



Universidade de Brasília  
Instituto de Ciências Biológicas  
Departamento de Biologia Celular  
Programa de Pós-Graduação em Biologia Molecular

# Genoma e Diversidade Genética de Populações de *Chrysodeixis includens* *nucleopolyhedrovirus* (ChinNPV)

**Saluana Rocha Craveiro**  
Matrícula (UnB): 12/0087120

Tese apresentada ao Departamento de  
Biologia Celular do Instituto de Ciências  
Biológicas da Universidade de Brasília como  
uma das exigências para a obtenção do título  
de Doutora em Biologia Molecular

**Orientadora:** Dra Maria Elita Batista de Castro  
Pesquisadora - Embrapa Recursos Genéticos e Biotecnologia

Brasília-DF, abril de 2016



Genoma e Diversidade Genética de Populações de  
*Chrysodeixis includens nucleopolyhedrovirus* (ChinNPV)

Pesquisa realizada (2012 / 2016) no Laboratório de Virologia de Insetos da Embrapa Recursos Genéticos e Biotecnologia (CENARGEN), com apoio financeiro da CAPES/EMBRAPA e FAPDF/CNPq.

**Banca Examinadora:**

Dra. Maria Elita B. Castro (CENARGEN) - Orientadora / Presidente

Dr. Marcelo de Macedo Brigido (UnB) - 1º Examinador

Dr. Fernando Hercos Valicente (Embrapa Milho e Sorgo) - 2º Examinador

Dra. Natalia Florencio Martins (CENARGEN)- 3º Examinador

Dra. Maria Cléria Valadares Inglis (CENARGEN)- 4º Examinador

**Suplente:**

Dra. Débora Pires Paula (CENARGEN)

À minha família, meu marido Fernando e meu amor maior, meu querido filho Guilherme.

Aos meus pais José Saulo e Rosana Craveiro.

**Dedico**

## AGRADECIMENTOS

A Deus a quem declaro toda glória e honra, que escuta meus pedidos e lamentações, que renova minhas forças toda manhã e que me dá perseverança para seguir em frente hoje e sempre.

À minha orientadora Dra. Maria Elita B. Castro pelo seu árduo trabalho e grande empenho estando ao meu lado em todas as etapas do desenvolvimento dessa tese. Por ser meu maior exemplo profissional de avidez à ciência, de ética e senso de justiça. Pela motivação diária e por se preocupar, até mesmo, em me ensinar a construir o meu caráter e conduta profissional. Obrigada por toda sua dedicação e paciência.

Ao Dr. Peter W. Inglis por estar sempre disposto a colaborar com esse projeto e estando sempre presente na idealização e discussão dos resultados no decorrer de todo trabalho.

Ao Dr. Bergmann M. Ribeiro pela sua disponibilidade ao projeto e, em particular, pela sua colaboração na correção da redação dos artigos científicos resultantes desta tese.

Ao Dr. Fernando Lucas Melo pela colaboração no início dos estudos relacionados à diversidade genética do gene *pif-2* e treinamento para utilização do programa *Geneious*.

Aos pesquisadores do Laboratório de Bioinformática da Embrapa Cenargen, Dr. Orzenil Silva Jr por participar no planejamento inicial do processamento dos dados obtidos do sequenciamento e, em especial, os meus sinceros agradecimentos ao Dr. Roberto C. Togawa e a Dra. Priscila Grynberg por me auxiliar quando eu mais precisei me dando todo suporte para as análises de bioinformática e auxiliando na exploração e discussão dos resultados.

À Dra. Débora Paula pela sua grande disposição e generosidade. Por ajudar na análise e discussão dos resultados sempre que foi necessário, em particular, com os dados obtidos do estudo do gene *p26*.

À Dra Sônia Nair Bão pela sua disponibilidade e colaboração prestada.

Aos membros do Laboratório de Virologia de Invertebrados (LVI), onde realizei a maior parte dos experimentos, Zilda Maria A. Ribeiro pelo suporte técnico, pela aprendizagem de técnicas e experimentos e pelo seu acolhimento e carinho de mamãe com que cuida de mim, William Sihler pelos ensinamentos

durante o trabalho diário dentro do laboratório e pela sua grande amizade e carinho, Raimunda C. Mesquita pelo auxílio nas tarefas diárias do laboratório provendo o material necessário para os experimentos e pela sua alegria e animação que nos inspira. Agradeço também aos estudantes do laboratório Samantha, Yasmin, Regio pelo convívio diário e, em especial, ao Luis Arthur que colaborou muito nas fases finais do trabalho principalmente durante o processo de anotação dos genomas.

Aos membros da banca de qualificação Dra. Débora P. Paula, Dr. Tatsuya Nagata, Dr. Dário Grattapaglia e Dr. Julio Carlyle M. Rodrigues, por aceitarem ao convite ajudando na avaliação dos dados e na redação da Tese.

À Embrapa Recursos Genéticos e Biotecnologia (Cenargen) pelo suporte técnico e financeiro e pela oportunidade de realização deste trabalho.

À secretaria do Prédio de Controle Biológico do Cenargen, na pessoa da Rosângela Zansávio, pela colaboração sempre que necessário.

Ao Decanato de Pós-Graduação da Universidade de Brasília (UnB) e, em particular, ao Programa de Pós-Graduação em Biologia Molecular do Departamento de Biologia Celular do Instituto de Ciências Biológicas da Universidade de Brasília pelo suporte acadêmico e oportunidade de realização do curso de doutorado.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da bolsa de estudo e de suporte financeiro a participação de eventos científicos.

À Fundação de Apoio à Pesquisa do Distrito Federal (FAPDF/ PRONEX-CNPq Edital 03-2009) pelo suporte financeiro durante a etapa de sequenciamento das amostras reais.

# AGRADECIMENTOS

## Familiares e Amigos

Aos meus queridos familiares, pais, irmãos, avós, cunhados, sobrinhos, sogros, tios e primos pela preocupação demonstrada durante toda a realização do trabalho e pelo carinho, amor e dedicação proporcionados todos esses anos.

À minha maravilhosa família construída ao lado do meu querido marido Fernando C. Fernandes, meu conselheiro e maior motivador, que suporta as recaídas e festeja as vitórias ao meu lado. Além disso, me presenteou com nosso bem mais valioso, meu bebê Guilherme, que se tornou minha maior inspiração.

Ao Papai Saulo e Mamãe Rosana obrigada por sempre estarem presentes me apoiando, auxiliando nos momentos difíceis, incentivando quando faltam forças e sendo exemplos a quem procuro-me espelhar todos os dias.

Ao meu querido irmão Allan Saul que me surpreende todos os dias pela sua resiliência e que me cativa com toda sua intensidade. Me proporcionou tanta alegria com o amor de tia dinda da nossa linda Emanuella e novamente me agrada com a oportunidade de fazer parte de sua família e conviver com minha cunhada Mariana e meu sobrinho de coração Davi.

A minha mais nova família meus sogros Lauriana e Carlos Alberto e meus cunhados Vítor, Laura e Vívian obrigada por me receberem tão bem. Sou muito grata por toda atenção e carinho e estou muito feliz por compartilhar minha vida com mais vocês.

A todos colegas e amigos, em especial, Briana, Daniela, Cristina e Lorena pela agradável convivência diária na Embrapa e UnB e pelos valiosos momentos de diversão.

Muito Obrigada!!

## SUMÁRIO

|   |    |
|---|----|
| LISTA DE FIGURAS .....  | 1  |
| LISTA DE TABELAS .....  | 2  |
| ABREVIATURAS E SIGLAS .....   | 3  |
| RESUMO.....   | 6  |
| ABSTRACT .....  | 7  |
| INTRODUÇÃO GERAL .....  | 8  |
| OBJETIVOS .....   | 11 |
| Objetivos específicos .....   | 11 |
| REVISÃO BIBLIOGRÁFICA.....  | 12 |
| <i>Chrysodeixis includens nucleopolyhedrovirus</i> (ChinNPV).....   | 12 |
| Lagarta falsa-medideira: <i>Chrysodeixis includens</i> (Walker, 1858) .....   | 14 |
| O que são os baculovírus? .....   | 16 |
| Aplicações biotecnológicas dos baculovírus .....  | 19 |
| Evolução da Família <i>Baculoviridae</i> .....  | 21 |
| Genoma dos baculovírus .....  | 23 |
| Diversidade de baculovírus intra e interespecífica.....   | 26 |
| CAPÍTULO 1: Diversidade genética de <i>Chrysodeixis includens</i> NPV baseada nos core genes: <i>pif-2</i> , <i>lef-9</i> e <i>helicase</i> .....               | 28 |
| 1.1. Introdução .....   | 29 |
| 1.2. Materiais e Métodos .....  | 30 |
| 1.2.1. Vírus.....   | 30 |
| 1.2.2. Purificação de OBs .....   | 31 |
| 1.2.3. Extração de DNA a partir de partículas OBs purificadas.....  | 32 |
| 1.2.4. Determinação da sequência nucleotídica dos genes <i>pif-2</i> , <i>lef-9</i> e <i>helicase</i> dos isolados de ChinNPV .....                             | 32 |
| 1.2.5. Genomas virais e análises de diversidade .....   | 33 |
| 1.2.6. Análise de pressão de seleção do gene <i>pif-2</i> de ChinNPV .....  | 35 |
| 1.3. Resultados e Discussão.....  | 37 |
| 1.3.1. Variações nucleotídicas em isolados de ChinNPV.....  | 37 |
| 1.3.2. Análise comparativa da variação genética dos genes <i>pif-2</i> , <i>lef-9</i> e <i>helicase</i> de isolados geográficos de diferentes baculovírus ..... | 40 |
| 1.3.3. Análise de seleção de <i>pif-2</i> em isolados de ChinNPV .....  | 43 |
| 1.4. Conclusões.....  | 46 |

|   |    |
|---|----|
| CAPÍTULO 2: Genoma completo e genômica comparativa de populações de <i>Chrysodeixis includens</i> NPV .....           | 47 |
| 2.1. Introdução .....   | 48 |
| 2.2. Materiais e Métodos .....  | 49 |
| 2.2.1. Sequenciamento e montagem do genoma de isolados de ChinNPV ..  | 49 |
| 2.2.2. Anotação do genoma .....   | 52 |
| 2.3. Resultados e Discussão .....   | 53 |
| 2.3.1. Descrição geral do genoma de ChinNPV-IE .....  | 53 |
| 2.3.2. Genômica comparativa de sete populações de ChinNPV .....   | 55 |
| 2.3.3. Comparação de ChinNPV-IE com outros <i>Alphabaculovirus</i> .....  | 57 |
| 2.3.4. Replicação, transcrição e genes estruturais .....  | 64 |
| 2.3.5. Metabolismo de nucleotídeos e reparação de DNA .....   | 65 |
| 2.3.6. Genes auxiliares .....   | 65 |
| 2.3.7. <i>Homologous regions (hrs)</i> são ausentes no genoma de ChinNPV-IE.  | 66 |
| 2.3.8. Genes <i>bro</i> de ChinNPV .....  | 66 |
| 2.3.9. ORFs únicas de ChinNPV .....   | 67 |
| 2.4. Conclusões .....   | 70 |
| CAPÍTULO 3: Identificação do gene <i>p26</i> e sua evolução na família <i>Baculoviridae</i> .....                     | 71 |
| 3.1. Introdução .....   | 72 |
| 3.2. Materiais e Métodos .....  | 73 |
| 3.2.1. Determinação e caracterização da sequência nucleotídica do gene <i>p26</i> de baculovírus .....                | 73 |
| 3.2.2. Análise filogenética da proteína P26 .....   | 73 |
| 3.3. Resultados e Discussão .....   | 74 |
| 3.3.1. As proteínas P26 de baculovírus se correlacionam de acordo com o contexto de seu posicionamento genômico ..... | 74 |
| 3.3.2. Eventos independentes de aquisição do gene <i>p26</i> dentro da família <i>Baculoviridae</i> .....             | 78 |
| 3.3.3. Presença de peptídeo sinal nas proteínas P26 de <i>Alphabaculovirus</i> com dois homólogos .....               | 81 |
| 3.4. Conclusões .....   | 82 |
| DISCUSSÃO GERAL .....   | 83 |
| CONCLUSÕES GERAIS E PERSPECTIVAS .....  | 88 |
| REFERÊNCIAS BIBLIOGRÁFICAS .....  | 90 |



## LISTA DE FIGURAS

|   |    |
|---|----|
| Figura 1. Corpos de oclusão de <i>Chrysodeixis includens</i> NPV.....   | 12 |
| Figura 2. Diversidade genética de isolados de ChinNPV-IA a -IG.....   | 13 |
| Figura 3. Fotos do inseto <i>C. includens</i> .....   | 15 |
| Figura 4. Desenho esquemático das partículas virais de <i>nucleopolyhedrovirus</i> (NPV) e <i>granulovirus</i> (GV).....  | 17 |
| Figura 5. Desenho esquemático das progênes infectiosas encontradas no ciclo de infecção dos baculovírus.....  | 18 |
| Figura 6. Filogenia da família <i>Baculoviridae</i> .....   | 22 |
| Figura 7. Alinhamento de sequências deduzidas de aminoácidos do gene <i>helicase</i> dos sete isolados de ChinNPV (IA a IG).....  | 38 |
| Figura 8. Perfil de hidrofobicidade da sequência de aminoácidos PIF2 dos isolados de ChinNPV (IA a IG).....   | 39 |
| Figura 9. Diversidade genética dos genes <i>pif-2</i> , <i>lef-9</i> e <i>helicase</i> de isolados geográficos de diferentes baculovírus: ChinNPV, HearNPV, MacoNPV e SfMNPV.....                   | 42 |
| Figura 10. Mapa circular do genoma de ChinNPV-IE.....   | 54 |
| Figura 11. Árvore filogenética (NJ) baseada em sequências de genomas completos de sete isolados de ChinNPV (IA a IG).....   | 57 |
| Figura 12. Matriz de similaridade entre os genomas de ChinNPV-IE e de ChchNPV (A), TnSNPV (B), AcMNPV (C) e MacoNPV-B (D).....  | 62 |
| Figura 13. Alinhamento múltiplo do genoma completo de ChinNPV-IE (A) com os <i>Alphabaculovirus</i> ChchNPV (B), TnSNPV (C), MacoNPV-B (D) e AcMNPV (E) usando o programa <i>Mauve v. 2.0</i> ..... | 63 |
| Figura 14. Hélices transmembrana preditas a partir das sequências deduzidas de aminoácidos das ORFs únicas de ChinNPV.....  | 69 |
| Figura 15. Média do ponto isoelétrico (pI) da proteína P26 nas posições adjacentes ao gene <i>p10</i> (P1) e ao gene <i>iap-2</i> (P2).....   | 78 |
| Figura 16. Cladograma baseado nas seqüências deduzidas de aminoácidos de P26.....   | 80 |
| Figura 17. Peptídeo sinal predito na proteína P26 de ChinNPV-IE utilizando <i>SignalP v.4.1</i> .....   | 82 |

## LISTA DE TABELAS

|  |    |
|--|----|
| Tabela 1. Genes conservados em todos os baculovírus ( <i>core genes</i> ).....   | 25 |
| Tabela 2. Isolados de ChinNPV obtidos de larvas <i>C. includens</i> infectadas .....   | 31 |
| Tabela 3. Sequências de baculovírus utilizados na análise comparativa da variação genética entre isolados virais.....  | 34 |
| Tabela 4. Análise de pressão de seleção do gene <i>pif-2</i> de ChinNPV realizada no PAML e dados estatísticos da razão de verossimilhança para a comparação dos modelos utilizados..... | 45 |
| Tabela 5. Análise de seleção positiva para o gene <i>pif-2</i> de ChinNPV utilizando diferentes modelos implementados no programa HyPhy v. 2.1.2 .....                                   | 46 |
| Tabela 6. Estatística dos resultados do sequenciamento de isolados de ChinNPV -454 GS FLX Titanium (Roche) .....   | 50 |
| Tabela 7. Dados estatísticos da montagem dos genomas de ChinNPV utilizando os programas: <i>Newbler Assembler v.2.8</i> , <i>MIRA v.4.0.2</i> e <i>Celera Assembler v.8.3</i> .....      | 51 |
| Tabela 8. Características gerais dos genomas de isolados de ChinNPV.....   | 55 |
| Tabela 9. Características do genoma de ChinNPV-IE comparadas com as de outros <i>Alphabaculovirus</i> .....  | 58 |
| Tabela 10. Comparação das 141 ORFs putativas de ChinNPV-IE com genes homólogos nos <i>Alphabaculovirus</i> .....   | 59 |
| Tabela 11. Sequências de aminoácidos utilizadas na análise filogenética da proteína P26.....   | 76 |

## ABREVIATURAS E SIGLAS

AcMNPV - *Autographa californica mutiple nucleopolyhedrovirus*

AgMNPV- *Anticarsia gemmatalis multiple nucleopolyhedrovirus*

BmNPV - *Bombyx mori nucleopolyhedrovirus*

CfMNPV- *Choristoneura fumiferana nucleopolyhedrovirus*

ChchNPV - *Chrysodeixis chalcites nucleopolyhedrovirus*

ChinNPV – *Chrysodeixis includens nucleopolyhedrovirus*

ChocNPV - *Choristoneura occidentalis nucleopolyhedrovirus*

ChroNPV - *Choristoneura rosaceana nucleopolyhedrovirus*

CuniNPV - *Culex nigripalpus nucleopolyhedrovirus*

EcobNPV - *Ectropis obliqua nucleopolyhedrovirus*

HearNPV - *Helicoverpa armigera nucleopolyhedrovirus*

LdMNPV - *Lymantria dispar multiple nucleopolyhedrovirus*

LeseNPV - *Leucania separata nucleopolyhedrovirus*

MacoNPV - *Mamestra configurata nucleopolyhedrovirus*

NeabNPV - *Neodiprion abietis nucleopolyhedrovirus*

NeleNPV - *Neodiprion lecontei nucleopolyhedrovirus*

NeseNPV - *Neodiprion sertifer nucleopolyhedrovirus*

OrleNPV - *Orgyia leucostigma nucleopolyhedrovirus*

PsinSNPV – *Pseudoplusia includens single nucleopolyhedrovirus*

PsunGV – *Pseudaletia unipuncta granulovirus*

SeNPV - *Spodoptera exigua nucleopolyhedrovirus*

SfMNPV- *Spodoptera frugiperda nucleopolyhedrovirus*

SpliNPV- *Spodoptera littoralis multiple nucleopolyhedrovirus*

TnSNPV - *Trichoplusia ni single nucleopolyhedrovirus*

%CG – porcentagem da composição nucleotídica composta por citosina e guanina

%ID – porcentagem de identidade da sequência nucleotídica

aa – aminoácido

BEVS- sistema de expressão de baculovírus (*baculovirus expression vector system*)

BLAST – *Basic Local Alignment Search Tool*

bro – ORF repetitiva de baculovirus (*baculovirus repeated ORF*)

BV – vírus extracelular (*budded virus*)

dN - substituições não-sinônimas por sítio não-sinônimo

dS - substituição sinônima por sítio sinônimo

GV – granulovirus

hrs – região homóloga (*homologous region*)

IB - Inferência filogenética Bayesiana

IFEL - *Internal Fixed Effects Likelihood*

kDa - quiloDaltons

M – molar

MCMC- Monte Carlo via Cadeia de Markov

MEME - *Mixed Effects Model of Evolution*

MIP – Manejo Integrado de Pragas (*Integrated Pest Management*)

ML- *Maximum Likelihood*

MNPV – multiple nucleopolyhedrovirus

MP – Máxima Parcimônia

MUSCLE - *Multiple Sequence Alignment*

NCBI – *National Center for Biotechnology Information*

NGS - *Next Generation Sequencing*

NJ – *Neighbour joining*

NPV – nucleopolyhedrovirus

nt – nucleotídeo

OB – corpo de oclusão (*occlusion body*)

ODV – vírus derivado de oclusão (*occlusion derived virus*)

ORF – fase de leitura aberta (*Open Reading Frame*)

PAML- *Phylogenetic Analysis by Maximum-Likelihood*

pb – pares de bases

PCR – reação em cadeia da polimerase (*polymerase chain reaction*)

pI -ponto isoelétrico

PM- peso molecular

REL - *Random Effect Likelihood*

RFLP – polimorfismo de tamanho de fragmento de restrição (*restriction fragment length polymorphism*)

rpm – rotações por minuto

SNPs - polimorfismos de base única ou polimorfismos com base na substituição de nucleotídeos únicos (*single nucleotide polymorphisms*)

SNPV – single nucleopolyhedrovirus

SP - sítios de clivagem de peptídeo sinal

TM - domínio hidrofóbico transmembrânico

UV - radiação ultravioleta

## RESUMO

*Chrysodeixis includens nucleopolyhedrovirus* (ChinNPV) é um vírus patogênico à lagarta falsa-medideira, *Chrysodeixis includens* (Walker, 1858), uma praga polífaga que ataca 174 plantas hospedeiras, incluindo soja e algodão. Os baculovírus formam um grande grupo de vírus de DNA fita dupla circular, que infectam insetos da ordem Lepidoptera, Hymenoptera e Diptera. Para uma melhor compreensão da virulência, evolução e biologia molecular de ChinNPV, este trabalho teve como objetivo determinar a sequência completa do genoma e a diversidade genética de sete isolados de ChinNPV (IA a IG). A diversidade genética foi inicialmente investigada por meio de análise das variações presentes nos *core genes pif-2*, *lef-9* e *helicase* de isolados de ChinNPV e de outros *Alphabaculovirus*. O gene *pif-2* de ChinNPV recebeu destaque por, em contraste com o esperado, apresentar um grande número de polimorfismos não-sinônimos com a identificação de dois sítios sob pressão diversificadora. Os genomas dos isolados de ChinNPV (IA a IG) possuem tamanho variando de 138.869 a 140.859 bp e conteúdo de CG ~ 39%. Um total de 142 ORFs foram identificadas incluindo 37 *core genes*, 102 genes encontrados em outros baculovírus, 3 genes *bro* e duas ORFs únicas (*Psin5* e *Psin8*). *Homologous repeats (hrs)*, típicas de baculovírus, estão ausentes no genoma de ChinNPV. Os genomas dos isolados de ChinNPV têm alta similaridade com identidade mínima de 96,4% e, polimorfismos, *indels* e pequenos fragmentos foram observados ao longo dos genomas. Dois homólogos ao gene *p26* (*p26a* e *p26b*) foram encontrados em ChinNPV. O padrão de agrupamento observado no cladograma da proteína P26 de NPVs indica que a aquisição do gene ocorreu em três eventos independentes de captura dentro da família *Baculoviridae*. ChinNPV tem uma sequência genômica distinta comparada a outros baculovírus sequenciados. As sete novas sequências do genoma completo de ChinNPV determinadas aqui constituem uma importante ferramenta para o melhor entendimento dos fatores de virulência, interação inseto-hospedeiro e da variação fenotípica observada em populações de baculovírus.

**Palavras-chave:** ChinNPV, análise genômica, variação genética, aquisição de *p26*.

## ABSTRACT

*Chrysodeixis includens* nucleopolyhedrovirus (ChinNPV) is a virus pathogenic to the soybean looper, *Chrysodeixis includens* (Walker, 1858), a polyphagous pest that attacks 174 host plants including soybean and cotton. Baculoviruses are a large group of circular double-stranded DNA viruses that infect insects of the orders Lepidoptera, Hymenoptera and Diptera. In order to better understanding the virulence, evolution and molecular biology of ChinNPV, this study aimed to determine the complete genome sequence and genetic diversity of seven ChinNPV (IA to IG) isolates. Genetic diversity was initially investigated by analyzing variation present in the *pif-2*, *lef-9* and *helicase* core genes in ChinNPV and other Alphabaculovirus isolates. The *pif-2* gene unexpectedly showed a large number of non-synonymous polymorphisms with two sites identified to be under diversifying selection. The ChinNPV (IA to IG) genomes range in length between 138,869 and 140,859 bp and GC content of ~39%. A total of 142 ORFs were identified including 37 baculovirus core genes, 102 genes found in other baculoviruses, three *bro* genes and two ORFs unique to ChinNPV (Psin5 and Psin8). The typical baculoviral homologous repeats (*hrs*) are absent in the ChinNPV genome. The ChinNPV genomes have high similarity with minimal identity of 96.4% and polymorphisms, indels and small fragment insertions throughout the genome were observed. Two *p26* gene homologs (*p26a* and *p26b*) were found in the ChinNPV genome. The clustering pattern seen in the cladogram of NPV P26 protein indicates that the acquisition of these genes occurred in three independent capture events within the family *Baculoviridae*. ChinNPV has a distinct genomic sequence compared to other baculoviruses sequenced so far. The seven new ChinNPV whole genome sequences determined herein constitute an important tool for a better understanding of virulence factors, insect-host interaction and phenotypic variation observed in baculovirus populations.

**Keywords:** ChinNPV, genomic analysis, genetic variation, *p26* acquisition

## INTRODUÇÃO GERAL

*Chrysodeixis includens nucleopolyhedrovirus* (ChinNPV) é um baculovírus patogênico à lagarta falsa-medideira, *Chrysodeixis* (= *Pseudoplusia*) *includens* (Walker, 1858) (Lepidoptera: Noctuidae), uma praga emergente da cultura da soja, e encontrada também em outras culturas de importância econômica como: algodão, feijão, tomate, batata, fumo, amendoim, girassol e outras (Bottimer, 1926; Folsom, 1936; Wolcott, 1936; Hensley *et al.*, 1964; Herzog e Todd, 1980; Bueno *et al.*, 2009; Bernardi *et al.*, 2012).

Por muitos anos, a lagarta falsa-medideira foi referida como *Pseudoplusia includens*. Entretanto, em 2003, a espécie foi reclassificada e *Chrysodeixis includens* passou a ser a classificação válida (Goater *et al.*, 2003; Moscardi *et al.*, 2012). Com isso, prévios trabalhos nomearam o vírus patogênico a *C. includens* como *Pseudoplusia includens nucleopolyhedrovirus* ocorrendo, inclusive, em artigos provenientes da tese em questão. Assim, para adequação as normas vigentes, passaremos a nomear o baculovírus como *Chrysodeixis includens nucleopolyhedrovirus* mesmo com os artigos já publicados estando nomeados com a nomenclatura antiga de *Pseudoplusia includens nucleopolyhedrovirus*.

O ataque de pragas é o principal motivo das perdas na produtividade sendo *C. includens* e *Anticarsia gemmatalis* as principais pragas desfolheadoras na cultura da soja brasileira. Apesar disso, o Brasil é o segundo maior produtor de soja do mundo e foi o maior exportador desse grão na safra 2010/2011 (Embrapa, 2011). Diante da importância econômica da soja, são vários os esforços em busca de eficientes métodos de controle dos insetos praga. Bioinseticidas a base do baculovírus de *A. gemmatalis* já foram desenvolvidos e amplamente utilizados (Moscardi, 2007). Até então, não há relatos da produção comercial de bioinseticidas a base do baculovírus de *C. includens* (Lacey *et al.*, 2015). Assim, a identificação, caracterização e produção de baculovírus para o controle de *C. includens* são de grande relevância, pois os isolados de *C. includens* NPV que vêm se apresentando como mais patogênicos a essa praga (Alexandre *et al.*, 2010), surgem como uma interessante alternativa para os programas de Manejo Integrado de Pragas (MIP) em plantações de soja no Brasil.



