	<b>APÊNDICE F – TABELA SUPLEMENTAR 3</b> .....	78

## 1 APRESENTAÇÃO

Esta dissertação aborda a montagem, descrição e caracterização do genoma de um novo baculovírus que encontra-se estruturada da seguinte forma: o item **INTRODUÇÃO**, contém uma revisão de literatura concisa sobre os temas abordados nesta dissertação; seguido dos **OBJETIVOS** deste trabalho. Os **RESULTADOS** estão apresentados em forma de artigo com o título: The complete genome sequence of the first hesperiid-infecting alphabaculovirus isolated from the leguminous pest *Urbanus proteus* (Lepidoptera: Hesperidae), nele contém a metodologia utilizada e os resultados obtidos que compõe esta dissertação. Neste artigo contém as seções: Introdução, Material e Métodos, Resultados e Discussão e Conclusão. Os itens **DISCUSSÃO** e **CONCLUSÃO**, encontrados no final desta dissertação, apresentam descrições, interpretações e comentários gerais sobre o artigo científico incluído neste trabalho. As **REFERÊNCIAS BIBLIOGRÁFICAS** correspondem somente às citações que aparecem nos itens **INTRODUÇÃO** e **DISCUSSÃO** da dissertação. As demais referências bibliográfica estão presentes no artigo publicado.

## 2 INTRODUÇÃO

### 2.1 HOMEOSTASE CELULAR E O VÍRUS

Células, a unidade fundamental de todos os seres vivos, permanecem em um estado estável e dinâmico em nível molecular, que para cada reação metabólica em uma via, o substrato é fornecido pela reação anterior na mesma velocidade em que é convertido a produto (NELSON e COX, 2014). Os vírus são parasitas intracelulares obrigatórios e causam modificações nos processos metabólicos de seus hospedeiros, como é o caso dos baculovírus, uma família de microrganismos causadores de infecções vários insetos durante sua fase larval (CLEM e PASSARELLI, 2013). Durante a infecção, o vírus manipula o ambiente celular, perturbando o estado dinâmico do hospedeiro por meio de diversas proteínas virais que alteram o seu comportamento e fisiologia (KAMITA et al., 2005; KATSUMA et al., 2012; O'REILLY e MILLER, 1989).

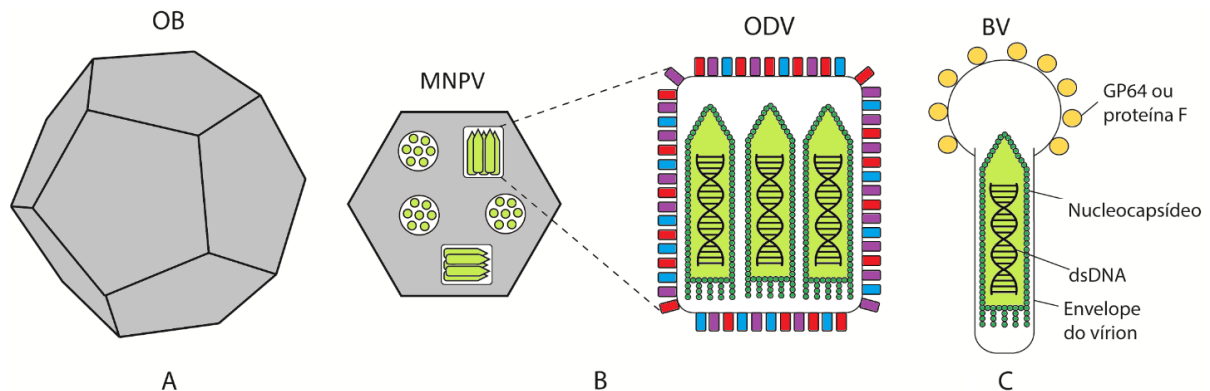
### 2.2 CARACTERÍSTICAS GERAIS DE BACULOVÍRUS

Os baculovírus pertencem à família *Baculoviridae* e caracterizam-se por ser vírus de DNA fita-dupla circular com tamanho de genoma variando entre 80 a 180 kbp (SLACK e ARIF, 2007). O nome baculovírus decorre do latim *baculo* que significa bastão, isso se deve à forma bacilar do nucleocapsídeo viral em que está envelopado numa membrana lipoproteica (ROHRMANN, 2013). Baculovírus formam o mais estudado grupo de vírus específicos de insetos com a característica de formar corpos de oclusão (OBs, do inglês 'occlusion bodies') para resistência dos vírions no ambiente. São capazes de infectar insetos das ordens Lepidoptera (estágio larval de mariposas e borboletas), Hymenoptera (estágio larval de vespas que se comportam como lagartas) e Diptera (estágio larval e aquático de mosquitos) (JEHLE et al., 2006b; ROHRMANN, 2013) com alta especificidade de hospedeiro. Esta característica de especificidade faz com que os baculovírus sejam importantes agentes de controle biológico, sendo considerados seguros para as pessoas e animais e uma alternativa viável aos inseticidas químicos (SUN e PENG, 2007; SZEWCZYK et al., 2006). Além disso, baculovírus são ferramentas amplamente utilizadas para expressão heteróloga de proteínas eucarióticas e apresentam uso potencial como entregadores para terapia gênica (SUMMERS, 2006).

### 2.3 ESTRUTURA DOS CORPOS DE OCLUSÃO E FENÓTIPOS VIRAIS

Para controle dos insetos-pragas da agricultura, os OBs (Figura 1A) de baculovírus são pulverizados nas folhas das plantações e ingeridos juntos com o alimento pelas lagartas, de modo a propiciar a multiplicação do vírus (MOSCARDI, 1999). Os baculovírus produzem dois tipos de fenótipos infectivos durante o ciclo infeccioso completo. O primeiro fenótipo é o vírion derivado de oclusão (ODV, do inglês ‘occlusion derived virus’) (Figura 1B), responsável pela infecção primária oral e está ocluído em um corpo cristalino proteico chamado de OB. O segundo fenótipo, o vírion brotado (BV, do inglês ‘budded virus’) (Figura 1C), tem função de promover a infecção secundária e espalhar a infecção ao longo do corpo do inseto hospedeiro (CLEM e PASSARELLI, 2013; ROHRMANN, 2013). Os BVs são produzidos após a infecção primária inicial e contêm um nucleocapsídeo dentro de um envelope que é derivado da membrana plasmática do hospedeiro, e que contém várias proteínas virais importantes para a infecção.

Figura 1 - Diagrama esquemático dos dois tipos de fenótipos infectivos de baculovírus.



Fonte: Acervo pessoal.

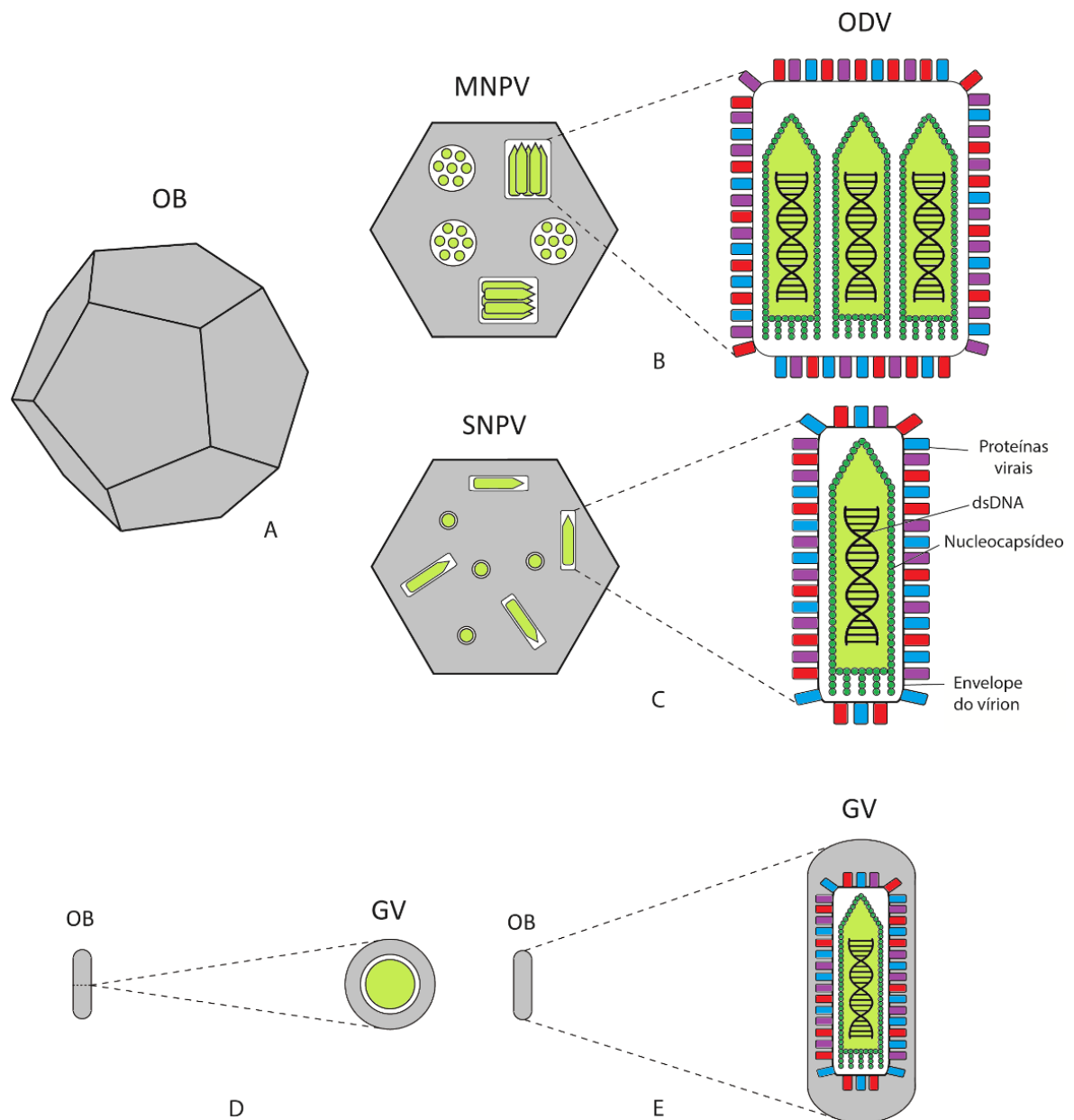
A estrutura dos corpos de oclusão (OB) de baculovírus são mostrados em (A), os vírions derivados de oclusão (ODV) (B), que são responsáveis pela infecção primária e o vírion brotado (BV) (C), que é responsável pela infecção secundária. O envelope de ODV e BV e seus nucleocapsídeos contêm inúmeras proteínas. As proteínas de fusão GP64 ou proteína F presentes em alfa, beta e deltabaculovírus formam uma estrutura peplomérica (círculos amarelos) no fenótipo BV (AU et al., 2013).

Os OBs são formados por uma matriz cristalina proteica que varia de acordo com o gênero viral, sendo a proteína poliedrina codificada pelo gene *polh* para *Alphabaculovirus* e *Gammabaculovirus* ou a proteína granulina codificada pelo gene *gran* quando for *Betabaculovirus*. Dentro destes OBs podem conter um (no caso de ‘granulovirus’ – GV) (Figura 2D) ou vários vírions por corpo de oclusão (no caso de ‘nucleopolyhedrovirus’ - NPV) (Figura 2A). Cada partícula infecciosa (vírion), constituída de DNA e de uma camada externa de



proteínas (capsídeo), pode conter um (no caso de GVs (Figura 2E) e ‘single nucleopolyhedrovirus’ – SNPV (Figura 2C) ou vários (no caso dos ‘multiple nucleopolyhedrovirus’ – MNPV (Figura 2B) nucleocapsídeos por vírion. Os ODVs consistem num único ou múltiplos nucleocapsídeos presentes em um envelope com origem e composição proteica diferentes comparado com o BV. Mesmo tendo duas formas distintas, os vírions são geneticamente idênticos e possuem papéis distintos na transmissão viral, ODV de inseto para inseto (inter-hospedeiro) e BV de célula para célula (intra-hospedeiro) (JEHLE et al., 2006b).

Figura 2 - Diagrama esquemático da estrutura dos corpos de oclusão (OB) e vírions derivados de oclusão (ODV) de baculovírus.



Fonte: Acervo pessoal.

Os ODVs estão embebidos numa matriz cristalina proteica para formar os OBs, que podem ter formato de poliedro quando os cristais são formados por poliedrina (A) ou formato de granulo, quando formados por granulina (D). Os OBs podem conter um (granulovírus – GV (D, E) ou ‘single nucleopolyhedrovirus’ – SNPV (C) ou mais (‘multiple nucleopolyhedrovirus’ - MNPV) (B) nucleocapsídeos por envelope. O envelope de ODV contém inúmeras proteínas representados pelos quadrados coloridos no envelope do vírion e pelos círculos de coloração verde nos nucleocapsídeos.

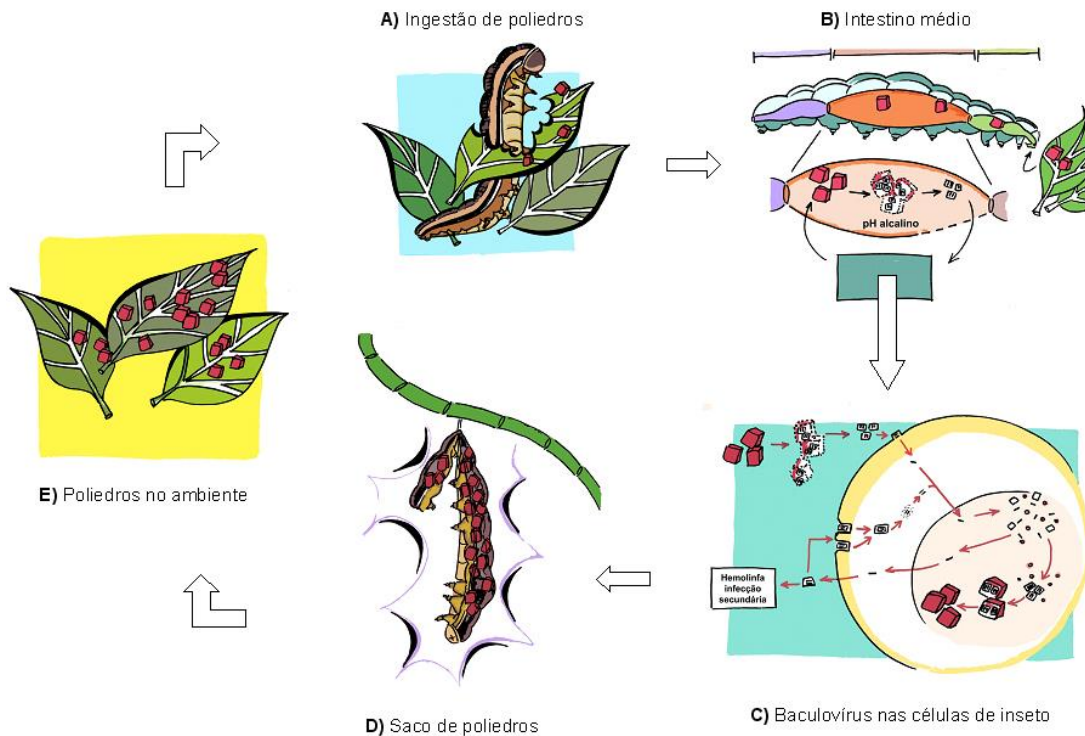
## 2.4 CICLO DE INFECÇÃO

A lagarta ingere os OBs ao se alimentar de substratos contaminados (Figura 3A), e para que ocorra o início do processo infeccioso, é necessária a liberação dos ODVs pela solubilização da proteína (granulina ou poliedrina) que envolve os vírions nos OBs. Essa solubilização só é possível porque o intestino médio da lagarta possui um pH alcalino (Figura 3B), variando entre 9,5 a 11,5, iniciando assim a etapa da infecção denominada infecção primária, que se estende até 12 h pós-infecção (p.i.) viral (HAAS-STAPLETON et al., 2004; SLACK e ARIF, 2007). As partículas virais atravessam a membrana peritrófica (uma rede de quitina, mucopolissacarídeos e proteínas) e os vírions envelopados se fundem às membranas das microvilosidades das células epiteliais do intestino médio, mediada por receptores desconhecidos presentes na membrana celular (HORTON e BURAND, 1993; HAAS-STAPLETON et al., 2004). O nucleocapsídeo é direcionado para o núcleo via filamentos de actina, passando pelo poro nuclear. A estrutura em forma de bastão é desfeita, o genoma viral exposto e os promotores de genes da fase precoce (responsáveis por subverter a célula hospedeira e preparar o ambiente para replicação) são apresentados para a maquinaria de transcrição do hospedeiro (RNA polimerase II) (ROHRMANN, 2013). Os ODVs promovem a infecção primária ou infecção oral devido à existência de proteínas de membrana chamadas fatores de infectividade oral (PIF, do inglês ‘*per os infectivity factors*’), que são responsáveis pelo reconhecimento do vírus ao seu hospedeiro (OHKAWA et al., 2005).

Em sequência, na fase tardia existe a transcrição e produção de genes estruturais e a replicação do genoma viral no núcleo, entre 12 e 24 h p.i. A célula perde sua morfologia natural devido a hipertrofia nuclear (Figura 4A). É possível observar uma região de intensa síntese de RNAs virais, replicação de genoma viral e montagem de nucleocapsídeos, denominada estroma virogênico. Os nucleocapsídeos montados saem por processo de brotamento do envelope nuclear para o citoplasma, onde adquirem um novo envelope na membrana plasmática para produzir BV (Figura 4B) (AU e PANTE, 2012; WILLIAMS e FAULKNER, 1997). No final desta fase, os BVs são liberados pelas inúmeras células intestinais infectadas, ganham acesso à

hemolinfa do inseto e infectam os demais órgãos como túbulos de Malpighi e tecido adiposo, disseminando-se por toda o corpo do inseto (Figura 3C) (BLISSARD e ROHRMANN, 1990).

Figura 3 - Esquema do ciclo de infecção baculoviral completo no inseto hospedeiro.



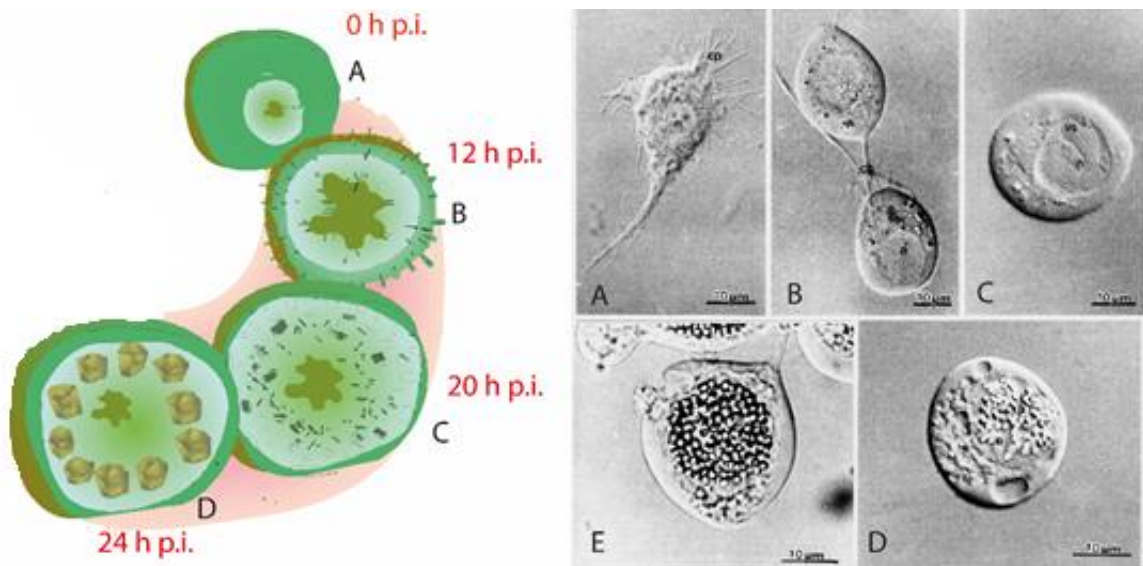
Fonte: Modificado de BARROS, 2012.

A larva ingere os corpos de oclusão (OBs) presentes no alimento (A), neste esquema são poliedros em folhas. Os OBs atingem o intestino médio e se dissolvem, liberando os vírus derivado de oclusão (ODVs) (B e C). Os vírus estabelecem a infecção nas células do intestino, replicam-se, brotam das células em direção a hemolinfa e espalham a infecção ao longo de todo o corpo do inseto (C). As células com infecção tardia produzem novos OBs (C). No final do ciclo, o inseto se torna um saco de OBs que são liberados no ambiente (D e E).

Na fase muito tardia, aproximadamente 24 h p.i., os genes com promotores muito tardios, responsáveis pela montagem do fenótipo ODV, começam a ser expressos. Entre os genes da fase tardia, pode-se destacar a proteína responsável pela formação da matriz dos OBs (poliedrina ou granulina) e a proteína P10, possivelmente responsável pela montagem do OB e lise nuclear (SLACK e ARIF, 2007; RUSSELL et al., 1991; VAN OERS e VLAK, 1997; CARPENTIER et al., 2008). Durante essa fase, o estroma virogênico se torna mais condensado e a zona de anel nuclear se expande para acomodar os nucleocapsídeos que se acumulam (SLACK e ARIF, 2007). Os nucleocapsídeos, deixam de ser destinados à produção de BV e permanecem na zona do anel nuclear para formação de ODVs (Figura 4C), onde adquirem um envelope da bicamada lipídica da membrana nuclear interna que invagina os nucleocapsídeos

como microvesículas no final da infecção (BRAUNAGEL et al., 1996; BRAUNAGEL e SUMMERS, 2007). O envelope de ODV é mais rígido que o envelope de BV devido à presença de fosfolipídios de ácidos graxos saturados. Os ODVs são oclusos para formação dos OBs, tornando o núcleo com uma aparência granulosa (Figura 4D). Os OBs são liberados no ambiente após a morte e liquefação da lagarta (Figura 3D), servindo de inóculo para a infecção de novas lagartas, de modo a reiniciar o ciclo infeccioso (Figura 3E) (BLISSARD e ROHRMANN, 1990; THIEM e CHENG, 2009).

Figura 4 - Ciclo de infecção com baculovírus em uma célula de inseto.