Desde 2008, isolados virais obtidos de larvas *Chrysodeixis includens*, coletadas em plantações de soja e algodão no Brasil e Guatemala, cedidos pela Embrapa Soja, estão sendo objetos de pesquisa científica e tecnológica no Laboratório de Virologia de Insetos da Embrapa Recursos Genéticos e Biotecnologia (Cenargen). Sete isolados virais foram identificados e estão depositados na Coleção de Vírus de Invertebrados (CVI - Cenargen).

Estudos de caracterização morfológica, molecular e de patogenicidade dos sete isolados virais (ChinNPV-IA a ChinNPV-IG), complementados com a identificação de polimorfismos em regiões altamente conservadas demonstraram a ocorrência de diversidade genética entre os isolados (Alexandre *et al.*, 2010; Craveiro, 2012).

Diante disso, pôde inferir que as variações genéticas identificadas podem estar ocasionando a variação fenotípica observada entre os isolados, demonstrada pela maior patogenicidade apresentada pelos isolados ChinNPV-IA, ChinNPV-IE e ChinNPV-IF comparados aos outros quatro isolados analisados (Alexandre *et al.*, 2010; Craveiro, 2012).

Assim, para o aprofundamento dos estudos de diversidade genética de ChinNPV, no presente trabalho foi conduzida uma investigação detalhada quanto as possíveis variações genéticas ocorridas nos genomas completos dos sete isolados virais (ChinNPV- IA a –IG) com destaque para regiões e genes com alta variação.

A tese aqui apresentada está descrita em três capítulos. No primeiro capítulo, a variação genética de ChinNPV é discutida e investigada com foco no gene *pif-2*. Na literatura, o gene *pif-2* é conhecido como o mais conservado entre os *core genes* (Miele *et al.*, 2011), entretanto, em estudos prévios, foi observada uma alta variação desse gene nos isolados de ChinNPV. Para análises de comparação, foram selecionados dois *core genes*, *lef-8* e *helicase*, genes com pouca e alta variação genética, respectivamente. Isolados de outros *Alphabaculovirus* descritos na literatura também foram comparados para a investigação da manutenção do padrão de variação genética em outras espécies virais.

O segundo capítulo apresenta as características gerais do genoma de ChinNPV. Nesse capítulo, foi realizado o sequenciamento, a montagem e anotação do genoma completo dos sete isolados de ChinNPV (IA a IG) e a

descrição gênica detalhada do isolado ChinNPV-IE, selecionado como representante da espécie. Além disso, foi realizada a análise comparativa dos genomas para a avaliação da diversidade genômica presente nas populações de ChinNPV.

Finalizando, o terceiro capítulo trata da caracterização da proteína P26 da família *Baculoviridae*. Durante a anotação do genoma de ChinNPV foi identificado duas cópias do gene *p26* o que despertou o interesse na investigação dos eventos de aquisição e da atividade funcional desses genes nos baculovírus. Há poucos relatos sobre a caracterização e função do gene *p26*, o que requer um melhor entendimento do modo de ação dessa proteína nos baculovírus.

## OBJETIVOS

O presente trabalho tem como objetivo auxiliar nos estudos de genética molecular relacionados a fatores determinantes da virulência, interação inseto-hospedeiro e variação fenotípica entre isolados virais, promovendo avanços no conhecimento da diversidade genética e genômica funcional dos baculovírus.

### Objetivos específicos

- Determinar a sequência genômica e realizar a descrição de genes de sete isolados de ChinNPV (IA a IG).
- Analisar a diversidade genética dos sete isolados de ChinNPV (IA a IG).
- Identificar e analisar as possíveis ORFs únicas do genoma de ChinNPV e comparar as diferentes ORFs encontradas entre o genomas dos sete isolados de ChinNPV.
- Analisar o genoma de ChinNPV em comparação com o genoma de outros *Alphabaculovirus*.
- Investigar as variações genéticas dos genes conservados *pif-2*, *lef-9* e *helicase* dos isolados de ChinNPV e de isolados de outros baculovírus.
- Analisar a pressão de seleção evolutiva sob o gene *pif-2* de ChinNPV.
- Identificar o gene *p26* no genoma dos baculovírus e analisar sua evolução na família *Baculoviridae*.

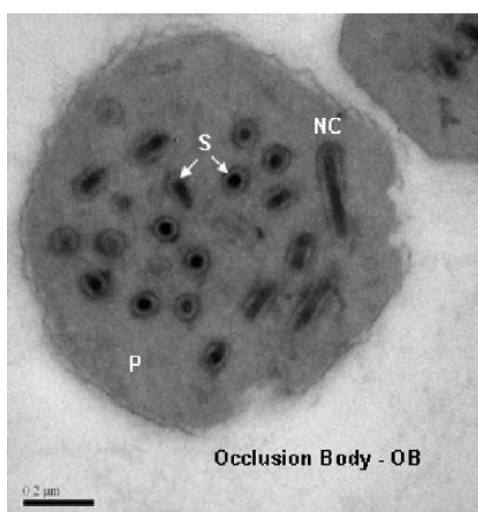
## REVISÃO BIBLIOGRÁFICA

### ***Chrysodeixis includens* nucleopolyhedrovirus (ChinNPV)**

*Chrysodeixis includens* nucleopolyhedrovirus (ChinNPV) é um baculovírus patogênico à lagarta falsa-medideira, *Chrysodeixis* (= *Pseudoplusia*) *includens* (Walker, 1858) (Lepidoptera: Noctuidae).

A partir da década de 80, foram publicados os primeiros estudos que relatam a ocorrência de baculovírus patogênicos a larvas *C. includens*. Esses estudos relatam vírus obtidos de larvas *C. includens* provenientes da cultura de algodão na Guatemala e abordam aspectos como dosagem e resposta da temperatura, efeitos de potenciadores virais, persistência no campo e eficácia de controle (Livingston e Yearian, 1972; Young e Yearian, 1979; McLeod *et al.*, 1982; Young e Yearian, 1982, 1988; Ali e Young, 1991; Zou e Young, 1996; Young, 2001).

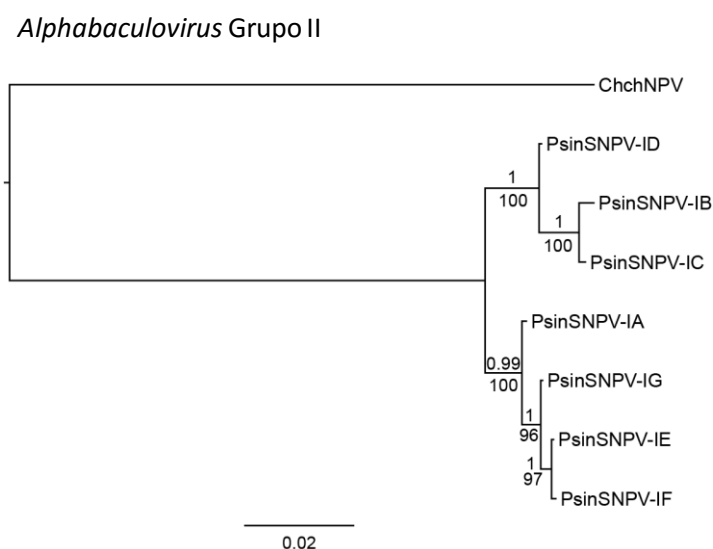
A partir de 2008, sete isolados de ChinNPV (-IA a -IG) obtidos de larvas *C. includens* coletadas em plantações de soja e algodão no Brasil e Guatemala foram identificados e caracterizados no Laboratório de Virologia de Insetos (Cenargen). Esses isolados foram identificados e analisados por microscopia eletrônica e apresentaram forma poliédrica (*Nucleopolyhedrovirus* - NPV) com vírions contendo apenas um nucleocapsídeo por envelope (*Single*) (Alexandre *et al.*, 2010) passando a ser nomeados como *Chrysodeixis includens* *single* *nucleopolyhedrovirus* (ChinSNPV) (Figura 1)



**Figura 1. Corpos de oclusão de *Chrysodeixis includens* NPV.** Poliedros (P) com vírions contendo um único nucleocapsídeo (NC) por envelope (Single – S) (Figura adaptada de (Alexandre *et al.*, 2010).

Análise de perfis de restrição de DNA indicou a presença de variações genéticas entre os sete isolados virais de ChinNPV e a ocorrência de variantes genotípicos dentro de um mesmo isolado (Alexandre *et al.*, 2010; Craveiro *et al.*, 2013). A diversidade genética entre esses isolados resultou na ocorrência de variação fenotípica demonstrada pela diferença significativa da infectividade e patogenicidade em que os isolados ChinNPV-IA, -IE e -IF apresentaram maior virulência que os outros quatro isolados analisados (Alexandre *et al.*, 2010; Craveiro *et al.*, 2013).

A relação filogenética do vírus ChinNPV com outros 56 baculovírus apresentou ChinNPV pertencente ao gênero *Alphabaculovirus* do Grupo II da família *Baculoviridae* (Craveiro *et al.*, 2013). Entre os baculovírus analisados, ChinNPV está proximamente relacionado a *Chrysodeixis chalcites* NPV (ChchNPV) e *Trichoplusia ni* SNPV (TnSNPV) (Craveiro *et al.*, 2013). A árvore filogenética dos sete isolados de ChinNPV (-IA a -IG) apresentou dois grupos monofiléticos, um contendo três isolados coletados no Paraná na safra de 2006 (ChinNPV-IB, -IC e -ID) e o outro contendo o isolado da Guatemala (ChinNPV-IA) agrupado aos isolados coletados no sul e centro-oeste do Brasil nas safras de 2007 e 2008 (ChinNPV-IE, -IF e IG) (Figura 2). Esse padrão de agrupamento indicou que, embora sejam similares, foi observada uma distinção entre os sete isolados (Craveiro *et al.*, 2013).



**Figura 2. Diversidade genética de isolados de ChinNPV-IA a -IG.** Árvore filogenética baseada na concatenação dos genes *late expression factor -8* e *-9* (*lef-8* e *lef-9*), *per os infectivity factor -2* (*pif-2*), *photolyase* (*phr*) e *polyhedrin* (*polh*). ChchNPV foi utilizado como grupo externo (Craveiro *et al.*, 2013).

Há poucos relatos na literatura sobre a caracterização de baculovírus patogênicos a *C. includens* além dos estudos referentes aos sete isolados descritos acima. Dentre eles, podem ser citados o estudo do gene *DNA photolyase* em genomas de baculovírus no qual é relatada a ocorrência de dois isolados de PsinNPV: um proveniente de Los Angeles – EUA e outro proveniente da Guatemala (Xu *et al.*, 2008); e estudos de isolados virais obtidos de larvas do complexo de Heliothinae que foi confirmada a presença do vírus PsinNPV-458 em amostra de larvas obtidas na Colômbia (Rowley *et al.*, 2011).

Recentemente, *C. includens* tem ganhado importância no cenário nacional por apresentar prejuízos significativos na produção de soja (Sosa-Gómez e Bassoi, 2010; Moscardi *et al.*, 2012), o que despertou um maior interesse em caracterizar o baculovírus ChinNPV e avaliar seu potencial para controle biológico de seu inseto hospedeiro.

#### **Lagarta falsa-medideira: *Chrysodeixis includens* (Walker, 1858)**

Por muito tempo, *Chrysodeixis includens* foi referida como *Pseudoplusia includens*. Entretanto, em 2003, ela foi reclassificada no gênero *Chrysodeixis* e, atualmente, a classificação válida dessa espécie é *Chrysodeixis includens* (Goater *et al.*, 2003; Moscardi *et al.*, 2012).

*Chrysodeixis includens*, popularmente conhecida como lagarta falsa-medideira, pertence à ordem Lepidoptera, família *Noctuidae* e subfamília *Plusiinae*. As lagartas dessa espécie apresentam coloração verde claro com linhas brancas longitudinais na lateral do corpo e duas linhas finas dorsais (Silvie *et al.*, 2007). Elas apresentam três pares de falsas pernas abdominais, fazendo com que durante seu deslocamento ocorra intenso movimento do corpo parecendo medir palmos, característica que lhe confere o nome popular de lagarta-medide-palmo ou lagarta falsa-medideira (Figura 3A). As larvas se alimentam das folhas causando um padrão típico de consumo em que são mantidas íntegras as nervuras foliares principais, o que confere um aspecto rendilhado às folhas atacadas (Figura 3B) (Herzog, 1980; Sosa-Gómez e Bassoi, 2010). Na fase adulta, *C. includens* são mariposas com envergadura de 35 mm com as asas dispostas de forma inclinada. As asas possuem coloração cinza-escura e apresentam um pequeno desenho prateado no centro (Figura 3C).



**Figura 3. Fotos do inseto *C. includens*.** (A) Larva sadia. (B) Dano foliar com o típico aspecto rendilhado causado pelas larvas após o consumo das folhas de soja. (C) Mariposa de *C. includens* ([www.bugguide.net](http://www.bugguide.net)).

*C. includens* é um inseto polífono com capacidade de se desenvolver em 174 plantas hospedeiras, pertencentes a 39 famílias (Specht *et al.*, 2015). Entretanto, ela possui preferência e está mais bem adaptada a cultura da soja (Martin *et al.*, 1976; Jost e Pitre, 2002). Durante muito tempo a *C. includens* foi considerada praga secundária da cultura da soja, no entanto, após as safras de 2000/2001 e 2001/2002, o desequilíbrio no agroecossistema provocou mudanças significativas no sistema produtivo tornando *C. includens* praga-chave da cultura da soja. Surtos dessa praga ocorrem individualmente ou associados à lagarta-da-soja (*Anticarsia gemmatilis*) em vários estados brasileiros como: MS, GO, SP e PR (Bueno *et al.*, 2007). Além da soja e algodão, a lagarta *C. includens* é encontrada em outras culturas de importância econômica para o Brasil como: feijão, tomate, batata, fumo, amendoim, girassol, maracujá, alface, couve-flor e outros (Bottimer, 1926; Folsom, 1936; Wolcott, 1936; Hensley *et al.*, 1964; Herzog e Todd, 1980; Bueno *et al.*, 2009; Bernardi *et al.*, 2012).

Esse lepidóptero está geograficamente distribuído apenas no hemisfério ocidental, ocorrendo desde o norte dos Estados Unidos da América (EUA) até o sul da América do Sul e, no Brasil, é encontrado em todas as regiões produtoras de soja do sul ao norte do país (Eichlin e Cunningham, 1978; Alford e Hammond, 1982; Marsaro Júnior *et al.*, 2010; Moscardi *et al.*, 2012; Palma *et al.*, 2015).

O consumo total médio de folhas de soja por lagartas de *C. includens* é bastante variável de 64 a 200 cm<sup>2</sup> (Trichilo e Mack, 1989; Bueno *et al.*, 2011). O aumento populacional e a voracidade no consumo das folhas são aspectos que vêm causando grande preocupação aos sojicultores, pois o ataque da lagarta falsa-medideira pode reduzir significativamente a área foliar e ocasionar intenso dano econômico principalmente durante o período reprodutivo da cultura. Além disso, o controle da lagartas falsa-medideira tem sido considerado difícil por

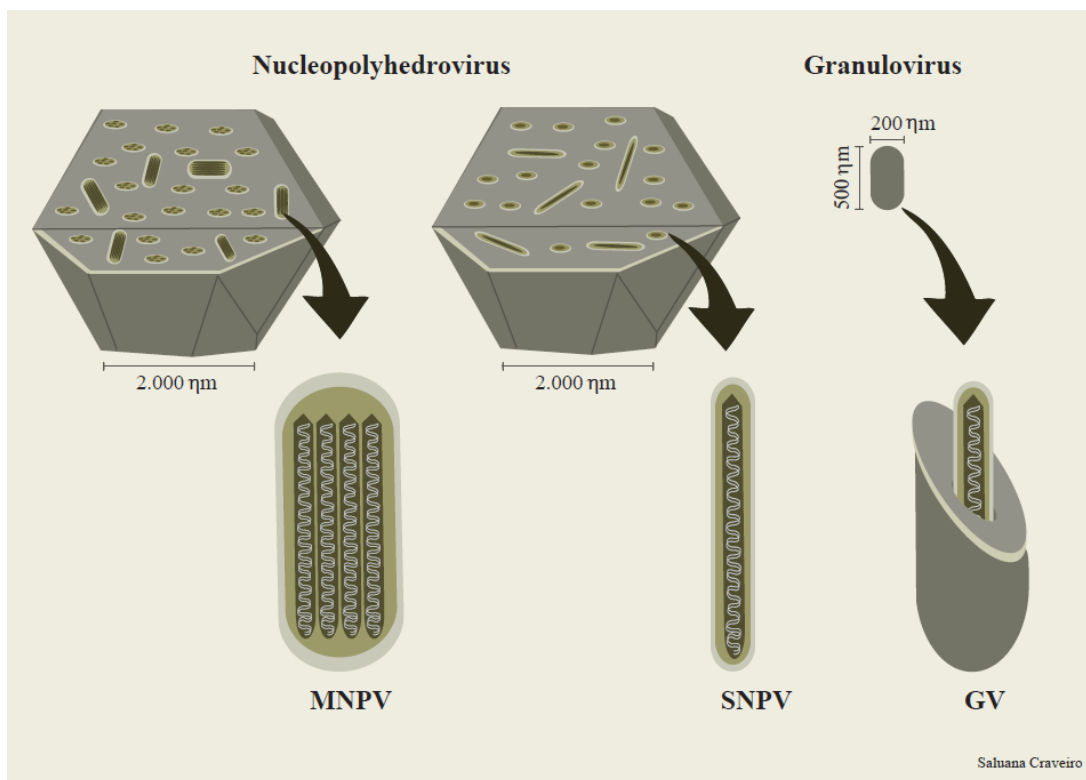
serem mais tolerantes a inseticidas que a lagarta-da-soja (Bernardi *et al.*, 2012) e o hábito de se alimentarem das folhas do terço inferior das plantas dificulta o contato e a ação do inseticida pulverizado no topo das plantas (Sosa-Gómez e Bassoi, 2010).

### **O que são os baculovírus?**

Baculovírus constituem o maior grupo de vírus específicos de artrópodes que infectam insetos da ordem Lepidoptera, Hymenoptera e Diptera. Esses vírus possuem vírions envelopados oclusos em uma matriz protéica formando a partícula viral. Nos vírions, o material genético viral é circundado por uma proteína básica, proteína p6.9, com massa molecular de 7 kDa que neutraliza a carga negativa do DNA e mantém o DNA super condensado. O DNA condensado fica envolto por um capsídeo protéico em forma de bastonete formando o nucleocapsídeo de 30-60 x 250-300 nm, estrutura correspondente à unidade infectiva do vírus chamada de vírus ocluso (*occluded virus* – OV) (Jehle *et al.*, 2006; Ferrelli *et al.*, 2012). Os vírus oclusos (OVs) estão imersos em uma matriz protéica constituindo os corpos de oclusão – *occlusion bodies* (OBs) que variam de 500 a 2000 nm de diâmetro dependendo do vírus (Adams e McClintock, 1991; Bilimoria, 1991; Boucias e Pendland, 2012).

As partículas virais que apresentam corpos de oclusão com forma poliédrica (poliedros) são os *nucleopolyhedrovirus* (NPVs) e os vírus que apresentam corpos de oclusão de forma ovocilíndrica (grânulos) são os *granulovirus* (GVs). Os NPVs possuem vários vírions por cristal protéico contendo apenas um (Single - SNPV) ou vários (Multiple - MNPV) nucleocapsídeos por envelope, enquanto que os GVs possuem um ou, raramente, dois a três vírions por grânulo com um único nucleocapsídeo (Figura 4) (Friesen e Miller, 2001; Theilmann, 2005; Jehle *et al.*, 2006; Herniou *et al.*, 2012; Rohrmann, 2013).



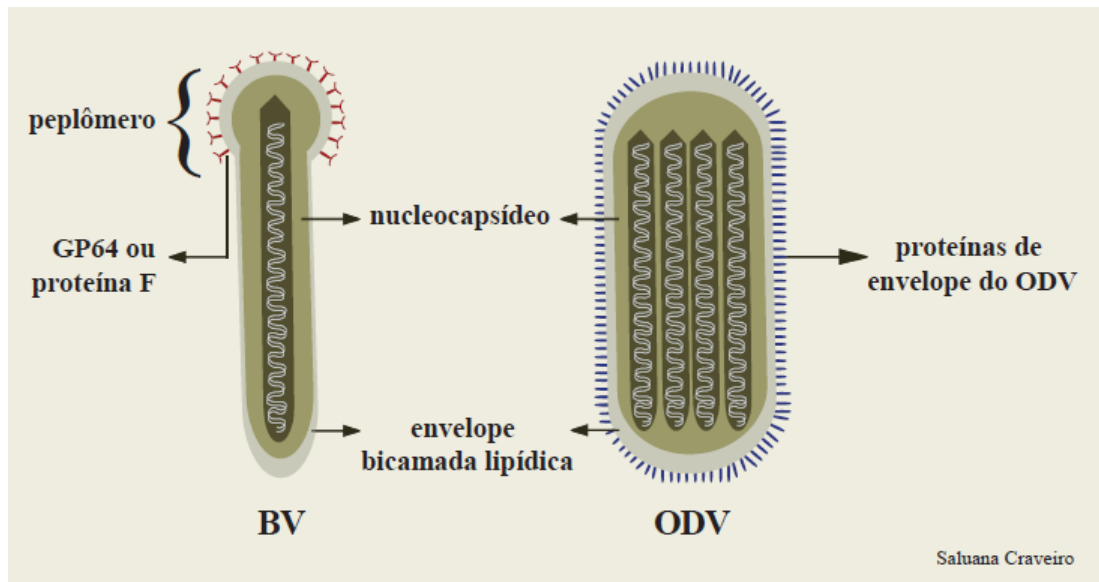


**Figura 4. Desenho esquemático das partículas virais de nucleopolyhedrovirus (NPV) e granulovirus (GV).**

As proteínas poliedrina e granulina possuem massa molecular variando de 29 a 33 kDa correspondendo a cerca de 95% do conteúdo protéico total dos corpos de oclusão de NPVs e GVs, respectivamente (Rohrmann, 2013). Uma exceção entre os baculovirus é o vírus *Culex nigripalpus* NPV (CuniNPV), pois sua principal proteína formadora do corpo de oclusão não é homóloga à poliedrina encontrada nos NPVs patogênicos a lepidópteros e himenópteros (Jehle *et al.*, 2006; Perera *et al.*, 2006).

A transmissão e replicação dos baculovirus ocorrem exclusivamente no estágio larval do inseto hospedeiro. Durante o ciclo de infecção, os baculovirus produzem dois tipos de progênie infecciosas, os vírus derivados de oclusão (*occlusion derived viruses* – ODVs) e as partículas extracelulares (*budded viruses* - BVs) que são idênticas na estrutura e informação genética do nucleocapsídeo, mas diferem na composição do envelope viral e propriedades biológicas (Braunagel e Summers, 1994; Funk, 1997; Friesen e Miller, 2001; Wang *et al.*, 2016). Os ODVs são responsáveis pela infecção oral, transmissão

horizontal entre insetos e são altamente infecciosos às células do intestino médio da larva hospedeira. Os BVs são responsáveis pela infecção sistêmica no inseto hospedeiro e são altamente infecciosos para cultura de células de insetos, mas não para células endoteliais do hospedeiro (Figura 5) (Slack *et al.*, 2007).



**Figura 5. Desenho esquemático das progênies infecciosas encontradas no ciclo de infecção dos baculovírus.** *Budded virus* (BV) responsável pela infecção sistêmica (célula-célula) e *occlusion derived virus* (ODV) responsável pela infecção oral (inseto-inseto).

Os insetos ingerem as partículas virais OBs presentes na superfície das folhas e essas partículas, ao chegarem ao intestino médio do inseto, são submetidas a pH alcalino e ação de proteases que dissolvem as proteínas dos OBs liberando as partículas oclusas – ODVs (Horton e Burand, 1993). Os ODVs atravessam a membrana peritrófica e penetram nas células colunares do intestino médio da lagarta por um processo de fusão de membranas que envolvem proteínas específicas (proteínas PIF) do envelope de ODVs com receptores específicos presentes na célula endotelial do inseto (Horton e Burand, 1993). Ao entrar nas células, os nucleocapsídeos são dirigidos para o núcleo celular ou podem atravessar a célula e atingir o sistema traqueal e outros tecidos do inseto. No núcleo, a replicação do DNA resulta, inicialmente, na produção de BVs que brotam da membrana plasmática da célula hospedeira para infectar as

células adjacentes (Engelhard *et al.*, 1994; Flipsen *et al.*, 1995). Posteriormente, no núcleo da célula infectada, ocorre a produção e oclusão das partículas infecciosas ODVs para a formação dos OBs (Thiem e Cheng, 2009).

A progressão da infecção e a multiplicação das partículas OBs direcionam para uma série de mudanças comportamentais e morfológicas no inseto como a redução na alimentação, retardo do crescimento e descoloração e rompimento do tegumento que culminam na morte da larva e liberação de grande quantidade de OBs no ambiente que servirão de inóculo para infecção de outras larvas (Granados e Williams, 1986; Volkman e Keddie, 1990; Hegedus *et al.*, 2009; Katsuma, 2015).

### **Aplicações biotecnológicas dos baculovírus**

Os baculovírus ocorrem naturalmente no campo e infectam as larvas que se alimentam de folhas contaminadas. A alta virulência e especificidade ao inseto hospedeiro tornam os baculovírus uma ferramenta potencial e em constante avanço no controle de pragas agrícolas. Assim, os baculovírus vêm ganhando importância como componentes integrantes dos programas de manejo integrado de pragas (MIP), sendo uma alternativa biológica aos inseticidas químicos (Moscardi *et al.*, 2011; Szewczyk *et al.*, 2011).

O controle de insetos-praga é realizado, principalmente, com uso de inseticidas químicos que muitas vezes são aplicados de forma abusiva e indiscriminada, desencadeando uma série de problemas, como desequilíbrio biológico com a eliminação de inimigos naturais e surgimento de pragas secundárias ou novas pragas, resistência de pragas, intoxicação humana e contaminação do meio ambiente (Pedigo e Rice, 2014).

Desta forma, a utilização de métodos alternativos como o uso do controle biológico torna-se essencial para a diminuição das consequências indesejáveis provenientes do uso de inseticidas químicos.

Os baculovírus estão sendo utilizados mundialmente como biopesticidas no controle de insetos-praga em agricultura, áreas florestais e sistemas de produção de vegetais (Copping e Menn, 2000; Szewczyk *et al.*, 2006; Souza *et al.*, 2007; Erlandson, 2008; Szewczyk *et al.*, 2009). O biocontrole baseado em baculovírus tem sido efetivo contra pragas como a lagarta-da-soja (*Anticarsia gemmatilis*) (Moscardi, 1999, 2007; Sosa-Gómez *et al.*, 2008), lagarta do

algodão (*Helicoverpa zea* e *Helicoverpa armigera*) (Sun e Peng, 2007; Srinivasa *et al.*, 2008), lagarta-do-cartucho do milho (*Spodoptera frugiperda*) (Valicente e Cruz, 1991), traça das maçãs (*Cydia pomonella*) (Lacey e Unruh, 2005; Fritsch *et al.*, 2007; Vincent *et al.*, 2007), mariposa cigana (*Lymantria dispar*) (Cook *et al.*, 2003), mandarová da mandioca (*Erinnyis ello*) (Bellotti *et al.*, 1999), lagarta do álamo (*Condylorrhiza vestigialis*) (Machado, 2006) (Braúlio Santos, comunicação pessoal), entre outras.

O programa que utiliza o vírus AgMNPV no controle da lagarta-da-soja nas plantações de soja na América Latina tem sido o mais expressivo de todo o mundo. No Brasil, esse programa chegou a atingir 2 milhões de hectares durante os anos 2003/2004. Posteriormente, o surgimento de pragas emergentes (exemplo: lagarta falsamedideira), ocasionado pelo desequilíbrio ecológico decorrente do uso irregular e indiscriminado de inseticidas químicos, provocou um declínio na utilização desse agente no controle da lagarta-da-soja (Moscardi *et al.*, 2011; Szewczyk *et al.*, 2011).

A alta virulência dos baculovirus em um amplo espectro de hospedeiros, a estabilidade no meio ambiente e sua alta especificidade aos insetos hospedeiros são vantagens da utilização desses vírus como bioinseticidas. Entretanto, desvantagens como morte lenta do inseto, espectro de hospedeiros limitado e complexa produção massal do vírus causam resistência para uma utilização mais ampla do baculovírus como agente de controle biológico.

Além de sua utilização como agentes de controle biológico, os baculovírus são altamente eficientes como vetores de expressão de proteínas em células de insetos (*baculovirus expression vector system* – BEVS) sendo utilizados para a produção de proteínas heterólogas, produção de vacinas e terapia gênica (van Oers *et al.*, 2015). Vacinas animais e humanas contra peste suína, papilomavirus e influenza e terapias gênicas contra câncer de próstata e deficiência de lipoproteína lipase estão sendo produzidas em sistemas de vetores de expressão de baculovírus e usadas comercialmente (Felberbaum, 2015). Atualmente, BEVS está sendo ativamente usado pelas mais importantes indústrias de biotecnologia para a produção de novas vacinas e produtos de terapia gênica (van Oers *et al.*, 2015; Kwang *et al.*, 2016). Além disso, estudos recentes demonstraram uma eficácia e segurança no silenciamento de genes

utilizando baculovírus para o endereçamento de RNA interferente (Makkonen *et al.*, 2015).

### **Evolução da Família *Baculoviridae***

A origem evolutiva dos baculovirus pode ser explicada por diferentes hipóteses. Rohrmann (1986) propôs que os baculovirus originaram com os insetos da ordem Lepidoptera e posteriormente, por transferência horizontal, atingiram outras ordens de insetos. Posteriormente, Federici (1997) propôs que a origem dos baculovirus remonta a origem dos artrópodes com a cocladogênese do vírus e seu hospedeiro. Em 2004, Herniou e colaboradores sugeriram uma terceira hipótese em que ancestrais dos baculovirus, por transferência horizontal, infectaram diferentes ordens de insetos propondo uma antiga coevolução do vírus com seu hospedeiro que levou, em seguida, ao progresso da especiação das diferentes linhagens de baculovirus para as diferentes ordens dos hospedeiros (Herniou *et al.*, 2004).

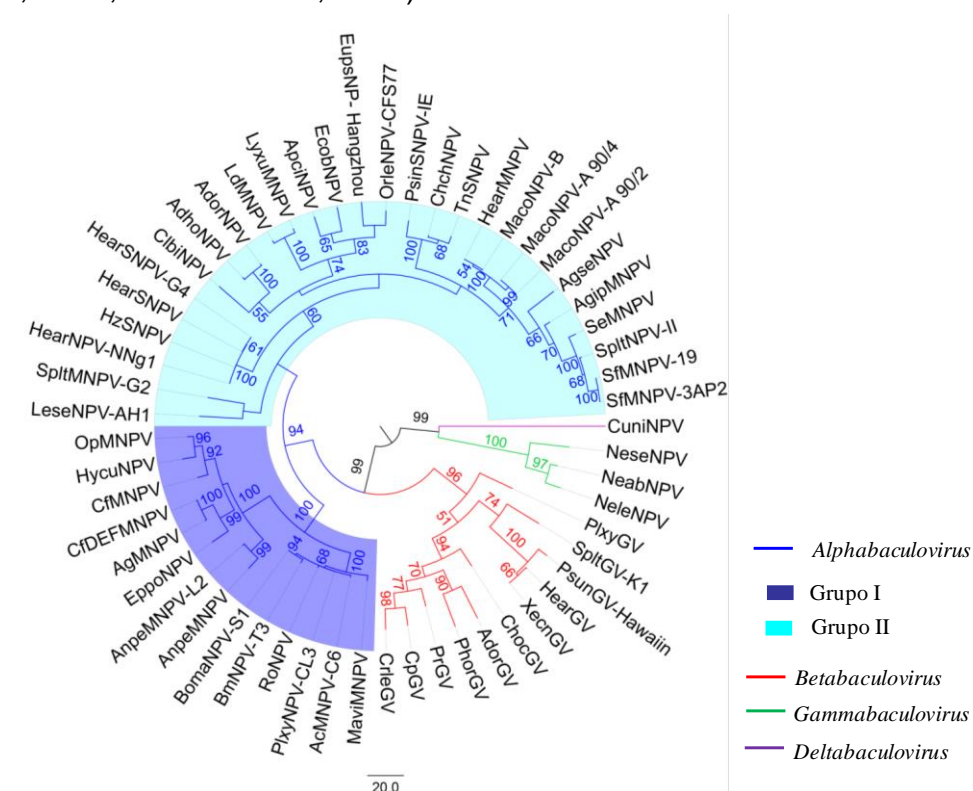
Thézé *et al.* (2011) relataram o surgimento e diversificação dos baculovirus no Carbonífero, período da era Paleozóica. Tendo em vista que os ancestrais dos vírus já infectavam os primeiros insetos que surgiram no período Devoniano, esses dados corroboram a terceira hipótese (Herniou *et al.*, 2004) que sugere uma coevolução do vírus com seu inseto hospedeiro.

Os baculovírus pertencem à família *Baculoviridae* e são genética e morfologicamente distintos das outras famílias de vírus estando mais proximamente relacionado aos nudivírus (Thézé *et al.*, 2011). A família *Baculoviridae* possui uma complexidade de forma e função sugerindo uma longa linhagem evolutiva (Slack e Arif, 2007).

Até 2005, o VIII Relatório do Comitê Internacional de Taxonomia de Vírus (Fauquet *et al.*, 2005) classificou taxonomicamente a família *Baculoviridae* em dois gêneros: *Nucleopolyhedrovirus* (NPV) e *Granulovirus* (GV). Por muitos anos, a filogenia dos baculovirus foi baseada no alinhamento de uma única sequência nucleotídica ou peptídica sendo os genes *dnapol*, *egt*, *gp41*, *chiA*, *cath*, *lef2*, *gp37* e, principalmente, o *polyhedrin/granulin* (*polh/gran*) os mais utilizados nas análises filogenéticas (Herniou *et al.*, 2003; Herniou e Jehle, 2007). Com o passar dos anos, a filogenia dos baculovirus pode ser mais bem elucidada

com o avanço das técnicas de seqüenciamento de genomas completos (Herniou e Jehle, 2007).

Herniou *et al.* (2003), em um estudo sobre a evolução dos baculovirus baseado na análise filogenética de seqüências genômicas completas de 13 baculovirus, apresentaram quatro grupos distintos: CuniNPV, GVs, NPVs Grupo I e NPVs Grupo II. Esses resultados estavam em discordância com a classificação taxonômica presente no momento, pois o vírus CuniNPV formava um clado independente não estando proxivamente relacionado aos outros NPVs de lepidópteros. Em continuidade, a análise filogenética baseada em 29 genes comuns a 29 baculovirus também apresentou discordância indicando que a classificação taxonômica da família *Baculoviridae* deveria ser revisada (Jehle *et al.*, 2006). Assim, foi proposta uma nova classificação que reorganiza a família *Baculoviridae* em quatro gêneros: *Alphabaculovirus*, *Betabaculovirus*, *Gammabaculovirus* e *Deltabaculovirus* (Figura 6) (Jehle *et al.*, 2006; Carstens e Ball, 2009; Herniou *et al.*, 2012).



**Figura 6. Filogenia da família *Baculoviridae*.** Relações filogenéticas entre os baculovirus divididos em quatro gêneros: *Alphabaculovirus* (Grupos I e II), *Betabaculovirus*, *Gammabaculovirus* e *Deltabaculovirus*.

O gênero *Alphabaculovirus* inclui NPVs específicos de lepidópteros, com OBs de forma poliédrica de 0,5 a 5 µm de tamanho e genoma de 100 a 180 kpb. O *Betabaculovirus* inclui os GVs específicos de lepidópteros com OBs de forma ovocilíndrica de aproximadamente 0,3 x 0,5 µm e genoma de tamanho similar aos *Alphabaculovirus*. O *Gammabaculovirus* inclui os vírus específicos de himenópteros e atualmente é constituído por *Neodiprion lecontei* NPV (NeleSNPV), *Neodiprion sertifer* NPV (NeseSNPV) e *Neodiprion abietis* NPV (NeabNPV), com OBs de 0,4-1,1 µm e genoma de 82 a 86 kpb (Jehle *et al.*, 2006). O *Deltabaculovirus* inclui os vírus específicos de dípteros atualmente representado pelo CuniNPV com OBs de 0,4 µm de diâmetro e genoma de 108 kpb (Afonso *et al.*, 2001). Os *Alphabaculovirus* diferem entre seus dois grupos: Grupo I e Grupo II (Monsma *et al.*, 1996; Hefferon *et al.*, 1999; Pearson *et al.*, 2000; Westenberg *et al.*, 2007). Os NPVs Grupo I usam a GP64 como a proteína de fusão dos BVs enquanto que os NPVs do Grupo II utilizam-se da proteína F para a transmissão de partículas BVs entre células do inseto hospedeiro (Ijkel *et al.*, 2000; Pearson *et al.*, 2000). Esses dois grupos diferem também pelo seu conteúdo de genes com a presença de 11 outros genes (ORFs de AcMNPV: Ac1 - *ptp*, Ac16 - BV-ODV26, Ac27 - *iap-1*, Ac30, Ac42 - *gta*, Ac72, Ac73, Ac114, Ac124, Ac132, Ac151 - *ie2*) que podem ser encontrados apenas nos baculovirus do Grupo I (Rohrmann, 2013). No gênero *Gammabaculovirus*, não foram identificados os genes que codificam proteínas constituintes de BVs, proteína F ou GP64, sugerindo a ausência desse fenótipo nesse grupo (Jehle *et al.*, 2006).

Miele e colaboradores (2011) demonstraram que os quatro gêneros dos baculovirus têm acumulado um grande número de genes durante a evolução. Uma menor diversidade de genes foi observada nos *Alphabaculovirus* do Grupo I e *Gammabaculovirus* quando comparados aos outros baculovirus. Isso ocorre, pois, desde o surgimento de seu ancestral comum, esses grupos tiveram menos tempo para incorporar novas sequências. Assim, na história evolutiva desses vírus, os *Alphabaculovirus* do Grupo I e *Gammabaculovirus* são linhagens mais recentes que os outros clados (Miele *et al.*, 2011).

### **Genoma dos baculovirus**

Os baculovirus possuem o genoma de DNA fita-dupla circular de, aproximadamente, 80 a 180 kpb dependendo da espécie. O genoma de um único

baculovirus possui de 89 a 183 ORFs preditas (*NeleNPV* e *Pseudaletia unipuncta* GV, respectivamente) em ambas as fitas e orientações (*forward* e *reverse*). O critério geralmente utilizado para identificação de ORFs preditas de baculovírus é a determinação de sequências nucleotídicas que codifiquem polipeptídeos maiores que 50 aminoácidos com o mínimo de sobreposição entre as ORFs (Ferrelli *et al.*, 2012). Além disso, como padrão, os genes *polyhedrin/granulin* são identificados como a primeira ORF predita e, a partir de então, os genes são numerados sequencialmente com a mesma orientação na direção horária (*forward*) e na orientação inversa na direção anti-horária (*reverse*).

Atualmente, 71 genomas de baculovírus estão sequenciados e depositados no *Genbank* do *National Center for Biotechnology Information* – NCBI. O conteúdo total de genes encontrados nos baculovírus é de, aproximadamente, 900 genes. Dentre eles, há 37 genes denominados de *core genes* que estão presentes no genoma de todos os baculovírus (Tabela 1). Os *core genes* são ancestrais e altamente conservados e representam 3% do conteúdo genético viral (Miele *et al.*, 2011; Rohrmann, 2013). Esses genes estão envolvidos nos diferentes estágios do ciclo viral como: replicação do DNA, transcrição do RNA, composição protéica das partículas virais, interação com proteínas dos hospedeiros, infectividade oral e outros. Quando comparados a outros, os *core genes* possuem menor tolerância às mutações que podem implicar na perda da viabilidade viral (Herniou *et al.*, 2003; Miele *et al.*, 2011; Ferrelli *et al.*, 2012).

Os genes de baculovírus não estão agrupados no genoma de acordo com sua função ou momento da transcrição durante os diferentes estágios do ciclo de infecção. Os genes de baculovírus transcritos no estágio precoce da infecção são precedidos por promotores TATA-box e/ou CAGT, que são motivos encontrados nos promotores do genoma dos hospedeiros e são transcritos pela RNA polimerase II das células de insetos. Os genes das fases tardias são expressos pela RNA polimerase viral e podem ser precedidos por promotores contendo o motivo DTAAG. Entretanto, muitos genes contêm nos promotores ambos os motivos das fases precoce e tardia que são expressos durante toda a infecção.

A primeira sequência completa de genoma de baculovírus publicado foi de *Autographa californica* NPV (AcMNPV) (Ayres *et al.*, 1994). Desde então,



muitos genomas de baculovírus estão sendo publicados, o que tem proporcionado um melhor entendimento da biologia molecular desses vírus.

**Tabela 1. Genes conservados em todos os baculovírus (*core genes*)**

| <b>Genes</b>                             | <b>AcMNPV (ORF)</b> |
|--|---------------------|
| <b>Replicação e processamento do DNA</b> |                     |
| <i>lef-1</i>                             | 14                  |
| <i>lef-2</i>                             | 6                   |
| <i>DNA polymerase</i>                    | 65                  |
| <i>helicase</i>                          | 95                  |
| <i>alcaline nuclease</i>                 | 133                 |
| <b>Transcrição e RNA polimerase</b>      |                     |
| <i>lef-4</i>                             | 90                  |
| <i>lef-5</i>                             | 50                  |
| <i>lef-8</i>                             | 62                  |
| <i>lef-9</i>                             | 40                  |
| <i>p47</i>                               | 99                  |
| <i>vlf-1</i>                             | 77                  |
| <b>Genes estruturais</b>                 |                     |
| <i>p6.9</i>                              | 100                 |
| <i>vp39</i>                              | 89                  |
| <i>vp1054</i>                            | 54                  |
| <i>vp91/p95</i>                          | 83                  |
| <i>gp41</i>                              | 80                  |
| <i>odv-ec43</i>                          | 109                 |
| <i>p49</i>                               | 142                 |
| <i>odv-e18</i>                           | 143                 |
| <i>desmoplakin</i>                       | 66                  |
| <i>odv-e27</i>                           | 144                 |
| <b>Fatores de infectividade oral</b>     |                     |
| <i>pif-0/p74</i>                         | 138                 |
| <i>pif-1</i>                             | 119                 |
| <i>pif-2</i>                             | 22                  |
| <i>pif-3</i>                             | 115                 |
| <i>pif-4/19k/odv-e28</i>                 | 96                  |
| <i>pif-5/odv-e56</i>                     | 148                 |
| <i>pif-6</i>                             | 68                  |
| <b>Enzimas</b>                           |                     |
| <i>38k phosphatase</i>                   | 98                  |
| <i>p33</i>                               | 92                  |
| <i>ubiquitin</i>                         | 53                  |
| <b>Outros</b>                            |                     |
| <i>Ac78</i>                              | 78                  |
| <i>Ac81</i>                              | 81                  |
| <i>Ac93</i>                              | 93                  |
| <i>Ac101</i>                             | 101                 |

### **baculovirus intra e interespecífica**

A análise comparativa da sequência genômica revelou alta diversidade em relação ao tamanho, organização e conteúdo gênico entre os genomas de baculovirus (Herniou *et al.*, 2003; Herniou e Jehle, 2007; van Oers e Vlak, 2007; Miele *et al.*, 2011). A alta variabilidade genômica reflete claramente na diversidade fenotípica observada entre os quatro gêneros da família *Baculoviridae*. Além da marcante diferença morfológica das partículas OBs de granulovirus (GV) e nucleopolyhedrovirus (NPV), variações fenotípicas são também observadas entre NPVs que infectam diferentes ordens de insetos. Enquanto os *Alphabaculovirus* apresentam infecção celular em praticamente todos os tecidos do inseto hospedeiro (Katsuma *et al.*, 2012), nos *Gamma-* e *Deltabaculovirus* a infecção e replicação do vírus se restringe às células do intestino médio do inseto (Becnel *et al.*, 2001; Moser *et al.*, 2001). Outro exemplo, é *Culex nigripalpus* NPV (CuniNPV) que não contém genes homólogos às proteínas poliedrina/granulina possuindo uma proteína de 90kDa revelando formas globulares para as partículas OBs e não poliédricas como nos demais nucleopolyhedrovirus (Afonso *et al.*, 2001).

Além da diversidade genética entre as diferentes espécies, variações genéticas são altamente frequentes em populações de baculovirus. Essas variações são facilmente mantidas devido à característica típica dos baculovirus de concentrar mais de um genótipo em uma única partícula viral, como o que ocorre nos nucleopolyhedrovirus que possuem vários nucleocapsídeos oclusos em um único poliedro (Herniou e Jehle, 2007; Clem e Passarelli, 2013).

Os variantes genotípicos, facilmente detectados por análises de perfis de restrição (*Restriction Fragment Length Polymorphism* - RFLP), geralmente exibem variações no fenótipo relacionadas, principalmente, a patogenicidade, tempo de morte e produção de partículas BVs e OBs (Stiles e Himmerich, 1998; Kamiya *et al.*, 2004; Cory *et al.*, 2005; Ogembo *et al.*, 2007; Harrison *et al.*, 2008). A heterogeneidade de fenótipos é comumente mantida em populações de campo e, devido a isso, acredita-se que a diversidade genética traz vantagens para a adaptação, evolução e tempo de sobrevivência do baculovirus no campo.

Uma pequena proporção do genoma viral codifica genes que são comuns a todos os baculovirus (Jehle *et al.*, 2006). No genoma dos baculovirus, como

nos outros vírus de DNA, a recombinação homóloga, a perda e duplicação de genes e a transferência lateral de genes para outros vírus, bactérias ou células eucarióticas são os principais mecanismos responsáveis pela formação e diversidade dos genomas (Shackelton e Holmes, 2004).

As variações são mutações pontuais, substituições, inserções e deleções que ocorrem por todo genoma, mas se concentram em regiões específicas que representam “*hot spots*” de hipervariabilidade ocorrendo, geralmente, nas regiões de repetição homóloga - *homologous repeat regions (hrs)* e nas ORFs repetidas dos baculovirus – *baculovirus repeated ORF (bro)* (Hayakawa *et al.*, 2000; Li *et al.*, 2002; De Jong *et al.*, 2005). As *hrs* são regiões intergênicas formadas por sequências repetitivas que podem funcionar como ativadores *in cis* de *enhancers* e de origens de replicação viral do DNA (Guarino *et al.*, 1986; Guarino e Summers, 1986; Pearson *et al.*, 1992; Hilton e Winstanley, 2007) e os genes *bro* constituem uma família de múltiplos genes encontrados nos baculovirus e em outros vírus de invertebrados de DNA fita-dupla (Bideshi *et al.*, 2003).

Os baculovírus consistem em uma das mais diversas famílias de vírus e essa alta variabilidade pode ser decorrente, principalmente, da co-evolução e co-diversificação do vírus com seu inseto hospedeiro. Essa diversidade pode representar um importante fator para o controle biológico de insetos-praga pois pode permitir a seleção de fenótipos com características específicas favoráveis para a produção comercial do vírus como princípio ativo de bioinseticidas.

**CAPÍTULO 1: Diversidade genética de *Chrysodeixis  
includens* NPV baseada nos core genes:  
*pif-2, lef-9 e helicase***

Este capítulo inclui resultados de um artigo publicado em 2013 na *J. Invertebr. Pathol.* 114:258-267:

***Pseudoplusia includens* single nucleopolyhedrovirus: Genetic diversity, phylogeny and hypervariability of the *pif-2* gene**

Saluana R. Craveiro, Fernando L. Melo, Zilda Maria A. Ribeiro, Bergmann M. Ribeiro, Sônia N. Bão, Peter W. Inglis, Maria Elita B. Castro\*

## 1.1. Introdução

Variações genéticas são altamente frequentes em populações de campo de baculovírus e uma diversidade intraespecífica foi claramente observada em análises de perfis de restrição de sete isolados de *Chrysodeixis includens* NPV (ChinNPV-IA a ChinNPV-IG) (Alexandre *et al.*, 2010).

Em estudos prévios que relatam as relações filogenéticas entre os sete isolados de ChinNPV, também foram observadas variações nas sequências parciais de cinco genes dentre os quais o gene *pif-2* recebeu destaque pela sua alta variação (Craveiro, 2012).

O gene *pif-2* codifica uma proteína expressa na membrana do envelope das partículas ODVs. Os corpos de oclusão (ODVs) são partículas virais responsáveis pela infecção oral e são altamente infecciosas às células do intestino médio do inseto hospedeiro. PIF2 pertence a uma família de proteínas que são essenciais para infectividade oral e são nomeadas *per os infectivity factors* (PIFs) (Mu *et al.*, 2014). Oito proteínas pertencem à família PIF: P74 (PIF0) (Kuzio *et al.*, 1989; Faulkner *et al.*, 1997; Yao *et al.*, 2004; Slack *et al.*, 2010), PIF1 (Kikhno *et al.*, 2002), PIF2 (Pijlman *et al.*, 2003; Fang *et al.*, 2006), PIF3 (Ohkawa *et al.*, 2005), PIF4 (Fang *et al.*, 2009), PIF5 (ODV-E56) (Harrison *et al.*, 2010; Sparks *et al.*, 2011; Xiang *et al.*, 2011), PIF6 (ac68) (Nie *et al.*, 2012) e PIF7 (ac83) (Zhu *et al.*, 2013).

A deleção de um dos genes *pif* provoca a interrupção da infecção e da expressão de genes virais nas células do intestino médio do hospedeiro (Haas-Stapleton *et al.*, 2004; Ohkawa *et al.*, 2005; Song *et al.*, 2008). Quatro das proteínas PIFs, PIF1, PIF2, PIF3 e P74, estão presentes no envelope dos ODVs na forma de um complexo (Peng *et al.*, 2010). As proteínas P74, PIF1 e PIF2 mediam a ligação específica dos ODVs com as células epiteliais e se apresentam como proteínas de ligação do envelope das ODVs (Haas-Stapleton *et al.*, 2004; Ohkawa *et al.*, 2005).

A proteína PIF2 com 44 kDa foi a primeira a ser identificada como um fator essencial da infecção *per os* de ODVs em *Spodoptera exigua* (Se) NPV (Pijlman *et al.*, 2003; Slack e Arif, 2006). PIF2 possui um motivo transmembrana N-terminal exposto na superfície dos ODVs indicando que essa proteína pode funcionar como âncora transmembrânica (Slack e Arif, 2006).

Os oito genes *pif* são *core genes* indicando que a infecção no intestino médio é um processo ancestral e altamente conservado (Herniou *et al.*, 2003; Braunagel e Summers, 2007; van Oers e Vlak, 2007; Fang *et al.*, 2009; Mu *et al.*, 2014). Outro gene conservado é o *late expression factor 9 (lef-9)* expresso na infecção precoce e essencial para expressão dos promotores dos genes tardios e muito tardios (Lu e Miller, 1995; Todd *et al.*, 1995; Li *et al.*, 1999). As proteínas LEF-8 e LEF-9 são subunidades da RNA polimerase viral e formam o sítio catalítico da enzima (Passarelli *et al.*, 1994; Guarino *et al.*, 1998).

O gene *pif-2* é a sequência ancestral mais conservada dos baculovírus e, junto com o gene *lef-9*, apresenta menor variabilidade comparada aos outros *core genes*, enquanto que os genes *desmoplaquina* e *helicase (p143)* são os mais variáveis (Miele *et al.*, 2011). P143 é essencial para a replicação do DNA viral. Essa proteína é capaz de se ligar inespecificamente ao DNA fita simples ou dupla e desta forma é transportada para dentro do núcleo pela proteína LEF-3 (McDougal e Guarino, 2000, 2001; Yu e Carstens, 2010). Além disso, o gene *helicase* foi identificado por ser um fator que influencia o espectro de hospedeiros dos baculovírus (Maeda *et al.*, 1993; Croizier *et al.*, 1994).

Neste estudo, foi realizada a análise comparativa da variação presente nos genes *pif-2*, *lef-9* e *helicase* de isolados geográficos de diferentes baculovírus: *Helicoverpa armigera* (Hear) NPV, *Mamestra configurata* (Maco) NPV, ChinNPV e *Spodoptera frugiperda* (Sf) MNPV. Além disso, foram realizadas análises de pressão de seleção do gene *pif-2* de ChinNPV para a identificação de sítios sob seleção positiva.

## 1.2. Materiais e Métodos

### 1.2.1. Vírus

Os isolados de *Chrysodeixis includens* -NPV (IA a IG), doados pelo Dr. Flávio Moscardi, da Embrapa Soja (Londrina-PR), foram obtidos a partir de larvas *C. includens* infectadas e coletadas em plantações de soja e algodão do Brasil e Guatemala (Tabela 2). Após a purificação das partículas virais, os isolados virais foram depositados na Coleção de Vírus de Invertebrados da Embrapa Recursos Genéticos e Biotecnologia e estão cadastrados no Sistema de Informação AleloMicro brasileiro e identificados sob os códigos listados na Tabela 2.

**Tabela 2. Isolados de ChinNPV obtidos de larvas *C. includens* infectadas****1.2.2. Purificação de OBs**

| <b>Isolados Virais</b> | <b>Data de Coleta</b> | <b>Local de Coleta</b> | <b>Instituição</b>         | <b>Cultura</b> | <b>AleloMicro Código</b> |
|------------------------|-----------------------|------------------------|----------------------------|----------------|--------------------------|
| IA                     | 1972                  | Guatemala              | University of Arkansas/EUA | Algodão        | BRM 005 102              |
| IB                     | Jan./2006             | Londrina, PR           | Embrapa Soja               | Soja           | BRM 005 103              |
| IC                     | Jan./2006             | Maringá, PR            | Embrapa Soja               | Soja           | BRM 005 104              |
| ID                     | Fev./2006             | Iguaraçu, PR           | Embrapa Soja               | Soja           | BRM 005 105              |
| IE                     | Fev./2007             | Iguaraçu, PR           | Embrapa Soja               | Soja           | BRM 005 106              |
| IF                     | Fev./2008             | Dourados, MS           | Embrapa Soja               | Soja           | BRM 005 107              |
| IG                     | Fev./2008             | Sertanópolis, PR       | Embrapa Soja               | Soja           | BRM 005 108              |

Os poliedros dos isolados de ChinNPV (IA a IG) foram purificados de acordo com protocolo descrito por Maruniak (1986) e utilizados para estoque das amostras e extração de DNA viral. Larvas de *C. includens* foram maceradas em tampão de homogeneização (ácido ascórbico 1%; SDS 2%; Tris-HCl 0,01 M, pH 7,8 e EDTA 0,001 M). O homogeneizado resultante foi filtrado em 6 camadas de gaze e centrifugado a 10.000 rpm (Sorvall RC-5B, rotor SS-34) por 10 min. Sucessivas centrifugações de 12.000 rpm (Sorvall RC-5B, rotor SS-34) por 12min foram realizadas primeiramente com o sedimento ressuspense em tampão TE (Tris 0,01M, pH 7,8 e EDTA 0,001 M) e SDS 0,5% e em seguida com o sedimento ressuspense em tampão TE e NaCl 0,5 M. O material resultante da centrifugação foi ressuspense em água deionizada Milli-Q e foram retirados 5mL para serem aplicados em gradiente de sacarose contínuo de 39-65% preparado em tampão TE a uma densidade de 1,17-1,32 g / mL. O material foi centrifugado a 24.000 rpm (Sorvall OTD 75 U, rotor AH-627) por 40min a 4°C, e a banda correspondente aos OBs, posicionada no terço inferior do tubo, foi coletada com pipeta Pasteur, diluída 5 vezes em tampão TE e centrifugada a 12.000 rpm (Sorvall RC-5B, rotor SS-34) por 15min a 4°C. A ressuspensão final foi feita em água deionizada Milli-Q autoclavada e os OBs foram armazenados a -20°C.