Fonte: Modificado de SLACK e ARIF, 2007; Modificado de POMBO et al., 1998

Em forma de esquema (a esquerda) e em microscopia de luz (a direita). Demonstra a mudança da morfologia celular nas horas pós-infecção (h p.i.), começando com a hipertrofia do núcleo e formação do estroma virogênico (A, B), produção do fenótipo BV (B) e fenótipo ODV (C), os quais são ocluídos por uma matriz proteica cristalina formando os OBs (D e E) e eventualmente liberados com a lise celular.

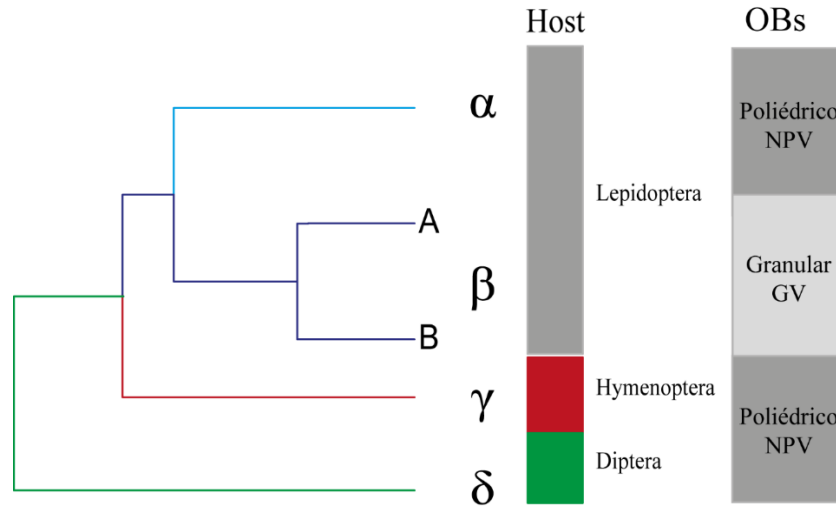
A morte por baculovírus é facilmente reconhecida em termos fenotípicos, pois estes têm a capacidade de bloquear a fase de muda e a empupação das larvas infectadas. Isso se deve à ação de uma enzima viral chamada ecdisteróide UDP-glicosiltransferase (EGT), que inativa o principal hormônio esteroide responsável pela muda do inseto, a ecdisona (O'REILLY e MILLER, 1989). Mais recentemente, foi visto que o gene *egt* também é responsável por induzir o comportamento de escalada para o topo da planta em alguns alphabaculovírus (HOOVER et al., 2011). Geralmente a maioria das lagartas vive no solo e só sai para alimentar-se à noite, para evitar predação. Em contrapartida, as larvas infectadas por baculovírus migram para a

folhagem superior da planta, onde morrem penduradas no topo da planta e se liquefazem (CLEM e PASSARELLI, 2013), aumentando assim o espalhamento viral.

## 2.5 FAMÍLIA *BACULOVIRIDAE* E OS GENOMAS SEQUENCIADOS

A família *Baculoviridae* está agrupada em quatro gêneros, com a análise dos 38 genes compartilhados (GARAVAGLIA et al., 2012; JAVED et al., 2017). Estes gêneros também se agrupam de acordo com a morfologia dos corpos de oclusão e espectro de hospedeiros, os *Alphabaculovirus* que infectam insetos da ordem Lepidoptera e tem seus OBs em formato de poliedro, com tamanho variando entre 800-2.000 nm (JEHLE et al., 2006b). Os *Betabaculovirus*, também infectam insetos da ordem Lepidoptera, porém apresentam OBs ovocilíndricos menores que os poliedros (aproximadamente 0,13 - 0,5 µm) e contêm um único vírion (HERNIOU et al., 2012). Estes são conhecidos como granulovírus (GVs) e são subdivididos em Clado A e B (MIELE et al., 2011). *Gammabaculovirus* são específicos de Hymenoptera, possuem OBs poliédricos de cerca de 0,4-1,1 µm de tamanho com os nucleocapsídeos envolvidos individualmente (SNPV) no seu vírion, além do fenótipo BV estar ausente (JEHLE et al., 2006b). Já os *Deltabaculovirus*, foram isolados em 27 larvas da ordem Diptera, porém apenas um foi sequenciado, o vírus isolado da espécie *Culex nigripalpus* (ROHRMANN, 2013). O CuniNPV possui OBs globulares com tamanho cerca de 400 nm de diâmetro. É importante destacar que a proteína do OB parece não ser homóloga aos genes *poliedrina* (*polh*) ou *granulina* (*gran*) dos baculovírus que infectam Lepidoptera e Hymenoptera e apresenta a característica de flutuar na água, permitindo que as larvas filtradoras dos mosquitos se infectem (Figura 5) (LUCAROTTI et al., 2012).

Figura 5 – Diagrama esquemático da família *Baculoviridae*.



Fonte: Acervo pessoal.

Os *Alphabaculovirus* são mostrados em azul claro, possuem hospedeiro da ordem Lepidoptera, corpo de oclusão (OB) em formato poliédrico. Os *Betabaculovirus* em azul escuro, tem hospedeiro da ordem Lepidoptera e OB em formato granular. *Gammabaculovirus* em vermelho, com seu hospedeiro pertencente a ordem Hymenoptera e o formato do OB poliédrico. *Deltabaculovirus* em verde, tem como hospedeiro a ordem Diptera e o formato do OB poliédrico.

O primeiro baculovírus sequenciado foi o da espécie tipo de *Baculoviridae*, *Autographa californica multiple nucleopolyhedrovirus* em 1994 pelo método de Sanger (AYRES et al., 1994). A partir de 2005, as plataformas de sequenciamento de alto desempenho – NGS, do inglês ‘Next Generation Sequencing’ foram criadas com o intuito de gerar dados em larga escala, usando métodos de alto desempenho viável economicamente (FERREIRA e NETO, 2003). Existem vários métodos de sequenciamento NGS, dentre eles pode-se destacar o de pirosequenciamento, que é uma técnica que tem como princípio a detecção de um pirofosfato liberado durante a incorporação do desoxinucleotídeo à cadeia, ao invés da terminação da cadeia de DNA com ddNTPs (dideoxy) (RONAGHI et al., 1998). Outro método é pelo sequenciador ‘Illumina Genome Analyzer’ que foi produzido em 2006 e baseia-se no conceito de “sequenciamento por síntese”, este requer que as sequências a serem determinadas sejam convertidas numa biblioteca de sequenciamento especial, permitindo à amplificação e imobilização das sequências para serem submetidas à sequenciamento. Para este propósito, dois adaptadores diferentes são adicionados às terminações 5’ e 3’ de todas as moléculas (KIRCHER et al., 2009). Já o método desenvolvido pela Ion Torrent utiliza uma abordagem indireta para sequenciamento de DNA com base em nucleotídeos fluorescente-etiquetados. Explorando o fato de que um íon de hidrogênio é liberado sempre que um nucleotídeo se atribui a uma fita de DNA, foram projetados sensores em microchip, para detectar esta assinatura química (MERRIMAN et al., 2012).

Até o presente momento existem 212 genomas sequenciados de baculovírus no banco de dados de nucleotídeos do NLM/NCBI ‘Genbank’, destes 40 foram descritos no Brasil, com apenas oito espécies inéditas: *Anticarsia gemmatalis multiple nucleopolyhedrovirus-2D* (AgMNPV-2D) (OLIVEIRA et al., 2006), *Mocis latipes granulovirus* (MolaGV) (ARDISSON-ARAÚJO et al., 2018), *Erinnyis ello granulovirus-1986* (ErelGV-1986) (ARDISSON-ARAÚJO et al., 2014), *Diatraea saccharalis granulovirus* (DisaGV) (ARDISSON-ARAÚJO et al., 2016a), *Chrysodeixis includens single nucleopolyhedrovirus-IE*, publicado como *Pseudoplusia includens single nucleopolyhedrovirus-IE* (PsinSNPV-IE) (CRAVEIRO et al., 2015), *Condylorrhiza vestigialis nucleopolyhedrovirus* (CoveNPV) (CASTRO et al., 2017), *Lonomia obliqua multiple nucleopolyhedrovirus-2000* (LoobMNPV-2000) (ARAGÃO-SILVA et al., 2016) e *Perigonia lusca single nucleopolyhedrovirus* (PeluSNPV) (ARDISSON-ARAÚJO et al., 2016b). Há 32 isolados brasileiros, destes, 17 são da espécie *Anticarsia gemmatalis multiple nucleopolyhedrovirus* (AgMNPV) (BRITO et al., 2016), seis da espécie *Chrysodeixis includens single nucleopolyhedrovirus* (CRAVEIRO et al., 2015), seis *Erinnyis ello granulovirus* (ErelGV) (ainda não publicado) um de *Spodoptera frugiperda multiple nucleopolyhedrovirus* (SfMNPV) (WOLFF et al., 2008), um de *Bombyx mori nucleopolyhedrovirus* (BmNPV) (ARDISSON-ARAÚJO et al., 2013) um de *Helicoverpa zea single nucleopolyhedrovirus* (HzSNPV) isolado da espécie *Helicoverpa armigera* (ARDISSON-ARAÚJO et al., 2015). Estes três últimos já foram também sequenciados em outros países.

## 2.6 BACULOVÍRUS COMO CONTROLADORES NATURAIS DE INSETOS PRAGA

Inerente ao fato de ser um vírus específico de insetos, os baculovírus são importantes controladores naturais de populações e são amplamente utilizados no controle biológico de insetos-pragas na agricultura. Sabe-se que o controle de insetos praga é realizado principalmente com uso de inseticidas químicos que, se usados de forma inadequada, podem favorecer o surgimento de pragas e sua expansão descontrolada, além de permitir o surgimento de desfolhadores secundários e promover contaminação ambiental e intoxicação humana (PEDIGO e RICE, 2005).

Uma alternativa aos inseticidas químicos que tem sido utilizado com sucesso é o controle biológico baseado em patógenos naturais de insetos, como o caso de baculovírus. Os baculovírus são usados desde 1930 nos EUA e Canadá, quando acidentalmente um vírus específico a *Diprion hercyniae* exportado da Europa, conseguiu controlar efetivamente a praga

do abeto (BIRD e ELGEE, 1957). Na China, o primeiro baculovírus a ser comercialmente vendido foi o vírus da espécie *Helicoverpa armigera single nucleopolyhedrovirus* (HaSNPV) no ano de 1993 (ZHANG et al., 1995; SUN e PENG, 2007). Atualmente, a China é o país com o maior número de produtos virais registrados voltados para o controle de pragas (SZEWCZYK et al., 2011; AHMAD et al., 2011; YANG et al., 2012).

No Brasil, na década de 1980 a lagarta que destrói plantações de mandioca *Erinnyis ello* foi controlada pelo uso do baculovírus *Erinnyis ello granulovirus* - ErelGV (SCHMITT, 1985; ARDISSON-ARAÚJO et al., 2014). Nos anos de 1982 e 1983 foi implementado no Brasil o maior programa de controle biológico do mundo usando o baculovírus da espécie *Anticarsia gemmatalis multiple nucleopolyhedrovirus* (AgMNPV) para pulverizar mais de dois milhões de hectares de soja no sul do país e conter a lagarta praga *Anticarsia gemmatalis* (MOSCARDI, 1999). Segundo a Empresa Brasileira de Pesquisa Agropecuária - EMBRAPA, 2017 o baculovírus AgMNPV é produzido por sete empresas privadas no Brasil e tem rigoroso controle de qualidade realizado pela Embrapa-Soja. No entanto, houve uma mudança no cenário, com a utilização descontrolada de fungicidas para contenção de fungos fitopatogênicos, *A. gemmatalis* deixou de ser a principal praga da soja e foi substituída pela lagarta *Chrysodeixis includens* que era controlada por fungos entomopatogênicos (SOSA-GOMÉZ et al., 2017). Atualmente, a utilização de AgMNPV é bem restrita, diminuindo a aplicação para 200 a 300 mil hectares de soja, devido ao sistema de tratamento fitossanitário que tem se modificado pelo surgimento do fungo da ferrugem asiática na soja (NAVA e NACHTIGAL, 2010) e pela especificidade do baculovírus que atua somente no controle de *A. gemmatalis*, não atingindo as demais espécies de lagartas que ganharam importância econômica nos últimos anos, como *C. includens* e algumas espécies do gênero *Spodoptera* (Bueno et al., 2012).

## 2.7 *Urbanus proteus* INSETO PRAGA DE LEGUMINOSA

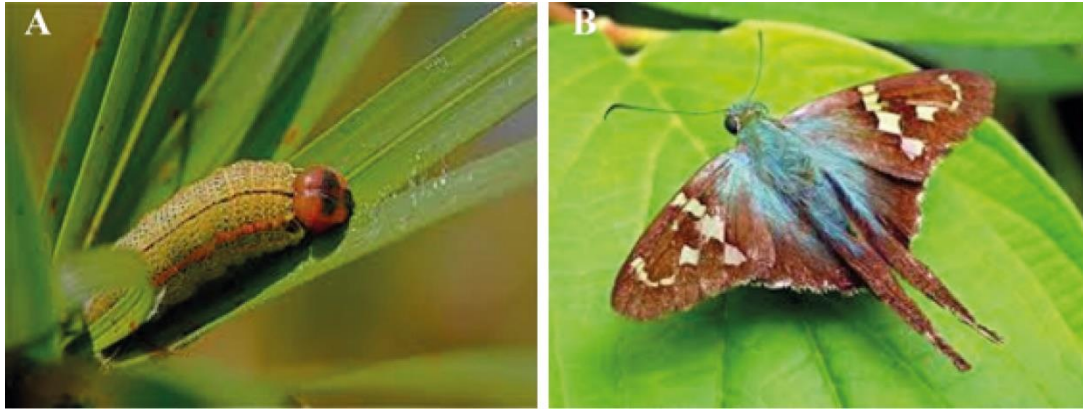
Assim como a soja, o feijão também é uma leguminosa amplamente cultivada nos trópicos e subtropicais tanto para consumo humano quanto para outros animais (JACKAI e DAOUST, 1986). Segundo o Ministério da Agricultura, Pecuária e Abastecimento - MAPA (2017), o Brasil é o principal produtor e consumidor de feijão no mundo. Só no ano passado foram consumidas 300 mil toneladas a mais do que é produzido, num total de 16,5 kg por habitante anualmente. Deste modo, a previsão da Companhia Nacional de Abastecimento – CONAB (2017) é de que os produtores aumentem a área plantada, passando de 2,8 milhões de hectares na safra 2015/16 para três milhões de hectares em 2016/17. Com a expansão da área



cultivada do feijoeiro sob irrigação, o cultivo sucessivo e intensivo das áreas, bem como o uso intensivo e sem o manejo apropriado de inseticidas e fungicidas químicos, favoreceram o aumento de pragas nas culturas. Dentre os principais invertebrados que são encontrados como pragas de culturas de leguminosas, destacam-se *Agrotis ipsilon* e *Spodoptera frugiperda*, como pragas do solo e *Omiodes indicata* e *Urbanus proteus* como pragas das folhas (QUINTELA, 2001). Durante a fase larval, insetos da espécie *U. proteus* são conhecidos como cabeça de fósforo (GALLO et al., 1988) ou rolos de folhas de feijão e podem ser facilmente identificados pela sua proeminente cabeça e coloração escura (Figura 6A) e porque têm o hábito de dobrar a folha para se proteger contra inimigos naturais (GREENE, 1971). Além do feijoeiro, essa lagarta é uma praga desfolhadora de várias outras leguminosas, como o feijão-caupi (*Vigna unguiculata* (L.) Walp.), ervilha (*Pisum sativum* L.) e fava (*Vicia faba* L.), ocorrendo nos estados brasileiros do Amazonas, Bahia, Ceará, Minas Gerais, Pará, Pernambuco, Rio de Janeiro, Rio Grande do Sul e São Paulo (COSTA LIMA, 1945). Este inseto também foi relatado para ocorrer na Colômbia, México, EUA e Canadá (VAN DAM e WILDE, 1977). No entanto, parece que esta praga de lepidópteros é dificilmente encontrada no campo atualmente, devido ao aumento da plantação de soja transgênica por *Bacillus thuringiensis* que eficientemente controla *U. proteus*.

*U. proteus* pertence à ordem Lepidoptera. A característica morfológica do inseto adulto é a presença de dois pares de asas membranosas cobertas de escamas e peças bucais adaptadas a sucção. O ciclo de vida dos lepidópteros é holometabólico e engloba quatro etapas: ovo, larva (denominada lagarta) (Figura 6A), pupa (crisálida) e imago (fase adulta) (Figura 6B) (EMBRAPA, 2016). *U. proteus* pertence à família Hesperidae, que é uma família de borboletas comumente conhecida como ‘skippers’ ou ‘skipper butterflies’, incluindo cerca de 4.000 espécies (BRIDGES, 1993), atualmente distribuídas em 567 gêneros (WARREN et al., 2008).

Figura 6 - *Urbanus proteus* em sua fase larval e adulta.



Fonte: adaptado de: <<https://thingsbiological.wordpress.com/2010/11/30/long-tailed-skipper-urbanus-proteus-larva/>> e <<https://www.flickr.com/photos/andreaskay/6973851042>>.

## 2.8 *Urbanus proteus* nucleopolyhedrovirus E A EMBRAPA-SOJA

No ano de 1984 foi caracterizado um baculovírus isolado de *U. proteus*, denominado *Urbanus proteus* nucleopolyhedrovirus. Temerak et al. (1984), descobriram um isolado de nucleopolyhedrovirus infectando a lagarta e causando um típico sinal de infecção por baculovírus. Em condições laboratoriais, este vírus foi capaz de matar *U. proteus* e foi considerado não-infeccioso para *Trichoplusia ni*, *Anticarsia gemmatalis* e *Spodoptera eridania*. Entretanto, as análises ficaram restritas a aspectos da patologia e estruturais do vírus e os pesquisadores não chegaram a caracterizar em nível molecular o vírus. Assim, a Embrapa-Soja atualmente dispõe de uma coleção com quase 100 extratos de lagartas que apresentavam características de infecção por baculovírus. Estas larvas foram isoladas em todo território brasileiro, mas principalmente no sul do país e foram enviados para este centro de pesquisa, e depositadas na coleção. Até 2012, a coleção estava sob a curadoria do pesquisador Dr. Flávio Moscardi, que após falecimento, foi transferida a responsabilidade para o Dr. Daniel Sosa-Gomez, colaborador deste presente projeto. Entre os extratos, encontrava-se um tubo identificado como *Urbanus proteus* nucleopolyhedrovirus (UrprNPV), que foi coletado no sul do Brasil nos anos 1980. Assim, em colaboração, foi-nos enviado parte deste extrato para caracterização do vírus putativo que poderia estar ali presente. O interesse primordial por este extrato deveu-se principalmente ao fato de que *U. proteus* pertence à família de borboletas HesperIIDae e nunca tinha sido isolado baculovírus desta família.

### 3 OBJETIVOS

#### 3.1 OBJETIVO GERAL

Sequenciar e analisar o genoma em nível molecular e filogenético do baculovírus isolado da lagarta praga de leguminosa *Urbanus proteus* de modo a ampliar o entendimento evolutivo da família *Baculoviridae*.

#### 3.2 OBJETIVOS ESPECÍFICOS

- Sequenciar e montar o genoma completo de *Urbanus proteus* nucleopolyhedrovirus;
- Descrever, caracterizar *in silico*, identificar e anotar todas as ORFs encontradas no genoma de UrprNPV;
- Estabelecer as relações filogenéticas da nova espécie de baculovírus com as demais espécies já descritas e promover análise de parâmetros para demarcação de nova espécie.
- Analisar peculiaridades do genoma de UrprNPV como a presença de ORFs únicas, a identidade e sintenia dos genes encontrados com outros baculovírus de espécies virais relacionadas.

## 4 RESULTADOS

A metodologia, os resultados e a discussão inseridos nesta dissertação apresentam-se sob a forma de manuscrito científico. O manuscrito foi publicado na revista *Virus Research*. SANTOS, E. R. et al. The complete genome sequence of the first hesperiid-infecting alphabaculovirus isolated from the leguminous pest *Urbanus proteus* (Lepidoptera: HesperIIDae). **Virus research**, v. 249, p. 76-84, 2018.

The complete genome sequence of the first hesperiid-infecting alphabaculovirus isolated from  
the leguminous pest *Urbanus proteus* (Lepidoptera: Hesperiiidae)

Ethiane R. Santos<sup>1</sup>, Lucas B. de Oliveira<sup>1</sup>, Lenen Peterson<sup>1</sup>, Daniel R. Sosa-Gómez<sup>2</sup>,  
Bergmann Morais Ribeiro<sup>3</sup>, Daniel M. P. Ardisson-Araújo<sup>1\*</sup>

<sup>1</sup>Laboratory of Insect Virology, Department of Biochemistry and Molecular Biology, Federal  
University of Santa Maria, Santa Maria, RS, Brazil. <sup>2</sup> Embrapa Soja, Londrina, PR, Brazil, <sup>3</sup>  
Laboratory of Baculovirus, Cell Biology Department, University of Brasilia, Brasília, DF,  
Brazil.

\*Corresponding author:

Dr. Daniel M. P. Ardisson-Araújo

Laboratory of Insect Virology

Department of Biochemistry and Molecular Biology

Federal University of Santa Maria

Santa Maria, RS, 97105-900

Email: [daniel.araujo@ufsm.br](mailto:daniel.araujo@ufsm.br)

## Abstract

Baculoviruses are insect viruses largely used as expression vectors and biopesticides. These viruses can efficiently infect the larval stage of several agricultural pests worldwide causing a lethal disease. In this work, we found a novel baculovirus isolated from the larval stage of *Urbanus proteus* (L.), the bean leafroller and characterized its complete genome. This is an important pest of several leguminous plants in Brazil and belongs to the butterfly family Hesperidae, from where no baculovirus genome sequence has been described. This new virus was shown to have the smallest genome among all alphabaculoviruses sequenced to date, with 105,555 bp and 119 putative ORFs. We found ten unique genes, seven *bro*, and the 38 baculovirus core genes. UrprNPV was found to be related to the *Adoxophyes*-infecting baculoviruses AdorNPV and AdhoNPV with high genetic distance and a long branch length. Interestingly, few individual core gene-based phylogenies were found to support the relationship of UrprNPV to both AdorNPV and AdhoNPV. Importantly, the increase in number of completely sequenced baculovirus points to a very exciting way to understand baculovirus and its evolution and could potentially help the use of baculovirus as both biopesticides and expression vectors.

**Keywords:** Baculovirus genome, alphabaculovirus, *Urbanus proteus*, evolution.

## 1. Introduction

Baculovirus belongs to the family *Baculoviridae*, a diverse family of rod-shaped enveloped viruses with circular double-stranded DNA genome ranging from 80 to 180 kbp (Rohrmann, 2013 a, b). As a hallmark, baculovirus infects orally the larval stage of insects by means of a crystalline protein matrix called occlusion body (OB). The OB carries the virions within and protects them from environmental adversities. The OB morphology used to be a taxonomic genus criterion to classify baculoviruses in those with polyhedral shape called by *Nucleopolyhedrovirus* (NPV) and those with granular shape called by *Granulovirus* (GV). Currently, the family is divided into four genera, the lepidopteran- specific *Alphabaculovirus* and *Betabaculovirus*, the hymenopteran- specific *Gammabaculovirus*, and the dipteran-specific *Deltabaculovirus* based on the alignment of 38 shared core genes (Jehle et al., 2006b; Javed et al., 2017). The complete infection cycle of alpha, beta, and deltabaculovirus infection produce two types of infectious phenotypes; the first one is the occlusion-derived virus (ODV), responsible for the primary oral infection, which is the occluded form into the OBs; the second one is the budded virus (BV), which plays a role to promoting the secondary infection and spreading the infection throughout the body of the host (Clem and Passarelli, 2013; Rohrmann, 2013 a, b). BVs seems to be lacked by the gammabaculoviruses, where the infection pathology is not clearly understood (Lucarotti et al., 2012).

Baculoviruses are specific to a narrow range of insects varying from one to dozens of hosts (Clem and Passarelli, 2013). This host specificity makes them widely used in the biological control of insect pests of forest and agricultural crops and they are also considered safe for humans and other animals and a viable alternative to chemical insecticides (Sun and Peng, 2007; Szewczyk et al., 2006). For instance, in 1982 and 1983 a baculovirus-based biological control program was established using the *Anticarsia gemmatalis* multiple nucleopolyhedrovirus isolate 2D (AgMNPV-2D) to control a prevalent soybean (*Glycine max*) pest at that time, the velvetbean caterpillar *Anticarsia gemmatalis* (Sosa-Gómez, 2017). Similar to soybean, other legumes are widely cultivated in the tropics and subtropics to feed human and other animals (Jackai and Daoust, 1986). *Urbanus proteus*, commonly known during the imago stage as “long-tailed skippers” and “skipper butterflies”, is a lepidopteran from the family HesperIIDae. This Family of butterflies includes around 4000 species (Bridges, 1993), distributed among 567 genera (Warren et al., 2008). During the larval stage, *U. proteus* (known by Brazilian growers as “match-head caterpillar” (Nava and Parra, 2002) and recognized by the Entomological Society of America Common Names Database

(<http://www.entsoc.org/commonnames>) as bean leafroller) can be easily identified by their prominent head, dark colored tegument, and by their habit of folding the leaf for protection against natural enemies (Greene, 1971). This caterpillar is a defoliating pest of several legumes, such as cowpea [*Vigna unguiculata* (L.) Walp.], pea (*Pisum sativum* L.), faba bean (*Vicia faba* L.), common bean, and also soybean, and occurs in the Brazilian states of Amazonas, Bahia, Ceará, Minas Gerais, Pará, Pernambuco, Rio de Janeiro, Rio Grande do Sul, and São Paulo (Costa Lima, 1945). This insect has been also reported to occur in Colombia, Mexico, the US, and Canada (van Dam and Wilde, 1977). However, it seems that this lepidopteran pest is uncommon in field nowadays and hard to be found due to the increase in *Bacillus thuringiensis* toxin-expressing transgenic soybean plantation.

In a previous work, Temerak et al. (1984) found a singly-embedded nucleopolyhedrovirus infecting the leafroller *Urbanus proteus* and causing a typical baculovirus wilt symptom. In a laboratory condition, this virus was able to kill *U. proteus* and found to be non-infectious to *Trichoplusia ni*, *Anticarsia gemmatalis*, and *Spodoptera eridania*. In this work, we sequenced the complete genome of a putative new species of baculovirus found in an *U. proteus* larva extract with symptoms of baculovirus infection obtained from the virus collection library of the Brazilian Agricultural Research Corporation (Portuguese acronym Embrapa, Empresa Brasileira de Pesquisa Agropecuária). This is the first baculovirus genome completely sequenced that was isolated from a member of the butterfly family HesperIIDae and was tentatively named *Urbanus proteus* nucleopolyhedrovirus (UrprNPV). We found that UrprNPV is an alphabaculovirus with the smallest genome found so far and that it is closely related to *Adoxophyes honmai* nucleopolyhedrovirus and *Adoxophyes orana* nucleopolyhedrovirus although with a high genetic diversity.



## 2. Material and methods

### 2.1. Virus purification.

Dead larvae of *Urbanus proteus* with symptoms of baculovirus infection were collected in soybean crops during 1984/85 season in Southern Brazil (the exactly location is missed). The cadavers were sent to the Embrapa and kept in a freezer for further analysis.

### 2.2. Host confirmation, viral genomic DNA extraction, and amplification.

One hundred  $\mu\text{l}$  of the OB-containing suspension ( $10^5$  OBs/mL of ddH<sub>2</sub>O) were heated for 20 min at 95°C, placed on ice for 5 min and treated with RQ1 RNase-Free DNase (Promega). The suspension was washed three times with SDS 0.5% and once with NaCl 0.5M by centrifuging ( $7000 \times g$  for 10 min) and resuspending with equal volumes. The last resulting pellet was resuspended in ddH<sub>2</sub>O and dissolved in alkaline solution to extract DNA as described elsewhere (O'Reilly et al., 1992). The DNA pellet was dissolved in 10  $\mu\text{l}$  of sterile ddH<sub>2</sub>O at 50°C for 1 h and directly subjected to a rolling cycle amplification (RCA) reaction using the phi29 DNA polymerase and a random 3' thiophosphate-protected hexamer primer according to the manufacturer's protocols (New England Biolabs). Both quantity and quality of the DNA were determined by electrophoresis on a 0.8% agarose gel (Sambrook and Russel, 2001), visualized, and photographed in AlphaImager® Mini (Alpha Innotech) (data not shown). For host identification at the species level, 10  $\mu\text{l}$  of the virus containing-extract were centrifuged to separate the OBs. We subjected the supernatant to phenol-chloroform DNA extraction. The extracted nucleic acid was used for PCR analysis to amplify a lepidopteran *cytochrome oxidase I (coi)* gene fragment according to a previously published protocol (Pinto et al., 2017). The PCR fragment was sequenced by standard Sanger Sequencing using the same set of primers and the resulting sequence compared to the Genbank database by BLASTN.

### 2.3. Genome sequencing, assembly, and annotation

The genomic DNA of the putative baculovirus was sequenced with the 454 Genome Sequencer (GS) Titanium at the Macrogen Company (South Korea). We obtained only one single contig by *de novo* assembly using the algorithm Geneious implemented in the Geneious 9.0.5 (Kearse et al., 2012). No other contig resembling baculovirus genomes were found. The open reading frames (ORFs) that started with a methionine codon (ATG) and encoded polypeptides of at least 50 amino acids were identified with the Geneious and annotated using

BLASTX (Altschul et al., 1997) and HMMER (Eddy and Wheeler, 2007). The genomic DNA sequence was submitted to the GenBank under the accession number KR011717.

#### **2.4. Phylogenetic analysis**

To perform the phylogenetic analysis, the MAFFT method (Katoh et al., 2002) was used to align individually the 38 baculovirus core genes from several baculovirus genomes (we selected only the unique species and avoid isolates, Table S1). A maximum likelihood tree was inferred using the Fast-Tree method (Stamatakis et al., 2008) and a Shimodaira-Hasegawa-like test for branch support (Anisimova et al., 2011). To perform that, we used the concatenated alignments of the 38 core genes. To construct the Fig. 6, we calculated for each core gene the best evolutionary model using ProtTest 2.4 (Abascal et al., 2005) and performed both the RaxML (Liu et al., 2011) and the PhyML (Guindon et al., 2010) analysis. The best evolutionary model for all the core genes were JTT+G+I+F. We did not find any substantial difference regarding the RaxML and the PhyML that could impair our analyses based on the FastTree method.

#### **2.5. Genome comparison**

The complete genome of UrprNPV was compared to other alphabaculovirus genomes through the construction of syntenic maps using the progressive Mauve algorithm implemented in the Genious 9.0.5 software. The gene parity plots were constructed based on the BLASTX results for each gene and a specific genome query (AdorNPV and, AcMNPV). To understand the reduction in the size of the UrprNPV genome, we re-annotated all the alphabaculovirus genomes according to the same criteria used for the UrprNPV. We analyzed the size of the intergenic space and the number of genes for the selected genomes (Table S1). A Venn Diagram was constructed to represent the number of ORFs shared among UrprNPV and the closest relatives.

### 3. Results and Discussion

#### 3.1. Sample analysis and host identification

We evaluated the insect extract labeled as *Urbanus proteus* nucleopolyhedrovirus (UrprNPV) from Embrapa virus collection library to check for the presence of baculoviral occlusion bodies (OBs) by optical microscopy and electron microscopy. We found OBs similar to those from the *Alphabaculovirus* with an irregular shape and size (Fig. S1A). The OBs were resistant to SDS and sensitive to alkaline pH. Moreover, we extracted the DNA of host cell fragments to identify the host at the species level based on the barcoding region of the mitochondrial *cytochrome c oxidase i (coi)* gene. We sequenced the amplified PCR fragment and confirmed the host as belonging to the species *Urbanus proteus* (Lepidoptera: Hesperidae).

#### 3.2. Genome sequencing analysis

We sequenced the viral DNA by the 454 Genome Sequencer (GS) FLX™ Titanium method (Macrogen Inc., Korea). The extracted DNA presented no small fragments and was found to be intact (Fig. S1B). Over 14,346 single-end reads were obtained after size and quality trimming (average size of  $734.2 \pm 227.9$  nt) and used for de novo assemble. We mapped 13,996 reads in one single circular contig of 105,555 bp long with a mean coverage of  $98 \pm 29.9$  X. The complete genome of UrprNPV (Genbank accession number: KR011717) is 105,555 bp in length, which makes it the smallest alphabaculovirus sequenced to date, followed by both *Adoxophyes orana* nucleopolyhedrovirus (AdorNPV) (111,724 bp) (Hilton and Winstanley, 2008) and *Maruca vitrata* multiple nucleopolyhedrovirus (MaviMNPV) (111,953 bp) (Chen et al., 2008). The novel virus genome presents a G+C content of 34.7% and is among the species with the lowest G+C content. The alphabaculovirus *Apocheima cinerarium* nucleopolyhedrovirus (ApciNPV) presents the lowest G+C content (33.4%) followed by AdorNPV (35%), AdhoNPV (35.6%), and *Lonomia oblique* multiple nucleopolyhedrovirus (LoobMNPV, 35.7%). The alphabaculoviruses with highest percentage of G+C content are *Lymantria dispar* multiple nucleopolyhedrovirus (LdMNPV, 57.5%), *Orgyia pseudotsugata* multiple nucleopolyhedrovirus (OpMNPV, 55.1%) and *Dasychra pudibunda* nucleopolyhedrovirus (DapuNPV, 54.4%) (Table S1). It is not clear the reason why the selective pressure drive for such a range of G+C content. For instance, high G+C content may protect the Herpes simplex virus 1 (HSV-1) genes from invasion by mobile genetic elements active during cell differentiation in the nervous system that could impair essential products for virus infection (Brown, 2007).

### 3.3. Gene content

We have annotated 119 putative ORFs encoding polypeptides of at least 50 amino acid residues in the UrprNPV genome (Table S2). Among the hypothetical ORFs annotated, ten were found to be unique in baculovirus (UrprNPV-ORF-5, UrprNPV-ORF-6, UrprNPV-ORF-18, UrprNPV-ORF-23, UrprNPV-ORF-34, UrprNPV-ORF-37, UrprNPV-ORF-40, UrprNPV-ORF-45, UrprNPV-ORF-63, and UrprNPV-ORF-118), *i.e.* not found in any other baculovirus genomes (Table S3). Only one hypothetical gene (UrprNPV-ORF-5) did not show any hit with no other gene from Genbank and no domains was found either by the HHpred. The 38 currently defined baculovirus core genes were found, including the new identified *ac110* (Javed et al., 2017). It also contains 25 of the 26 ORFs identified by Garavaglia et al. (2012) as present in genomes of alpha- and betabaculoviruses. The UrprNPV is missing only the *ac64* (*gp37*) as previously observed for other baculovirus genomes (Ardisson-Araujo et al., 2014; Harrison et al., 2017a).

### 3.4. UrprNPV phylogeny and gene content comparison

We investigated the evolutionary relationship of UrprNPV to other baculoviruses. UrprNPV belongs to the genus *Alphabaculovirus*. The closest relatives were *Adoxophyes honmai* nucleopolyhedrovirus (AdhoNPV) and *Adoxophyes orana* nucleopolyhedrovirus (AdorNPV), sharing a common hypothetical ancestor (Fig. 1). The average nucleotide identities based on the concatenated core gene pairwise alignment of UrprNPV and both AdhoNPV and AdorNPV were 53.3% and 53.5% respectively, followed by *Perigonia lusca* single nucleopolyhedrovirus (PeluSNPV) with 53.2%. The most distant alphabaculoviruses were found to be *Orgyia pseudotsugata* multiple nucleopolyhedrovirus (OpMNPV) with 45.4%, *Philosamia cynthia ricini* nucleopolyhedrovirus (PhcyNPV) 45.6%, and *Antheraea pernyi* nucleopolyhedrovirus (AnpeNPV) 45.8% (Table S1). The branch length separating this virus from its closest relatives is in a range that is comparable to the branch lengths separating viruses in other recognized species. Moreover, the *lef-8*, *lef-9*, and *polh* pairwise distances of the viral sequences of UrprNPV to other alphabaculoviruses for both single locus and concatenated alignment are well in excess of 0.05 substitutions/site, fulfilling the criteria for a novel baculovirus species (Jehle et al., 2006a).

We also performed a gene content comparison among UrprNPV and its closest relatives (AdorNPV and AdhoNPV), and the alphabaculovirus that is found in a basal position in the phylogenetic tree, *Operophtera brumata* nucleopolyhedrovirus (OpbuNPV) (Harrison et al.,

2017a), and we plotted the result in a Venn Diagram (Fig. 2A). A total of 200 different genes were found considering all the four species. Only 79 genes were shared among the four species. Twenty-one ORFs were found only in UrprNPV with 15 being hypothetical proteins (10 out of 15 were unique in baculovirus), *three baculovirus repeat ORFs (bro)*; one *cg30*, one *ac111-like*, and one *p26b* (a common and ancestral duplication of the p26 gene (Craveiro et al., 2015)). Sixteen ORFs were shared solely by UrprNPV and both AdorNPV and AdhoNPV when we compared the four species (among them, six hypothetical proteins, *hoar*, *p26a*, *bjdp*, *iap-2*, *bro*, *ac17-like*, *ac19-like*, *parg*, *gp16*, and *pkip*). The genus Adoxophyes, from where AdorNPV and AdhoNPV were isolated, belongs to the family Tortricidae. Conversely, there is no baculovirus described infecting the Hesperidae so far. Both Hesperidae and Tortricidae belong to the same Clade called by Ditrysia whose characteristic is the presence of two distinct sexual openings, one for mating and the other one for laying eggs (Minet, 1991). These baculovirus species belong likely to a novel and distinct lineage inside alphabaculovirus.

OphuNPV was found to be related to the ancestor of all alphabaculoviruses (Harrison et al., 2017a). The regular tree topology based on the 38 baculovirus core genes clusters alphabaculovirus into two monophyletic groups, called by group I and II. However, this is not consistent anymore to the current topology observed in this work and other recent studies (Harrison et al., 2017a, b). Alphabaculovirus group II was found to be polyphyletic, besides alphabaculovirus group I remains monophyletic. Therefore, the terminology alphabaculovirus group II is not used to classify UrprNPV. Importantly, the absence of UrprNPV genes from a core gene-based baculovirus phylogeny did not change the polyphyletic topology of alphabaculovirus group II (Harrison et al., 2017b).

The OphuNPV-ORF61 codes for an *iap-3* (*inhibitor of apoptosis protein- 3*) homolog that is the unique gene shared only between UrprNPV and OphuNPV. Interestingly, this ORF is not the unique *iap-3* homolog present in this genome. The OphuNPV genome presents the peculiarity of harboring three copies of the *iap-3* gene including ORF-61, OphuNPV-ORF-24, OphuNPV-ORF-26. Multiple copies of the same *iap* gene is an uncommon feature among baculoviruses (Clem, 2015). For instance, the *iap-2* gene was found duplicated in the alphabaculovirus *Hemileuca* sp. nucleopolyhedrovirus (Rohrmann, 2013a, b). Both the OphuNPV and the UrprNPV *iap-3* genes were found to be related to insect genes (Harrison et al., 2017a). The prediction of conserved domains in the deduced amino acid sequence found two BIR (Baculovirus Inhibitor of apoptosis protein Repeat) domains, which is involved in protein-protein interactions at the N-terminal region and one RING (Really Interesting New Gene) domain at the C-terminal portion, which is related to the ubiquitination pathway (Fig.

2B) (Hinds et al., 1999). This is consistent to the hypothesis of Harrison et al. (2017a) that places this *iap-3* gene as a possible product of a horizontal gene transfer from insects. The UrprNPV genome, differently from OpbuNPV presents also an *iap-2* gene. Both *iap-2* (UrprNPV-ORF50) and *iap-3* (UrprNPVORF39) genes, which are usually present in the genomes of alphabaculoviruses (not the monophyletic alpha-I), are involved in the antiapoptotic response induced by virus infection.

### 3.5. The *bro* genes

The UrprNPV genome codes for seven *bro* (baculovirus repeat ORF) genes. Homologs of UrprNPV *bro-a* are not present in the genome of the UrprNPV closest relatives AdorNPV and AdhoNPV and presents high identity to a *Heliothis virescens* ascovirus *bro* gene (Table S3). In order to understand this identity, we collected all the *bro-a*-related genes by BLASTX and performed a phylogenetic analysis (Fig. 3A). Homologs to the UrprNPV *bro-a* were also found in the genome of some alphabaculoviruses and a single betabaculovirus, that clustered together in the phylogenetic tree with the hypothetical ancestor related to ascovirus. This is suggestive of a horizontal gene transfer from ascovirus to alphabaculovirus and from alphabaculovirus to betabaculovirus, which may indicate a coinfection event of the same host by different insect viruses. However, the UrprNPV *bro-a* itself was directly clustered to a branch formed only by ascovirus genes, depicting a second event of gene acquisition by the ancestor of UrprNPV from ascovirus during the virus evolution (Fig. 3A). We also analyzed the *bro-b*, *bro-c*, and *bro-d* that are in tandem in the UrprNPV genome, near the *bro-a*. We found that these three *bro* genes seem to be a split of an ancestor *bro* gene related to the *bro-a* (ORF-60) of HaNPV-G4. AdorNPV and AdhoNPV present in their genome the same homolog with an N-terminal deletion, that corresponds to the location of both the UrprNPV *bro-c* and *bro-d* (Fig. 3A). Homologs of *bro* gene are present in all alpha, beta and, deltabaculovirus genomes but are not in gammabaculovirus (Rohrmann, 2013a,b). Although there is only one copy of the *bro* gene in AcMNPV, the number of copies may vary in different baculoviruses (Harrison and Lynn, 2007). Duplication/loss events in *bro* is rather common in baculovirus and ascovirus, ranging even in different isolates from the same virus, as seen in BmNPV (Ferber et al., 2001; Zhou et al., 2012). In a similar way to baculovirus, ascoviruses are large dsDNA viruses infectious to the larval stage of some insects (Federici, 1983). These viruses have ovoid, bacilliform or allantoid virions with 130 nm in diameter and 200 to 400 nm in length (Bigot et al., 2011). The ascovirus genome has 150–190 kbp and this variation in size is due to differences in repeated structural elements and the variable copy numbers of *bro* gene that can range from

3 to 23 (Bigot et al., 2000). The BRO proteins can be present in both the cytoplasm and nuclei during virus infection (Kang et al., 1999; Gong et al., 2003). The specific functions of *bro* genes and their protein products are still unknown, besides of presenting some plasticity for baculovirus evolution. The function is not clear indeed and may present nucleic acid and nucleosome association capabilities (Zemskov et al., 2000), single-stranded DNA (ssDNA) binding motif (Zemskov et al., 2000) or even being related to the virion structure (Wang et al., 2010). Some specific *bro* genes seem to be essential for the virus replication (Kang et al., 1999).

### 3.6. The UrpvNPV genomic structure

Several inversions, deletions, and insertions were found in the UrprNPV by gene parity plot compared to the genomes of AdorNPV (Fig. 4A), and the type-species of baculovirus AcMNPV-C6 (Fig. 4B). In general aspects, the gene order was not strictly conserved among the closely related species and the AcMNPV-C6. However, a genomic block containing most of the UrprNPV core genes (Fig. 4 A and 4B, in pink dots) were syntenic when compared to the other viruses. We also performed a whole genome comparison by the Mauve analysis between the UrprNPV and AdorNPV genomes to understand possible genomic block evolution. We found twelve Locally Collinear Blocks (LCB) (Fig. 4C). The blocks were well conserved between AdhoNPV and AdorNPV (not shown), but when compared with the UrprNPV genome, several losses occurred in almost all blocks, except for the LCB2 (composed by two genes, *per os infectivity factor-0 – pif-0/p74* and the *major early transcription gene 53 – me53*) and the LCB8 (composed by two hypothetical proteins, the *late expression factor-2 - lef-2*, the *viral capsid protein - p24* and the *polyhedron envelope protein - pep*). The LCB6 contained the region with the highest content of core genes and seemed to be strictly collinear across the virus species (Garavaglia et al., 2012).