### 1.2.3. Extração de DNA a partir de partículas OBs purificadas

A extração de DNA viral foi baseada no protocolo descrito por O'Reilly *et al.* (1994). OBs purificados, na concentração de  $1 \times 10^9$  OB/mL, foram solubilizados com solução alcalina 1x, pH 10,9 (estoque 3X: Na<sub>2</sub>CO<sub>3</sub> 0,3 M; NaCl 0,51 M e EDTA 0,03 M) e incubados a 37°C por 30min. Após a verificação da solubilização por microscópio óptico, foram adicionados SDS 1% e 500µg/mL de proteinase K e incubados a 37°C por 16h. Para a extração do DNA viral, foi adicionado ao sobrenadante o mesmo volume de fenol saturado com tampão TE. As fases foram homogeneizadas invertendo-se delicadamente os tubos por 3-5min e o material centrifugado a 12.000 rpm (Eppendorf 5410 – rotor fixo), por 2min. A fase aquosa foi transferida para outro microtubo e o mesmo procedimento foi então realizado para a extração com fenol:clorofórmio:álcool isoamílico (25:24:1) e clorofórmio:álcool isoamílico (24:1). Após a extração, o DNA foi precipitado com dois volumes de etanol absoluto gelado e 10% do volume inicial de acetato de sódio 3 M, pH 5,2 e colocado a -20 °C por 16h. O DNA precipitado foi centrifugado a 12.000 rpm (Eppendorf 5410 – rotor fixo) por 30 min, lavado com etanol 70% gelado e centrifugado a 12.000 rpm por mais 10 min. O DNA foi solubilizado em tampão TE autoclavado e, após adição de RNase (10 µg / mL), incubado a 37°C por 1h. A qualidade do DNA extraído foi determinado por electroforese em gel de agarose 0,5% e quantificado usando Qubit v. 2,0 Fluorometer (Invitrogen) de acordo com as instruções do fabricante.

### 1.2.4. Determinação da sequência nucleotídica dos genes *pif-2*, *lef-9* e *helicase* dos isolados de ChinNPV

A técnica de pirosequenciamento (454 – Roche) foi utilizada para o sequenciamento individual do genoma completo de sete isolados de ChinNPV (IA a IG). Com o auxílio do programa *Geneious Pro v.6.1.6* (Biomatters Ltd., New Zealand) e utilizando como referência os genes *pif-2*, *lef-9* e *helicase* de ChchNPV, foi determinada a sequência nucleotídica completa desses três genes para os isolados de ChinNPV utilizando as leituras obtidas do pirosequenciamento. Após a determinação das sequências, os genes *pif-2*, *lef-9* e *helicase* de cada isolado de ChinNPV foram depositados no *GenBank* (Tabela 3).



### 1.2.5. Genomas virais e análises de diversidade

Para análise da variação genética e o cálculo das taxas evolutivas, foram utilizadas as sequências completas dos genes *pif-2*, *lef-9* e *helicase* dos isolados de ChinNPV. Adicionalmente, foram analisadas as sequências de outros baculovírus obtidas do *GenBank*: 4 isolados de *Helicoverpa armigera* (Hear) NPV, 4 isolados de *Spodoptera frugiperda* (Sf) MNPV e 3 isolados de *Mamestra configurata* (Maco) NPV (Tabela 3). Dentre esses isolados, clones de um mesmo isolado geográfico foram analisados: HearNPV-C1 e HearNPV-G4 (China), SfMNPV-B e SfMNPV-G (Nicarágua) e MacoNPV-A90/2 e MacoNPV-A90/4 (Canadá). ChinNPV-IE foi também comparado com as sequências gênicas de *Trichoplusia ni* (Tn) SNPV e *Chrysodeixis chalcites* (Chch) NPV (Tabela 3). As sequências *pif-2*, *lef-9* e *helicase* foram alinhadas com auxílio do programa *MUSCLE* v.3.5 (Edgar, 2004). Para a análise de variações nos genes *pif-2*, *lef-9* e *helicase* nos diferentes baculovírus analisados, foram feitos os cálculos da Diversidade nucleotídica ( $\pi$ ) (*DNASP* v.5) e Distância PAM (*Point accepted mutation*) (*MEGA* v. 5.0.5 - (Tamura *et al.*, 2011). A Diversidade nucleotídica ( $\pi$ ) é o grau de polimorfismos dentro de populações em que, por meio da análise do alinhamento da sequência nucleotídica dos indivíduos, é calculada a média do número dos diferentes nucleotídeos por sítio. A Distância PAM é a distância evolutiva entre pares de sequência, ou seja, é a probabilidade de duas sequências terem divergido de um ancestral comum e é calculada utilizando a matriz de similaridade (*Dayhoff matrix*) baseada no alinhamento da sequência de aminoácidos das proteínas de interesse.

**Tabela 3. Sequências de baculovírus utilizados na análise comparativa da variação genética entre isolados virais**

| Vírus                  | Isolado               | Origem                      | Nº de acesso <i>GenBank</i> |              |                 | Referência  |
|------------------------|-----------------------|-----------------------------|-----------------------------|--------------|-----------------|---|
|                        |                       |                             | <i>lef-9</i>                | <i>pif-2</i> | <i>helicase</i> |   |
| ChinNPV                | IA*                   | América Central - Guatemala | KC136321                    | KC136314     | KC136328        | Craveiro <i>et al.</i> , 2013                     |
|                        | IB                    | Brasil - Londrina (PR)      | KC136322                    | KC136315     | KC136329        | Craveiro <i>et al.</i> , 2013                     |
|                        | IC*                   | Brasil - Maringá (PR)       | KC136323                    | KC136316     | KC136330        | Craveiro <i>et al.</i> , 2013                     |
|                        | ID                    | Brasil - Iguaraçu (PR)      | KC136324                    | KC136317     | KC136331        | Craveiro <i>et al.</i> , 2013                     |
|                        | IE*                   | Brasil – Iguaraçu (PR)      | KC136325                    | KC136318     | KC136332        | Craveiro <i>et al.</i> , 2013                     |
|                        | IF                    | Brasil - Dourados (MS)      | KC136326                    | KC136319     | KC136333        | Craveiro <i>et al.</i> , 2013                     |
|                        | IG*                   | Brasil – Sertãoópolis (PR)  | KC136327                    | KC136320     | KC136334        | Craveiro <i>et al.</i> , 2013                     |
| <b>Genoma completo</b> |                       |                             |                             |              |                 |   |
| SfMNPV                 | 3AP2                  | EUA                         | EF035042                    |              |                 | Harrison <i>et al.</i> , 2008                     |
|                        | 19                    | Brasil                      | EU258200                    |              |                 | Wolff <i>et al.</i> , 2008                        |
|                        | B <sup>(a)</sup>      | América Central-Nicarágua   | HM595733                    |              |                 | Símon <i>et al.</i> , 2011                        |
|                        | G <sup>(a)</sup>      | América Central-Nicarágua   | JF899325                    |              |                 | Símon <i>et al.</i> , 2012                        |
| HearNPV                | C1 <sup>(b)</sup>     | China                       | AF303045                    |              |                 | Zhang <i>et al.</i> , 2005                        |
|                        | G4 <sup>(b)</sup>     | China                       | AF271059                    |              |                 | Zhang <i>et al.</i> , 2005                        |
|                        | NNG1                  | Quênia                      | AP010907                    |              |                 | Ogembo <i>et al.</i> , 2009                       |
|                        | AUST                  | Austrália                   | JN584482                    |              |                 | Unpublished, 2012                                 |
| MacoNPV                | A-90/2 <sup>(c)</sup> | Canadá-Western Canada       | U59461                      |              |                 | Li <i>et al.</i> , 1997; Li <i>et al.</i> , 2002a |
|                        | A-90/4 <sup>(c)</sup> | Canadá-Western Canada       | AF539999                    |              |                 | Li <i>et al.</i> , 2005                           |
|                        | B                     | Canadá-Southern Alberta     | AY126275                    |              |                 | Li <i>et al.</i> , 2002b                          |
| ChchNPV                | -                     | -                           | AY864330                    |              |                 | van Oers <i>et al.</i> , 2005                     |
| TnSNPV                 | -                     | -                           | DQ017380                    |              |                 | Willis <i>et al.</i> , 2005                       |

\*ChinNPV – isolados utilizados na análise de seleção do gene *pif-2*.

Clones virais de um mesmo isolado de <sup>(a)</sup>SfMNPV, <sup>(b)</sup>HearNPV e <sup>(c)</sup>MacoNPV.

### 1.2.6. Análise de pressão de seleção do gene *pif-2* de ChinNPV

O programa *CodeML*, implementado no *PAML- Phylogenetic Analysis by Maximum-Likelihood* v.4 (Yang, 2007), foi utilizado para avaliar a presença de seleção positiva no gene *pif-2* de ChinNPV. Usando uma abordagem de máxima verossimilhança, os valores de dN (substituições não-sinônimas por sítio não-sinônimo), dS (substituição sinônima por sítio sinônimo) e relação  $\omega$  (dN/dS) foram calculados. A proporção  $\omega$  fornece uma medida sensível da seleção que atua sobre um gene. Os genes que mostram valor de  $\omega=1$  são submetidos à seleção neutra quando as mutações não-sinônimas não provocam vantagem adaptativa. Para um valor de  $\omega<1$ , em que as mutações não-sinônimas são eliminadas a uma taxa mais rápida do que as mutações sinônimas (silenciosa), o gene está sob seleção negativa ou purificadora e, para valores de  $\omega>1$ , o gene está sob seleção positiva ou diversificadora onde as mutações não-sinônimas são fixadas mais rapidamente do que as mutações sinônimas. Para análise das sequências foram utilizados modelos de substituição baseados nos códons permitindo  $\omega$  variar entre os sítios: modelo M0 (taxa única), que assume que a relação  $\omega$  é uma média ao longo de todos os sítios; modelo M1a (neutro), que considera sítios conservados quando  $0<\omega<1$  e completamente neutros quando  $\omega=1$ ; modelo M2a (seleção positiva), que acrescenta uma terceira classe para M1a em que  $\omega$  pode assumir valores maiores que 1; modelo M3 (discreto) apresenta três classes com proporções P0, P1 e P2 e valores  $\omega_0$ ,  $\omega_1$  e  $\omega_2$  estimados a partir dos dados; modelo M7 ( $\beta$ ), que assume uma distribuição de probabilidade beta ( $\beta$ ) e não permite sítios com  $\omega>1$ ; modelo M8 ( $\beta$  e  $\omega$ ) que adiciona uma classe extra de sítios para o modelo M7, em que  $\omega$  pode ser maior do que 1. O teste da razão de verossimilhança foi realizada utilizando jModelTest v.0.1.1 (Posada, 2008) contrapondo os modelos M0 e M1a com M2a e M3 e o modelo M7 com M8. A hipótese nula (modelos M0, M1a e M7) foi rejeitada e a seleção positiva (modelos M2a, M3 ou M8) foi inferida para um valor  $p$  significativo ( $p<0,05$ ). Finalmente, uma abordagem *Bayes Empirical Bayes* (BEB) foi utilizada para identificar os locais que potencialmente estão sob seleção positiva na proteína (probabilidade $\geq 0,5$ ).

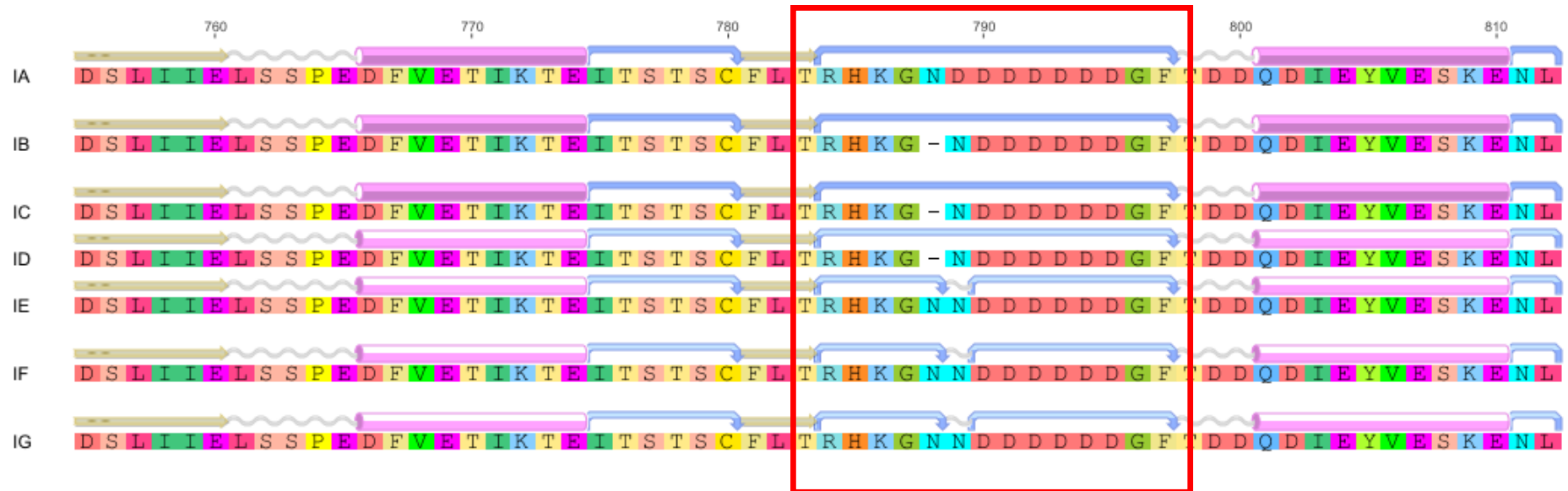
A taxa  $\omega$  (dN/dS) também foi calculada utilizando algoritmo da Máxima Verossimilhança (ML- *Maximum Likelihood*) implementado no pacote *HyPhy*

(Pond *et al.*, 2005) do *Datamonkey* ([www.datamonkey.org/](http://www.datamonkey.org/)) (Pond e Frost, 2005a; Delpont *et al.*, 2010). O pacote *HyPhy* foi utilizado para realizar os testes de taxa global assumindo uma única pressão de seleção para todos os ramos. Sítios específicos sob seleção positiva foram avaliados utilizando os métodos *Internal Fixed Effects Likelihood* (IFEL) (Pond *et al.*, 2006), *Random Effect Likelihood* (REL) (Pond e Frost, 2005b) e *Mixed Effects Model of Evolution* (MEME) (Murrell *et al.*, 2012). O modelo *HKY85* (Hasegawa M. *et al.*, 1985) selecionado pelo HyPhy como o modelo mais apropriado de substituição foi utilizado para a construção da árvore de distância (*Neighbor Joining* - NJ). Sítios sob pressão de seleção foram selecionados considerando um valor de  $p \leq 0,1$  ou um fator de Bayes  $\geq 0,9$ .

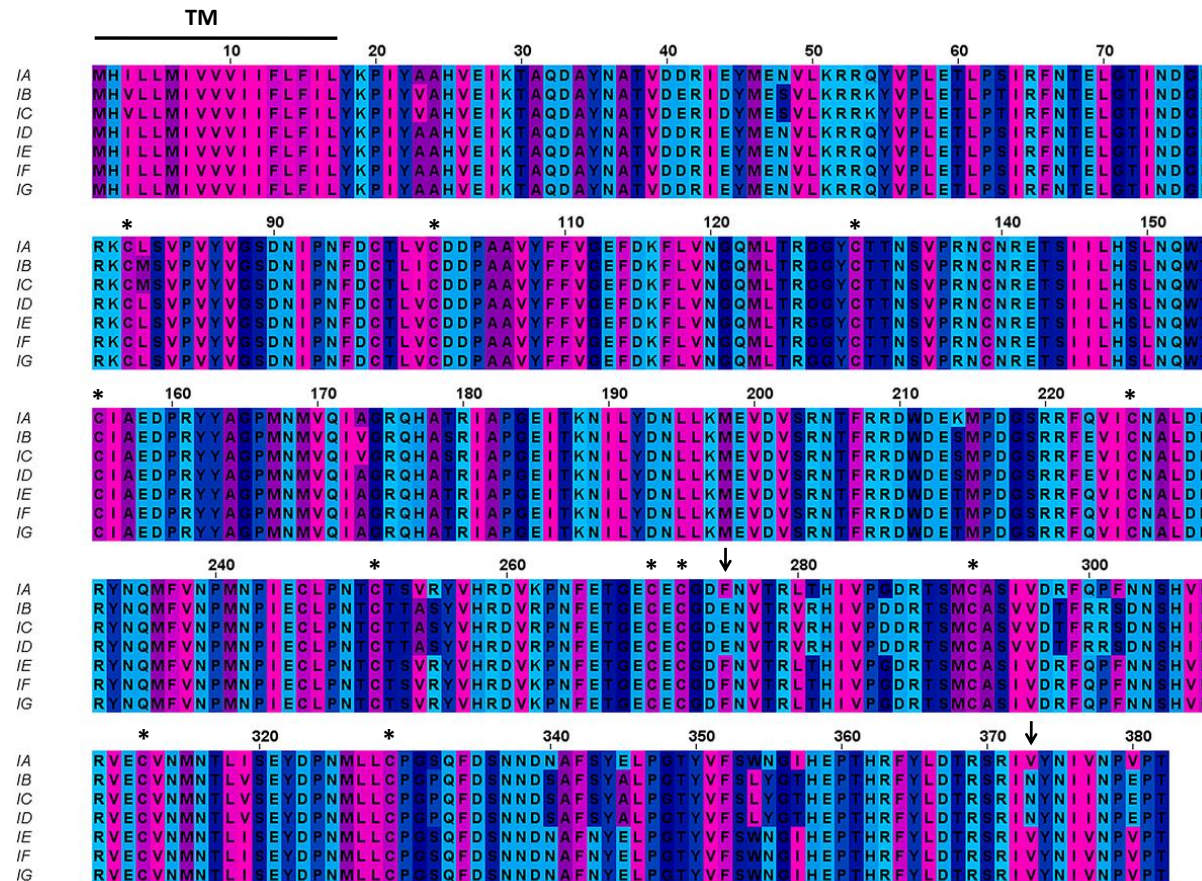
## 1.3. Resultados e Discussão

### 1.3.1. Variações nucleotídicas em isolados de ChinNPV

Amostras individuais de DNA dos isolados de ChinNPV (IA a IG) foram submetidas ao pirosequenciamento e uma cobertura de mais de 30x foi obtida para os genes *lef-9*, *helicase* e *pif-2*. Sequências de 1.491, 3.633 e 1.149 pb foram determinadas por serem correspondentes às ORFs *lef-9*, *helicase* e *pif-2*, respectivamente. Os alinhamentos dos genes *lef-9* e *helicase* apresentaram poucos polimorfismos de base única (*Single Nucleotide Polymorphisms* - SNPs), com 8 SNPs para 1.491pb e 22 SNPs para 3.633 pb, respectivamente. Polimorfismos do gene *lef-9* não provocaram diferenças nas sequências deduzidas de aminoácidos nos sete isolados de ChinNPV analisados (IA a IG). O gene *helicase* apresentou uma exclusão ATG na posição 2381-2383 nucleotídeos dos isolados IB, IC e ID. Estes isolados estão intimamente relacionados e apresentaram menor virulência entre os sete isolados analisados (Alexandre *et al.*, 2010). Esta deleção no gene *helicase* produziu alteração na sequência deduzida de aminoácidos provocando uma modificação na estrutura secundária prevista da proteína P143 que apresentou uma interrupção do domínio alça (*looping domain*) por um domínio hélice (*coil domain*) (Figura 7). Diferente dos genes *lef-9* e *helicase*, o alinhamento das sequências do gene *pif-2* apresentou um alto número de polimorfismos com 158 SNPs para 1.149 pb. Esses polimorfismos causaram alterações na sequência deduzida de aminoácidos, com resíduos (F275E e V37N) mostrando diferenças previstas em hidrofobicidade (Figura 8) e potencialmente causadoras de alterações significativas na estrutura secundária da proteína correspondente.



**Figura 7. Alinhamento de seqüências deduzidas de aminoácidos do gene *helicase* dos sete isolados de ChinNPV (IA a IG).** O gráfico acima das seqüências representa a predição da estrutura secundária das proteínas helicase e a caixa em vermelho destaca a interrupção do domínio alça (*turn domain*) por um domínio hélice (*coil domain*). Representação gráfica: seta bege = *beta strand*; linha ondulada cinza = *coil*; cilindro rosa = *alpha helix*; seta azul = *turn*.



**Figura 8. Perfil de hidrofobicidade da sequência de aminoácidos PIF2 dos isolados de ChinNPV (IA a IG).** As setas indicam os códons selecionados positivamente com alterações radicais na sequência de aminoácidos. Os asteriscos (\*) indicam resíduos de cisteína conservados (Pijlman *et al.*, 2003) e a linha acima da sequência indica domínio hidrofóbico transmembrânico (TM). Aminoácidos hidrofóbicos estão coloridos em rosa e aminoácidos hidrofílicos em azul.

### 1.3.2. Análise comparativa da variação genética dos genes *pif-2*, *lef-9* e *helicase* de isolados geográficos de diferentes baculovírus

Vários métodos têm sido utilizados para estudar a variação genotípica em baculovírus como: análise de polimorfismo dos fragmentos de restrição (RFLP) (Cory *et al.*, 2005; Alexandre *et al.*, 2010), sequenciamento de DNA de clones virais (Khan *et al.*, 2003; Zhang *et al.*, 2005) e sequenciamento de produtos de PCR usando primers degenerados para amplificação de genes conservados (Lange *et al.*, 2004; Rowley *et al.*, 2010). A tecnologia de sequenciamento de nova geração (*Next Generation Sequencing* - NGS) fornece uma ferramenta poderosa para a análise da variação genética dentro de uma população viral e esse método foi utilizado pela primeira vez por Baillie e Bouwer (2012) para avaliar a variação genotípica em populações de HearNPV. Estudos têm relacionado pequenas alterações nos genes às mudanças no fenótipo viral. Por exemplo, a substituição individual de aminoácido na proteína codificada pelo gene *helicase* ampliou o espectro de hospedeiros de AcMNPV (Kamita e Maeda, 1997; Argaud *et al.*, 1998). Isto demonstra que substituições nucleotídicas em genes essenciais podem influenciar o espectro de hospedeiros ou modificar a virulência.

Estudos filogenéticos mostram que ChinNPV-IE é mais proximamente relacionado com TnSNPV e ChchNPV (Craveiro *et al.*, 2013). O alinhamento do gene *pif-2* destes três vírus apresentou 91 polimorfismos, menos do que os SNPs encontrados entre os isolados de ChinNPV indicando uma maior variação intraespecífica do que interespecífica para esse gene.

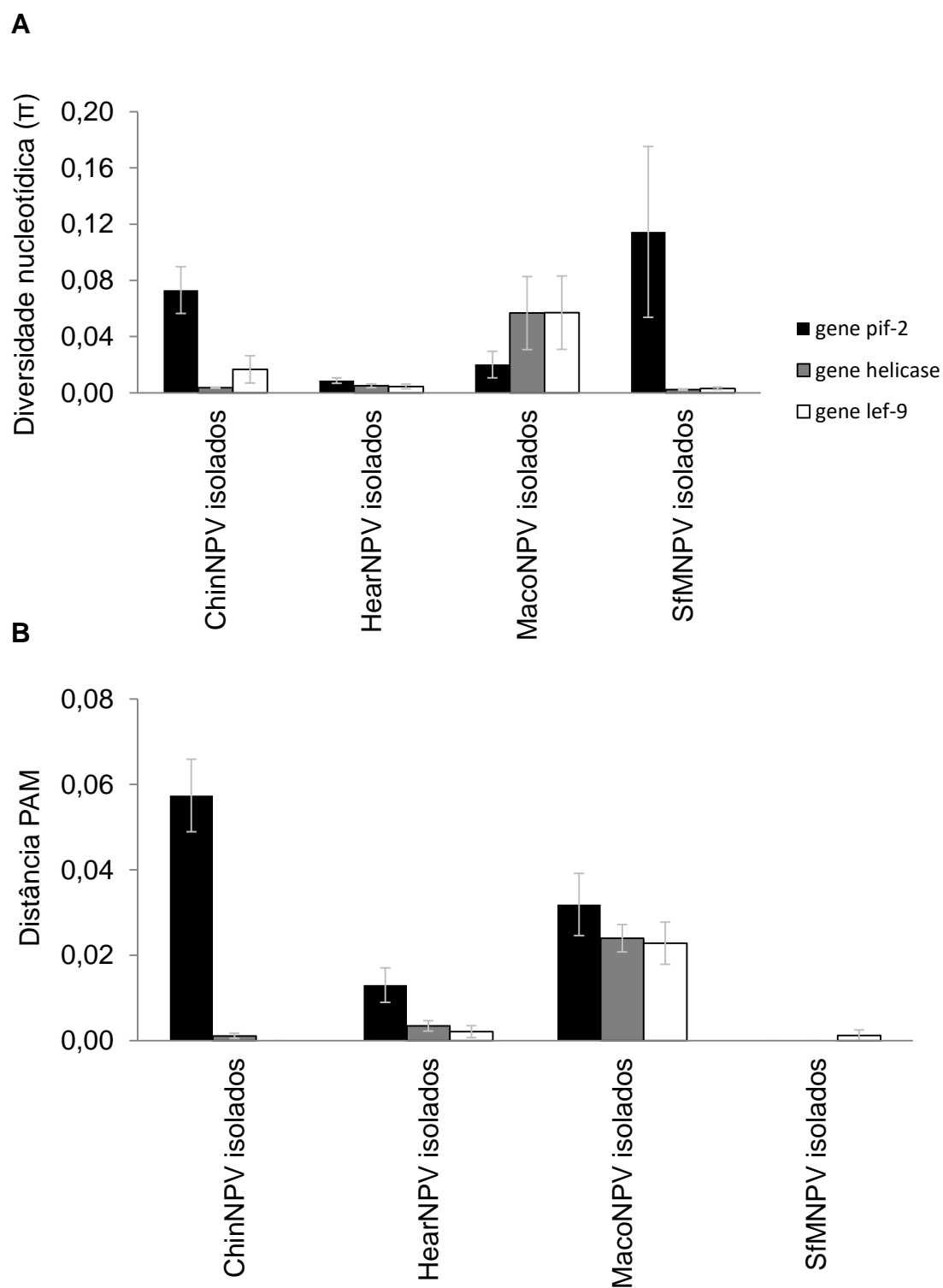
As sequências de *pif-2* dos isolados de HearNPV e MacoNPV também foram analisadas e exibiram poucos polimorfismos. Os valores de diversidade nucleotídica ( $\pi$ ) calculados utilizando o programa *DNASP v.5* mostraram que os genes *pif-2* dos isolados de ChinNPV e SfMNPV possuem maior variação genética do que a dos outros vírus analisados (Figura 9A). No entanto, a grande variação observada no gene *pif-2* dos isolados de SfMNPV corresponde apenas ao clone defeutivo SfMNPV-G, clone obtido de isolado viral de ocorrência natural no campo. Entretanto, os polimorfismos encontrados em SfMNPV-G não causaram diferença na sequência de aminoácidos de PIF2 (Figura 9B). Provavelmente, o clone SfMNPV-G foi retido na população de campo por outros genótipos virais e, portanto, pôde acumular mutações que não são mantidas na



população em geral. Os polimorfismos do gene *pif-2* de ChinNPV são variações não-sinônimas e foi observada uma Distância PAM média de ~0,06 o que indica uma distância evolutiva entre os pares de sequências protéicas PIF2 dos isolados de ChinNPV.