### 3.7. The UrprNPV genome size

UrprNPV presents the smallest genome among all alphabaculovirus sequenced to date with 105,555 bp. In order to understand this short genomic size, we re-annotated several selected completely sequenced alphabaculovirus genomes following the same criteria. Then, we compared in base pairs (bp) both the total size of intergenic spaces correlated to the genome size for each alphabaculovirus genome (Fig. 5A). As expected, the larger the genome, the more intergenic spaces were found. *Apocheima cinerarium* nucleopolyhedrovirus (ApciNPV), *Clanis bilineata* nucleopolyhedrovirus (ClbiNPV), *Spodoptera litura* nucleopolyhedrovirus II

(SpliNPV-II), *Orgyia leucostigma* nucleopolyhedrovirus isolate CFS-77 (OrleNPV), and *Leucania separata* nuclear polyhedrovirus strain AH1 (LeseNPV) were outliers with small genome size and larger intergenic space. The UrprNPV genome was close to the trend line. Interestingly, no homologous regions (hrs) were found in the UrprNPV genome. Hrs are very common repeat sequences in baculoviruses. They commonly present an imperfect palindromic nucleus and can be distributed into the genomes arranged *in tandem* or as a single repeating nucleus (van Oers and Vlak, 2007). In a second attempt to understand the UrprNPV genome-shortening, we analyzed the correlation of both number of ORFs and genome size (Fig. 5B). We found, as expected, that the higher the genome, the more ORFs were present. UrprNPV remained also close to the trend line with the smaller number of annotated genes (119 genes) followed by both AdorNPV (with 123) and *Maruca vitrata* multiple nucleopolyhedrovirus (MaviNPV, with 126). As expected, all the alphabaculovirus genomes remained near the trend line with no exceptions. Moreover, when we looked at the average size of baculovirus genes for all alphabaculoviruses, we found no difference compared to that of UrprNPV (data not shown). UrprNPV presents the smallest alphabaculovirus genome due to both its smaller intergenic spaces and its smaller number of auxiliary genes. The essential genes described in alphabaculovirus are present. Interestingly, we found that 89% of the UrprNPV genome are composed by genes shared with other baculoviruses whereas 3% and 8% are composed by unique genes and intergenic spaces, respectively. Similar percentages were found in the genome of the closest relatives AdorNPV and AdhoNPV. Conversely, OpbuNPV presents higher percentage of unique genes (12% of the genome) and smaller percentage of both shared genes (83%) and intergenic spaces (5%), in comparison to the novel described genome (Fig. S2).

However, it was not clear whether the ancestor lineage of UrprNPV decreased in genome size was due to independent losses or no acquisition of auxiliary genes and intergenic spaces during evolution. In addition, observing the BLASTX results for each of the 119 genes, we found that most of the genes presented identity to a different baculovirus species (Table S3) other than AdorNPV and AdhoNPV, indicating a broad genetic diversity. This fact is evidenced in the average percentage of identity using the alignment of the concatenated core genes (i.e. about 50%), as well as in the phylogenetic tree (Fig. 1), as a long branch in relation to its closest relatives.

### **3.8. Genetic diversity and the use of individual core gene for the UrprNPV phylogeny**



We speculate that UrprNPV could be potentially a multiple recombinant or a novel lineage with restricted amounts of representatives with fully sequenced genomes. Therefore, we carried out a reconstruction analysis based on each individual core gene to support the phylogeny shown in Fig. 1. We analyzed all the 38 nodes generated from that phylogeny and scored in Fig. 6A each of these nodes from 0 to 38 based on the 38 individual core genes (the individual trees are summarized in the Fig. S3, highlighting the position of UrprNPV). The genes that most reconstructed the nodes were *vlf-1* (28/38), *p47* (28/38), *helicase* (27/38), *odv-ec43* (27/38), *p40* (27/38), and *49k* (27/38). The gene *p6.9* reconstructed 14 nodes (14/38) followed by *pif-1* and *pif-2* genes (16/38). In Fig. 1, UrprNPV was found to share a common ancestor with AdorNPV and AdhoNPV. However, this relationship was supported only by 13 core genes including *vlf-1* (*very late factor-1*), *vp1054*, *38k*, *ak-exo* (*alkaline exonuclease*), *desmoplakin*, *helicase*, *dna-pol* (*dna polymerase*), *lef-4* (*late expression factor-4*), *lef-8*, *lef-9*, *p33*, *p40*, and *pif-3* (*'per os' infectivity factor-3*). UrprNPV varied greatly its position for the other gene reconstructions (Fig. S3). The genome of OpbuNPV is 13,5 kbp longer than UrprNPV besides presenting a very small intergenic space compared to the trend line observed in the Fig. 5A. UrprNPV shares most of the genes with OpbuNPV, albeit the opposite is not true; OpbuNPV presents 53 genes that are not present in UrprNPV. Indeed, we speculate that UrprNPV could have lost genes compared to the last common ancestor of alphabaculovirus and retained the feature of having a short intergenic space. Importantly, OpbuNPV presents several HRs that are completely lacked by the UrprNPV genome. Overall, both the diverse topology of UrprNPV and the distance to OpbuNPV point together to the hypothesis of this distinct hesperiidisolated virus as a multiple recombinant virus (Rohrmann, 2013a,b). Therefore, UrprNPV is likely forming a novel lineage into the *Alphabaculovirus* with the autapomorphic feature of having a compacted genome. As new related species are sequenced, the topology will be clarified.

To check whether the gene size is influencing the phylogeny result biasing the tree topology of the concatenated core gene alignment-based phylogeny reconstruction, we carried out a correlation analysis between the average gene size for all baculovirus (alpha, beta, gamma and delta) and the number of node reconstructions in UrprNPV-related alphabaculoviruses based on the concatenated core gene alignments (Fig. 1). We found no positive correlation for most of the genes. Most of the core genes ranges from 250 to 1,500 bp and are grouped into discrete populations. Some genes are small in size and strongly (more than 25/38) reconstruct the concatenated-based phylogeny (Fig. 6B) (e.g. as *vlf-1*, *p47*, *49k*, *odv-ec43*, *p40*, *lef-4*, *odv-e27*, *gp41*, and *vp1054*). On the other hand, some genes are small and weakly reconstruct the

concatenated-based phylogeny (Fig. 6B). These nine genes individually or together could much better recount the concatenated core genes alignment-based phylogeny than do both *lef-8* and *lef-9*, that are traditionally used to solve the baculovirus phylogeny. Interestingly, *lef-8* and *lef-9* reconstruct only 23 nodes for UrprNPV-related alphabaculovirus out from 38 total nodes found in the concatenated gene alignment phylogeny. Both genes are components of the viral RNA polymerase important for the late gene expression (Rohrmann, 2013a,b) and used as a diversity marker to demarcate a new species into *Baculoviridae*, based on the ICTV criteria. In a worse mode, *pif-1* and 2 recount only 16 nodes out from 38 total nodes. Both genes are related envelope genes related to the oral infectivity of the ODV. Indeed, our findings suggest that *lef-8*, *lef-9*, *pif-2*, and *pif-1* are not the better genes to solve the baculovirus phylogeny in a similar result observed for the phylogeny based on the concatenated data set of core genes. The same gene tendency was observed for betabaculoviruses and alphabaculovirus group I (data not shown). Herniou et al. (2003) found that baculovirus singlegene alignments yield trees that are incongruent with each other based on 13 completely sequenced baculovirus genomes available at that time. This is suggestive of the fact that a single-gene-based tree may reflect the evolutionary history of that particular baculovirus gene, and requires some caution with respect to its interpretation.

#### 4. Conclusion

We found in the extract of dead *Urbanus proteus* larva a novel and distinct alphabaculovirus, tentatively named *Urbanus proteus* nucleopolyhedrovirus (UrprNPV). After complete genome sequencing, the virus was found to present the smallest genome so far among alphabaculovirus with 105,555 bp. As expected, the genome shortening was likely due to loss of both auxiliary genes and intergenic spaces. UrprNPV related itself to the *Adoxophyes*-infecting baculoviruses AdorNPV and AdhoNPV with high synteny but low nucleotide identity. Moreover, the genome contained ten unique ORFs, seven *bro* genes, and all the 38 baculovirus core genes. Few core genes were found to cluster UrprNPV with AdorNPV and AdhoNPV depicting a multiple recombinant virus or a distinct virus lineage with restricted number of representatives completely sequenced. Importantly, we found that *vlf-1*, *p47*, *49k*, *odv-ec43*, *p40*, *lef-4*, *odv-e27*, *gp41*, and *vp1054* may depict individually or together in a better way the phylogeny based on the concatenated data set of core genes. Overall, the increase in number of completely sequenced baculovirus points to a very exciting way to understand baculovirus and its evolution, including the fact that alphabaculovirus group II was recently shown to be a paraphyletic group and should be abandoned as a usual term in baculovirology.

## 5. Acknowledgments

We thank *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq), grant number 407908/2013-7 and *Fundação de Apoio à Pesquisa do Distrito Federal* (FAPDF) grant number 193.001.532/2016 to BMR for financial support.

## 6. References

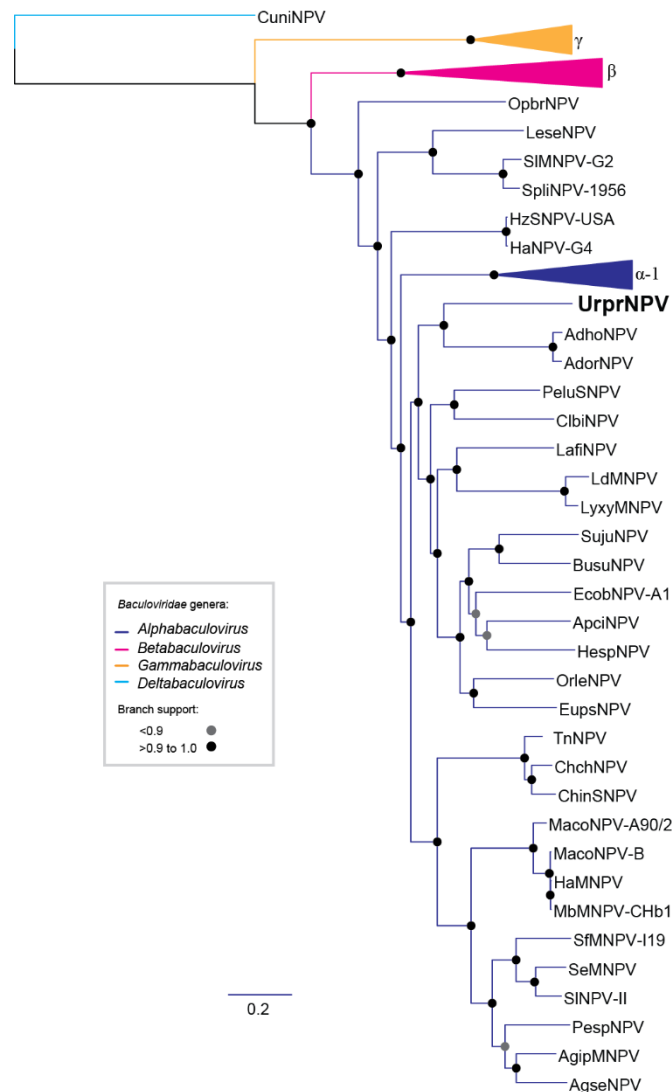
- Abascal, F., Zardoya, R., & Posada, D. (2005). ProtTest: selection of best-fit models of protein evolution. *Bioinformatics*, *21*(9), 2104-2105.
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research*, *25*(17), 3389-3402.
- Anisimova, M., Gil, M., Dufayard, J. F., Dessimoz, C., & Gascuel, O. (2011). Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. *Systematic biology*, *60*(5), 685-699.
- Ardisson-Araújo, D. M. P., de Melo, F. L., de Souza Andrade, M., Sihler, W., Bão, S. N., Ribeiro, B. M., & de Souza, M. L. (2014). Genome sequence of Erinnyis ello granulovirus (ErelGV), a natural cassava hornworm pesticide and the first sequenced sphingid-infecting betabaculovirus. *BMC genomics*, *15*(1), 856.
- Bigot, Y., Stasiak, K., Rouleux-Bonnin, F., & Federici, B. A. (2000). Characterization of repetitive DNA regions and methylated DNA in ascovirus genomes. *Journal of General Virology*, *81*(12), 3073-3082.
- Bigot, Y., Asgari, S., Bideshi, D.K., Cheng, X., Federici, B.A. & Renault, S. (2011). Family *Ascoviridae*. In: King, A. M., Lefkowitz, E., Adams, M. J., & Carstens, E. B. (Eds.). (2011). *Virus taxonomy: ninth report of the International Committee on Taxonomy of Viruses*. Elsevier.
- Bridges, C. A. (1993). Catalogue of the Family-Group, Genus-Group and Species-Group Names of the HesperIIDae (Lepidoptera) of the World. Published by Author, *Urbana, IL*.
- Brown, J. C. (2007). High G+ C content of herpes simplex virus DNA: proposed role in protection against retrotransposon insertion. *The open biochemistry journal*, *1*, 33.
- Chen, Y. R., Wu, C. Y., Lee, S. T., Wu, Y. J., Lo, C. F., Tsai, M. F., & Wang, C. H. (2008). Genomic and host range studies of Maruca vitrata nucleopolyhedrovirus. *Journal of general virology*, *89*(9), 2315-2330.
- Clem, R. J., & Passarelli, A. L. (2013). Baculoviruses: sophisticated pathogens of insects. *PLoS pathogens*, *9*(11), e1003729.
- Clem, R. J. (2015, March). Viral IAPs, then and now. In *Seminars in cell & developmental biology* (Vol. 39, pp. 72-79). Academic Press.
- Craveiro, S. R., Inglis, P. W., Togawa, R. C., Grynberg, P., Melo, F. L., Ribeiro, Z. M. A., ... & Castro, M. E. B. (2015). The genome sequence of Pseudoplusia includens single nucleopolyhedrovirus and an analysis of p26 gene evolution in the baculoviruses. *BMC genomics*, *16*(1), 127.

- Costa Lima, A. (1945). Insetos do Brasil–Lepidópteros–Série Didática nº 7–5º tomo. *Rio de Janeiro, Escola Nacional de Agronomia*.
- Eddy, S. R., & Wheeler, T. J. (2007). HMMER-biosequence analysis using profile hidden Markov models. URL <http://hmmer.janelia.org>.
- Federici, B. A. (1983). Enveloped double-stranded DNA insect virus with novel structure and cytopathology. *Proceedings of the National Academy of Sciences*, 80(24), 7664-7668.
- Ferber, M. L., Argaud, O., Croizier, L., & Croizier, G. (2001). Diversity, distribution, and mobility of bro gene sequences in *Bombyx mori* nucleopolyhedrovirus. *Virus Genes*, 22(3), 247-254.
- Garavaglia, M. J., Miele, S. A. B., Iserte, J. A., Belaich, M. N., & Ghiringhelli, P. D. (2012). The ac53, ac78, ac101, and ac103 genes are newly discovered core genes in the family Baculoviridae. *Journal of virology*, 86(22), 12069-12079.
- Gong, Y., Li, Z., Wang, L., Pan, L., Yang, K., & Pang, Y. (2003). Characterization of bro-b gene of *Spodoptera litura* multicapsid nucleopolyhedrovirus. *Virus genes*, 27(2), 115-123.
- Greene, G. L. (1971). Economic damage levels of bean leaf roller populations on snap beans 1, 2. *Journal of Economic Entomology*, 64(3), 673-674.
- Guindon S., Dufayard J.F., Lefort V., Anisimova M., Hordijk W. & Gascuel O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol*, 59(3), 307-321.
- Herniou, E. A., Olszewski, J. A., Cory, J. S., & O'Reilly, D. R. (2003). The genome sequence and evolution of baculoviruses. *Annual review of entomology*, 48(1), 211-234.
- Harrison, R. L., & Lynn, D. E. (2007). Genomic sequence analysis of a nucleopolyhedrovirus isolated from the diamondback moth, *Plutella xylostella*. *Virus Genes*, 35(3), 857.
- Harrison, R. L., Rowley, D. L., Mowery, J. D., Bauchan, G. R., & Burand, J. P. (2017). The *Operophtera brumata* Nucleopolyhedrovirus (OpbuNPV) Represents an Early, Divergent Lineage within Genus *Alphabaculovirus*. *Viruses*, 9(10), 307a.
- Harrison, R. L., Mowery, J. D., Rowley, D. L., Bauchan, G. R., Theilmann, D. A., Rohrmann, G. F., & Erlandson, M. A. (2017). The complete genome sequence of a third distinct baculovirus isolated from the true armyworm, *Mythimna unipuncta*, contains two copies of the lef-7 gene. *Virus genes*, 1-14b.
- Hilton, S., & Winstanley, D. (2008). Genomic sequence and biological characterization of a nucleopolyhedrovirus isolated from the summer fruit tortrix, *Adoxophyes orana*. *Journal of general virology*, 89(11), 2898-2908.
- Hinds, M. G., Norton, R. S., Vaux, D. L., & Day, C. L. (1999). Solution structure of a baculoviral inhibitor of apoptosis (IAP) repeat. *Nature Structural & Molecular Biology*, 6(7), 648-651.
- Jackai, L. E. N., & Daoust, R. A. (1986). Insect pests of cowpeas. *Annual review of entomology*, 31(1), 95-119.
- Javed, M. A., Biswas, S., Willis, L. G., Harris, S., Pritchard, C., van Oers, M. M., ... & Theilmann, D. A. (2017). *Autographa californica multiple nucleopolyhedrovirus* AC83 is a per os infectivity factor (PIF) protein required for occlusion-derived virus (ODV) and budded virus nucleocapsid assembly as well as assembly of the PIF complex in ODV envelopes. *Journal of virology*, 91(5), e02115-16.
- Jehle, J. A., Lange, M., Wang, H., Hu, Z., Wang, Y., & Hauschild, R. (2006). Molecular identification and phylogenetic analysis of baculoviruses from Lepidoptera. *Virology*, 346(1), 180-193a.
- Jehle, J. A., Blissard, G. W., Bonning, B. C., Cory, J. S., Herniou, E. A., Rohrmann, G. F., ... & Vlak, J. M. (2006). On the classification and nomenclature of baculoviruses: a proposal for revision. *Archives of virology*, 151(7), 1257-1266b.

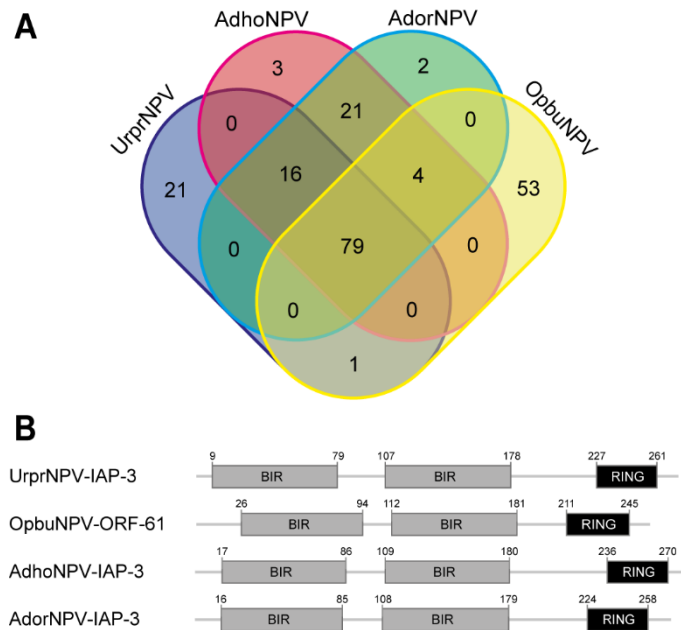
- Kang, W., Suzuki, M., Zemskov, E., Okano, K., & Maeda, S. (1999). Characterization of Baculovirus Repeated Open Reading Frames (bro) in *Bombyx mori* nucleopolyhedrovirus. *Journal of virology*, 73(12), 10339-10345.
- Katoh, K., Misawa, K., Kuma, K. I., & Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic acids research*, 30(14), 3059-3066.
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., & Thierer, T. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28(12), 1647-1649.
- Liu K, Linder CR, Warnow T. (2011). RAxML and FastTree: comparing two methods for large-scale maximum likelihood phylogeny estimation. *PLoS One*, 6(11):e27731.
- Lucarotti, C. J., Whittome-Waygood, B. H., Lapointe, R., Morin, B., & Levin, D. B. (2012). Pathology of a Gammabaculovirus in its natural balsam fir sawfly (*Neodiprion abietis*) host. *Psyche: A Journal of Entomology*, 2012.
- Minet, J. (1991). Tentative reconstruction of the ditrysian phylogeny (Lepidoptera: Glossata). *Insect Systematics & Evolution*, 22(1), 69-95
- Nava, D. E., & Parra, J. R. P. (2002). Development and soybean leaf consumption by *Urbanus proteus proteus* (L.). *Scientia Agricola*, 59(4), 661-663.
- O'Reilly, D. R., Miller, L. K., & Luckow, V. A. (1992). *Baculovirus expression vectors: a laboratory manual*. Oxford University Press on Demand.
- Pinto, F. A., Mattos, M. V., Silva, F. W., Rocha, S. L., & Elliot, S. L. (2017). The Spread of *Helicoverpa armigera* (Lepidoptera: Noctuidae) and Coexistence with *Helicoverpa zea* in Southeastern Brazil. *Insects*, 8(3), 87.
- Rohrmann, G. F. (2013). *Baculovirus molecular biology*. Third Edition [Internet] edn. Bethesda (MD): National Center for Biotechnology Information (US). Available from: [https://www.ncbi.nlm.nih.gov/books/NBK114593/pdf/Bookshelf\\_NBK114593.pdf](https://www.ncbi.nlm.nih.gov/books/NBK114593/pdf/Bookshelf_NBK114593.pdf)
- Rohrmann, G. F., Erlandson, M. A., & Theilmann, D. A. (2013). The genome of a baculovirus isolated from *Hemileuca* sp. encodes a serpin ortholog. *Virus genes*, 47(2), 357-364.
- Sambrook, J., Russel, D. W. (2001). *Molecular Cloning: a laboratory manual*. Cold Spring Harbor, New York.
- Sosa-Gómez, D. R. (2017). Microbial Control of Soybean Pest Insects and Mites. In *Microbial Control of Insect and Mite Pests* (pp. 199-208). From Theory to Practice. Lacey, Lawrence (Ed.). Academic Press Elsevier. Chapter 13. ISBN 978-0-121-803527-6. 461pp.
- Stamatakis, A., Hoover, P., & Rougemont, J. (2008). A rapid bootstrap algorithm for the RAxML web servers. *Systematic biology*, 57(5), 758-771.
- Sun, X. L., & Peng, H. Y. (2007). Recent advances in biological control of pest insects by using viruses in China. *Virologica Sinica*, 22(2), 158-162.
- Szewczyk, B., Hoyos-Carvajal, L., Paluszek, M., Skrzecz, I., & De Souza, M. L. (2006). Baculoviruses—re-emerging biopesticides. *Biotechnology advances*, 24(2), 143-160.
- Temerak, S. A., Boucias, D. G., & Whitcomb, W. H. (1984). A singly-embedded nuclear polyhedrosis virus and entomophagous insects associated with populations of the bean leafroller *Urbanus proteus* L.(Lepid., Hesperidae). *Journal of Applied Entomology*, 97(1-5), 187-191.
- van Dam, W., & Wilde, G. (1977). Biology of the Bean Leafroller *Urbanus proteus* (Lepidoptera: Hesperidae). *Journal of the Kansas Entomological Society*, 157-160.
- van Oers, M. M., & Vlak, J. M. (2007). Baculovirus genomics. *Current drug targets*, 8(10), 1051-1068.