A mesma análise foi realizada para os genes *lef-9* e *helicase* de ChinNPV. Entre os baculovírus, o gene *helicase* é considerado mais variável do que os genes *lef-9* e *pif-2* (Miele *et al.*, 2011). Entretanto, nossos resultados apresentaram valores menores de Diversidade nucleotídica e Distância PAM para o gene *helicase* do que os valores obtidos para o gene *pif-2* de ChinNPV (Figura 9).

Estudos prévios destacam o gene *pif-2* como o mais conservado entre os *core genes* dos baculovírus (Miele *et al.*, 2011). Em contraste, os resultados aqui obtidos mostram uma elevada variação genética com um grande número de mutações não-sinônimas na sequência de aminoácidos de PIF2 de ChinNPV.



**Figura 9.** Diversidade genética dos genes *pif-2*, *lef-9* e *helicase* de isolados geográficos de diferentes baculovírus: ChinNPV, HearNPV, MacoNPV e SfMNPV. (A) Diversidade nucleotídica ( $\pi$ ) (DNASP v.5). (B) Distância PAM média (MEGA v.5.0.5).

### 1.3.3. Análise de seleção de *pif-2* em isolados de ChinNPV

A análise da evolução molecular adaptativa usando  $dN/dS$  ( $\omega$ ) é geralmente informativa no estudo dos genes sob seleção positiva. Substituições de nucleotídeos não-sinônimas podem alterar as propriedades da proteína codificada e, somente quando estas mutações conferem vantagem adaptativa, elas são fixadas na população a uma taxa maior do que as substituições sinônimas, que geralmente são invisíveis para a seleção natural. Assim, a seleção positiva ou diversificadora ( $\omega > 1$ ) pode ser detectada quando a taxa de substituições não-sinônimas ( $dN$ ) em um gene é maior do que a taxa de substituições sinônimas ( $dS$ ) (Yang, 2001). Quando mutações não sinônimas são deletérias e são eliminadas a uma taxa mais rápida do que as mutações sinônimas, o gene é submetido à seleção negativa ou purificadora com  $\omega$  inferior a 1.

Para melhor compreender a força evolutiva que dirige a alta variação genética encontrada no gene *pif-2* de ChinNPV, foi utilizado o método *CodeML* implementado no pacote do software *PAML* para testar diferentes modelos de evolução de códons que permitem a variação de  $\omega$  ( $dN/dS$ ) entre os sítios (modelos M1a, M2a, M3, M7 e M8) (Yang e Bielawski, 2000). Como mostrado na Tabela 4, o modelo neutro (M1a) não foi rejeitado por modelos que suportam a seleção positiva (M2a, M3 e M8). No entanto, na análise de BEB, os modelos M2a e M8 identificaram três posições de aminoácidos na proteína PIF2 (resíduos 214, 275 e 373) com  $\omega \approx 2$ , sugerindo que estes resíduos podem ter sido submetidos a seleção positiva.

Estes resultados foram confirmados usando os métodos *REL*, *IFEL* e *MEME* implementados no pacote do software *Hyphy*. Os códons previamente identificados (214, 275 e 373) foram confirmados como sendo selecionados positivamente para os resultados *HyPhy* com um valor de  $p < 0,1$  (*MEME*) ou um fator de Bayes  $> 50$  (*REL*) (Tabela 5). Além disso, o algoritmo de *MEME* também encontrou o códon 343 sob seleção positiva. Como mostrado na Tabela 5, os códons 214 e 275 foram confirmados por mais de um método enquanto que os códons 343 e 373 foram detectados por *MEME* e *REL*, respectivamente. É importante ressaltar que as análises *REL* e *IFEL* são conhecidas pela mal execução em pequenos conjuntos de dados e na análise de poucos táxons, detectando incorretamente sítios sob seleção positiva (Pond e Frost, 2005b). No

entanto, MEME, método recomendado para a identificação de sítios sob seleção no pacote *HyPhy*, confirmou a maioria dos códons detectados pelos métodos do *PAML*. Apesar de todos esses métodos serem amplamente utilizados para a detecção de seleção positiva, pequenos conjuntos de dados, poucos táxons e baixa divergência genética pode limitar a precisão e o poder de inferência de seleção positiva (Anisimova *et al.*, 2002). Com a análise de apenas sete isolados, nossos resultados podem sofrer com estas limitações e uma maior amostragem poderia melhorar a precisão e o poder de nossas análises. No entanto, é difícil aumentar o tamanho da amostra, devido à inerente dificuldade na obtenção de diferentes isolados geográficos. Além disso, os nossos resultados foram confirmados por MEME, que detecta indícios de seleção positiva pervasiva ou episódica, mesmo na presença de um sinal de seleção purificadora forte (Murrell *et al.*, 2012).

Dois dos códons identificados sob pressão de seleção positiva apresentaram alterações que modificam as propriedades dos aminoácidos. A alteração no códon 275 (F275E) envolve a alteração de um aminoácido não-polar para um aminoácido ácido e no códon 373 (V373N) envolve a alteração de um aminoácido neutro e grande para um aminoácido neutro e pequeno, o que, de acordo com a classificação proposta por Hanada *et al.* (2007), são alterações radicais. Consequentemente, essas variações podem causar diferenças na estrutura secundária da proteína ou na interação proteína-proteína. A proteína PIF2 é parte de um complexo protéico da membrana dos ODVs envolvido na infecção oral do inseto hospedeiro, desta forma, é razoável considerar que as alterações radicais dos aminoácidos nesta proteína podem causar mudanças na capacidade de infecção e transporte do vírus através da membrana.

Além disso, outros genes do envelope das ODVs como *odv-e66* e *pif-3* foram identificados por estarem sob seleção positiva em isolados de SfMNPV (Simon *et al.*, 2011), corroborando a hipótese de que mudanças em genes envolvidos na infecção oral podem estar relacionados com adaptação à novos hospedeiros, superação à variação fenotípica do hospedeiro ou para o aumento da eficiência de transmissão do vírus (Simon *et al.*, 2011). Variantes fenotípicas foram previamente relatados em ChinNPV, onde os isolados ChinNPV-IA, ChinNPV-IE e ChinNPV-IF foram mais virulentos entre sete isolados analisados (Alexandre *et al.*, 2010). Portanto, o gene *pif-2*, sob pressão de seleção positiva,

pode estar contribuindo para as diferenças na eficiência de transmissão dos isolados de ChinNPV.

**Tabela 4. Análise de pressão de seleção do gene *pif-2* de ChinNPV realizada no PAML e dados estatísticos da razão de verossimilhança para a comparação dos modelos utilizados**

| Modelo                       | Valor Log-likelihood (lnL) | Parâmetro estimado  | Sítio seleção positiva (probabilidade)* |   | Teste Likelihood ratio  |
|------------------------------|----------------------------|---|---|---|---|
|                              |                            |   | NEB                                     | BEB   |   |
| M0<br>(taxa única)           | -2214,459                  | $\omega=0,10389$  |   |   |   |
| M1a<br>(neutro)              | 2212,771                   | $p_0=0,96175$<br>$\omega_0=0,078$<br>$p_1=0,03825$<br>$\omega_1=1,000$                                      |   |   |   |
| M2a<br>(positivo)            | -2212,684                  | $p_0=0,98557$<br>$\omega_0=0,089$<br>$p_1=0,00000$<br>$\omega_1=1,000$<br>$p_2=0,01443$<br>$\omega_2=2,043$ | 214 K<br>(0,800)*<br>275 F<br>(0,893)*  | 214 K<br>(0,662)*<br>275 F<br>(0,682)*<br>373 V<br>(0,504)* | M0 vs. M2a (df=2):<br>P=0,169603<br>M1a vs. M2a (df=2):<br>P=0,917554 |
| M3<br>(discreto)             | -2212,684                  | $p_0=0,00006$<br>$\omega_0=0,000$<br>$p_1=0,98551$<br>$\omega_1=0,089$<br>$p_2=0,01443$<br>$\omega_2=2,043$ | 214 K<br>(0,800)*<br>275 F<br>(0,893)*  |   | M0 vs. M3 (df=4):<br>P=0,470529<br>M1a vs. M3 (df=4):<br>P=0,996504   |
| M7<br>( $\beta$ )            | -2213,017                  | $p=0,33439$<br>$q=2,54477$  |   |   |   |
| M8<br>( $\beta$ e $\omega$ ) | -2212,686                  | $p=9,78067$<br>$q=99,00000$<br>$p_0=0,98652$<br>$p_1=0,01348$<br>$\omega_1=2,114$                           | 214 K<br>(0,780)*<br>275 F<br>(0,878)*  | 214 K<br>(0,774)*<br>275 F<br>(0,820)*<br>373 V<br>(0,578)* | M7 vs. M8 (df=2):<br>P=0,71806  |

df: graus de liberdade correspondentes ao número de parâmetros entre os dois modelos.

**Tabela 5. Análise de seleção positiva para o gene *pif-2* de ChinNPV utilizando diferentes modelos implementados no programa HyPhy v. 2.1.2**

| Seleção                             | IFEL modelo                       | REL modelo                   | MEME modelo                  |
|-------------------------------------|-----------------------------------|------------------------------|------------------------------|
|                                     | Códons                            | Códons                       | Códons                       |
| <b>Positiva</b>                     |                                   | 214 (T → S ou K);            | 214 (T → S ou K);            |
| <b>(<math>\omega &gt; 1</math>)</b> | 214 (T → S ou K)                  | 275 (F → E);<br>373 (V → N). | 275 (F → E);<br>343 (V → N). |
| <b>Negativa</b>                     |                                   | -                            | N/A                          |
| <b>(<math>\omega &lt; 1</math>)</b> | 43(I); 51(K);<br>175(R); 376 (I). |                              |                              |

#### 1.4. Conclusões

- Ao contrário do esperado, o gene *pif-2* de ChinNPV apresentou um grande número de polimorfismos com uma elevada variação genética.
- Polimorfismos observados no gene *pif-2* de ChinNPV são variações não-sinônimos com resíduos de aminoácidos apresentando diferenças previstas em hidrofobicidade e potencialmente causadoras de alterações significativas na estrutura secundária da proteína correspondente.
- ChinNPV, ChchNPV e TnSNPV apresentaram menos polimorfismos do que os SNPs encontrados entre os isolados de ChinNPV demonstrando uma maior variação genética intra-específica do que inter-específica.
- Sítios sob pressão de seleção positiva (diversificadora) foram identificados na proteína PIF2 de ChinNPV o que pode estar envolvido na variação fenotípica, como a diferença na virulência observada entre os isolados de ChinNPV analisados.

## **CAPÍTULO 2: Genoma completo e genômica comparativa de populações de *Chrysodeixis includens* NPV**

Este capítulo inclui resultados apresentados nos artigos:

1. Artigo publicado em 2015 na *BMC Genomics* 16 (1), DOI: 10.1186/s12864-015-1323-9  
**The genome sequence of *Pseudoplusia includens* single nucleopolyhedrovirus and an analysis of *p26* gene evolution in the baculoviruses**

Saluana R. Craveiro, Peter W. Inglis, Roberto C. Togawa, Priscila Grynberg, Fernando L. Melo, Zilda Maria A. Ribeiro, Bergmann M. Ribeiro, Sônia N. Bão, Maria Elita B. Castro\*

2. Artigo submetido em abril/2016 na *Genome Announcement* :  
**Complete genome sequences of six *Chrysodeixis includens* nucleopolyhedrovirus isolates from Brazil and Guatemala**

Saluana R. Craveiro, Luis Arthur V. Santos, Roberto C. Togawa, Peter W. Inglis, Maria Elita B. Castro\*

## 2.1. Introdução

Sete populações de baculovírus obtidos de larvas infectadas de *C. includens*, coletadas em plantações de soja e algodão no Brasil e Guatemala e caracterizadas como isolados geográficos de *Chrysodeixis includens* NPV, apresentaram diversidade genética e fenotípica em eventos observados inicialmente baseadas em análises comparativas de perfil de restrição e avaliação de patogenicidade (Alexandre *et al.*, 2010).

Atualmente, 71 genomas de baculovírus estão depositados no *Genbank*, do NCBI - *National Center for Biotechnology Information* (<http://www.ncbi.nlm.nih.gov/genomes/GenomesGroup.cgi?taxid=10442>) (acessado em: 15/02/2016). Nessa busca, a sequência completa de genoma de vírus patogênico a *C. includens* não foi encontrada. Portanto, este trabalho constitui o primeiro registro da descrição genômica de *Chrysodeixis includens* NPV.

Os genomas de baculovírus possuem em torno de 90 a 181 genes dos quais 37, denominados *core genes*, são ancestrais e altamente conservados entre todos os baculovírus e representam 3% do conteúdo genético viral (Miele *et al.*, 2011; Rohrmann, 2011).

*Chrysodeixis includens nucleopolyhedrovirus* isolado IE (ChinNPV-IE) foi obtido de larvas infectadas coletadas em plantações de soja da região de Iguaraçu – PR, Brasil, na safra de 2007. Dentre os sete isolados analisados, ChinNPV-IE foi um dos isolados que apresentou maior infectividade (Alexandre *et al.*, 2010).

Neste estudo, foi realizado o sequenciamento, montagem e anotação do genoma completo de sete isolados de ChinNPV (IA a IG) e a descrição gênica detalhada do genoma de ChinNPV-IE, isolado selecionado como representante da espécie e utilizado como referência para análise genômica dos outros seis isolados de ChinNPV. Além disso, o genoma dos isolados de ChinNPV foram comparados entre si e com outros baculovírus para a determinação das *open reading frames* (ORFs), dos polimorfismos de base única (SNPs - *Single Nucleotide Polymorphisms*) e das sequências únicas desse vírus.

Uma proposta taxonômica intitulada “One new species *Chrysodeixis includens nucleopolyhedrovirus*, in the genus *Alphabaculovirus*” (Harrison, 2015.



Code 2015.007aD), fundamentada essencialmente nos estudos aqui apresentados e com a indicação de PsinNPV-IE (ChinNPV-IE) como isolado representativo da espécie, foi recentemente aprovada e deverá ser incluída na ocasião da atualização do Ninth ICTV Report do International Committee on Taxonomy of Viruses (ICTV) (King *et al.*, 2011).

## 2.2. Materiais e Métodos

### 2.2.1. Sequenciamento e montagem do genoma de isolados de ChinNPV

Os procedimentos da purificação das partículas OBs e extração do DNA viral estão detalhados no capítulo 1, juntamente com a lista das amostras virais de *C. includens* analisadas neste estudo (Tabela 2).

A técnica de pirosequenciamento 454 (*Roche*) foi utilizada para o sequenciamento individual do genoma completo de sete isolados de ChinNPV (IA a IG). O sequenciamento foi realizado na Plataforma Genômica DF (Distrito Federal – Brasil) e aproximadamente 5 µg de DNA foi utilizado para a reação de *shotgun* pirosequenciamento utilizando *GS FLX Titanium sequencer (454 Life Sciences Technologies)*. Após a corrida do sequenciamento 454, o programa *Newbler v. 2.8 (Roche Applied Science)* foi utilizado para o processamento das leituras (*reads*) e, posteriormente, criação dos arquivos *FastQ*. O programa *FastQC v. 0.11.4* foi utilizado para avaliação da qualidade das leituras e o programa *Coral v. 1.4 (Salmela e Schröder, 2011)* para correção dos erros de sequenciamento. O programa *PrinSeq v. 0.20.3 (Schmieder e Edwards, 2011)* foi aplicado para retirar regiões de baixa qualidade nas leituras ( $Phred \leq 20$ ) e remover sequências curtas (comprimento  $\leq 50$  pb) e foi permitida uma probabilidade de erro de 0,1%. O valor de *Phred* foi medido e uma qualidade média da seqüência  $> 30$  foi estimada, exibindo uma precisão de 99,9%. Os dados brutos do sequenciamento dos sete isolados de ChinNPV e os dados após o pré-processamento das leituras estão descritos na Tabela 6.

Tabela 6. Estatística dos resultados do sequenciamento de isolados de ChinNPV -454 GS FLX Titanium (Roche)

|  | <b>Dados brutos (<i>raw</i>) / pré-processados (<i>trimmed</i>)</b> |                   |                   |                   |                   |                   |                   |
|--|---|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
|  | <b>ChinNPV-IA</b>   | <b>ChinNPV-IB</b> | <b>ChinNPV-IC</b> | <b>ChinNPV-ID</b> | <b>ChinNPV-IE</b> | <b>ChinNPV-IF</b> | <b>ChinNPV-IG</b> |
| <b>Total de sequências</b>               | 34.988  | 98.111            | 22.049            | 27.315            | 38.281            | 28.602            | 26.295            |
|  | 30.731  | 85.950            | 16.979            | 21.148            | 33.684            | 22.083            | 23.021            |
| <b>Total de bases</b>                    | 18.976.875  | 53.386.536        | 11.987.411        | 14.885.363        | 20.751.970        | 15.573.433        | 14.270.445        |
|  | 11.035.431  | 30.667.888        | 7.001.048         | 8.623.971         | 12.051.794        | 9.047.100         | 8.265.091         |
| <b>Tamanho médio das sequências (pb)</b> | 542.38 ± 65.33  | 544.14 ± 67.16    | 543.67 ± 57.19    | 544.95 ± 62.07    | 542.10 ± 67.48    | 544.49 ± 58.64    | 542.71 ± 64.68    |
|  | 359,10 ± 123,30   | 356,81 ± 123,62   | 412,34 ± 107,57   | 407,79 ± 110,30   | 357,79 ± 123,33   | 409,69 ± 110,47   | 359,02 ± 123,03   |
| <b>Tamanho mínimo (pb)</b>               | 63  | 56                | 67                | 61                | 55                | 58                | 57                |
|  | 60  | 60                | 60                | 60                | 60                | 60                | 60                |
| <b>Tamanho máximo (pb)</b>               | 1.200   | 1.200             | 1.200             | 1.200             | 1.200             | 1.200             | 1.200             |
|  | 500   | 500               | 500               | 500               | 500               | 500               | 500               |
| <b>Conteúdo médio de GC (%)</b>          | 40.95 ± 4.07  | 40.22 ± 4.22      | 39.77 ± 4.30      | 40.27 ± 4.76      | 40.56 ± 4.07      | 39.83 ± 4.41      | 40.68 ± 4.17      |
|  | 39,86 ± 5,10  | 39,59 ± 5,24      | 39,62 ± 4,93      | 40,13 ± 5,38      | 39,57 ± 5,15      | 39,75 ± 5,08      | 39,75 ± 5,24      |
| <b>Total de sequências com N</b>         | 14.058 (40.18 %)  | 39.335 (40.09 %)  | 4.688 (21.26 %)   | 6.382 (23.36 %)   | 15.645 (40.87 %)  | 6.193 (21.65 %)   | 10.579 (40.23 %)  |
|  | 162 (0,53 %)  | 749 (0,87 %)      | 40 (0,24 %)       | 275 (1,30 %)      | 131 (0,39 %)      | 149 (0,67 %)      | 98 (0,43 %)       |
| <b>Máximo de Ns por sequência (%)</b>    | 48  | 50                | 19                | 57                | 36                | 53                | 48                |
|  | 6   | 1                 | 1                 | 2                 | 3                 | 8                 | 2                 |

Dados pré-processados (*trimmed*) estão sombreados.

Inicialmente, as leituras pré-processadas dos sete isolados ChinNPV-IA a IG foram utilizados conjuntamente para a montagem *De novo assembly* utilizando montador MIRA v. 4.0.2 (Chevreux *et al.*, 2004). Os contigs obtidos foram ordenados por *Abacas* (Assefa *et al.*, 2009) utilizando como referência o genoma de *Chrysodeixis chalcites* (Chch) NPV. O *scaffold* de aproximadamente 142 Kpb foi realinhado com as leituras pré-processadas de ChinNPV-IE utilizando ferramenta *Map to reference* da plataforma *Geneious v.6.1.8* (Drummond *et al.*, 2014), a sequência consenso foi corrigida mantendo um pareamento de bases de 50% e a sequência genômica de ChinNPV-IE foi determinada.

A montagem dos outros seis isolados foi realizada diferente da montagem do isolado ChinNPV-IE. *De novo assembly* foi realizado individualmente para cada isolado utilizando os programas *Newbler Assembler v.2.8* (Roche) *MIRA v.4.0.2* (Chevreux *et al.*, 1999) e *Celera Assembler v.8.3* (Myers *et al.*, 2000) (Tabela 7). Os contigs obtidos de cada montagem foram compilados utilizando o programa *Geneious v.6.1.8* e então os *scaffolds* foram alinhados com o genoma de ChinNPV-IE utilizando ferramenta *Mauve v.2.0* do *Geneious v.6.1.8* para avaliação da montagem correta do genoma de cada isolado.

**Tabela 7. Dados estatísticos da montagem dos genomas de ChinNPV utilizando os programas: *Newbler Assembler v.2.8*, *MIRA v.4.0.2* e *Celera Assembler v.8.3***

|               | ChinNPV |         |        |         |        |        | Proramas |
|---------------|---------|---------|--------|---------|--------|--------|----------|
|               | IA      | IB      | IC     | ID      | IF     | IG     |          |
| N° de contigs | 30      | 52      | 21     | 101     | 26     | 12     | Newbler  |
|               | 12      | 79      | 9      | 6       | 10     | 8      | MIRA     |
|               | 33      | 26      | 12     | 5       | 7      | 6      | Celera   |
| N50 (pb)      | 90.270  | 135.889 | 89.450 | 34.022  | 32.387 | 21.691 | Newbler  |
|               | 30.508  | 8.823   | 28.902 | 29.209  | 22.728 | 31.413 | MIRA     |
|               | 42.117  | 3.238   | 36.086 | 128.357 | 74.660 | 40.654 | Celera   |
| L50           | 1       | 1       | 1      | 2       | 2      | 2      | Newbler  |
|               | 2       | 7       | 2      | 2       | 3      | 3      | MIRA     |
|               | 2       | 7       | 2      | 1       | 1      | 2      | Celera   |

N50: tamanho do último contig utilizado na somatória dos tamanhos de contigs (ordenados de forma crescente) calculada até que o resultado alcance valor igual ou maior que a metade do tamanho do genoma.

L50: número de contigs com tamanho igual ou maior que N50

### 2.2.2. Anotação do genoma

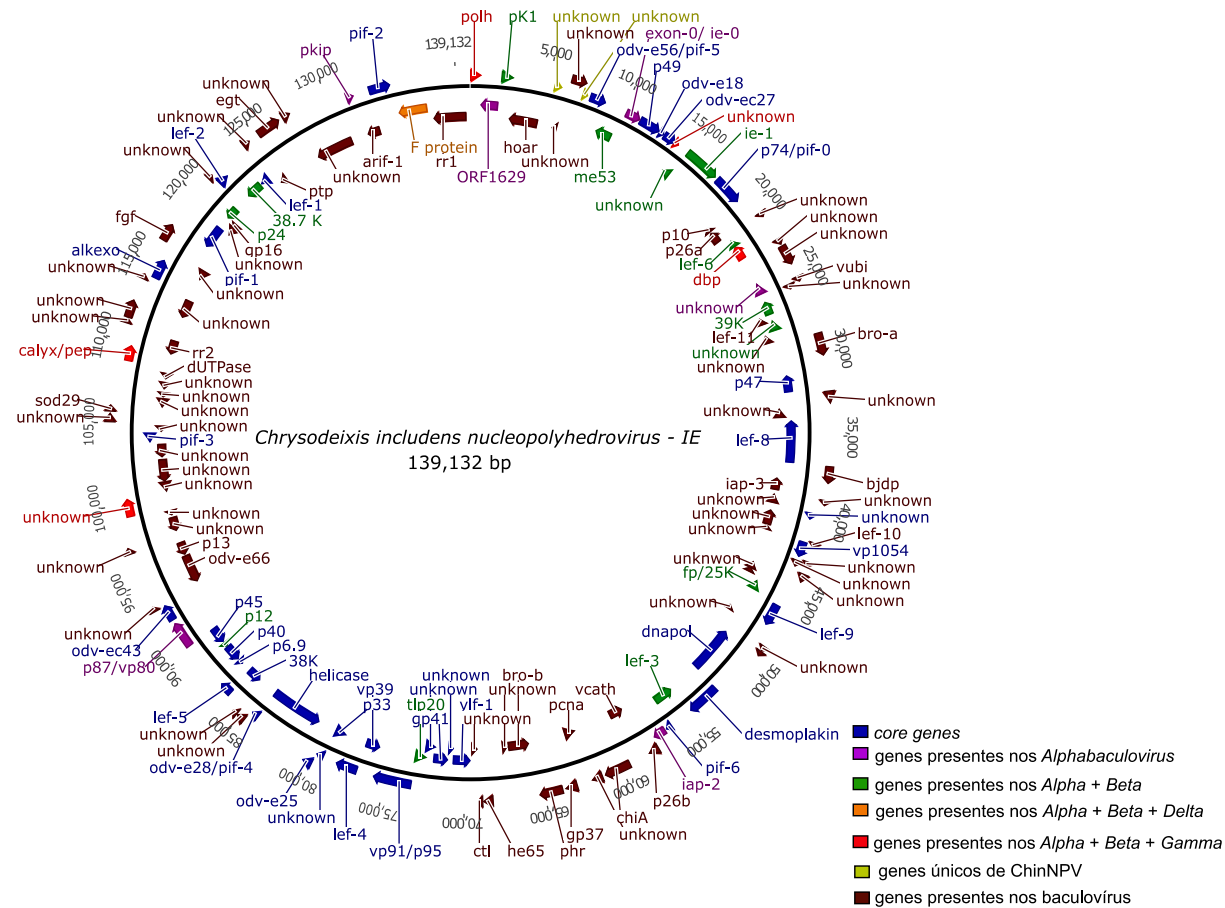
A predição das ORFs foi realizada com *ORF Finder* (NCBI) e *Geneious v. 6.1.8*. As regiões codantes foram selecionadas com mais de 50 aminoácidos a partir do códon de início ATG e o mínimo de sobreposições entre as ORFs. O software *Artemis* (Rutherford *et al.*, 2000) foi utilizado para visualização e anotação das ORFs preditas que foram confirmadas utilizando algoritmo *BLASTx* (Altschul *et al.*, 1997). Para a construção da Tabela 9 comparativa do genoma de ChinNPV-IE com outros *Alphabaculovirus*, as porcentagens de identidade entre os genes homólogos foram calculadas por meio do alinhamento utilizando o programa *tBLASTn* com valor de corte de *e-value* 10e-50 (Altschul *et al.*, 1997). O genoma de ChinNPV foi comparado com os vírus proximamente relacionados ChchNPV e TnSNPV, com o vírus *Mamestra configurata* (Maco) NPV-B, que também pertence aos *Alphabaculovirus* do Grupo II e com o protótipo dos baculovirus, *Autographa californica* (Ac) MNPV, que pertence aos *Alphabaculovirus* do Grupo I. O alinhamento global múltiplo e o alinhamento par-a-par do genoma de ChinNPV com o genoma dos quatro *Alphabaculovirus* citados acima foram realizados utilizando software *Mauve v. 2.0* implementado no pacote *Geneious v. 6.1.8* e *LBDotView v.1.0* (Huang e Zhang, 2004), respectivamente.

Seqüências deduzidas de aminoácidos foram analisadas usando *SignalP v.4.1* (Petersen *et al.*, 2011) e *TMHMM v.2.0* (Sonnhammer *et al.*, 1998; Krogh *et al.*, 2001) para a predição dos sítios de clivagem de peptídeo sinal (SP) e domínios transmembrana (TM), respectivamente.

## 2.3. Resultados e Discussão

### 2.3.1. Descrição geral do genoma de ChinNPV-IE

O tamanho do genoma do DNA circular de ChinNPV-IE foi determinado como sendo 139.132 pb (cobertura 30x) com um teor de GC de 39,3%, o que está em concordância com o teor médio de GC dos Alphabaculoviruses do Grupo II (GC = 41,6%) (Miele *et al.*, 2011). Foram identificados um total de 141 ORFs preditas perfazendo ao todo 80% do genoma, que inclui 37 *core genes* e duas ORFs únicas de ChinNPV (Psin5 e Psin8) (Figura 10). Um grande número de ORFs, em torno de 54, nomeadas como *unknown* são proteínas hipotéticas. Essas ORFs foram encontradas no genoma de outros baculovírus, entretanto, sua atividade funcional não foi descrita até o momento e, portanto, possuem função desconhecida. As ORFs foram numeradas sequencialmente a partir do gene da poliedrina com 69 ORFs no sentido horário e 72 ORFs no sentido anti-horário.



**Figura 10.** Mapa circular do genoma de ChinNPV-IE. A direção de transcrição das 141 ORFs preditas estão representadas por setas e, por convenção, a posição das ORFs no genoma está ordenada (pb) a partir do gene de *polh*.

### 2.3.2. Genômica comparativa de sete populações de ChinNPV

As sequências do genoma completo de sete isolados de ChinNPV (IA a IG) foram sequenciadas, montadas, anotadas e o alinhamento global apresentou heterogeneidade genética entre essas sete populações. Além disso, uma mistura de genótipos dentro de um mesmo isolado foi observado, principalmente, no genoma de ChinNPV-IB.

Os genomas dos sete isolados possuem entre 138.899 pb (ChinNPV-IB) a 140.859 pb (ChinNPV-IC) e conteúdo de CG de, aproximadamente, 39% (Tabela 8). Comparado ao genoma de ChinNPV-IE, isolado viral utilizado como referência, os genomas dos sete isolados possuem alta similaridade com identidade de sequência nucleotídica entre 96,3% (ChinNPV-ID) a 99,4% (ChinNPV-IF) (Tabela 8). Rearranjos estruturais como inversões e translocações não foram observados. Entretanto, *indels* (deleções ou inserções), polimorfismos de base única (SNPs) e inserções de pequenos fragmentos foram comuns. Assim, foram contabilizados 1.995 mutações pontuais ao longo de todo o genoma utilizando ChinNPV-IE como referência. Quanto a essas variações, é interessante considerar que erros de sequenciamento, em especial, inserções de adeninas em regiões homopoliméricas, são inerentes à tecnologia 454 (Roche) e podem causar uma superestimativa do número de variações genéticas entre o genoma dos sete isolados.

**Tabela 8. Características gerais dos genomas de isolados de ChinNPV**

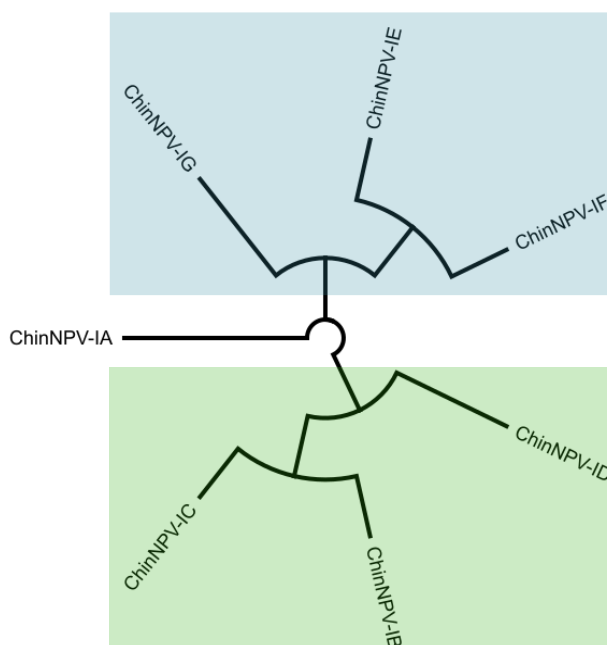
| Características               | ChinNPV  |          |          |          |          |          |          |
|-------------------------------|----------|----------|----------|----------|----------|----------|----------|
|                               | IA       | IB       | IC       | ID       | IE       | IF       | IG       |
| <b>Tam. genoma (pb)</b>       | 140.808  | 138.869  | 140.859  | 140.787  | 139.132  | 139.181  | 139.116  |
| <b>Conteúdo GC (%)</b>        | 39,2     | 39,2     | 39,2     | 39,2     | 39,3     | 39,2     | 39,2     |
| <b>Total de ORFs</b>          | 142      | 141      | 142      | 142      | 141      | 141      | 141      |
| <b>% ID ChinNPV-IE</b>        | 96,8     | 97,5     | 96,4     | 96,4     | -        | 99,4     | 99,4     |
| <b>Nº de genes <i>bro</i></b> | 3        | 2        | 3        | 3        | 2        | 2        | 2        |
| <b>GenBank acesso</b>         | KU669289 | KU669290 | KU669291 | KU669292 | KJ631622 | KU669293 | KU669294 |

Embora as variações estejam bem distribuídas ao longo de todo o genoma, 4 regiões se destacaram por concentrar um grande número de polimorfismos e estão apresentadas de acordo com sua localização no genoma de ChinNPV-IE. A primeira região (3.300 - 5.700 pb) corresponde, em grande parte, ao gene *hoar*. Semelhante ao que foi observado nos isolados de ChinNPV, Bernal *et al.* (2013) também relatam uma alta variação genética presente no gene *hoar* de cinco isolados de ChchNPV, vírus proximoamente relacionado filogeneticamente a ChinNPV (Craveiro *et al.*, 2013). A segunda região (29.500 - 29.850 pb) corresponde a região intra-gênica do gene *bro-a*. Os genes *bro* são considerados regiões *hot spots* para alta variabilidade e, apesar de *bro-a* apresentar alta variação, os outros 2 genes *bro* (*bro-b* e *bro-c*) encontrados no genoma de ChinNPV não apresentaram diversidade genética significativa. A terceira região (103.975 e 105.780 pb) corresponde a região inter-gênica dos genes *pif-3* e *sod29*. Nessa região, uma inserção de aproximadamente 1.700 pb introduziu um gene *bro* nos genomas dos isolados ChinNPV-IA, ChinNPV-IC e ChinNPV-ID. Desta forma, o genoma dos isolados ChinNPV-IB, -IE, -IF e IG apresentam duas cópias do gene *bro* (*bro-a* e *bro-b*) enquanto que os outros isolados possuem um gene *bro-a* mais com a aquisição de *bro-c* em seus genomas (Tabela 8). A quarta região (132.320 a 135.670 pb) corresponde a região inter-gênica dos genes *arif-1* e *F protein*. Nessa região está contido o gene *pif-2* que foi descrito por ser altamente variável entre os isolados de ChinNPV (Craveiro *et al.*, 2013). Esses dados indicam que, apesar do gene *pif-2* ser altamente conservado nos baculovirus (Miele *et al.*, 2011), a hipervariabilidade desse gene nos genomas dos isolados de ChinNPV pode estar sendo favorecida pela sua localização em uma região genômica suscetível a variações genéticas.

A árvore filogenética construída com base na sequência genômica dos isolados de ChinNPV utilizando método *Neighbor Joining* - NJ (Figura 11) apresentou topologia similar a árvore filogenética construída com base em genes conservados descrita em estudos anteriores (Craveiro *et al.*, 2013). Os isolados ChinNPV-IB, -IC e -ID agruparam em um clado separado dos isolados ChinNPV-IE, IF e -IG, topologia semelhante a apresentada por Craveiro *et al.* (2013). Entretanto, o isolado ChinNPV-IA, que previamente foi agrupado aos isolados ChinNPV-IE, -IF e -IG (Craveiro *et al.*, 2013), não agrupou aos outros seis isolados formando um clado em separado. Considerando que este isolado,



ChinNPV-IA, difere acentuadamente quanto à região geográfica (Guatemala), cultura (algodão) e safra (1972) dos outros seis isolados obtidos de larvas coletadas em plantações de soja em diferentes regiões brasileiras, safras de 2006 a 2008, é de se esperar que ChinNPV-IA não agrupe aos outros isolados. Assim, a topologia obtida sugere que os isolados brasileiros apresentem ancestral comum diferente do isolado ChinNPV-IA.



**Figura 11. Árvore filogenética (NJ) baseada em sequências de genomas completos de sete isolados de ChinNPV (IA a IG).** Topologia similar a descrita por Craveiro *et al.* (2013).

### 2.3.3. Comparação de ChinNPV-IE com outros *Alphabaculovirus*

O genoma de ChinNPV-IE foi comparado com os *Alphabaculovirus* ChchNPV, TnSNPV, MacoNPV-B (Grupo II) e com a espécie-tipo AcMNPV (Grupo I) (Tabela 9 e Tabela 9). Como esperado, os vírus proximalmente relacionados ChinNPV-IE, ChchNPV e TnSNPV compartilham tamanho de genomas semelhantes possuindo alta similaridade de sequências nucleotídicas (Craveiro *et al.*, 2013). No entanto, o genoma de ChinNPV-IE é 4.738 pb maior que TnSNPV e 10.490 pb menor que ChchNPV (Tabela 9). O alinhamento total do genoma de ChinNPV-IE comparado com ChchNPV, TnSNPV, MacoNPV-B e AcMNPV revelou ChinNPV-IE colinear a ChchNPV e TnSNPV, mas não a MacoNPV-B e AcMNPV (Figura 12). O alinhamento múltiplo do genoma completo de ChinNPV-IE utilizando o programa *Mauve v. 2.0* apresentou 18

regiões de colinearidade local com sequências homólogas comuns a outros *Alphabaculovirus* analisados (Figura 13). Regiões invertidas relativas ao genoma de ChinNPV-IE foram observadas na maior parte da sequência genômica interna do genoma de MacoNPV-B (Figura 12D). Os genes de ChinNPV-IE apresentaram homologia com 82 ORFs de AcMNPV, 110 ORFs de MacoMNPV-B, 134 ORFs de ChchNPV e 122 ORFs de TnSNPV (Tabela 9). As ORFs relatadas como únicas de ChchNPV (Chch-24, -34, -36 e -90) e TnSNPV (Tn-36 e -62) apresentaram homologia com ChinNPV-IE ORF-25, -35, -36, -88 e ChinNPV-IE ORF-40 e -65, respectivamente (Tabela 9).

**Tabela 9. Características do genoma de ChinNPV-IE comparadas com as de outros *Alphabaculovirus***

| Características                | <i>Alphabaculovirus</i> |         |           |          |          |
|--------------------------------|-------------------------|---------|-----------|----------|----------|
|                                | ChinNPV-IE              | AcMNPV  | MacoNPV-B | ChchNPV  | TnSNPV   |
| Tamanho do genoma (pb)         | 139,132                 | 133,894 | 158,482   | 149,622  | 134,394  |
| Conteúdo GC (%)                | 39.3                    | 40.7    | 40.0      | 39.1     | 39.0     |
| Total de ORFs                  | 140                     | 156     | 168       | 151      | 145      |
| Sequência codante (%)          | 80                      | 91      | 89        | 83       | 95       |
| Sequências <i>hr</i>           | -                       | 9       | 4         | -        | -        |
| Genes <i>bro</i>               | 2                       | 1       | 7         | 4        | 2        |
| Média % ID a.a com ChinNPV     | -                       | 48.6    | 47.8      | 73.1     | 74.5     |
| Homólogos em ChinNPV           | -                       | 82      | 110       | 134      | 122      |
| ORFs únicas de ChinNPV         | 2                       | -       | -         | -        | -        |
| Nº de acesso do <i>GenBank</i> | KJ631622                | L22858  | AY126275  | AY864330 | DQ017380 |

**Tabela 10. Comparação das 141 ORFs putativas de ChinNPV-IE com genes homólogos nos *Alphabaculovirus***

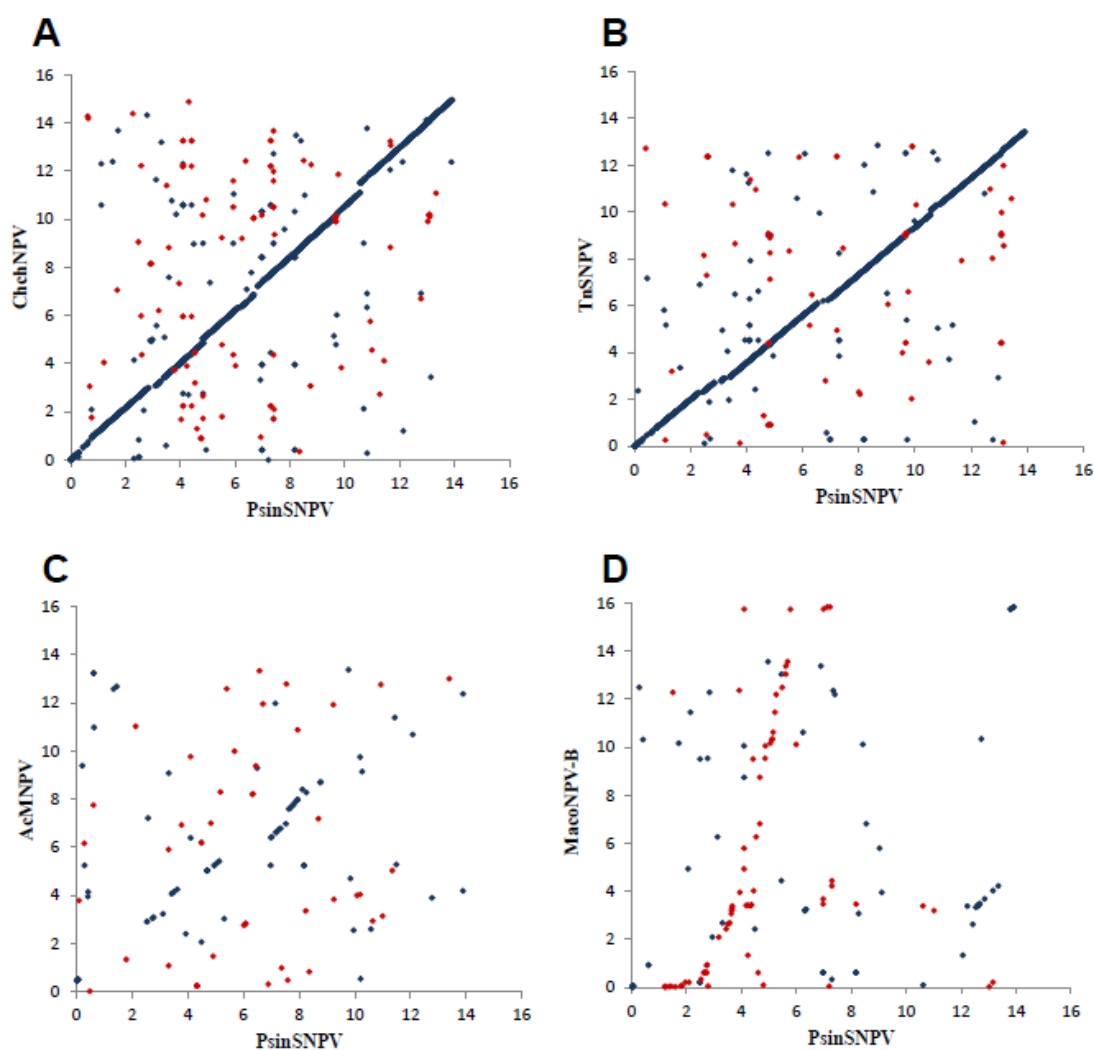
| ChinNPV |                      |               |              | Homólogos (ORF; %ID) |       |         |       |        |        |           |       |
|---------|----------------------|---------------|--------------|----------------------|-------|---------|-------|--------|--------|-----------|-------|
| ORF     | Nome                 | Posição       | Tamanho (nt) | AcMNPV               |       | ChchNPV |       | TnSNPV |        | MacoNPV-B |       |
| 1       | <i>polh</i>          | 1→741         | 741          | 8                    | 88,84 | 1       | 96,34 | 1      | 95,12  | 1         | 92,28 |
| 2       | <i>orf1629</i>       | 738←1,937     | 1,200        | 9                    | 39,6  | 2       | 76,98 | 2      | 62,23  | 2         | 28,16 |
| 3       | <i>pK1</i>           | 1,961→2,779   | 819          | 11                   | 42,57 | 3       | 95,22 | 3      | 92,65  | 3         | 50,00 |
| 4       | <i>hoar</i>          | 2,804←4,822   | 2,019        | -                    |       | 4       | 57,20 | 4      | 55,17  | 4         | 22,62 |
| 5       | <i>unknown</i>       | 5,401→5,916   | 516          | -                    |       | -       |       | -      |        | -         |       |
| 6       | <i>unknown</i>       | 6,016←6,195   | 180          | 152                  | 59,62 | -       |       | -      |        | -         |       |
| 7       | <i>unknown</i>       | 6,338→7,312   | 975          | -                    |       | -       |       | 7      | 31     | -         |       |
| 8       | <i>unknown</i>       | 7,325→7,624   | 300          | -                    |       | -       |       | -      |        | -         |       |
| 9       | <i>odv-e56/pif-5</i> | 7,721→8,797   | 1,077        | 148                  | 51,22 | 6       | 75,98 | 8      | 81,84  | 6         | 51,57 |
| 10      | <i>me53</i>          | 8,896←9,999   | 1,104        | 139                  | 24,57 | 7       | 89,67 | 9      | 86,68  | 7         | 40,72 |
| 11      | <i>exon-0/ie-0</i>   | 10,298→11,248 | 951          | 141                  | 34,72 | 8       | 79,25 | 10     | 74,05  | 167       | 39,08 |
| 12      | <i>p49</i>           | 11,265→12,689 | 1,425        | 142                  | 51,44 | 11      | 96,84 | 11     | 96,42  | 166       | 62    |
| 13      | <i>odv-e18</i>       | 12,699→12,944 | 246          | 143                  | 56,9  | 12      | 53,09 | 12     | 54,32  | 165       | 56,7  |
| 14      | <i>odv-ec27</i>      | 12,985→13,866 | 882          | 144                  | 52,54 | 13      | 93,81 | 13     | 89,19  | 164       | 61,72 |
| 15      | <i>unknown</i>       | 13,872→14,153 | 282          | 145                  | 46,75 | 14      | 89,25 | 14     | 82,80  | 163       | 55    |
| 16      | <i>unknown</i>       | 14,203←14,826 | 624          | 146                  | 33,33 | 15      | 80,19 | 15     | 76,33  | 162       | 44    |
| 17      | <i>ie-1</i>          | 14,865→17,207 | 2,343        | 141a                 | 34,72 | 16      | 66,38 | 16     | 82,60  | 161       | 36,10 |
| 18      | <i>p74/pif-0</i>     | 17,320→19,296 | 1,977        | 138                  | 53,54 | 17      | 91,34 | 17     | 90,47  | 159       | 54,68 |
| 19      | <i>p10</i>           | 19,350←19,616 | 267          | -                    |       | 18      | 86,67 | 18     | 90,91  | 158       | 51,35 |
| 20      | <i>p26a</i>          | 19,668←20,519 | 852          | 136                  | 33,73 | 19      | 79,93 | 19     | 78,09  | 157       | 57,32 |
| 21      | <i>unknown</i>       | 20,680→20,979 | 300          | 29                   | 37    | 20      | 92,93 | 20     | 95,96  | 156       | 53    |
| 22      | <i>lef-6</i>         | 21,007←21,478 | 472          | 28                   | 35    | 21      | 69,51 | 21     | 75,30  | 155       | 44,03 |
| 23      | <i>dbp</i>           | 21,49←22,473  | 984          | 25                   | 30,26 | 22      | 74,14 | 22     | 84,29  | 154       | 32,64 |
| 24      | <i>unknown</i>       | 22,572→23,003 | 432          | 26                   | 33    | 23      | 61,70 | 23     | 64,79  | 153       | 37    |
| 25      | <i>unknown</i>       | 23,058→24,329 | 1,272        | -                    |       | 24      | 48,36 | -      |        | -         |       |
| 26      | <i>unknown</i>       | 24,423←25,165 | 743          | 34                   | 32,39 | 25      | 67,33 | 24     | 62,19  | 152       | 59    |
| 27      | <i>vubi</i>          | 25,175→25,408 | 234          | 35                   | 68,42 | 26      | 97,40 | 25     | 100,00 | 151       | 70,15 |
| 28      | <i>unknown</i>       | 25,405→25,635 | 231          | -                    |       | 27      | 80,82 | 25b    | 74,32  | 150       | 34    |
| 29      | <i>39K</i>           | 25,719←26,696 | 978          | 36                   | 34,04 | 28      | 73,93 | 26     | 71,56  | 149       | 45,21 |
| 30      | <i>lef-11</i>        | 26,671←27,138 | 468          | 37                   | 41,38 | 29      | 77,42 | 27     | 76,36  | 148       | 60,00 |
| 31      | <i>unknown</i>       | 27,090←27,821 | 732          | 38                   | 60,64 | 30      | 83,51 | 28     | 85,02  | 147       | 71    |
| 32      | <i>unknown</i>       | 27,962←28,489 | 528          | 63                   | 22,75 | 31      | 58,29 | 29     | 68,00  | -         |       |
| 33      | <i>bro-a</i>         | 28,758→30,152 | 1,395        | -                    |       | 55      | 51    | -      |        | 20        | 53,54 |
| 34      | <i>p47</i>           | 30,905←32,098 | 1,194        | 40                   | 54,76 | 33      | 92,73 | 31     | 92,70  | 144       | 65,00 |
| 35      | <i>unknown</i>       | 32,298→33,053 | 756          | -                    |       | 34      | 53,11 | -      |        | -         |       |
| 36      | <i>unknown</i>       | 33,168←33,758 | 591          | -                    |       | 36      | 37,61 | -      |        | -         |       |
| 37      | <i>lef-8</i>         | 34,016←36,844 | 2,829        | 50                   | 53,10 | 37      | 80,91 | 33     | 80,91  | 140       | 67,15 |
| 38      | <i>bjdp</i>          | 36,868→37,905 | 1,038        | -                    |       | 38      | 76,19 | 34     | 80,71  | 139       | 30,11 |
| 39      | <i>iap-3</i>         | 37,972←38,820 | 849          | 27                   | 31,60 | 39      | 71,48 | 35     | 74,37  | 138       | 33,33 |
| 40      | <i>unknown</i>       | 38,978→39,154 | 177          | -                    |       | -       |       | 36     | 52,54  | -         |       |