- Wang, R., Deng, F., Hou, D., Zhao, Y., Guo, L., Wang, H., & Hu, Z. (2010). Proteomics of the *Autographa californica* nucleopolyhedrovirus budded virions. *Journal of virology*, *84*(14), 7233-7242.
- Warren, A. D., Ogawa, J. R., & Brower, A. V. (2008). Phylogenetic relationships of subfamilies and circumscription of tribes in the family Hesperiidae (Lepidoptera: Hesperioidea). *Cladistics*, *24*(5), 642-676.
- Zemskov, E. A., Kang, W., & Maeda, S. (2000). Evidence for nucleic acid binding ability and nucleosome association of *Bombyx mori* nucleopolyhedrovirus BRO proteins. *Journal of virology*, *74*(15), 6784-6789.
- Zhou, J. B., Li, X. Q., De-Eknamkul, W., Suraporn, S., & Xu, J. P. (2012). Identification of a new *Bombyx mori* nucleopolyhedrovirus and analysis of its bro gene family. *Virus genes*, *44*(3), 539-547.

## Figure captions

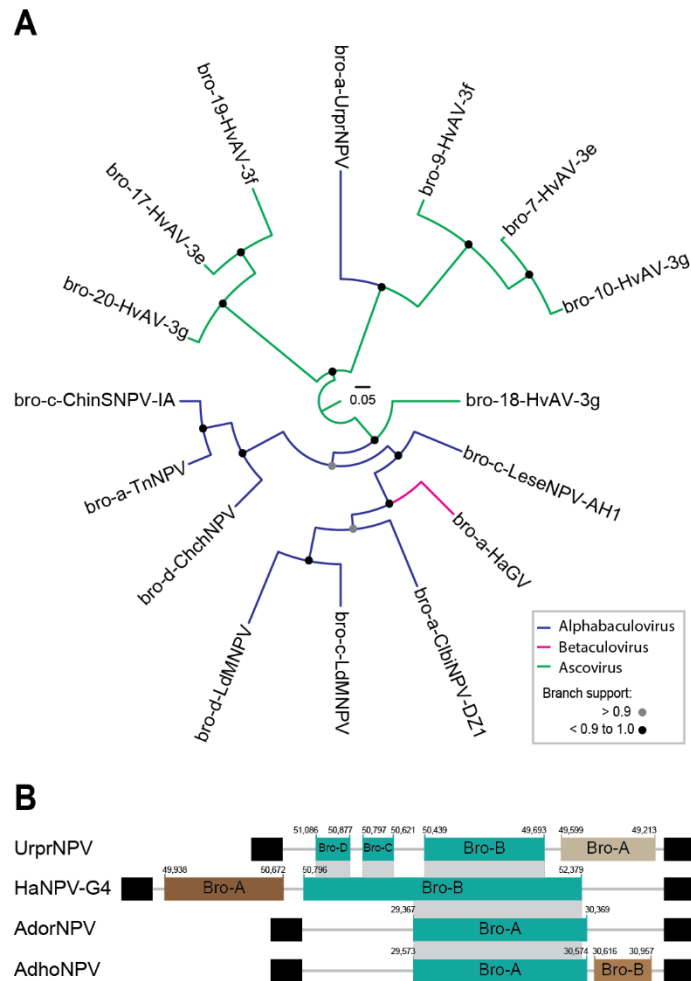


**Fig. 1.** UrprNPV is an alphabaculovirus. Maximum likelihood inference based on the concatenated nucleotide sequences of the 38 core genes from several selected baculovirus genomes (Table S1) using the FastTree method. The branch support was determined by a SH-like method (black and grey closed circles). 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. The genus *Alphabaculovirus* contained UrprNPV (boldface) that clustered together with both AdhoNPV and AdorNPV (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

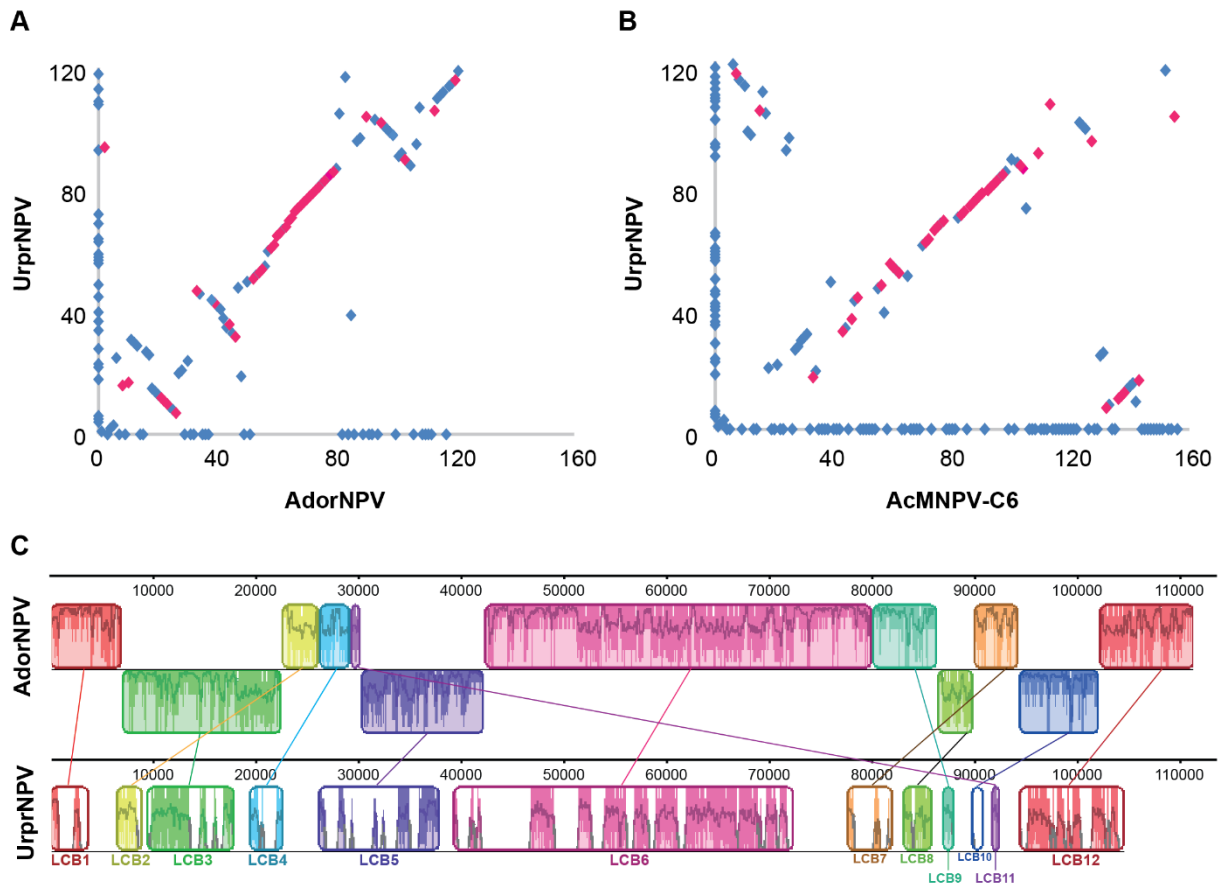


**Fig. 2.** The gene content among UrprNPV and its relatives and the genomic context of the *iap-3* of UrprNPV and OpbuNPV. (A) The whole gene content of four alphabaculoviruses (UrprNPV, AdhoNPV, AdorNPV and OpbuNPV) were compared by BLASTX to find homologs. A total of 200 genes were found: 79 were shared among all the four viruses, 16 among UrprNPV, AdhoNPV, and AdorNPV; one between UrprNPV and OpbuNPV (*iap-3*); and 21 were unique to the UrprNPV genome. (B) Comparison of the predicted amino acid sequence of the *iap-3* genes of UrprNPV and its relatives. We show the predicted domains found using HHpred and the encoded amino acid sequences for each gene and found both the BIR and the RING domains, characteristics of baculoviral IAP proteins. OpbuNPV-ORF-61 and UrprNPV-IAP-3 have two BIR domains and one RING domain. In the OpbuNPV-ORF-26, it appears that the C-terminal portion corresponding to the RING domain was lost.

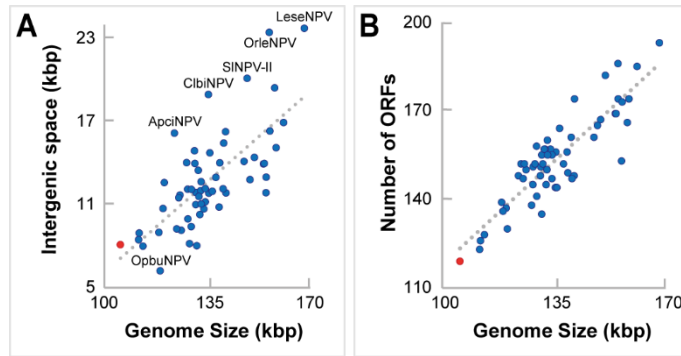




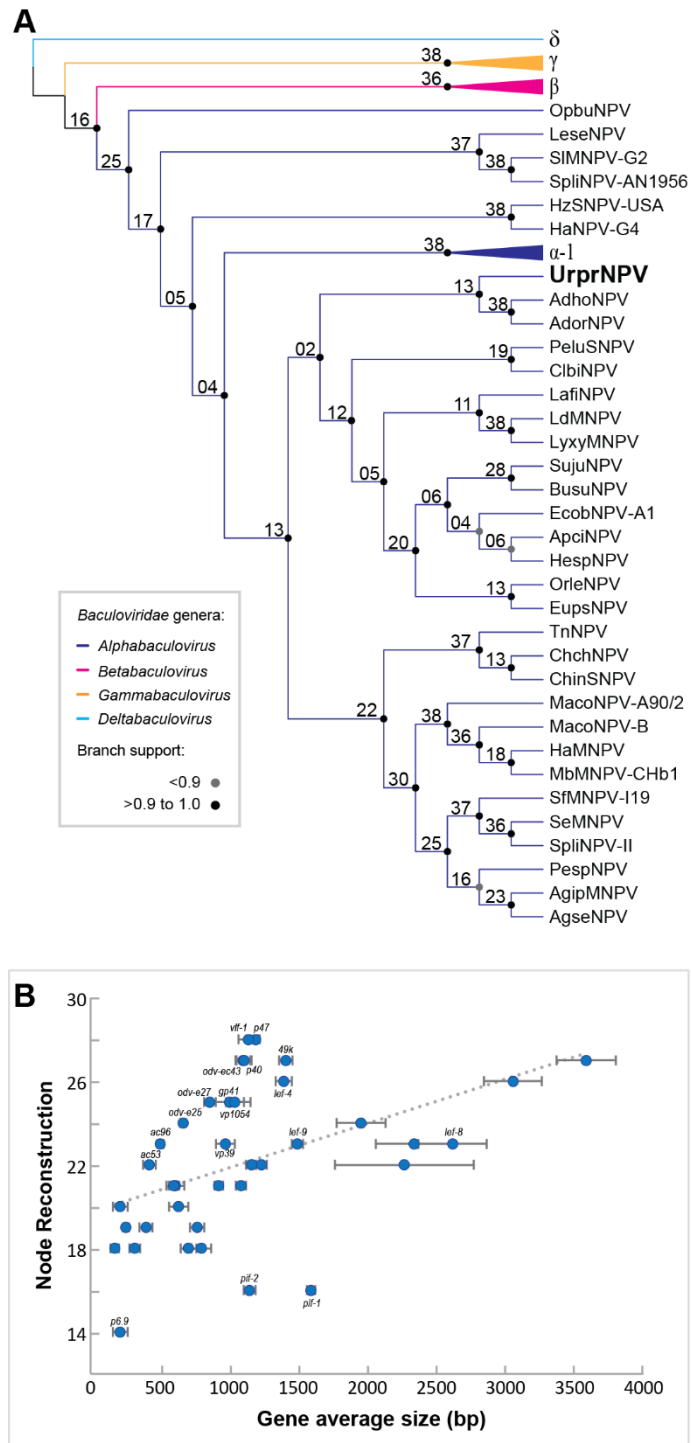
**Fig. 3.** The UrprNPV *bro* gene analyses. (A) Phylogenetic analysis of the UrprNPV *bro-a* gene and some homologs reveals an association with ascovirus genes. The tree was inferred by the FastTree and the branch support was determined by a SH-like method. (B) UrprNPV *bro-a* to *bro-d* locus is compared to other closely related organisms (UrprNPV, HaNPV-G4, AdorNPV, and AdhoNPV). An ancient HaNPV-G4-*bro-b*-related gene seemed to undergo several splits in the genome of UrprNPV, AdorNPV, and AdhoNPV. In the AdorNPV and AdhoNPV genomes, a small region at the amino-terminal portion is missed that corresponded to the *bro-c* and *bro-d* in the UrprNPV. The genome present three other *bro* genes (*e*, *f*, and *g*) in different loci.



**Fig. 4.** Gene parity plots and locally collinear blocks of the UrprNPV genome against AdorNPV and AcMNPV-C6. Plots show the gene content and position in the UrprNPV genome with those from (A) the AdorNPV and (B) the AcMNPV-C6 genomes. Each dot represents a homolog. The pink colored dots represent the 38 baculovirus core genes. (C) Genomic comparison of the UrprNPV genome and the AdorNPV. Twelve Locally Collinear Blocks (LCB) numbered from 1 to 12 were found. Same colors depict same LCBs across the genomes. Rearrangement can be seen between the analyzed genomes.



**Fig. 5.** Analysis of the baculovirus genome sizes across alphabaculoviruses and other features. (A) Comparison of the size of intergenic spaces and the total genomic size for all the re-annotated alphabaculovirus genomes. The six outliers are indicated (names above blue circles). (B) Comparison of the total number of ORFs and the total genomic size for all the reannotated alphabaculovirus genomes. There is a positive correlation among the genomes. UrprNPV is represented by the red closed circle. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).



**Fig. 6.** UrprNPV varies its relationship to other baculovirus based on the individual core gene analyses. (A) Most of the individual the FastTree-based phylogenies of each core gene did not reconstruct the UrprNPV phylogeny observed for the inference based on the concatenated dataset. Thirty eight nodes were found and scored from 0 to 38 depicting all the currently defined baculovirus core genes. Some branches were collapsed for clarity. (B) Correlation of the number of node reconstruction for each core gene and the core gene sizes for several

selected baculovirus genomes. We highlighted some outliers and the error bars represent the standard deviations for the gene size.

## 5 DISCUSSÃO

O extrato cedido pela viroteca da EMBRAPA nomeado como *Urbanus proteus* NPV, foi coletado nos anos 1980 no sul do Brasil e apresentava OBs em formato de poliedro característicos de alphabaculovírus. Os OBs foram visualizados em microscopia óptica e eletrônica de varredura. Para confirmação da espécie hospedeira, foi feito sequenciamento baseado numa região bem conservada do gene *citocromo c oxidase (coi)* de insetos vindo do DNA presente no extrato, confirmando assim que o extrato pertencia a espécie *U. proteus* (Lepidoptera: Hesperidae). Foi sequenciado pela primeira vez um baculovírus desta família de borboletas. O genoma viral foi sequenciado pelo método 454 de pirosequenciamento e como resultado foram obtidos 14.346 ‘reads’, com tamanhos entre  $734,2 \pm 227,9$  nucleotídeos e cobertura de  $98 \pm 29,9$  X. Os ‘reads’ foram mapeados e montados usando o método *de novo* ‘assembly’, sem um genoma de referência com algoritmo implementado no programa Geneious 9 (KEARSE et al., 2012). Obteve-se um único ‘contig’ circular de 105.555 pb e a sequência foi depositada no GenBank (KR011717). Esse método de montagem pode produzir genomas aberrantes, quando há presença de regiões repetidas, porém este não é o caso de UrprNPV, que não possui regiões homólogas, o que é incomum em baculovírus, salientando ainda mais a diferença deste vírus com os demais.

O genoma foi então anotado e apresentou 119 ORFs (do inglês ‘open read frames’); destas, nove eram únicas em baculovírus e apenas uma referência em nenhuma outra espécie do banco de dados mundial. Verificou-se também, a presença dos 38 genes compartilhados por todos os baculovírus até hoje sequenciados e com base nestes genes alinhados e concatenados foi inferida uma árvore filogenética de máxima verossimilhança usando o método Fast-Tree (STAMATAKIS et al., 2008). UrprNPV pertencia ao clado de *Alphabaculovirus*, mais proximamente relacionado às espécies *Adoxophyes honmai nucleopolyhedrovirus* (AdorNPV) e *Adoxophyes orana nucleopolyhedrovirus* (AdhoNPV), porém externo ao grupo e com um comprimento de ramo longo. Além disso, como característica marcante, UrprNPV apresentou o menor genoma entre os *Alphabaculovirus* até hoje sequenciados. Com relação a filogenia de *Baculoviridae*, a partir do genoma publicado por Harrison et al. (2017b), *Operophtera brumata nucleopolyhedrovirus* (OpbuNPV) foi encontrado como relacionado ao ancestral comum mais recente dos *Alphabaculovirus*, havendo assim uma mudança na filogenia atual, que anteriormente dividia alphabaculovírus em grupo I e II (JEHLE et al., 2006b). A partir da inserção deste genoma ancestral, o grupo II de *Alphabaculovirus* se tornou parafilético,

sugerindo assim que o termo Alphabaculovirus grupo II está incorreto e que deve ser evitado por baculovirologistas.

Para entender o porquê desta distância genética entre UrprNPV e os baculovírus proximamente relacionados AdorNPV e AdhoNPV, primeiramente comparou-se o conteúdo gênico entre eles e com o alphabaculovírus encontrado mais recentemente como basal ao clado, OpbuNPV (HARRISON et al., 2017b) e plotado em um Diagrama de Venn, foi verificado um total de 200 genes presentes nos quatro genomas, destes apenas 79 eram compartilhados entre eles, 16 entre UrprNPV e os vírus que infectam o gênero *Adoxophyes* e somente um entre UrprNPV e OpbuNPV, interessante que o gene compartilhado é *iap-3* (do inglês, ‘Inhibitor-of-Apoptosis-3’) um inibidor de apoptose celular característico de baculovírus, em contrapartida, ambos são relacionados com genes de inseto e não com de outros baculovírus, afirmando a hipótese de HARRISON et al. (2017b), que esse gene pode ter sido produto de uma transferência horizontal vinda de inseto. UrprNPV apresentou peculiaridades em relação ao gene *bro* (do inglês, ‘baculovirus repeat ORF’), que são comumente presentes em regiões homólogas de baculovírus, encontrou-se sete genes *bro* no genoma de UrprNPV, mas são muito diferentes entre si e não foram relacionados a regiões repetidas, isto já foi observado para vários outros genomas de baculovírus, tais como a espécie tipo de baculovírus *Autographa californica multiple nucleopolyhedrovirus* (AcMNPV), AdorNPV, AdhoNPV e OpbuNPV. Uma BRO foi mais proximamente relacionada com ascovírus, e uma árvore filogenética mostrou um possível evento de transferência horizontal de ascovírus (vírus de inseto) para alphabaculovírus e de alpha para betabaculovírus, podendo indicar um evento de co-infecção do mesmo hospedeiro por um vírus diferente. Além de um segundo evento de aquisição do gene pelo ancestral de UrprNPV a partir de ascovírus durante o evolução do vírus. Verificou-se também que três genes *bro* de UrprNPV parecem ter sofrido uma divisão do gene *bro* ancestral de *Helicoverpa armigera nucleopolyhedrovirus*- G4.

UrprNPV apresenta o menor genoma entre os *Alphabaculovirus* até hoje sequenciados. Diante disso, perguntou-se no presente trabalho se a linhagem ancestral que deu origem a UrprNPV diminuiu no tamanho do genoma por perdas independentes ou através da não-aquisição de genes auxiliares e espaços intergênicos durante a evolução. Verificou-se que este vírus possui uma ampla diversidade genética, visto que quando comparada cada umas das 119 ORFs de UrprNPV no banco de dados BLASTX, a maioria delas apresentava identidade com uma ampla variedade de baculovírus e não somente com os seus parentes mais proximamente relacionados, o que é geralmente mais comum. Desta forma, foi especulado se UrprNPV poderia ser uma linhagem gerada a partir da recombinação de múltiplas regiões de outros

baculovírus ancestrais (como um álbum de recortes) ou uma nova linhagem com quantidades restritas de representantes com genomas totalmente sequenciados. Para entender-se essa problemática, foi feita uma análise de reconstrução filogenética individual com cada um dos 38 genes compartilhados por todos os baculovírus, para verificar se havia congruência com a árvore filogenética inferida com estes genes alinhados e concatenados. Em trabalho anterior, Herniou et al. (2003), mostrou que o alinhamento de um único gene era incongruente comparado com os treze genomas de baculovírus publicados na época. Os genes que mais inferem a história do concatenado são *very late factor-1 (vlf-1)* e *p47*, seguidos por *helicase*, *occlusion-derived virion envelope capsid -43 (odv-ec43)*, *p40* e *49k (27/38)*. Por outro lado, os genes que menos reconstróem a história evolutiva baseada no concatenado foram *p6.9*, *per os infectivity factors-1 e 2 (pif-1)* e *(pif-2)*. Já, a relação filogenética de UrprNPV comparada com as dos parentes mais próximos AdorNPV e AdhoNPV, foi recontada por apenas 13 dos 38 genes core. UrprNPV está provavelmente formando uma nova linhagem em *Alphabaculovirus* com a característica automórfica de ter um genoma compacto. À medida que novas espécies relacionadas sejam sequenciadas, essa topologia será esclarecida.

Sequências parciais dos genes *late expression factor-8 e 9 (lef-8)*, *(lef-9)* e *poliedrina* são utilizados como critérios de demarcação de espécie, que consiste nas distâncias de nucleotídeos emparelhados estimada com o modelo de substituição Kimura-2-parameter, onde acima de 0,05 substituições/ sítio é considerado uma nova espécie (JEHLE et al., 2006a). O novo genoma de UrprNPV foi submetido a este critério e pôde-se verificar que as substituições variaram entre 0,193 a 0,340, mostrando que UrprNPV, além de ser considerado uma nova espécie de baculovírus, é um vírus realmente diferente comparado com os demais baculovírus reconhecidos como espécie pelo International Committee on Taxonomy of Viruses (ICTV).