|    |                    |               |       |     |       |    |       |     |       |     |       |
|----|--------------------|---------------|-------|-----|-------|----|-------|-----|-------|-----|-------|
| 41 | <i>unknown</i>     | 39,151←39,759 | 609   | -   |       | 40 | 68,98 | 37  | 64,22 | 137 | 32    |
| 42 | <i>unknown</i>     | 39,821→40,237 | 417   | 53  | 48,48 | 41 | 94,20 | 38  | 91,39 | 136 | 59    |
| 43 | <i>unknown</i>     | 40,242←41,354 | 1,113 | -   |       | 42 | 64,81 | 39  | 63,93 | 135 | 38    |
| 44 | <i>unknown</i>     | 41,395←41,628 | 234   | -   |       | 43 | 81,82 | 40  | 81,82 | 134 | 47    |
| 45 | <i>lef-10</i>      | 41,588→41,815 | 228   | 53a | 48,48 | 44 | 85,53 | 41  | 94,67 | 133 | 62,67 |
| 46 | <i>vp1054</i>      | 41,676→42,680 | 1,005 | 54  | 36,96 | 45 | 86,90 | 42  | 83,93 | 132 | 57,19 |
| 47 | <i>unknown</i>     | 42,804→43,040 | 237   | -   |       | 46 | 67,06 | 43  | 67,03 | 131 | 49    |
| 48 | <i>unknown</i>     | 42,916→43,314 | 399   | -   |       | 47 | 73,64 | 44  | 80,30 | 130 | 50    |
| 49 | <i>unknown</i>     | 43,556→44,068 | 513   | 57  | 44,03 | 48 | 82,35 | 45  | 82,46 | 129 | 45    |
| 50 | <i>unknown</i>     | 44,076←44,574 | 499   | 59  | 50,91 | 49 | 76,27 | 46  | 69,05 | 128 | 76    |
| 51 | <i>unknown</i>     | 44,595←44,871 | 277   | 60  | 50,82 | 50 | 94,03 | 47  | 95,52 | 127 | 60    |
| 52 | <i>fp25k</i>       | 45,140←45,828 | 689   | 61  | 62,50 | 51 | 86,85 | -   |       | 116 | 76,88 |
| 53 | <i>lef-9</i>       | 45,910→47,397 | 1,488 | 62  | 68,72 | 52 | 96,51 | 49  | 96,51 | 123 | 77,41 |
| 54 | <i>unknown</i>     | 47,553←47,762 | 210   | -   |       | 53 | 53    | -   |       | -   |       |
| 55 | <i>unknown</i>     | 48,574→48,828 | 255   | -   |       | 56 | 80,00 | 52  | 76,47 | -   |       |
| 56 | <i>unknown</i>     | 48,835→49,224 | 390   | -   |       | 57 | 93,02 | 53  | 93,02 | 115 | 62    |
| 57 | <i>dnapol</i>      | 49,246←52,425 | 3,18  | 65  | 42,95 | 58 | 80,91 | 54  | 81,72 | 114 | 60,86 |
| 58 | <i>desmoplakin</i> | 52,424→54,556 | 2,133 | 66  | 39,58 | 59 | 73,64 | 55  | 77,64 | 113 | 65,31 |
| 59 | <i>lef-3</i>       | 54,672←56,075 | 1,404 | 67  | 25,83 | 60 | 74,93 | 56  | 72,51 | 112 | 32,52 |
| 60 | <i>unknown</i>     | 56,074→56,469 | 396   | 68  | 45,83 | 61 | 75,57 | 57  | 88,55 | 111 | 62    |
| 61 | <i>iap-2</i>       | 56,521→57,414 | 894   | 71  | 25,71 | 62 | 79,22 | 58  | 72,85 | 109 | 45,21 |
| 62 | <i>p26b</i>        | 57,460→58,197 | 738   | -   |       | 63 | 83,68 | 59  | 85,36 | 108 | 32,92 |
| 63 | <i>vcath</i>       | 58,313←59,350 | 1,038 | 127 | 56,35 | 64 | 93,93 | 60  | 92,49 | 28  | 62,87 |
| 64 | <i>chitinase</i>   | 59,464→61,173 | 1,71  | 126 | 68,23 | 65 | 85,59 | 61  | 83,66 | 19  | 67,90 |
| 65 | <i>unknown</i>     | 61,291→62,037 | 747   | -   |       | -  |       | 62  | 61,75 | -   |       |
| 66 | <i>pcna</i>        | 62,009←62,797 | 789   | 49  | 24,09 | 66 | 70,61 | 63  | 73,28 | -   |       |
| 67 | <i>gp37</i>        | 62,929→63,747 | 819   | 64  | 60,50 | 67 | 93,01 | 64  | 91,91 | 32  | 69,77 |
| 68 | <i>phr</i>         | 63,844→65,376 | 1,533 | -   |       | 68 | 60,28 | 65  | 59,00 | -   |       |
| 69 | <i>bro-b</i>       | 65,452←66,912 | 1,461 | -   |       | 69 | 71    | 108 | 40,62 | 58  | 41,52 |
| 70 | <i>unknown</i>     | 67,059←67,344 | 286   | -   |       | 71 | 51,02 | 66  | 51,92 | -   |       |
| 71 | <i>he65</i>        | 68,173→68,877 | 705   | 105 | 40,12 | 73 | 82,21 | -   |       | 27  | 37,94 |
| 72 | <i>ctl</i>         | 68,972→69,121 | 150   | 3   | 45    | 74 | 85,71 | -   |       | 106 | 66    |
| 73 | <i>unknown</i>     | 69,153←69,539 | 387   | 84  | 37,50 | 75 | 39,42 | -   |       | -   |       |
| 74 | <i>vlf-1</i>       | 69,648←70,823 | 1,176 | 77  | 71,90 | 76 | 81,95 | 70  | 81,95 | 105 | 70,33 |
| 75 | <i>unknown</i>     | 70,820←71,194 | 375   | 78  | 33    | 77 | 78,23 | 71  | 78,57 | 104 | 41    |
| 76 | <i>gp41</i>        | 71,217←72,176 | 960   | 80  | 54,40 | 78 | 99,67 | 72  | 95,33 | 103 | 61,87 |
| 77 | <i>unknown</i>     | 72,118←72,834 | 717   | 81  | 50,70 | 79 | 83,25 | 74  | 82,68 | 102 | 66    |
| 78 | <i>tlp20</i>       | 72,722←73,441 | 720   | -   |       | 80 | 71,37 | 75  | 57,19 | 101 | 68    |
| 79 | <i>vp91/p95</i>    | 73,410→75,891 | 2,482 | 83  | 39,71 | 81 | 77,18 | 76  | 78,41 | 100 | 49    |
| 80 | <i>vp39</i>        | 75,970←76,980 | 1,011 | 89  | 40,80 | 82 | 93,13 | 77  | 93,73 | 98  | 50,16 |
| 81 | <i>lef-4</i>       | 76,955→78,370 | 1,416 | 90  | 41,37 | 83 | 78,97 | 78  | 76,06 | 97  | 55,70 |
| 82 | <i>p33</i>         | 78,466←79,227 | 762   | 92  | 51,16 | 84 | 92,03 | 79  | 93,23 | 95  | 66    |
| 83 | <i>p18</i>         | 79,220→79,702 | 483   | 93  | 52,23 | 85 | 97,50 | 80  | 98,12 | 94  | 79    |
| 84 | <i>odv-e25</i>     | 79,699→80,358 | 660   | 94  | 39,25 | 86 | 72,89 | 81  | 73,64 | 93  | 66,36 |
| 85 | <i>helicase</i>    | 80,480←84,109 | 3,63  | 95  | 40,32 | 87 | 84,60 | 82  | 81,59 | 92  | 59    |

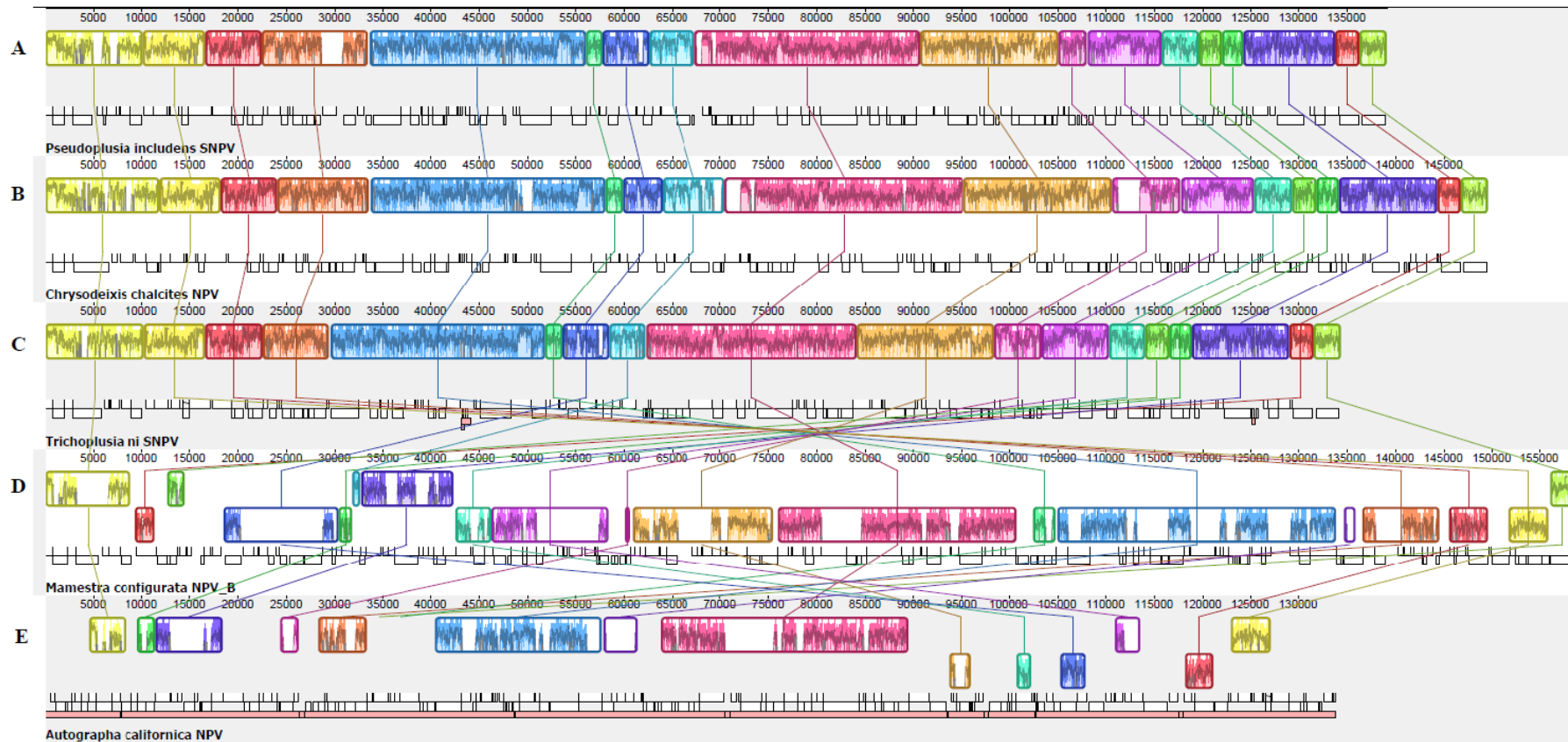
|     |                      |                 |       |     |       |     |       |     |       |     |       |
|-----|----------------------|-----------------|-------|-----|-------|-----|-------|-----|-------|-----|-------|
| 86  | <i>odv-e28/pif-4</i> | 84,036→84,584   | 549   | 96  | 52,45 | 88  | 91,67 | 83  | 81,40 | 91  | 67    |
| 87  | <i>unknown</i>       | 84,618→85,226   | 609   | -   |       | 89  | 58,82 | 84  | 57,06 | -   |       |
| 88  | <i>unknown</i>       | 85,253→85,483   | 231   | -   |       | 90  | 51,25 | -   |       | -   |       |
| 89  | <i>38K</i>           | 85,514←86,470   | 957   | 98  | 40,84 | 91  | 84,54 | 86  | 86,98 | 87  | 63    |
| 90  | <i>lef-5</i>         | 86,363→87,226   | 864   | 99  | 41,24 | 92  | 72,47 | 87  | 72,07 | 86  | 53    |
| 91  | <i>p6.9</i>          | 87,220←87,492   | 273   | 100 | 40,9  | 93  | 83    | 88  | 74,6  | 85  | 54,2  |
| 92  | <i>p40</i>           | 87,579←88,721   | 1,143 | 101 | 42,52 | 94  | 83,51 | 89  | 87,93 | 84  | 63    |
| 93  | <i>p12</i>           | 88,744←88,995   | 252   | 102 | 31,33 | 95  | 86,00 | 90  | 90,11 | 83  | 56    |
| 94  | <i>p48/p45</i>       | 89,098←90,237   | 1,14  | 103 | 46,46 | 96  | 91,56 | 91  | 89,97 | 82  | 69    |
| 95  | <i>p87/vp80</i>      | 90,272→91,981   | 1,71  | -   |       | 97  | 74,87 | 92  | 82,67 | 81  | 45,54 |
| 96  | <i>odv-ec43</i>      | 92,145→93,221   | 1,077 | 109 | 47,57 | 99  | 99,44 | 94  | 99,15 | 79  | 60    |
| 97  | <i>unknown</i>       | 93,257→93,535   | 279   | -   |       | 100 | 55,68 | 95  | 84,78 | 78  | 57    |
| 98  | <i>odv-e66</i>       | 93,579←95,606   | 2,028 | 46  | 36,08 | 101 | 79,91 | 96  | 80,09 | 77  | 57,94 |
| 99  | <i>p13</i>           | 95,692←96,606   | 915   | -   |       | 102 | 89,47 | 97  | 85,53 | 75  | 55,52 |
| 100 | <i>unknown</i>       | 96,925→97,374   | 450   | -   |       | 103 | 57,43 | 99  | 40,26 | -   |       |
| 101 | <i>unknown</i>       | 97,403←98,398   | 996   | -   |       | 104 | 40,06 | 100 | 44,41 | -   |       |
| 102 | <i>unknown</i>       | 98,569←98,961   | 393   | -   |       | 105 | 53,85 | -   |       | -   |       |
| 103 | <i>unknown</i>       | 99,155→100,27   | 1,116 | -   |       | 106 | 63,95 | 101 | 67,44 | 71  | 35    |
| 104 | <i>unknown</i>       | 100,279←100,965 | 687   | 106 | 62,50 | 107 | 88,21 | 102 | 83,77 | 70  | 65    |
| 105 | <i>unknown</i>       | 101,008←102,591 | 1,584 | -   |       | 108 | 73,91 | 103 | 77,14 | 69  | 25    |
| 106 | <i>unknown</i>       | 102,683←103,636 | 954   | -   |       | 109 | 53,63 | 104 | 48,73 | -   |       |
| 107 | <i>pif-3</i>         | 103,738←104,376 | 639   | 115 | 44,25 | 110 | 89,29 | 105 | 86,15 | 67  | 44    |
| 108 | <i>unknown</i>       | 104,461←104,889 | 429   | -   |       | 111 | 49,64 | -   |       | 66  | 24    |
| 109 | <i>unknown</i>       | 105,014→105,589 | 576   | -   |       | 112 | 87,30 | 107 | 86,24 | 118 | 34    |
| 110 | <i>sod29</i>         | 105,654→106,109 | 456   | 33  | 74,50 | 115 | 93,33 | 109 | 91,33 | 65  | 80,67 |
| 111 | <i>unknown</i>       | 106,182←106,823 | 642   | -   |       | 116 | 43,37 | 110 | 44,19 | -   |       |
| 112 | <i>unknown</i>       | 106,979←107,251 | 273   | -   |       | 117 | 70,30 | 111 | 67,37 | -   |       |
| 113 | <i>unknown</i>       | 107,485←107,997 | 513   | -   |       | 118 | 61,76 | 112 | 64,71 | -   |       |
| 114 | <i>dUTPase</i>       | 108,162←108,626 | 465   | -   |       | 119 | 43,51 | -   |       | 73  | 36,61 |
| 115 | <i>calyx/pep</i>     | 108,903→109,892 | 990   | 131 | 43    | 121 | 98,06 | 113 | 43,40 | 61  | 68    |
| 116 | <i>rr2</i>           | 109,984←110,937 | 954   | -   |       | 122 | 92,45 | 114 | 87,42 | 52  | 58,73 |
| 117 | <i>unknown</i>       | 111,057→111,455 | 399   | -   |       | 123 | 88,64 | 115 | 87,12 | 51  | 36    |
| 118 | <i>unknown</i>       | 111,483→112,607 | 1,125 | -   |       | 124 | 88,50 | 116 | 89,63 | -   |       |
| 119 | <i>unknown</i>       | 112,649←113,938 | 1,29  | 18  | 23,16 | 125 | 75,34 | 117 | 78,21 | 50  | 42    |
| 120 | <i>unknown</i>       | 113,940→114,320 | 381   | -   |       | 126 | 69,92 | 118 | 69,84 | 49  | 40    |
| 121 | <i>alkexo</i>        | 114,292→115,536 | 1,245 | 133 | 37,50 | 127 | 81,55 | 119 | 77,48 | 48  | 42    |
| 122 | <i>unknown</i>       | 115,555←116,295 | 741   | -   |       | 128 | 78,97 | 120 | 72,18 | 47  | 37    |
| 123 | <i>fgf</i>           | 116,574→117,686 | 1,113 | 35  | 31,67 | 130 | 62,31 | 122 | 68,94 | 46  | 39,52 |
| 124 | <i>pif-1</i>         | 117,791←119,347 | 1,557 | 119 | 49,05 | 131 | 89,42 | 123 | 88,40 | 44  | 50    |
| 125 | <i>unknown</i>       | 119,338←119,970 | 633   | -   |       | 132 | 63,58 | 124 | 63,10 | -   |       |
| 126 | <i>gp16</i>          | 119,992←120,279 | 288   | 130 | 33,68 | 133 | 73,68 | 125 | 88,42 | 10  | 44,09 |
| 127 | <i>p24</i>           | 120,215←121,030 | 816   | 129 | 48,15 | 134 | 80,40 | 126 | 84,90 | 11  | 57,78 |
| 128 | <i>unknown</i>       | 121,156→121,560 | 405   | -   |       | 135 | 68,57 | 127 | 62,59 | -   |       |
| 129 | <i>lef-2</i>         | 121,490→122,161 | 672   | 6   | 37,62 | 136 | 75,66 | 128 | 80,00 | 13  | 47,66 |
| 130 | <i>38.7 K</i>        | 122,255←123,394 | 1,14  | 13  | 45,45 | 137 | 70,00 | 129 | 74,26 | 31  | 37    |

|     |                  |                 |       |    |       |     |       |     |       |     |       |
|-----|------------------|-----------------|-------|----|-------|-----|-------|-----|-------|-----|-------|
| 131 | <i>lef-1</i>     | 123,440←124,081 | 642   | 14 | 44,39 | 138 | 79,91 | 130 | 88,32 | 30  | 51,15 |
| 132 | <i>unknown</i>   | 124,121→124,582 | 462   | -  |       | 139 | 60,00 | 131 | 52,98 | -   |       |
| 133 | <i>ptp2</i>      | 124,584←125,081 | 498   | -  |       | 140 | 85,28 | 132 | 87,12 | 33  | 48,65 |
| 134 | <i>egt</i>       | 125,241→126,806 | 1,566 | 17 | 49,31 | 141 | 93,78 | 133 | 90,79 | 34  | 65,83 |
| 135 | <i>unknown</i>   | 127,010→127,588 | 579   | -  |       | 142 | 86,26 | 135 | 79,12 | 35  | 33    |
| 136 | <i>unknown</i>   | 127,771←130,542 | 2,772 | -  |       | 143 | 75,19 | -   |       | 37  | 33    |
| 137 | <i>pkip</i>      | 131,117→131,620 | 504   | -  |       | 146 | 85,63 | 139 | 78,44 | 40  | 39,52 |
| 138 | <i>arif-1</i>    | 131,695←132,675 | 981   | 21 | 26,06 | 147 | 59,27 | 140 | 58,91 | 42  | 34,15 |
| 139 | <i>pif-2</i>     | 132,670→134,091 | 1,422 | 22 | 59,33 | 148 | 93,15 | 141 | 88,77 | 43  | 63    |
| 140 | <i>f protein</i> | 134,217←136,193 | 1,977 | 23 | 21,90 | 150 | 94,10 | 143 | 92,44 | 8   | 41    |
| 141 | <i>rr1</i>       | 136,539←138,884 | 2,346 | -  |       | 151 | 83,80 | 144 | 83,92 | 168 | 54,12 |

As linhas sombreadas em cinza indicam os *core genes*.



**Figura 12.** Matriz de similaridade entre os genomas de ChinNPV-IE e de ChchNPV (A), TnSNPV (B), AcMNPV (C) e MacoNPV-B (D). Análise de *dot plot* utilizando o programa LBDotView software v.1.0. Os pontos representam as regiões homólogas entre os genomas, sendo a cor azul na orientação direta e vermelha na orientação do complemento inverso do genoma de ChinNPV-IE.



**Figura 13.** Alinhamento múltiplo do genoma completo de ChinNPV-IE (A) com os *Alphabaculovirus* ChchNPV (B), TnSNPV (C), MacoNPV-B (D) e AcMNPV (E) usando o programa *Mauve v. 2.0*. As linhas e blocos coloridos conectam regiões de colinearidade local (*Local Collinear Blocks* – LCB) em que blocos de cores idênticas indicam regiões homólogas compartilhadas por dois ou mais genomas. LCBs abaixo da linha preta horizontal indicam possíveis rearranjos do tipo inversão em relação à ChinNPV-IE. Os números mostram a posição dos nucleotídeos no genoma de cada vírus (pb).

#### 2.3.4. Replicação, transcrição e genes estruturais

Os genes de baculovírus são classificados com base em suas funções durante o ciclo de infecção viral como: replicação de DNA, transcrição de RNA, proteínas estruturais de ODVs e BVs e proteínas envolvidas na infectividade oral (Miele *et al.*, 2011). Os mecanismos de replicação do genoma de baculovírus não são totalmente compreendidos. Vários estudos têm sido desenvolvidos para tentar identificar os genes responsáveis pela replicação e tradução do DNA viral. Genes homólogos aos fatores essenciais de replicação do DNA como *late expression factor 1 (lef-1)*, *lef-2*, *lef-3*, *DNA polimerase (dnapol)*, *p6.9*, *38K*, *helicase (hel)* e *immediate early 1 (ie-1)* estão todos presentes no genoma de ChinNPV-IE. Além disso, o genoma de ChinNPV-IE contém genes homólogos aos genes *proliferating cell nuclear antigen (pcna)*, *major early-transcribed protein 53 (me53)*, *DNA binding protein (dbp)*, *alkaline exonuclease (alkexo)* e *exon-0 / ie-0*, que são conhecidos por estarem envolvidos na replicação do DNA viral. Entretanto, estes genes não estão descritos em todos os baculovírus (Mikhailov *et al.*, 2003; Ferrelli *et al.*, 2012).

O sistema de transcrição de AcMNPV é ativado em duas principais fases. Na primeira fase, os genes *lef-4*, *lef-8*, *lef-9* e *p47* são transcritos para codificar as quatro subunidades do complexo da RNA polimerase viral (Guarino *et al.*, 1998). Este complexo atua sobre a transcrição de genes e processamento do mRNA, incluindo a adição do 5'-cap e da cauda poli-A. Em seguida, os fatores de ativação de transcrição *lef-5* e *very late factor 1 (vlf-1)* são transcritos (van Oers e Vlak, 2007). Homólogos a todos estes genes também foram encontrados no genoma de ChinNPV-IE. Além disso, alguns genes que, supostamente, não são essenciais e estão envolvidos na regulação da transcrição de AcMNPV, estão também presentes no genoma de ChinNPV-IE: *lef-6*, *lef-11*, *39K*, *lef-10* e *protein kinase 1 (pk1)*.

O genoma de ChinNPV-IE possui 28 genes pertencentes a baculovírus que codificam proteínas estruturais. As proteínas estruturais das partículas virais são codificadas pelos genes *poliedrina (polh)*, *orf1629*, *pk1*, *occlusion derived virus envelope protein 18 (odv-e18)*, *occlusion derived virus enveloped capsid protein 27 (odv-ec27)*, *p10*, *viral protein 1054 (vp1054)*, *few polyhedra protein/25k (fp25k)*, *desmoplakin*, *41-kDa glycoprotein (gp41)*, *telokin-like*



*peptide 20 (tlp20)*, *viral protein 91 (vp91/p95)*, *vp39*, *p33*, *odv-e25*, *p87/vp80*, *odv-ec43*, *odv-e66*, *p13*, *calyx/polyhedrin enveloped protein (calyx/pep)*, *p24*, *per os infectivity factor 0 (p74/pif-0)*, *pif-1*, *pif-2*, *pif-3*, *odv-e28/pif-4*, *odv-e56/pif-5* e proteína de fusão (*F protein*). Os seis genes PIF que são constituintes das ODVs e estão envolvidos na infectividade oral apresentaram alta similaridade com as seqüências PIFs de ChchNPV e TnSNPV. O gênero *Alphabaculovirus* é subdividido em Grupos I e II, com base no conteúdo de genes, em particular, a proteína de fusão das partículas BVs: GP64 e proteína F, respectivamente. ChinNPV-IE, um *Alphabaculovirus* Grupo II, possui, como esperado, o gene homólogo a proteína F.

### 2.3.5. Metabolismo de nucleotídeos e reparação de DNA

Os *Alphabaculovirus* Grupo II e *Betabaculovirus* codificam genes envolvidos na biossíntese de nucleotídeos. ChinNPV possui os genes que codificam a ribonucleotídeo reductase (RR) grande (RR1) e pequena (RR2) e as subunidades da proteína dUTPase. As proteínas RR são enzimas envolvidas na formação de desoxirribonucleotídeos e ribonucleotídeos (Rohrmann, 2013). A proteína dUTPase é responsável por impedir a incorporação de dUTP no DNA mutado (Herniou *et al.*, 2003; Rohrmann, 2013). Poly (ADP-ribose) polimerase (PARP) e glico-hidrolase poli ADP-ribose (PARG) são enzimas envolvidas na síntese de riboses ADP que ativam e recrutam enzimas de reparo de DNA (Miwa *et al.*, 1974; Chen *et al.*, 2001). Embora tenha sido relatado que os *Alphabaculovirus* Grupo II codificam homólogos de PARG (Rohrmann, 2013), esse gene não foi identificado no genoma de ChinNPV.

A CPD fotoliase, codificada pelo gene *DNA fotoliase (phr)*, atua no reparo de danos no DNA do tipo dímeros de pirimidina ciclobutano induzidos pela radiação ultravioleta (UV). O gene *phr* foi identificado no genoma de ChinNPV entretanto mais estudos serão necessários para confirmar e investigar a atividade dessa enzima.

### 2.3.6. Genes auxiliares

Os genes auxiliares *ubiquitina viral (vubi)*, *catepsina (v-cath)*, *quitinase (chiA)*, *37 kDa-glicoproteína (gp37)*, *conotoxina (ctl)*, *superóxido dismutase 29 (sod29)*, *fatores de crescimento de fibroblastos (fgf)*, *fosfotirosina fosfatase (ptp)*, *transferase de glucose ecdisona (egt)*, *fator de infectividade do rearranjo de*

*actina 1 (arif-1)*, *inibidor de apoptose 2 (iap-2)*, *iap-3* e *p35/p49* foram encontrados no genoma de ChinNPV. Os genes auxiliares não são essenciais para a replicação do DNA, tradução ou formação de partículas virais. No entanto, estes genes conferem vantagens seletivas ao vírus como foi observado em homólogos dos genes auxiliares de ChinNPV descritos na literatura. A atividade dos genes *chiA* e *v-cath* é claramente perceptível em larvas de *C. includens* infectadas com ChinNPV, em que essas proteínas causam degradação e liquefação do cadáver do hospedeiro (Zhu *et al.*, 2013; Mu *et al.*, 2014). Os genes *fgf*, *ptp* e *egt* foram relatados por estar envolvidos no comportamento hiperativo do inseto-hospedeiro aumentando a motilidade larval, impedindo a muda e prolongando a vida do inseto (Nguyen *et al.*, 2013). O gene *ptp* foi previamente relatado apenas nos *Alphabaculovirus* Grupo I (Nguyen *et al.*, 2013; Rohrmann, 2013), no entanto, esse gene também está presente no genoma de ChinNPV.

### **2.3.7. Homologous regions (hrs) são ausentes no genoma de ChinNPV-IE**

*Homologous regions (hrs)* são sequências repetitivas com núcleo de palíndromes imperfeitos que se distribuem no genoma como únicos ou dispostos em *tandem*. Estas sequências repetitivas estão presentes no genoma de baculovírus e outros vírus de invertebrados proximamente relacionados (van Oers e Vlak, 2007). Estas regiões atuam como fatores de transcrição de genes precoce em NPVs e/ou podem agir como origens de replicação em NPVs e GVs (Pearson e Rohrmann, 1995; Kuzio *et al.*, 1999; Pang *et al.*, 2001; Liu *et al.*, 2015). Além disso, as *hrs* estão envolvidas com sítios de recombinação homóloga sendo regiões com alta suscetibilidade a variações genéticas (Hayakawa *et al.*, 1999). A presença de *hrs* é uma característica comum encontrada nos genomas dos vírus pertencentes aos quatro gêneros da família *Baculoviridae* (Ferrelli *et al.*, 2012). No entanto, regiões *hrs* não foram encontradas no genoma de ChinNPV, como também relatado em *Buzura supressaria* NPV (BusuNPV), ChchNPV, TnSNPV e *Agrotis segetum* GV (AgseGV) (Faulkner *et al.*, 1997; Willis *et al.*, 2005; Hilton e Winstanley, 2008; Zhu *et al.*, 2014).