Alguns pontos comprovam que o baculovírus UrprNPV é uma entidade biológica real e não um artefato das análises *in silico*. Primeiramente, o vírus foi isolado de vários indivíduos da espécie *U. proteus* que morreram com sintomas característicos de infecção letal por baculovírus. A EMBRAPA dispõem de mais três extratos de lagartas com sintomas de infecção por baculovírus, do início dos anos 90 em locais de coletas diferentes. Segundo, o vírus foi purificado e na sequência, realizou-se microscopia ótica e eletrônica para confirmar a presença de corpos de oclusão. O DNA viral foi purificado dos OBs e estava intacto. O genoma foi montado pelo método *de novo* ‘assembly’ e os 38 genes compartilhados por todos os baculovírus foram encontrados em UrprNPV. Além disso, UrprNPV não apresenta regiões homólogas nem outras regiões repetitivas, que tornariam difícil a montagem do genoma. Vários outros genomas de baculovírus publicados recentemente foram sequenciados de bibliotecas



virais antigas e os autores também não realizaram o experimento de infectividade ao hospedeiro, provavelmente devido à quantidade de amostra viral e ao acesso a colônias de laboratórios de insetos (ROHRMANN et al., 2013; WANG et al., 2016; HARRISON et al., 2017a,c; WANG et al., 2018).

## 6 CONCLUSÃO

O presente trabalho dedicou-se a descrever e caracterizar o genoma da nova espécie de baculovírus *Urbanus proteus nucleopolyhedrovirus* (UrprNPV), isolado da família de borboletas Hesperidae. Este vírus, apresentou todos os 38 genes compartilhados entre os baculovírus, entretanto sem regiões homólogas e mostrou-se peculiar em relação as outras espécies de *Alphabaculovirus* até hoje sequenciadas, possuindo o menor genoma entre elas. UrprNPV também apresentou sintenia, porém com pouca identidade nucleotídica com as espécies mais proximamente relacionadas *Adoxophyes orana nucleopolyhedrovirus* e *Adoxophyes honmai nucleopolyhedrovirus*, entretanto com uma distância filogenética grande. Foi verificado também, que a filogenia dos genes virais *vlf-1*, *p47*, *49k*, *odv-ec43*, *p40*, *lef-4*, *odv-e27*, *gp41* e *vp1054* podem individualmente recontar a filogenia baseada nos genes compartilhados. Dessa forma, o aumento no número de genomas sequenciados, está tornando cada vez mais compreensível a evolução da família *Baculoviridae* e a relação vírus-hospedeiro. Além de facilitar o estudo da genômica destes organismos, como a aquisição e perda de genes e mecanismos de transferência horizontal e coinfeção. A caracterização desses vírus é de suma importância para uma possível utilização de baculovírus como controladores biológicos na agricultura e vetores de expressão gênica.

## REFERÊNCIAS BIBLIOGRÁFICAS

- AHMAD, I.; AHMAD, F.; PICHTEL, J. (Ed.). **Microbes and microbial technology: agricultural and environmental applications**. Springer Science & Business Media, 2011.
- ARAGÃO-SILVA, C. W. et al. The complete genome of a baculovirus isolated from an insect of medical interest: *Lonomia obliqua* (Lepidoptera: Saturniidae). **Scientific reports**, v. 6, 2016.
- ARDISSON-ARAÚJO, Daniel Mendes Pereira et al. Complete genome sequence of the first non-Asian isolate of *Bombyx mori* nucleopolyhedrovirus. **Virus genes**, v. 49, n. 3, p. 477-484, 2013.
- ARDISSON-ARAÚJO, D. M. P. et al. Genome sequence of *Erinnyis ello granulovirus* (ErelGV), a natural cassava hornworm pesticide and the first sequenced sphingid-infecting betabaculovirus. **BMC genomics**, v. 15, n. 1, p. 856, 2014.
- ARDISSON-ARAÚJO, D. M. P. et al. Characterization of *helicoverpa zea single nucleopolyhedrovirus* isolated in brazil during the first old world bollworm (noctuidae: *helicoverpa armigera*) nationwide outbreak. **Virus Reviews & Research**, v. 20, n. 1, p. 2, 2015.
- ARDISSON-ARAÚJO, D. M. P. et al. A betabaculovirus encoding a gp64 homolog. **BMC genomics**, v. 17, n. 1, p. 94, 2016a.
- ARDISSON-ARAÚJO, D. M. P. et al. Genome sequence of *Perigonia lusca single nucleopolyhedrovirus*: insights into the evolution of a nucleotide metabolism enzyme in the family Baculoviridae. **Scientific reports**, v. 6, 2016b.
- ARDISSON-ARAÚJO, D. M. P. et al. A Novel Betabaculovirus Isolated from the Monocot Pest *Mocis latipes* (Lepidoptera: Noctuidae) and the Evolution of Multiple-Copy Genes. **Viruses**, v. 10, n. 3, p. 134, 2018.
- AU S.; PANTE N. Nuclear transport of baculovirus: revealing the nuclear pore complex passage. **Journal of structural biology**, n. 177, p. 90–98. 2012.
- AU, S.; WU, W.; PANTÉ, N. Baculovirus nuclear import: open, nuclear pore complex (NPC) sesame. **Viruses**, v. 5, n. 7, p. 1885-1900, 2013.
- AYRES, Martin D. et al. The complete DNA sequence of *Autographa californica* nuclear polyhedrosis virus. **Virology**, v. 202, n. 2, p. 586-605, 1994.
- BARROS, M. C. E. S. **Atividade de baculovírus selvagens em camundongos in vivo e in vitro e expressão da proteína do envelope do vírus da Febre Amarela (YFE) e da glicoproteína do vírus da raiva (RVGP) em células de inseto**. 2012. 26 p. Tese (Doutorado em Biologia Molecular) – Universidade de Brasília, Brasília, 2012.
- BIRD, F. T.; ELGEE, D. E. A virus disease and introduced parasites as factors controlling the European spruce sawfly, *Diprion hercyniae* (Htg.), in central New Brunswick. **The Canadian Entomologist**, v. 89, n. 8, p. 371-378, 1957.

BLISSARD, G. W.; ROHRMANN, G. F. Baculovirus diversity and molecular biology. **Annual Review of Entomology**. n. 35, p. 127–155. 1990.

BRAUNAGEL, S. C. et al. Identification and analysis of an *Autographa californica* nuclear polyhedrosis virus structural protein of the occlusion-derived virus envelope: ODV-E56. **Virology**, v. 217, n. 1, p. 97-110, 1996.

BRAUNAGEL, S. C.; SUMMERS, M. D. Molecular biology of the baculovirus occlusion-derived virus envelope. **Current drug targets**, v. 8, n. 10, p. 1084-1095, 2007.

BRIDGES, C. A. Catalogue of the Family-Group, Genus-Group and Species-Group Names of the Hesperidae (Lepidoptera) of the World. **Urbana, IL**. 1993.

BRITO, A. F. et al. The pangenome of the *Anticarsia gemmatalis multiple nucleopolyhedrovirus* (AgMNPV). **Genome biology and evolution**, v. 8, n. 1, p. 94-108, 2016.

BUENO, A. F. et al. Inimigos naturais das pragas da soja. In: HOFFMANN-CAMPO, C. B.; CORRÊA-FERREIRA, B. S.; MOSCARDI, F. (Ed.). **Soja: manejo integrado de insetos e outros Artrópodes-praga**. Brasília, DF: Embrapa, p. 493-630. 2012.

CARPENTIER, D. C. J. et al. The baculovirus P10 protein of AcMNPV forms two distinct cytoskeletal-like structures and associates with polyhedral occlusion bodies during infection. **Virology**. n. 371, p. 278-291. 2008.

CASTRO, M. E. B. et al. The genome sequence of *Condylorrhiza vestigialis* NPV, a novel baculovirus for the control of the Alamo moth on *Populus* spp. in Brazil. **Journal of invertebrate pathology**, v. 148, p. 152-161, 2017.

CLEM, R. J.; PASSARELLI, A. L. Baculoviruses: sophisticated pathogens of insects. **PLoS Pathogens**, v. 9, n. 11, p. e1003729, 2013.

CONAB. **Companhia Nacional de Abastecimento**. Disponível em: <<http://www.conab.gov.br/>>. Acesso em maio de 2017.

COSTA LIMA, A. M. Insetos do Brasil: Lepidópteros. **Escola Nacional de Agronomia**, Rio de Janeiro, p. 379. 1945.

CRAVEIRO, S. R. et al. The genome sequence of *Pseudoplusia includens single nucleopolyhedrovirus* and an analysis of p26 gene evolution in the baculoviruses. **BMC genomics**, v. 16, n. 1, p. 127, 2015.

EMBRAPA. **Empresa Brasileira de Pesquisa Agropecuária**. Fauna de Lepidópteros do Bioma Cerrado. Disponível em: <[http://www.agencia.cnptia.embrapa.br/Agencia16/AG01/arvore/AG01\\_81\\_911200585235.html](http://www.agencia.cnptia.embrapa.br/Agencia16/AG01/arvore/AG01_81_911200585235.html)>. Acesso em maio de 2016.

EMBRAPA. **Empresa Brasileira de Pesquisa Agropecuária**. Baculovírus anticarsia, um inseticida biológico. Disponível em: <<https://www.embrapa.br/en/busca-de-noticias/>>.

/noticia/17914027/baculovirus-anticarsia-um-inseticida-biologico>. Acesso em junho de 2017.

FERREIRA, M. E.; NETO, C. R. B. A importância da pesquisa genômica e o sequenciamento de DNA. **Embrapa Recursos Genéticos e Biotecnologia. Comunicado Técnico**, 2003.

GALLO, D., et al. Manual de entomologia agrícola. **Ceres**. São Paulo, 2 ed., p. 649. 1988.

GARAVAGLIA, M. J. et al. The ac53, ac78, ac101, and ac103 genes are newly discovered core genes in the family Baculoviridae. **Journal of virology**, v. 86, n. 22, p. 12069-12079, 2012.

GREENE, G. L. Economic damage levels of bean leaf roller populations on snap beans. **Journal of Economic Entomology**. v.64, p. 673-674. 1971.

HAAS-STAPLETON, E. J. et al. P74 mediates specific binding of *Autographa californica* *M nucleopolyhedrovirus* occlusion-derived virus to primary cellular targets in the midgut epithelia of *Heliothis virescens* Larvae. **Journal of Virology**, v. 78, n. 13, p. 6786-6791. 2004.

HARRISON, R. L. et al. The complete genome sequence of a second distinct betabaculovirus from the true armyworm, *Mythimna unipuncta*. **PLoS one**, v. 12, n. 1, p. e0170510, 2017a.

HARRISON, R. L. et al. The Operophtera brumata Nucleopolyhedrovirus (OpbuNPV) Represents an Early, Divergent Lineage within Genus Alphabaculovirus. **Viruses**, v. 9, n. 10, p. 307, 2017b.

HARRISON, R. L. et al. The complete genome sequence of a third distinct baculovirus isolated from the true armyworm, *Mythimna unipuncta*, contains two copies of the *lef-7* gene. **Virus genes**, v. 54, n. 2, p. 297-310, 2017c.

HERNIOU, E. A. et al. The genome sequence and evolution of baculoviruses. **Annual review of entomology**, v. 48, n. 1, p. 211-234, 2003.

HERNIOU, E. A. et al. In: KING, Andrew MQ et al. (Ed.). **Virus taxonomy: ninth report of the International Committee on Taxonomy of Viruses**. Elsevier, p.163–174. 2012.

HOOVER, K. et al. A gene for an extended phenotype. **Science**, v. 333, n. 6048, p. 1401-1401, 2011.

HORTON, H. M.; BURAND, J. P. Saturable attachment sites for polyhedron-derived baculovirus on insect cells and evidence for entry via direct membrane fusion. **Journal of Virology**, v. 67, n. 4, p. 1860- 1868. 1993.

JACKAI, L. E. N.; DAOUST, R. A. Insect pests of cowpeas. **Annual review of entomology**, v. 31, n. 1, p. 95-119. 1986.

JAVED, M. A. et al. *Autographa californica multiple nucleopolyhedrovirus* AC83 is a *per os* infectivity factor (PIF) protein required for occlusion-derived virus (ODV) and budded virus

nucleocapsid assembly as well as assembly of the PIF complex in ODV envelopes. **Journal of Virology**, v. 91, n. 5, p. e02115-16, 2017.

JEHLE, J. A. et al. Molecular identification and phylogenetic analysis of baculoviruses from Lepidoptera. **Virology**, v. 346, n. 1, p. 180-193, 2006a.

JEHLE, J. A., et al. On the classification and nomenclature of baculoviruses: A proposal for revision. **Archives of Virology**, v. 151, n. 7, 2006b.

KAMITA, S. G. et al. A baculovirus-encoded protein tyrosine phosphatase gene induces enhanced locomotory activity in a lepidopteran host. **Proceedings of the National Academy of Sciences of the United States of America**, v. 102, n. 7, p. 2584-2589, 2005.

KATSUMA, S. et al. The baculovirus uses a captured host phosphatase to induce enhanced locomotory activity in host caterpillars. **PLoS Pathogens**, v. 8, n. 4, p. e1002644, 2012.

KEARSE, M. et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. **Bioinformatics**, v. 28, n. 12, p. 1647-1649, 2012.

KIRCHER, M.; STENZEL, U.; KELSO, J. Improved base calling for the Illumina Genome Analyzer using machine learning strategies. **Genome biology**, v. 10, n. 8, p. R83, 2009.

LUCAROTTI, C. J. et al. Pathology of a Gammabaculovirus in its natural balsam fir sawfly (*Neodiprion abietis*) host. **Psyche: A Journal of Entomology**, v. 2012, 2012.

MAPA. **Ministério da Agricultura, Pecuária e Abastecimento**. Disponível em: <<http://www.agricultura.gov.br/>>. Acesso em: maio de 2017.

MERRIMAN, B.; R&D TEAM, Ion Torrent; ROTHBERG, J. M. Progress in ion torrent semiconductor chip based sequencing. **Electrophoresis**, v. 33, n. 23, p. 3397-3417, 2012.

MIELE, S. A. B. et al. Baculovirus: molecular insights on their diversity and conservation. **International journal of evolutionary biology**, v. 2011, 2011.

MOSCARDI, F. Assessment of the application of baculoviruses for control of Lepidoptera. **Annual review of entomology**, v. 44, n. 1, p. 257-289, 1999.

NAVA, D. E.; NACHTIGAL, G. de F. **Controle biológico no Sul do Brasil**. 2010.

NELSON, D. L.; COX, M. M. **Princípios de bioquímica de Lehninger** – 6ª edição – Porto Alegre: Artmed, 2014.

OHKAWA, T. et al. Specific binding of *Autographa californica M nucleopolyhedrovirus* occlusion-derived virus to midgut cells of *Heliothis virescens* larvae is mediated by products of pif genes Ac119 and Ac022 but not by Ac115. **Journal of virology**, v. 79, n. 24, p. 15258-15264, 2005.

OLIVEIRA, J. V. et al. Genome of the most widely used viral biopesticide: *Anticarsia gemmatalis multiple nucleopolyhedrovirus*. **Journal of General Virology**, v. 87, n. 11, p. 3233-3250, 2006.

O'REILLY, D. R. O.; MILLER, L. K. A baculovirus blocks insect molting by producing ecdysteroid UDP-glucosyl transferase. **Science**, v. 245, n. 4922, p. 1110, 1989.

PEDIGO, L. P.; RICE, M. E. Entomology and Pest Management, **Prentice Hall**, New Jersey, p. 784. 2005.

POMBO, V. et al. Structural and ultrastructural changes during the infection of UFL-AG-286 cells with the baculovirus AgMNPV. **Journal of invertebrate pathology**, v. 72, n. 3, p. 239-245, 1998.

QUINTELA, E. D. Manejo integrado de pragas do feijoeiro. **Embrapa Arroz e Feijão. Circular técnica**, 2001.

ROHRMANN, George F.; ERLANDSON, Martin A.; THEILMANN, David A. The genome of a baculovirus isolated from *Hemileuca* sp. encodes a serpin ortholog. **Virus genes**, v. 47, n. 2, p. 357-364, 2013.

ROHRMANN, G. F. **Baculovirus molecular biology**. National Center for Biotechnology Information (US), Bethesda (MD), 2013. Disponível em: <  
<http://www.ncbi.nlm.nih.gov/books/NBK49500/>>. Acesso em: 03 de maio, 2018.

RONAGHI, Mostafa; UHLÉN, Mathias; NYRÉN, Pål. A sequencing method based on real-time pyrophosphate. **Science**, v. 281, n. 5375, p. 363-365, 1998.

RUSSELL, R. L. Q.; PEARSON, M. N.; ROHRMANN, G. F. Immunoelectron microscopic examination of *Orgyia pseudotsugata multicapsid nuclear polyhedrosis virus*-infected *Lymantria dispar* cells: time course and localization of major polyhedron-associated proteins. **Journal of general virology**, v. 72, n. 2, p. 275-283, 1991.

SCHMITT, A. T. **Eficiência da aplicação de Baculovirus erinnyis no controle do mandarová da mandioca**. Volume, 1985.

SLACK, J.; ARIF, B. M. The Baculoviruses Occlusion-Derived Virus: Virion Structure and Function. **Advances in virus research**, v. 69, p. 99-165, 2007.

SOSA-GÓMEZ, D. R. Microbial control of soybean pest insects and mites. In: **Microbial Control of Insect and Mite Pests**. p. 199-208. 2017.

STAMATAKIS, Alexandros; HOOVER, Paul; ROUGEMONT, Jacques. A rapid bootstrap algorithm for the RAxML web servers. **Systematic biology**, v. 57, n. 5, p. 758-771, 2008.

SUMMERS, M. D. Milestones leading to the genetic engineering of baculoviruses as expression vector systems and viral pesticides. **Advances in virus research**, v. 68, p. 3-73, 2006.

- SUN, X.; PENG, H. Recent advances in biological control of pest insects by using viruses in China. **Virologica sinica**, v. 22, n. 2, p. 158-162, 2007.
- SZEWCZYK, B. et al. Baculoviruses—re-emerging biopesticides. **Biotechnology advances**, v. 24, n. 2, p. 143-160, 2006.
- SZEWCZYK, B. et al. **Baculovirus Biopesticides Pesticides - Formulation, Effects, Fate**. InTech, New York, v. 25, n. 36. 2011.
- TEMERAK, S. A.; BOUCIAS, D. G.; WHITCOMB, W. H. A singly-embedded nuclear polyhedrosis virus and entomophagous insects associated with populations of the bean leafroller *Urbanus proteus* L.(Lepid., Hesperidae). **Journal of Applied Entomology**, v. 97, n. 1-5, p. 187-191, 1984.
- THIEM, S. M., CHENG, X. W. Baculovirus host-range. **Virologica Sinica**, v. 24, p. 436-457. 2009.
- VAN DAM, W.; WILDE, Gerald. Biology of the Bean Leafroller *Urbanus proteus* (Lepidoptera: Hesperidae). **Journal of the Kansas Entomological Society**, p. 157-160, 1977.
- VAN OERS, M. M.; VLAK, J. M. The baculovirus 10-kDa protein. **Journal of invertebrate pathology**, v. 70, n. 1, p. 1-17, 1997.
- WANG, J. et al. Genome sequencing and analysis of *Catopsilia pomona* nucleopolyhedrovirus: a distinct species in group I Alphabaculovirus. **PloS one**, v. 11, n. 5, p. e0155134, 2016.
- WANG, J. et al. Genome analysis of a novel Group I alphabaculovirus obtained from *Oxyplax ochracea*. **PloS one**, v. 13, n. 2, p. e0192279, 2018.
- WARREN, A. D.; OGAWA, J. R.; BROWER, A. V. Z. Phylogenetic relationships of subfamilies and circumscription of tribes in the family Hesperidae (Lepidoptera: Hesperioidea). **Cladistics**, v. 24, n. 5, p. 642-676, 2008.
- WILLIAMS, Greg V.; FAULKNER, Peter. Cytological changes and viral morphogenesis during baculovirus infection. In: **The baculoviruses**. Springer, Boston, MA, p. 61-107, 1997.
- WOLFF, J. L. C. et al. Analysis of the genome of *Spodoptera frugiperda* nucleopolyhedrovirus (SfMNPV-19) and of the high genomic heterogeneity in group II nucleopolyhedroviruses. **Journal of General Virology**, v. 89, n. 5, p. 1202-1211, 2008.
- YANG, M. M. et al. Baculoviruses and insect pests control in China. **African Journal of Microbiology Research**, v. 6, n. 2, p. 214-218, 2012.
- ZHANG, G.Y et al. Production and effectiveness of the new formulation of *Helicoverpa* virus pesticide-emulsifiable suspension. **Virologica Sinica**, v. 10, p. 242–247. 1995.



## APÊNDICE A – FIGURA SUPLEMENTAR 1

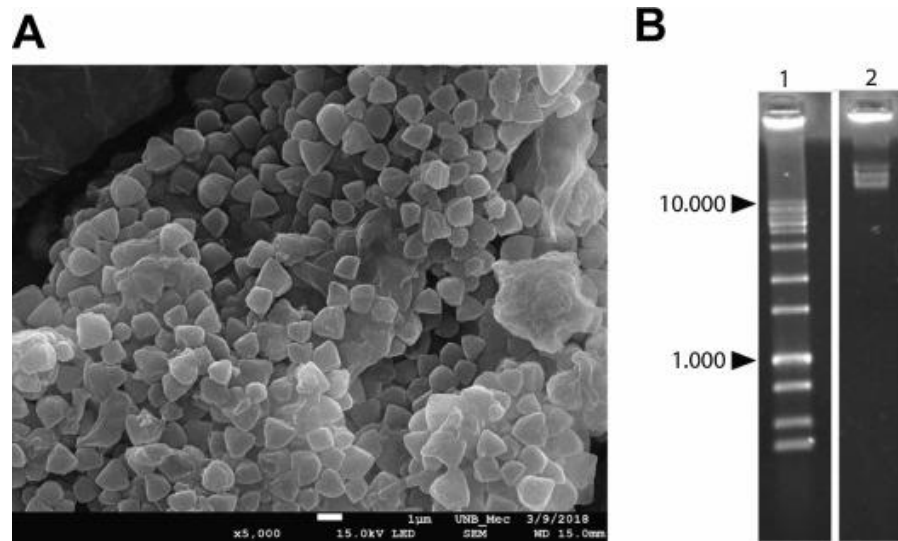


Figure S1. Occlusion bodies and ODV-extracted viral DNA of UrprNPV. (A) Scanning electron micrograph reveals several occlusion bodies with non-regular shape, which are characteristic of alphabaculoviruses. (B) UrprNPV genomic DNA subjected to 0.8% Agarose gel electrophoresis.

## APÊNDICE B – FIGURA SUPLEMENTAR 2

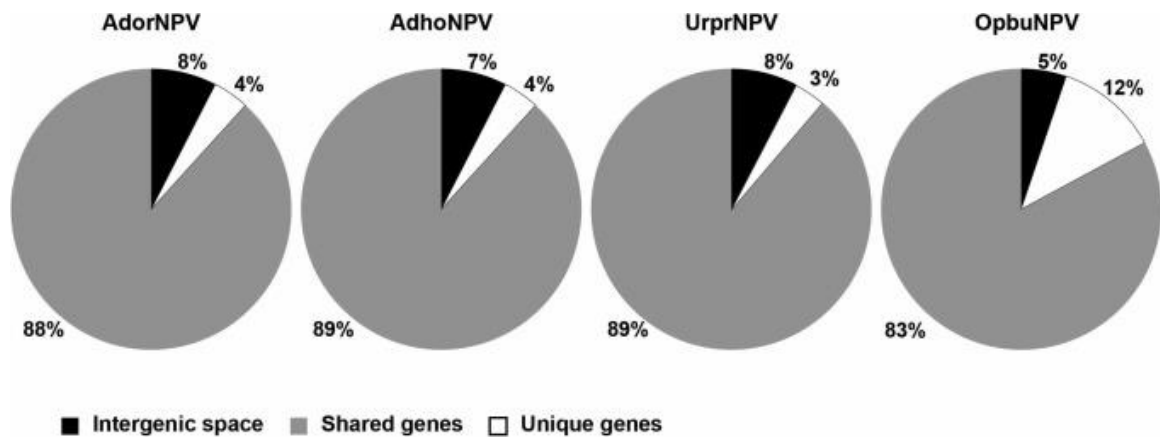


Figure S2. Genome content of some selected aphabaculovirus genomes. We calculated the percentage of shared genes (grey), unique genes (white), and the intergenic space (black) for AdorNPV, AdhoNPV, UrprNPV, and OpbuNPV.

### APÊNDICE C – FIGURA SUPLEMENTAR 3

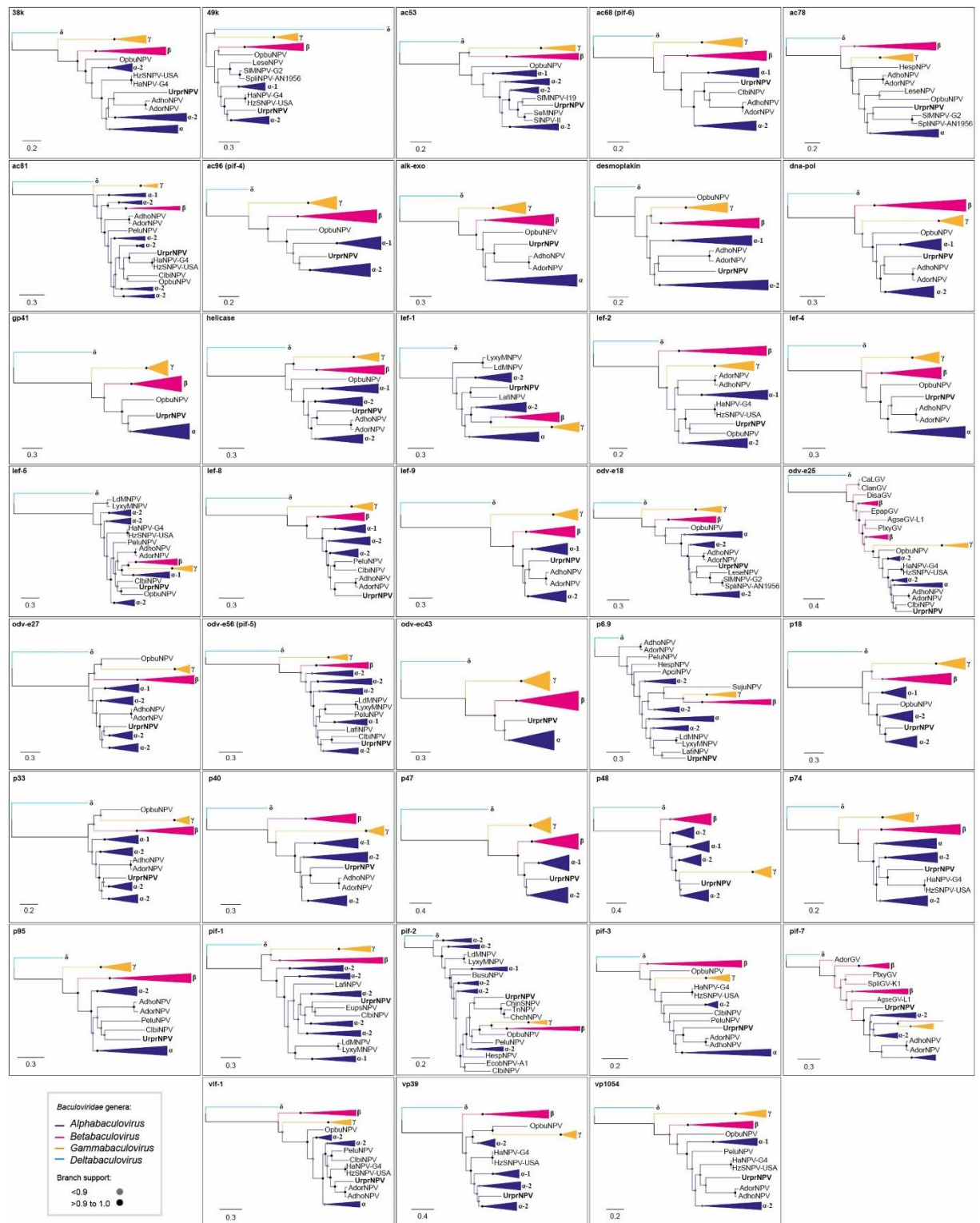


Figure S3. UrprNPV groups in different positions in the evolutionary history of each core gene. Individual trees from each of the 38 core genes are summarized to better show the relationship of UrprNPV with other baculoviruses: deltabaculovirus (light blue), gammabaculovirus

(orange), betabaculovirus (pink), and alphabaculovirus (dark blue). **UrprNPV** is in boldface. The branch support was determined by a SH-like method (black and grey closed circles).

## APÊNDICE D – TABELA SUPLEMENTAR 1

**Table S1.** Genomes used in this paper for reconstruction of the baculovirus phylogeny in both the Fig. 1 and 6. The isolates placed into the genera *Alphabaculovirus* (dark blue), *Betabaculovirus* (pink), *Gammabaculovirus* (orange), and *Deltabaculovirus* (light blue) are presented together with the acronym used in the main text, the host family where the virus was isolated from, the G+C content, the Genbank accession number, the genome size, and the global identity in relation to UrprNPV (using a pairwise alignment of the concatenated nucleotide sequence of the 38 baculovirus core genes).

	<b>Baculovirus</b>	<b>Acronym</b>	<b>Host family</b>	<b>G + C content (%)</b>	<b>Accession number</b>	<b>Genome Size (bp)</b>	<b>Id (%)</b>
1	<i>Adoxophyes honmai</i> nucleopolyhedrovirus	AdhoNPV	Tortricidae	35.6	AP006270	113220	53.3
2	<i>Adoxophyes orana</i> nucleopolyhedrovirus	AdorNPV	Tortricidae	35	EU591746	111724	53.5
3	<i>Agrotis ipsilon</i> multiple nucleopolyhedrovirus strain illinois	AgipMNPV	Noctuidae	48.6	EU839994	155122	51.2
4	<i>Agrotis segetum</i> nucleopolyhedrovirus	AgseNPV	Noctuidae	45.7	DQ123841	147544	51.5
5	<i>Antheraea pernyi</i> nucleopolyhedrovirus isolate L2	AnpeNPV-L2	Saturniidae	53.5	EF207986	126246	45.8
6	<i>Anticarsia gemmatalis</i> multiple nucleopolyhedrovirus	AgMNPV	Noctuidae	44.5	DQ813662	132239	47.5
7	<i>Apocheima cinerarium</i> nucleopolyhedrovirus	ApciNPV	Geometridae	33.4	FJ914221	123876	48.9
8	<i>Autographa californica</i> multiple nucleopolyhedrovirus clone C6	AcMNPV-C6	Noctuidae	40.7	L22858	133894	49.1
9	<i>Bombyx mandarina</i> nucleopolyhedrovirus S2	BomaNPV-S2	Bombycidae	40.4	JQ071499	129646	49.2
10	<i>Bombyx mori</i> nucleopolyhedrovirus strain T3	BmNPV-T3	Bombycidae	40.4	L33180	128413	49.2
11	<i>Buzura suppressaria</i> nucleopolyhedrovirus	BusuNPV	Geometridae	36.8	KF611977	120420	52.9
12	<i>Catopsilia pomona</i> nucleopolyhedrovirus	CapoNPV	Pieridae	39.7	KU565883	128058	48.3
13	<i>Choristoneura fumiferana</i> defective multiple nucleopolyhedrovirus	CfDEFMNPV	Tortricidae	45.8	AY327402	131160	47.2
14	<i>Choristoneura fumiferana</i> multiple nucleopolyhedrovirus	CfMNPV	Tortricidae	50.1	AF512031	129593	46.2
15	<i>Choristoneura murinana</i> nucleopolyhedrovirus	ChmuNPV	Tortricidae	50	KF894742	124688	46.3

16	<i>Choristoneura occidentalis</i> nucleopolyhedrovirus	ChocNPV	Tortricidae	50.1	KC961303	128446	46.2
17	<i>Choristoneura rosaceana</i> nucleopolyhedrovirus	ChroNPV	Tortricidae	48.6	KC961304	129052	46.6
18	<i>Chrysodeixis chalcites</i> nucleopolyhedrovirus	ChchNPV	Noctuidae	39	AY864330	149622	52.8
19	<i>Chrysodeixis includens</i> single nucleopolyhedrovirus IF	ChinSNPV	Noctuidae	39.2	KU669293	139181	52.3
20	<i>Clanis bilineata</i> nucleopolyhedrovirus	ClbiNPV	Sphingidae	37.7	DQ504428	135454	52.9
21	<i>Condylorrhiza vestigialis</i> multiple nucleopolyhedrovirus	CoveMNPV	Crambidae	42.9	KJ631623	125767	47.8
22	<i>Cyclophragma undans</i> nucleopolyhedrovirus	CyunNPV	Lasiocampidae	45.1	KT957089	140418	47.6
23	<i>Dasychra pudibunda</i> nucleopolyhedrovirus	DapuNPV	Lymantriidae	54.4	KP747440	136761	45.4
24	<i>Dendrolimus kikuchii</i> nucleopolyhedrovirus	DekiNPV	Lasiocampidae	48	JX193905	141454	47.1
25	<i>Ectropis obliqua</i> nucleopolyhedrovirus strain A1	EcobNPV-A1	Geometridae	37.6	DQ837165	131204	52.9
26	<i>Epiphyas postvittana</i> nucleopolyhedrovirus	EppoNPV	Tortricidae	40.7	AY043265	118584	48
27	<i>Euproctis pseudoconspersa</i> nucleopolyhedrovirus	EupsNPV	Lymantriidae	40.3	FJ227128	141291	52.4
28	<i>Helicoverpa armigera</i> multiple nucleopolyhedrovirus	HaMNPV	Noctuidae	40.1	EU730893	154196	53
29	<i>Helicoverpa armigera</i> nucleopolyhedrovirus G4	HaNPV-G4	Noctuidae	39	AF271059	130759	51.8
30	<i>Helicoverpa zea</i> single nucleopolyhedrovirus USA	HZSNPV-USA	Noctuidae	39.1	AF334030	130869	52.2
31	<i>Hemileuca</i> sp. Nucleopolyhedrovirus	HespNPV	Saturniidae	38.1	KF158713	140633	51
32	<i>Hyphantria cunea</i> nucleopolyhedrovirus	HycuNPV	Arctiidae	45.5	AP009046	132959	47.1
33	<i>Lambdina fiscellaria</i> nucleopolyhedrovirus	LafiNPV	Geometriidae	43.7	KP752043	157977	49.2
34	<i>Leucania separata</i> nuclear polyhedrovirus strain AH1	LeseNPV	Noctuidae	48.6	AY394490	168041	47.9
35	<i>Lonomia obliqua</i> multiple nucleopolyhedrovirus	LoobMNPV	Saturniidae	35.7	KP763670	120022	49.1
36	<i>Lymantria dispar</i> multiple nucleopolyhedrovirus	LdMNPV	Lymantriidae	57.5	AF081810	161046	49
37	<i>Lymantria xyliana</i> multiple nucleopolyhedrovirus	LyxyMNPV	Lymantriidae	53.5	GQ202541	156344	50.2
38	<i>Mamestra brassicae</i> multiple nucleopolyhedrovirus strain Chb1	MbMNPV-Chb1	Noctuidae	40.1	JX138237	154451	49.2

39	<i>Mamestra configurata</i> nucleopolyhedrovirus B	MacoNPV-B	Noctuidae	40	AY126275	158482	52.4
40	<i>Mamestra configurata</i> nucleopolyhedrovirus-A strain 90/2	MacoNPV-A 90/2	Noctuidae	41.7	U59461	155060	52.4
41	<i>Maruca vitrata</i> multiple nucleopolyhedrovirus	MaviMNPV	Crambidae	38.6	EF125867	111953	49.3
42	<i>Operophtera brumata</i> nucleopolyhedrovirus	OpbuNPV	Geometridae	38.9	-	119054	48.5
43	<i>Orgyia leucostigma</i> nucleopolyhedrovirus isolate CFS-77	OrleNPV	Lymantriidae	39.9	EU309041	156179	52.7
44	<i>Orgyia pseudotsugata</i> multiple nucleopolyhedrovirus	OpMNPV	Lymantriidae	55.1	U75930	131995	45.4
45	<i>Peridroma</i> sp. nucleopolyhedrovirus	PespNPV	Noctuidae	53.2	KM009991	151109	50.7
46	<i>Perigonia lusca</i> single nucleopolyhedrovirus	PeluSNPV	Sphingidae	39.6	KM596836	132831	53.2
47	<i>Philosamia cynthia ricini</i> nucleopolyhedrovirus	PhcyNPV	Saturniidae	53.7	JX404026	125376	45.6
48	<i>Plutella xylostella</i> multiple nucleopolyhedrovirus isolate CL3	PlxyMNPV	Plutellidae	40.7	DQ457003	134417	49.2
49	<i>Rachiplusia ou</i> multiple nucleopolyhedrovirus	RoMNPV	Noctuidae	39.1	AY145471	131526	49.5
50	<i>Spodoptera exigua</i> nucleopolyhedrovirus	SeMNPV	Noctuidae	43.8	AF169823	135611	52
51	<i>Spodoptera frugiperda</i> multiple nucleopolyhedrovirus isolate 19	SfMNPV-I19	Noctuidae	40.3	EU258200	132565	52.7
52	<i>Spodoptera litoralis</i> nucleopolyhedrovirus isolate AN1956	SpliNPV-1956	Noctuidae	44.7	JX454574	137998	48.9
53	<i>Spodoptera litura</i> nucleopolyhedrovirus G2	SINPV-G2	Noctuidae	42.8	AF325155	139342	49.2
54	<i>Spodoptera litura</i> nucleopolyhedrovirus II	SINPV-II	Noctuidae	45	EU780426	148634	51.5
55	<i>Sucra jujuba</i> nucleopolyhedrovirus	SujuNPV	Geometridae	38.7	KJ676450	135952	52
56	<i>Thysanoplusia orichalcea</i> nucleopolyhedrovirus	ThorNPV	Noctuidae	39.5	JX467702	132978	49.3
57	<i>Trichoplusia ni</i> single nucleopolyhedrovirus	TnSNPV	Noctuidae	39	DQ017380	134394	52.7
58	<b>Urbanus proteus</b> nucleopolyhedrovirus	<b>UrprNPV</b>	<b>Hesperiidae</b>	<b>34.7</b>	<b>KR011717</b>	<b>105555</b>	<b>100</b>
59	<i>Adoxophyes orana</i> granulovirus	AdorGV	Tortricidae	34.5	AF547984	99657	40.9
60	<i>Agrotis segetum</i> granulovirus-L1	AgseGV-L1	Noctuidae	37.3	KC994902	131442	41.1
61	<i>Choristoneura occidentalis</i> granulovirus	ChocGV	Tortricidae	32.7	DQ333351	104710	41.3

62	<i>Clostera anastomosis</i> granulovirus	CaLGV	Notodontidae	46.7	KC179784	101818	38.9
63	<i>Clostera anastomosis</i> granulovirus isolate ClanGV-B	ClanGV-B	Notodontidae	37.8	KR091910	107409	40.7
64	<i>Clostera anachoreta</i> granulovirus	ClanGV	Notodontidae	44.4	HQ116624	101487	39.3
65	<i>Cnaphalocrocis medinalis</i> granulovirus	CnmeGV	Crambidae	35.2	KP658210	112060	39.8
66	<i>Cryptophlebia leucotreta</i> granulovirus isolate CV3	CrleGV	Tortricidae	32.4	AY229987	110907	41.2
67	<i>Cydia pomonella</i> granulovirus	CpGV	Tortricidae	45.3	U53466	123500	39.2
68	<i>Diatraea saccharalis</i> granulovirus	DisaGV	Crambidae	34.9	KP296186	98392	41
69	<i>Epinotia aporema</i> granulovirus	EpapGV	Tortricidae	41.5	JN408834	119082	40.6
70	<i>Erinnyis ello</i> granulovirus	ErelGV	Sphingidae	38.7	KJ406702	102759	40.6
71	<i>Helicoverpa armigera</i> granulovirus	HaGV	Noctuidae	40.8	EU255577	169794	40.2
72	<i>Mocis latipes</i> granulovirus	MolaGV	Noctuidae	38.3	KR011718	134272	40.6
73	<i>Mythimna unipuncta</i> granulovirus	MyunGV#8	Noctuidae	49.9	KX855660	144.673	39
74	<i>Phthorimaea operculella</i> granulovirus	PhopGV	Gelechiidae	35.7	AF499596	119217	40.5
75	<i>Pieris rapae</i> granulovirus E3	PiraGV-E3	Pieridae	33.2	GU111736	108476	41.6
76	<i>Plodia interpunctella</i> granulovirus	PiGV	Pyalidae	44.2	KX151395	112536	39.6
77	<i>Plutella xylostella</i> granulovirus	PlxyGV	Plutellidae	40.7	AF270937	100999	40.1
78	<i>Pseudaletia unipuncta</i> granulovirus	PsunGV-Hawaii	Noctuidae	39.8	EU678671	176677	40.3
79	<i>Spodoptera frugiperda</i> granulovirus	SpfrGV	Noctuidae	46.2	KM371112	140913	39.9
80	<i>Spodoptera litura</i> granulovirus isolate K1	SpliGV	Noctuidae	38.8	DQ288858	124121	39.9
81	<i>Trichoplusia ni</i> granulovirus	TnGV	Noctuidae	39.8	KU752557	175360	40.3
82	<i>Xestia c-nigrum</i> granulovirus	XcGV	Noctuidae	40.7	AF162221	178733	40.1
83	<i>Neodiprion sertifer</i> nucleopolyhedrovirus	NeseNPV	Diprionidae	33.8	AY430810	86462	37
84	<i>Neodiprion lecontei</i> nucleopolyhedrovirus	NeleNPV	Diprionidae	33.4	AY349019	81755	36.9



---

85	Neodiprion abietis nucleopolyhedrovirus	NeabNPV	Diprionidae	33.4	DQ317692	84264	36.8
86	Culex nigripalpus nucleopolyhedrovirus	CuniNPV	Culicidae	50.9	AF403738	108252	29.3

---

## APÊNDICE E – TABELA SUPLEMENTAR 2

**Table S2.** Characteristics of the *Urbanus proteus* nucleopolyhedrovirus (UrprNPV) genome: number, position, nucleotide size, and amino acid size of each ORF and homology search. Predicted ORFs are compared with homolog genes in two related genomes: *Adoxophyes orana* nucleopolyhedrovirus (AdorNPV), *Adoxophyes honmai* nucleopolyhedrovirus (AdhoNPV), *Helicoverpa armigera* nucleopolyhedrovirus G4 (HaNPV-G4), and the type species of *Baculoviridae*, *Autographa californica* multiple nucleopolyhedrovirus clone C6 (AcMNPV-C6).

ORF	Name	Position	Size (nt)	Size (aa)	AdorNPV		AdhoNPV		HaNPV-G4		AcMNPV-C6	
					ORF	Max ID (%)	ORF	Max ID (%)	ORF	Max ID (%)	ORF	Max ID (%)
1 <sup>a</sup>	<i>polh</i>	1 > 741	741	246	1	89	1	90	1	92	8	88
2	<i>pp78/83</i>	767 < 2.179	1.413	470	4	32	4	36	2	23	9	27
3 <sup>a</sup>	<i>pk-1</i>	2.197 > 2.997	801	266	5	51	5	51	3	52	10	44
4	<i>hoar</i>	3.148 < 4.389	1.242	413	6	11	6	10	4	12	-	-
5*	<i>UrprNPV-ORF-5</i>	4.786 < 5.007	222	73	-	-	-	-	-	-	-	-
6*	<i>UrprNPV-ORF-6</i>	5.201 > 6.349	1.149	382	-	-	-	-	-	-	-	-
7 <sup>b</sup>	<i>p74(pif-0)</i>	6.441 < 8.402	1.962	653	26	59	27	58	20	53	138	61
8 <sup>a</sup>	<i>me-53</i>	8.426 < 9.325	900	299	25	23	26	23	16/17	18	139	20
9 <sup>a</sup>	<i>ie-0</i>	9.683 > 10.456	774	257	24	39	25	38	8	28	141	30
10 <sup>b</sup>	<i>p49</i>	10.437 > 11.882	1.446	481	23	55	24	55	9	53	142	48
11 <sup>b</sup>	<i>odv-e18</i>	11.882 > 12.136	255	84	22	62	23	63	10	62	143	58
12 <sup>b</sup>	<i>odv-ec27</i>	12.155 > 13.009	855	284	21	48	22	49	11	53	144	48
13 <sup>a</sup>	<i>chtBD2</i>	13.035 > 13.313	279	92	20	48	21	50	12	53	145	45
14 <sup>a</sup>	<i>ep23</i>	13.302 < 13.913	612	203	19	23	20	23	13	26	146	30

15 <sup>a</sup>	<i>ie-1</i>	13.928	>	15.682	1.755	584	18	32	19	32	14	33	147	27
16 <sup>b</sup>	<i>odv-e56 (pif-5)</i>	15.762	<	16.835	1.074	357	8	57	8	57	15	49	148	53
17 <sup>b</sup>	<i>p47</i>	16.892	<	18.106	1.215	404	10	56	10	55	35	47	40	53
18*	<i>UrprNPV-ORF-18</i>	18.147	>	18.410	264	87	-	-	-	-	-	-	-	-
19	<i>lef-12</i>	18.415	>	18.951	537	178	48	24	51	21	36	22	41	25
20 <sup>a</sup>	<i>dbp</i>	19.150	>	20.106	957	318	27	40	28	42	25	40	25	34
21 <sup>a</sup>	<i>lef-6</i>	20.115	>	20.594	480	159	28	42	29	40	24	25	28	25
22	<i>UrprNPV-ORF-22</i>	20.645	<	20.860	216	71	-	-	-	-	-	-	-	-
23*	<i>UrprNPV-ORF-23</i>	21.087	>	21.818	732	243	-	-	-	-	-	-	-	-
24	<i>p26a</i>	21.841	>	22.692	852	283	30	40	31	40	22	34	136	32
25	<i>p10</i>	22.704	>	23.000	297	98	31	44	32	53	21	46	137	27
26	<i>UrprNPV-ORF-26</i>	23.030	<	23.467	438	145	17	58	18	57	27	60	34	35
27 <sup>a</sup>	<i>ubiquitin</i>	23.611	>	23.865	255	84	16	80	16	80	28	70	35	80
28	<i>UrprNPV-ORF-28</i>	23.852	>	24.061	210	69	-	-	-	-	-	-	-	-
29 <sup>a</sup>	<i>39k</i>	24.102	<	24.902	801	266	13	41	13	42	31	37	36	40
30 <sup>a</sup>	<i>lef-11</i>	24.958	<	25.359	402	133	12	44	12	44	32	32	37	40
31 <sup>a</sup>	<i>nudix</i>	25.290	<	26.000	711	236	11	47	11	46	33	47	38	52
32 <sup>b</sup>	<i>lef-8</i>	26.067	<	28.679	2.613	870	46	63	49	63	38	62	50	58
33	<i>bjdp</i>	28.739	>	29.557	819	272	45	35	48	35	39	28	51	20
34*	<i>UrprNPV-ORF-34</i>	29.464	<	29.814	351	116	-	-	-	-	-	-	-	-
35	<i>ac52-like</i>	29.881	<	30.387	507	168	43	36	46	36	42	29	52	18
36 <sup>b</sup>	<i>ac53-like</i>	30.404	>	30.814	411	136	44	55	47	54	43	53	53	50

37*	<i>UrprNPV-ORF-37</i>	30.896	>	31.294	399	132	-	-	-	-	-	-	-	-
38	<i>UrprNPV-ORF-38</i>	31.403	<	32.113	711	236	42	33	45	31	44	18	-	-
39	<i>iap-3</i>	32.557	<	33.378	822	273	85	44	88	42	103	53	-	-
40*	<i>UrprNPV-ORF-40</i>	33.473	>	33.631	159	52	-	-	-	-	-	-	-	-
41	<i>UrprNPV-ORF-41</i>	33.734	<	33.973	240	79	41	27	44	27	45	30	-	-
42	<i>lef-10</i>	33.924	>	34.163	240	79	40	57	43	54	46	49	53-a	48
43 <sup>b</sup>	<i>vp-1054</i>	34.021	>	35.034	1.014	337	39	52	42	53	47	47	54	39
44	<i>UrprNPV-ORF-44</i>	35.141	>	35.359	219	72	38	47	41	23	48	32	-	-
45*	<i>UrprNPV-ORF-45</i>	35.519	>	35.830	312	103	-	-	-	-	-	-	-	-
46 <sup>a</sup>	<i>fp-25k</i>	35.949	<	36.551	603	200	34	60	36	60	53	53	61	65
47 <sup>b</sup>	<i>lef-9</i>	36.641	>	38.137	1.497	498	33	71	35	71	55	67	62	66
48	<i>odv-e66</i>	38.141	<	40.003	1.863	620	47	26	50	26	96	25	46	25
49	<i>p26b</i>	40.110	<	40.847	738	245	-	-	-	-	-	-	-	-
50	<i>iap-2</i>	40.894	<	41.595	702	233	50	32	53	31	62	30	71	25
51 <sup>b</sup>	<i>ac68-like (pif-6)</i>	41.568	<	41.966	399	132	52	53	55	53	64	68	68	40
52 <sup>a</sup>	<i>lef-3</i>	42.025	>	43.164	1.140	379	53	32	56	30	65	31	67	27
53 <sup>b</sup>	<i>desmoplakin</i>	43.272	<	45.668	2.397	798	54	49	57	29	66	55	66	41
54 <sup>b</sup>	<i>dna-pol</i>	45.667	>	48.678	3.012	1003	55	48	58	49	67	50	65	42
55 <sup>a</sup>	<i>UrprNPV-ORF-55</i>	48.697	<	49.080	384	127	56	36	59	36	69	37	-	-
56	<i>bro-a</i>	49.213	<	49.599	387	128	-	-	-	-	-	-	-	-
57	<i>bro-b</i>	49.693	<	50.439	747	248	32	41	33	41	60	27	-	-
58	<i>bro-c</i>	50.621	<	50.797	177	58	-	-	-	-	60	15	-	-

59	<i>bro-d</i>	50.877	<	51.086	210	69	-	-	-	-	60	55	-	-
60	<i>ac76-like</i>	51.197	<	51.448	252	83	57	46	60	69	70	41	76	60
61 <sup>b</sup>	<i>vlf-1</i>	51.524	<	52.618	1.095	364	58	75	61	72	71	70	77	71
62 <sup>b</sup>	<i>ac78-like</i>	52.623	<	52.955	333	110	59	36	62	39	72	31	78	34
63*	<i>UrprNPV-ORF-63</i>	52.976	>	53.128	153	50	-	-	-	-	-	-	-	-
64	<i>UrprNPV-ORF-64</i>	53.240	>	53.836	597	198	-	-	-	-	-	-	-	-
65 <sup>b</sup>	<i>gp41</i>	53.953	<	54.885	933	310	60	63	63	62	73	53	80	56
66 <sup>b</sup>	<i>ac81-like</i>	54.875	<	55.516	642	213	61	60	64	58	74	62	81	53
67 <sup>a</sup>	<i>ac82-like</i>	55.401	<	55.925	525	174	62	28	65	26	75	31	82	26
68 <sup>b</sup>	<i>vp91</i>	55.918	>	58.434	2.517	838	63	40	66	41	76	37	83	36
69	<i>cg30-like</i>	58.450	<	58.806	357	118	-	-	-	-	77	19	88	32
70 <sup>b</sup>	<i>vp39</i>	58.806	<	59.840	1.035	344	64	46	67	47	78	49	89	42
71 <sup>b</sup>	<i>lef-4</i>	59.842	>	61.203	1.362	453	65	48	68	48	79	48	90	43
72	<i>ac111-like</i>	61.317	>	61.553	237	78	-	-	-	-	116 <sup>+</sup>	33	111	56
73 <sup>b</sup>	<i>p33</i>	61.588	<	62.364	777	258	66	50	69	50	80	47	92	41
74 <sup>b</sup>	<i>p18</i>	62.369	>	62.857	489	162	67	56	70	56	81	58	93	49
75 <sup>b</sup>	<i>odv-e25</i>	62.850	>	63.548	699	232	68	56	71	56	82	56	94	43
76 <sup>b</sup>	<i>dna-helicase</i>	63.613	<	67.176	3564	1187	69	41	72	51	84	43	95	42
77 <sup>b</sup>	<i>pif-4</i>	67.148	>	67.639	492	163	70	51	73	52	85	55	96	52
78 <sup>b</sup>	<i>38k</i>	67.652	<	68.611	960	319	71	41	74	41	86	40	98	38
79 <sup>b</sup>	<i>lef-5</i>	68.504	>	69.322	819	272	72	55	75	54	87	45	99	50
80 <sup>b</sup>	<i>p6.9</i>	69.356	<	69.616	261	86	73	42	76	42	88	47	100	41

81 <sup>b</sup>	<i>p40</i>	69.655	<	70.767	1.113	370	74	40	77	40	89	42	101	31
82 <sup>a</sup>	<i>p12</i>	70.786	<	71.160	375	124	75	36	78	37	90	31	102	24
83 <sup>b</sup>	<i>p48</i>	71.153	<	72.268	1.116	371	76	50	79	50	91	47	103	43
84	<i>vp80</i>	72.298	>	73.605	1.308	435	77	31	80	31	92	23	104	25
85 <sup>a,b</sup>	<i>pif-7</i>	73.605	>	73.793	189	62	78	37	81	36	93	37	110	33
86 <sup>b</sup>	<i>odv-ec43</i>	73.771	>	74.841	1.071	356	79	56	82	57	94	54	109	44
87 <sup>a</sup>	<i>ac108-like</i>	74.855	>	75.127	273	90	80	32	83	34	95	28	108	38
88 <sup>a</sup>	<i>ac106/107-like</i>	75.122	<	75.748	627	208	105	46	110	48	101	55	106	57
89	<i>parg</i>	75.777	<	77.093	1.317	438	104	21	109	21	100	18	-	-
90 <sup>b</sup>	<i>pif-3</i>	77.097	<	77.708	612	203	103	51	108	52	98	45	115	45
91	<i>sod</i>	77.763	<	78.221	459	152	101	69	106	69	106	66	31	68
92	<i>UrprNPV-ORF-92</i>	78.267	>	78.584	318	105	102	26	107	24	-	-	-	-
93	<i>UrprNPV-ORF-93</i>	78.642	<	79.055	414	137	-	-	-	-	-	-	-	-
94 <sup>b</sup>	<i>alk-exo</i>	79.404	>	80.672	1.269	422	2	39	2	39	114	37	133	34
95	<i>fgf</i>	80.802	>	81.632	831	276	107	20	112	22	113	21	32	28
96	<i>ac19-like</i>	81.637	<	81.924	288	95	87	33	91	35	115	38	19	33
97	<i>ac18-like</i>	81.939	>	83.099	1.161	386	88	29	92	29	-	-	18	23
98 <sup>a</sup>	<i>pep</i>	83.130	<	84.038	909	302	99	46	104	43	120	33	131	28
99	<i>gp-16</i>	84.079	<	84.450	372	123	98	27	103	30	119	21	130	25
100 <sup>a</sup>	<i>p24</i>	84.466	<	85.140	675	224	97	51	102	52	118	49	129	38
101	<i>UrprNPV-ORF-101</i>	85.228	>	85.578	351	116	96	31	101	31	118A <sup>+</sup>	21	-	-
102 <sup>b</sup>	<i>lef-2</i>	85.529	>	86.182	654	217	95	34	100	34	117	34	6	36

103	<i>pkip</i>	86.185	>	86.715	531	176	93	22	98	24	130	34	24	23
104 <sup>b</sup>	<i>pif-2</i>	86.824	>	87.972	1.149	382	90	63	95	64	132	66	22	60
105	<i>bro-e</i>	88.089	<	89.549	1.461	486	81	55	84	55	105	68	-	-
106 <sup>b</sup>	<i>pif-1</i>	89.703	>	91.259	1.557	518	113	50	117	50	111	45	119	49
107	<i>UrprNPV-ORF-107</i>	91.256	>	91.489	234	77	108	32	113	31	-	-	-	-
108	<i>bro-f</i>	91.783	>	92.820	1.038	345	-	-	-	-	-	-	-	-
109	<i>bro-g</i>	93.003	<	94.574	1.572	523	-	-	-	-	-	-	-	-
110 <sup>a</sup>	<i>fp</i>	95.095	<	97.131	2.037	678	114	39	118	40	133	41	23	19
111	<i>UrprNPV-ORF-111</i>	97.156	>	99.891	2.736	911	115	27	119	26	129	26	-	-
112	<i>ac17-like</i>	99.939	<	100.613	675	224	116	36	120	37	128	27	17	34
113	<i>UrprNPV-ORF-113</i>	100.606	<	101.187	582	193	-	-	-	-	-	-	-	-
114	<i>egt</i>	101.334	<	102.863	1.530	509	118	45	122	46	126	48	15	42
115	<i>UrprNPV-ORF-115</i>	102.921	<	103.343	423	140	119	34	123	32	125	40	-	-
116 <sup>b</sup>	<i>lef-1</i>	103.373	>	104.101	729	242	120	41	124	41	124	38	14	36
117	<i>ctl</i>	104.156	<	104.317	162	53	83	48	86	48	-	-	3	63
118*	<i>UrprNPV-ORF-118</i>	104.361	>	104.510	150	49	-	-	-	-	-	-	-	-
119 <sup>a</sup>	<i>38.7 kDa</i>	104.531	>	105.523	993	330	121	25	125	25	123	24	13	26

\*: unique gene to the UrprNPV genome

a: 25 genes shared among alphabaculoviruses (the ac110-like or pif-7 is now recognized as a core gene).

b: 38 baculovirus core genes

†: not annotated in the Genbank database genome

### APÊNDICE F – TABELA SUPLEMENTAR 3

**Table S3.** Best hits from the BLASTX for the *Urbanus proteus* nucleopolyhedrovirus genes and their respective e-value.

<b>Best hit</b>	<b>E. value</b>	<b>Genes</b>
<i>Adoxophyes honmai nucleopolyhedrovirus</i>	3,00E-161	<i>vlf-1</i>
	1,00E-121	<i>odv-ec43</i>
<i>Adoxophyes orana nucleopolyhedrovirus</i>	1,00E-128	<i>gp41</i>
	1,00E-06	<i>UrprNPV-ORF-107</i>
<i>Agrotis ipsilon multiple nucleopolyhedrovirus</i>	5,00E-86	<i>nudix</i>
	3,00E-65	<i>ac53-like</i>
	7,00E-38	<i>p26b</i>
	4,00E-48	<i>UrprNPV-ORF-55</i>
	5,00E-08	<i>parg</i>
	4,00E-67	<i>ac18-like</i>
	0,00E+00	<i>bro-e</i>
<i>Agrotis segetum nucleopolyhedrovirus</i>	4,00E-90	<i>pk-1</i>
	1,00E-120	<i>vp-1054</i>
	4,00E-43	<i>ac68-like (pif-6)</i>
	7,00E-40	<i>lef-3</i>
	0,00E+00	<i>vp91</i>
	3,00E-89	<i>p33</i>
	2,00E-07	<i>pif-7</i>
<i>Apocheima cinerarium nucleopolyhedrovirus</i>	1,00E-155	<i>polh</i>
	3,00E-05	<i>UrprNPV-ORF-28</i>
	0,00E+00	<i>lef-8</i>
	4,00E-06	<i>cg30-like</i>
	1,00E-156	<i>p48</i>
<i>Autographa californica nucleopolyhedrovirus</i>	3,00E-16	<i>ac111-like</i>
<i>Bombyx mori nucleopolyhedrovirus</i>	3,00E-73	<i>sod</i>
<i>Buzura suppressaria nucleopolyhedrovirus</i>	1,00E-63	<i>ac106/107-like</i>
	0,00E+00	<i>pif-2</i>
	2,00E-62	<i>UrprNPV-ORF-111</i>
<i>Candidatus Nomurabacteria bacterium - GW2011_GWB1_43_7</i>	4,40E-01	<i>UrprNPV-ORF-18</i>
<i>Cellulosilyticum lentocellum</i>	7,10E-10	<i>UrprNPV-ORF-118</i>
<i>Chrysodeixis chalcites nucleopolyhedrovirus</i>	5,00E-159	<i>p47</i>
<i>Chrysodeixis chalcites single nucleopolyhedrovirus - TF1-A</i>	3,00E-105	<i>p40</i>
<i>Chrysodeixis includens nucleopolyhedrovirus</i>	2,00E-04	<i>UrprNPV-ORF-113</i>
<i>Clanis bilineata nucleopolyhedrovirus</i>	0,00E+00	<i>p49</i>
	9,00E-24	<i>UrprNPV-ORF-38</i>
	3,00E-17	<i>vp80</i>
	1,80E-01	<i>UrprNPV-ORF-92</i>
<i>Clostera anastomosis granulovirus</i>	5,00E-16	<i>UrprNPV-ORF-93</i>
<i>Dendrolimus kikuchii nucleopolyhedrovirus</i>	2,00E-73	<i>odv-e66</i>
<i>Ectropis obliqua nucleopolyhedrovirus</i>	2,00E-71	<i>lef-5</i>
	5,00E-36	<i>lef-2</i>
	1,00E-19	<i>pkip</i>
<i>Euproctis pseudoconsersa nucleopolyhedrovirus</i>	4,00E-59	<i>pep</i>
	3,00E-52	<i>ie-1</i>



<i>Helicoverpa armigera nucleopolyhedrovirus</i>	3,00E-81 3,00E-99 1,20E-01	<i>dbp</i> <i>alk-exo</i> <i>UrprNPV-ORF-101</i>
<i>Helicoverpa armigera single nucleopolyhedrovirus - H25EA1</i>	9,00E-22	<i>iap-2</i>
<i>Heliothis virescens ascovirus - 3f</i>	3,00E-28 4,00E-158	<i>bro-a</i> <i>bro-g</i>
<i>Hemileuca sp. nucleopolyhedrovirus</i>	3,00E-69 3,00E-06	<i>pif-3</i> <i>ac19-like</i>
<i>Leucania separata nucleopolyhedrovirus</i>	1,00E-07 3,00E-49	<i>odv-e18</i> <i>pif-4</i>
<i>Lonomia obliqua multiple nucleopolyhedrovirus</i>	4,00E-14	<i>ctl</i>
<i>Lymantria dispar multiple nucleopolyhedrovirus</i>	3,00E-121 6,00E-25 2,00E-07 7,00E-20 1,00E-07 2,00E-129	<i>odv-e56 (pif-5)</i> <i>lef-11</i> <i>bjdp</i> <i>bro-c</i> <i>ac76-like</i> <i>vp39</i>
<i>Malacosoma sp. alphabaculovirus</i>	1,00E-03 9,00E-25 3,00E-63 0,00E+00	<i>pp78/83</i> <i>chtBD2</i> <i>ac81-like</i> <i>fp</i>
<i>Mamestra configurata nucleopolyhedrovirus A</i>	6,00E-113 3,00E-09 6,00E-71	<i>odv-ec27</i> <i>UrprNPV-ORF-22</i> <i>ac34-like</i>
<i>Mamestra configurata nucleopolyhedrovirus B</i>	3,00E-26	38.7 kDa
<i>Methanocaldococcus jannaschii</i>	4,00E-01	<i>UrprNPV-ORF-6</i>
<i>Moniliophthora roreri</i>	2,10E+00	<i>UrprNPV-ORF-23</i>
<i>Niabella soli DSM 19437</i>	2,90E-01	<i>UrprNPV-ORF-37</i>
No hit	- -	<i>UrprNPV-ORF-5</i> <i>p6.9</i>
<i>Operophtera brumata</i>	8,00E-01	<i>UrprNPV-ORF-45</i>
<i>Orgyia leucostigma nucleopolyhedrovirus</i>	3,50E-02 1,00E-06 3,00E-07 2,00E-179 2,00E-06 2,00E-63	<i>UrprNPV-ORF-64</i> <i>p12</i> <i>fgf</i> <i>pif-1</i> <i>UrprNPV-ORF-115</i> <i>lef-1</i>
<i>Papilio xuthus</i>	4,00E-92	<i>iap-3</i>
<i>Penicillium italicum</i>	5,40E-09	<i>UrprNPV-ORF-63</i>
<i>Peridroma alphabaculovirus</i>	3,00E-25 5,00E-64	<i>ac52-like</i> <i>fp-25k</i>
<i>Perigonia lusca single nucleopolyhedrovirus</i>	8,00E-42 1,00E-10 2,00E-73	<i>ubiquitin</i> <i>ac78-like</i> <i>p24</i>
<i>Pseudoplusia includens single nucleopolyhedrovirus - IE</i>	5,00E-20 1,00E-25 3,00E-16 8,00E-03 3,00E-121	<i>lef-6</i> <i>p10</i> <i>lef-10</i> <i>UrprNPV-ORF-44</i> <i>bro-b</i>
<i>Pseudoplusia includens single nucleopolyhedrovirus - IF</i>	0,00E+00	<i>dna-helicase</i>
<i>Spodoptera exigua multiple nucleopolyhedrovirus</i>	1,20E+00 1,00E-20	<i>hoar</i> <i>ep23</i>

	8,00E-30	<i>ac82-like</i>
<i>Spodoptera frugiperda ascovirus 1a</i>	6,00E-19 0,00E+00	<i>bro-d</i> <i>bro-f</i>
<i>Spodoptera frugiperda multiple nucleopolyhedrovirus</i>	6,00E-50 2,00E-20 0,00E+00 6,00E-134 1,00E-89 5,00E-14 4,00E-157	<i>ie-0</i> <i>desmoplakin</i> <i>dna-pol</i> <i>lef-4</i> <i>38k</i> <i>ac108-like</i> <i>egt</i>
<i>Spodoptera littoralis nucleopolyhedrovirus</i>	6,00E-29	<i>ac17-like</i>
<i>Spodoptera litura nucleopolyhedrovirus II</i>	2,00E-10 3,00E-03	<i>me-53</i> UrprNPV-ORF-41
<i>Sucra jujuba nucleopolyhedrovirus</i>	0,00E+00 2,00E-10 2,00E-61	<i>p74(pif-0)</i> <i>lef-12</i> <i>p26a</i>
<i>Toxocara canis</i>	5,20E-04	UrprNPV-ORF-40
<i>Trichoplusia ni single nucleopolyhedrovirus</i>	2,00E-35 0,00E+00 2,00E-74 2,00E-84 5,00E-08	<i>39k</i> <i>lef-9</i> <i>p18</i> <i>odv-e25</i> <i>gp-16</i>
<i>Xenopus tropicalis</i>	3,20E-07	UrprNPV-ORF-34