### **2.3.8. Genes *bro* de ChinNPV**

O genoma de ChinNPV contém dois *baculovirus repeated ORFs* (genes *bro*), nomeados de acordo com sua ordem no genoma: *bro-a* (ORF-33) e *bro-b*

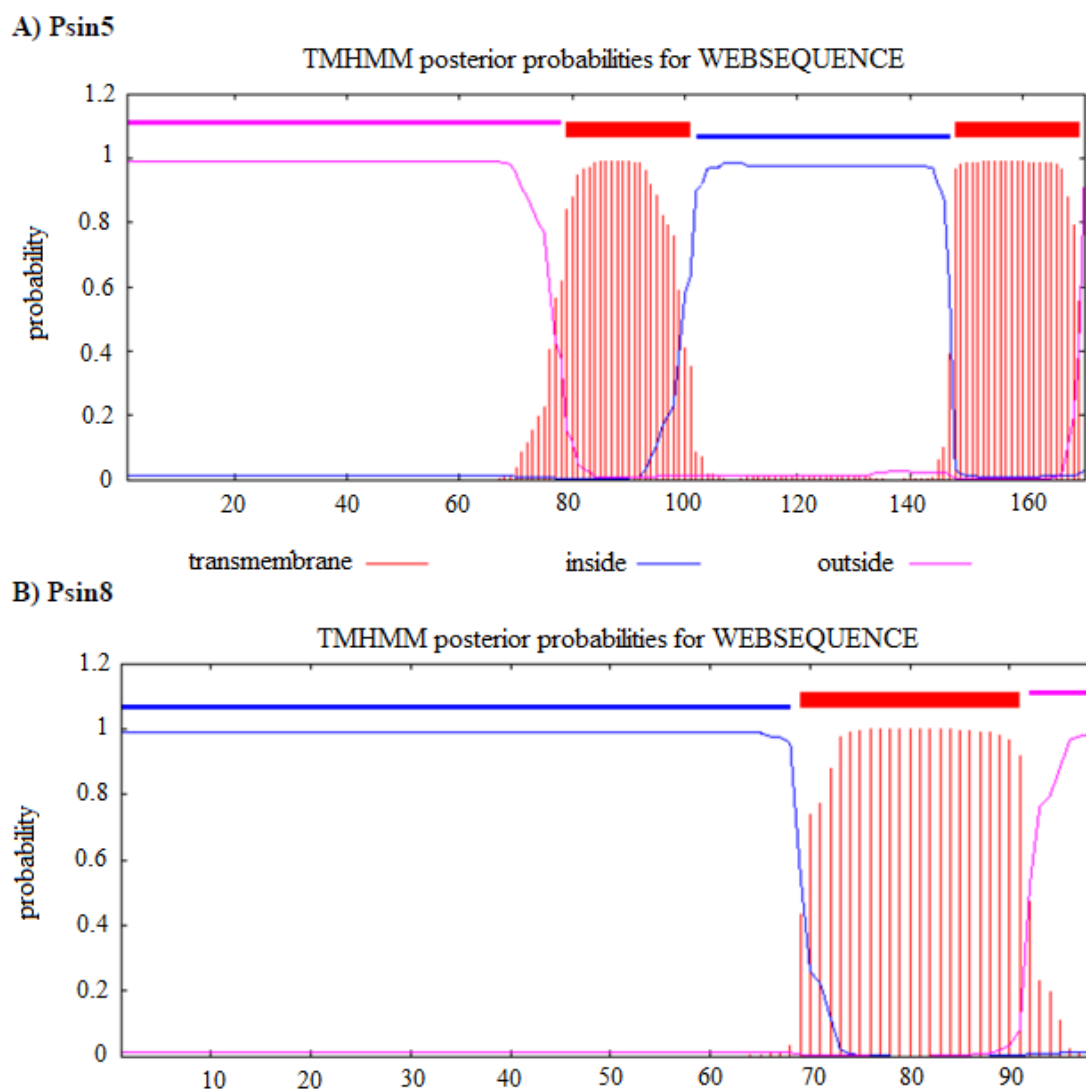
(ORF-69). Os genes *bro* geralmente ocorrem em *Alpha*-, *Beta*- e *Gammabaculovirus*, variando em número de cópias e tamanho entre os vírus (Kang *et al.*, 1999; Kuzio *et al.*, 1999; Iyer *et al.*, 2002; Bideshi *et al.*, 2003; Zhou *et al.*, 2012b). Estes genes foram inicialmente relatados nos baculovírus e, posteriormente, homólogos aos genes *bro* foram identificados em outros vírus de insetos de dsDNA como os entomopoxvirus e entomoiridovirus (Afonso *et al.*, 1999; Bawden *et al.*, 2000; Jakob *et al.*, 2001). Proteínas BRO apresentam o domínio altamente conservado N-terminal (BRO-N) de ligação ao DNA nos primeiros 100-150 a.a e o domínio C-terminal (BRO-C) variável (Zemskov *et al.*, 2000; Iyer *et al.*, 2002). As funções das proteínas BRO não estão bem elucidadas, entretanto foi proposto que essa proteína pode estar envolvida na replicação de DNA e/ou regulação da transcrição do hospedeiro e atuar como fator de replicação viral na fase tardia (Kang *et al.*, 1999; Zemskov *et al.*, 2000; Iyer *et al.*, 2002; Bideshi *et al.*, 2003). Mesmo tendo uma deleção de 425 pb (386-811 nt) (~ 140 aa), o gene *bro-b* de ChinNPV compartilha 72% de identidade com o gene *bro-b* de ChchNPV. O gene *bro-a* de ChinNPV apresentou maior similaridade com o gene *bro-m* de *Lymantria xyli*na MNPV e *bro-m* de *Mamestra brassicae* MNPV com 58 e 53% de identidade, respectivamente. Em contraste com a proteína BRO-B que apresentou um único domínio BRO-N, a proteína BRO-A de ChinNPV contém dois domínios BRO-N [Pfam: PF02498] na posição de 14 a 118 e 139 a 234 aminoácidos. Além disso, a proteína BRO-A de ChinNPV contém um domínio de função desconhecida DUF3627 [Pfam: PF12299] na posição de 334 a 423 aminoácidos. Embora ChinNPV, ChchNPV e TnSNPV estejam intimamente relacionados, os seus genes *bro* não apresentaram alta similaridade.

### 2.3.9. ORFs únicas de ChinNPV

Duas ORFs putativas, Psin5 e Psin8, foram identificadas e são exclusivas do genoma de ChinNPV. Essas ORFs não apresentaram similaridade significativa com outras ORFs de baculovírus até então descritas e exibem domínios conservados preditos por InterProScan 5. Psin5 codifica uma proteína de 172 aminoácidos (a.a.) com massa molecular de 18,73 kDa que apresenta baixa similaridade com um transdutor de sinal e ativador de transcrição de proteínas nas espécies aviárias, *Pseudopodoces humilis* [GenBank:

XP\_005533966] (%ID=80%, cobertura=40% e e-value=3E-12). *Psin8* codifica uma proteína de 150 a.a. e 11,9 kDa e não apresenta similaridade significativa ( $p > 0,01$ ) com nenhum gene presente no banco de dados *GenBank*. O motivo promotor tardio TAAG combinado com o promotor precoce TATA (TATAAGG motivo) foi identificado a cerca de 100 pb a montante do códon de início de ambos os genes *Psin5* e *Psin8* (5031 e 7219 pb, respectivamente). Acredita-se que esses promotores podem ser transcritos tanto pela RNA polimerase II do hospedeiro quanto pela RNA polimerase viral, sugerindo que estes genes podem ser expressos na fase precoce e tardia da infecção. A busca por famílias de proteínas, domínios e sítios funcionais identificou domínios transmembrana nas duas proteínas *Psin5* e *Psin8*. Usando TMHMM v 2.0, foram preditas hélices transmembrana nas posições 79-101aa e 148-170aa na proteína *Psin5* e na posição 69-91aa na proteína *Psin8* (Figura 14). Mais estudos serão necessários para caracterizar a expressão desses genes e a função dessas duas novas proteínas encontradas em ChinNPV.

Este é o primeiro relato da organização e descrição genômica do baculovírus patogênico a *C. includens*. A sequência do genoma completo de sete isolados de ChinNPV apresentaram entre 138.869 pb e 140.859 pb e são altamente similares com identidade maior que 96%. Entretanto, variações pontuais foram identificadas e foi observada uma inserção correspondente ao gene *bro* nos genomas de ChinNPV-IA, ChinNPV-IC e ChinNPV-ID totalizando 142 ORFs nos genomas desses isolados, uma ORF a mais do que o número encontrado nos outros isolados analisados. O genoma de ChinNPV não contém sequências *hrs* típicas de baculovírus e possui duas ORFs com domínios transmembrana preditos (*Psin5* e *Psin8*) que são exclusivas desse vírus.



**Figura 14. Hélices transmembrana previstas a partir das seqüências deduzidas de aminoácidos das ORFs únicas de ChinNPV.** As hélices transmembrana foram previstas da seqüências proteicas das (A) ORF-5 e (B) ORF-8 de ChinNPV utilizando o programa *TMHMM v. 2.0* e estão indicadas em vermelho entre 79 e 101, 148 e 170 aminoácidos em Psin5 e entre 69 e 91 aminoácidos em Psin8.

## 2.4. Conclusões

- Este estudo é o primeiro relato da descrição genômica detalhada de ChinNPV.
- Os genomas de ChinNPV (-IA a -IG) possuem tamanho variando de 138.869 pb (ChinNPV-IB) a 140.859 pb (ChinNPV-IC) com conteúdo de CG de ~39%.
- O genoma de ChinNPV-IE possui 141 ORFs putativas: 37 *core genes* de baculovírus; 100 genes encontrados em outros baculovírus; 2 genes *bro* (*bro-a* ORF-33 e *bro-b* ORF-69); duas ORFs únicas de ChinNPV (Psin5 e Psin8).
- *Homologous regions (hrs)* não foram identificadas no genoma de ChinNPV, embora essas regiões repetidas estejam presentes no genoma da maioria dos baculovírus.
- As duas ORFs exclusivas do genoma de ChinNPV (Psin5 e Psin8) codificam proteínas que apresentam hélices transmembrana preditas pelo programa TMHMM v 2.0.
- Os genomas dos sete isolados possuem alta similaridade com identidade mínima de 96,4% (ChinNPV-IC e ID).
- Há uma heterogeneidade genética entre os isolados de ChinNPV com a presença de *indels*, polimorfismos e pequenas inserções ao longo de todo o genoma.
- Os genomas de ChinNPV-IA, -IC e -ID apresentaram, além das ORFs de *bro-a* e *bro-b*, uma ORF a mais correspondente ao gene *bro-c*.
- O genoma de ChinNPV é colinear e apresenta alta similaridade com os genomas de ChchNPV e TnSNPV.
- ORFs relatadas como únicas a ChchNPV e TnSNPV foram encontradas no genoma de ChinNPV.

### **CAPÍTULO 3: Identificação do gene *p26* e sua evolução na família *Baculoviridae***

Este capítulo é parte do estudo de um artigo publicado em *BMC Genomics* (2015) 16: 27:

#### **The genome sequence of *Pseudoplusia includens* single nucleopolyhedrovirus and an analysis of *p26* gene evolution in the baculoviruses**

Saluana R. Craveiro, Peter W. Inglis, Roberto C. Togawa, Priscila Grynberg, Fernando L. Melo, Zilda Maria A. Ribeiro, Bergmann M. Ribeiro, Sônia N. Bão, Maria Elita B. Castro\*

### 3.1. Introdução

No vírus *Autographa californica* MNPV, o gene *p26* é um gene expresso na fase precoce da infecção e codifica uma proteína de 240 resíduos de aminoácidos (Liu *et al.*, 1986). Esse gene se localiza na região montante ao gene *p10* ao qual se sobrepõe compartilhando o mesma sequência 5' terminal (Raykin *et al.*, 1986).

Até o momento, há poucos estudos que relatam a caracterização e a atividade funcional do gene *p26* nos baculovírus. Portanto, a função do gene *p26* não é bem conhecida na família *Baculoviridae*.

Estudos iniciais relataram que a deleção da região final 3' do gene *p26* não causou efeitos na replicação viral *in vitro* e *in vivo* (Rodems e Friesen, 1993; Simón *et al.*, 2008). Além disso, não foi identificada a presença da proteína P26 nas partículas ODVs indicando que essa proteína não está envolvida no processo de infecção oral de baculovírus (Braunagel *et al.*, 2003).

Posteriormente, Simón *et al.* (2008) mostraram que o gene *p26* não é requerido para a replicação do DNA e não é essencial para a infecção oral em AcMNPV.

Em estudo de um mutante de AcMNPV com os genes *p10*, *p26* e *p74* deletados foram observadas partículas OBs sem a presença de vírions oclusos. Nos ensaios de deleção individual de genes foi sugerido que *p10*, *p26* e *p74* estariam, em ação conjunta, envolvidos na regulação do processo de oclusão dos vírions nas partículas OBs (Wang *et al.*, 2009).

Desta forma, foi proposto que a deleção do gene *p26* não influencia significativamente na infectividade, virulência e produção das partículas OBs atuando como gene auxiliar com efeitos sutis na replicação e transmissão do vírus.

Recentemente, estudos em *Spodoptera exigua* NPV relataram os efeitos da deleção individual dos genes *Se87* e *Se129*, duas cópias homólogas do gene *p26* presentes no genoma desse vírus. A deleção do gene *Se87* não apresentou efeito significativo, entretanto, a deleção de *Se129* resultou na diminuição da patogenicidade das partículas OBs em *Spodoptera exigua* NPV (Serrano *et al.*, 2015).



Neste trabalho, o gene *p26* foi identificado em todos os genomas sequenciados de NPVs e investigada sua evolução na família *Baculoviridae* utilizando ferramentas de bioinformática.

## 3.2. Materiais e Métodos

### 3.2.1. Determinação e caracterização da sequência nucleotídica do gene *p26* de baculovírus

As sequências de aminoácidos do gene *p26* de todos os baculovírus sequenciados e depositados no *GenBank* foram determinadas e suas posições genômicas foram identificadas sendo categorizadas em 3 diferentes posições: posição 1 adjacente ao gene *p10*, posição 2 adjacente ao gene *iap-2* e posição 3 adjacente aos genes *ptp1* e *ptp2*. As propriedades físico-químicas como ponto isoelétrico (pI), peso molecular (Altschul *et al.*) e composição de aminoácidos foram calculadas utilizando a ferramenta computacional *ProtParam Tool v. 3.2.1* (<http://www.expasy.org/tools/protparam.html>) para todas as proteínas P26 de baculovírus. Para a predição de hélices transmembrana e predição da presença e localização de sítios de clivagem de peptídeo sinal foram utilizados softwares *TMHMM Server v.2.0* (Sonnhammer *et al.*, 1998; Krogh *et al.*, 2001) e *SignalP v.4.1* (Petersen *et al.*, 2011), respectivamente.

### 3.2.2. Análise filogenética da proteína P26

As sequências de aminoácidos das proteínas P26 de baculovírus foram alinhadas utilizando *MUSCLE v.3.5* (Edgar, 2004) e analisadas pelo método da Inferência filogenética Bayesiana (IB) utilizando o programa *MrBayes v.3.1.2* (Huelsenbeck e Ronquist, 2001) para a realização da análise evolutiva do gene *p26* dentro da família *Baculoviridae*. Para determinação do modelo de evolução mais apropriado às sequências peptídicas foi utilizado o programa *ProtTest* (Abascal *et al.*, 2005). O modelo *LG* (Le e Gascuel, 2008) foi selecionado por apresentar o melhor ajuste para o conjunto de dados das sequências de P26. O programa *MrBayes* não suporta o modelo de evolução *LG*, portanto, definições de probabilidade foram obtidas pelo comando `aamodel=mixed rates=invgamma`, o que permitiu a determinação do melhor modelo de substituição para o conjunto de dados em análise. O algoritmo *Monte Carlo via Cadeia de Markov* (MCMC)

foi iniciado a partir de uma árvore aleatória e cinco Cadeias de Markov foram processadas para 600.000 gerações ( $p < 0,01$ ) amostradas a cada 100 gerações. Nas análises, foram descartadas 25% das amostras iniciais (*burn-in*) e o restante utilizado para determinar as distribuições dos valores de probabilidade posterior e para o cálculo da árvore consenso. A edição das árvores foi realizada com o auxílio dos programas *FigTree v.1.3.1* e *Inkscape v. 0.48*.

### 3.3. Resultados e Discussão

#### 3.3.1. As proteínas P26 de baculovírus se correlacionam de acordo com o contexto de seu posicionamento genômico

O gene *p26* está presente em número variado no genoma de todos os *Alphabaculovirus* Grupos I e II, exceto no vírus *Spodoptera littoralis multiple nucleopolyhedrovirus*, SpliNPV [GenBank: AF325155] em que não foi identificado o gene *p26*. No entanto, a maioria dos vírus com mais de uma cópia do gene *p26* pertence aos *Alphabaculovirus* Grupo II. Além dos baculovírus, esse gene foi também identificado em *Anomala cuprea* entomopoxvirus [GenBank:BAO49539] (Mitsubishi *et al.*, 2014)

Os genes *p26* foram identificados nos genomas dos baculovírus até então sequenciados e depositados no *GenBank* e verificou-se que, dentre os 27 *Alphabaculovirus* Grupo II analisados, 18 apresentaram mais de uma cópia do gene *p26*. No entanto, nas sequências genômicas dos *Alphabaculovirus* Grupo I, apenas *Choristoneura fumiferana* (Cf) MNPV, *Choristoneura occidentalis* (Choc) NPV e *Choristoneura rosaceana* (Salmela e Schröder) NPV apresentaram duas sequências homólogas ao gene *p26* (Tabela 11).

A análise de posicionamento do gene *p26* no genoma dos baculovirus mostrou que as posições desse gene nos *Alphabaculovirus* coincidiram com o grupo (Grupos I e II) ao qual o vírus pertence: nos *Alphabaculovirus* contendo uma única cópia, o gene *p26* é conservado na posição adjacente ao gene *p10* (posição 1); a segunda cópia de *p26* é conservada adjacente aos genes *ptp1* e *ptp2* nos *Alphabaculovirus* Grupo I (posição 3) e na posição adjacente ao gene *iap-2* nos *Alphabaculovirus* Grupo II (posição 2) (Tabela 11).

O ponto isoelétrico (pI) e peso molecular (Altschul *et al.*) da sequência deduzida de aminoácidos da proteína P26 dos *Alphabaculovirus* com genomas

contendo duas cópias foram calculados e apresentaram valores coincidentes com seu posicionamento genômico. Proteínas P26 encontradas na posição 1 possuem  $pI$  médio de  $6,7 \pm 0,76$  e  $PM$  médio de  $30.526 \pm 2.387$ ; proteínas P26 na posição 2 possuem  $pI$  médio de  $8,2 \pm 0,71$  e  $PM$  médio de  $27.354 \pm 1.306$ ; e proteínas P26 na posição 3 possuem  $pI$  médio de  $9,5 \pm 0,07$  e  $PM$  médio de  $30.113 \pm 24$ . Os valores médios foram analisados utilizando o teste *t*-Student e as médias dos  $pI$  apresentaram diferença significativa ( $p < 0,05$ ) entre as três posições (Figura 15). A diferença de ponto isoelétrico sugere que a sequência de P26 na posição 3 apresenta os aminoácidos mais básicos do que nas posições 1 e 2.

**Tabela 11. Sequências de aminoácidos utilizadas na análise filogenética da proteína P26**

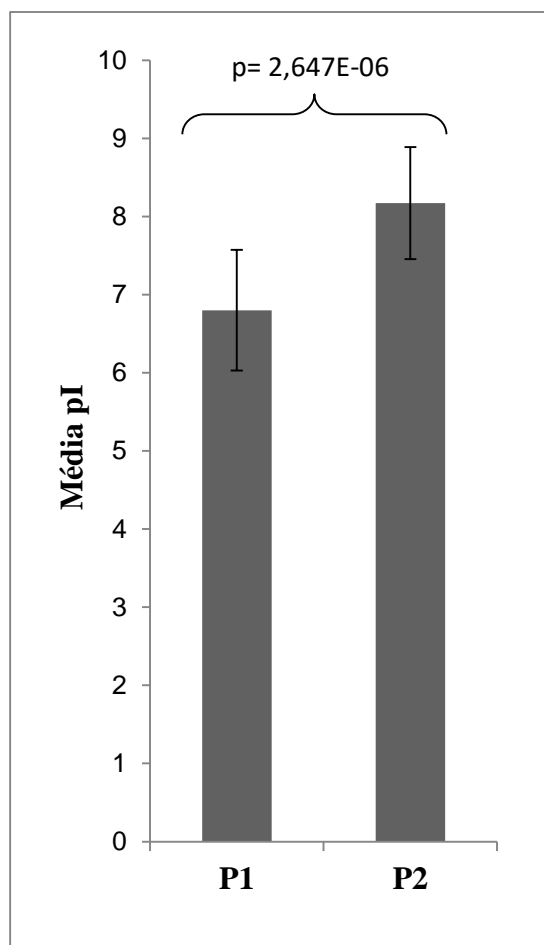
| <i>Alphabaculovirus</i>                     | Abreviatura | GenBank  | Posição<br>(ORF/Peptídeo sinal) |                |                |
|---|-------------|----------|---------------------------------|----------------|----------------|
|   |             |          | 1 <sup>a</sup>                  | 2 <sup>b</sup> | 3 <sup>c</sup> |
| <b>Grupo I</b>                              |             |          |                                 |                |                |
| <i>Autographa californica</i> NPV           | AcMNPV      | L22858   | 136 /<br>Não                    | -              | -              |
| <i>Bombyx mandarina</i> NPV                 | BomaNPV     | FJ882854 | 120 /<br>Não                    | -              | -              |
| <i>Bombyx mori</i> NPV                      | BmNPV       | L33180   | 121 /<br>Não                    | -              | -              |
| <i>Rachiplusia ou</i> MNPV                  | RoMNPV      | AY14547  | 129 /<br>1                      | -              | -              |
| <i>Antheraea pernyi</i> NPV                 | AnpeNPV     | DQ48603  | 127 /<br>0                      | -              | -              |
| <i>Anticarsia gemmatalis</i> MNPV           | AgMNPV      | DQ81366  | 132 /<br>2                      | -              | -              |
| <i>Choristoneura fumiferana</i> DEF<br>MNPV | CfDEFMNPV   | AY32740  | 130 /<br>2                      | -              | -              |
| <i>Choristoneura fumiferana</i> MNPV        | CfMNPV      | AF512031 | 128 /<br>Não                    | -              | 7 / Sim        |
| <i>Choristoneura murinana</i> NPV           | ChmuNPV     | KF894742 | 22 / Não                        | -              | -              |
| <i>Choristoneura occidentalis</i> NPV       | ChocNPV     | KC961303 | 21 / Não                        | -              | 143 /<br>Sim   |
| <i>Choristoneura rosaceana</i> NPV          | ChroNPV     | KC961304 | 22 / Não                        | -              | 145 /<br>Sim   |
| <i>Epiphyas postvittana</i> NPV             | EppoNPV     | AY04326  | 119 /<br>5                      | -              | -              |
| <i>Hyphantria cunea</i> NPV                 | HycuNPV     | AP009046 | 21 / Não                        | -              | -              |
| <i>Maruca vitrata</i> NPV                   | MaviNPV     | EF125867 | 104 /<br>Não                    | -              | -              |
| <i>Orgyia pseudotsugata</i> MNPV            | OpMNPV      | U75930   | 132 /<br>Não                    | -              | -              |
| <i>Plutella xylostella</i> MNPV             | PlxyMNPV    | DQ45700  | 132 /<br>3                      | -              | -              |
| <i>Thysanoplusia orichalcea</i> NPV         | ThorNPV     | JX467702 | 133 /<br>Não                    | -              | -              |
| <b>Grupo II</b>                             |             |          |                                 |                |                |
| <i>Adoxopyes honmai</i> NPV                 | AdhoNPV     | AP006270 | 31 / Sim                        | -              | -              |
| <i>Adoxophyes orana</i> NPV                 | AdorNPV     | EU591746 | 30 / Sim                        | -              | -              |
| <i>Agrotis ipsilon</i> MNPV                 | AgipMNPV    | EU839994 | 149 / Sim                       | 100 /<br>Não   | -              |
| <i>Agrotis segetum</i> NPV                  | AgseNPV     | DQ12384  | 1                               | 142 / Sim      | 94 / Não       |
| <i>Apocheima cinerarium</i> NPV             | ApciNPV     | FJ914221 | 102 / Sim                       | 40 / Não       | -              |
| <i>Buzura suppressaria</i> NPV              | BusuNPV     | KF611977 | 8 / Sim                         | 52 / Não       | -              |
| <i>Chrysodeixis chalcites</i> NPV           | ChchNPV     | AY86433  | 0                               | 19 / Não       | 63 / Não       |
| <i>Clanis bilineata</i> NPV                 | ClbiNPV     | DQ50442  | 8                               | 19 / Sim       | 61 / Não       |
| <i>Ectropis obliqua</i> NPV                 | EcobNPV     | DQ83716  | 5                               | 18 / Não       | 52 / Não       |
| <i>Euproctis pseudoconsersa</i> NPV         | EupsNPV     | FJ227128 | 51 / Sim                        | 61 / Não       | -              |

|                                      |                  |                     |           |                   |   |
|--------------------------------------|------------------|---------------------|-----------|-------------------|---|
| <i>Helicoverpa armigera</i> MNPV     | HearMNPV         | EU730893            | 151 / Sim | 101 / Não         | - |
| <i>Helicoverpa armigera</i> NPV      | HearNPV          | AF303045            | 22 / Sim  | -                 | - |
| <i>Helicoverpa armigera</i> NPV G4   | HearNPV-G4       | AF271059            | 22 / Sim  | -                 | - |
| <i>Helicoverpa armigera</i> NPV NNg1 | HearNPV-<br>NNg1 | AP010907            | 21 / Sim  | -                 | - |
| <i>Helicoverpa zea</i> SNPV          | HZSNPV           | AF334030            | 21 / Sim  | -                 | - |
| <i>Hemileuca sp.</i> NPV             | HespNPV          | KF158713<br>AY39449 | 22 / Sim  | 60 / Não          | - |
| <i>Leucania separata</i> NPV         | LeseNPV          | 0                   | 20 / Não  | -                 | - |
| <i>Lymantria dispar</i> MNPV         | LdMNPV           | AF081810<br>GQ20254 | 40 / Sim  | -                 | - |
| <i>Lymantria xyliana</i> MNPV        | LyxyMNPV         | 1                   | 36 / Sim  | -                 | - |
| <i>Mamestra brassicae</i> MNPV       | MabrMNPV         | JQ798165            | 147 / Sim | 99 / Não<br>109 / | - |
| <i>Mamestra configurata</i> NPV A    | MacoNPV-A        | U59461<br>AY12627   | 158 / Sim | Não<br>108 /      | - |
| <i>Mamestra configurata</i> NPV B    | MacoNPV-B        | 5                   | 157 / Sim | Não               | - |
| <i>Orgyia leucostigma</i> NPV        | OrleNPV          | EU309041            | 20 / Não  | 62 / Não          | - |
| <i>Spodoptera exigua</i> MNPV        | SeMNPV           | AF169823            | 129 / Sim | 87 / Não          | - |
| <i>Spodoptera frugiperda</i> MNPV    | SfMNPV           | EF035042            | 131 / Sim | 86 / Não          | - |
| <i>Spodoptera litura</i> NPV II      | SpliNPV-II       | EU780426<br>DQ01738 | 135 / Sim | 91 / Não          | - |
| <i>Trichoplusia ni</i> SNPV          | TnSNPV           | 0                   | 19 / Sim  | 59 / Não          | - |

<sup>a</sup>P26 ORF number adjacent to *p10* gene

<sup>b</sup>P26 ORF number adjacent to *iap-2* gene

<sup>c</sup>P26 ORF number adjacent to *ptp -1* and *-2* genes



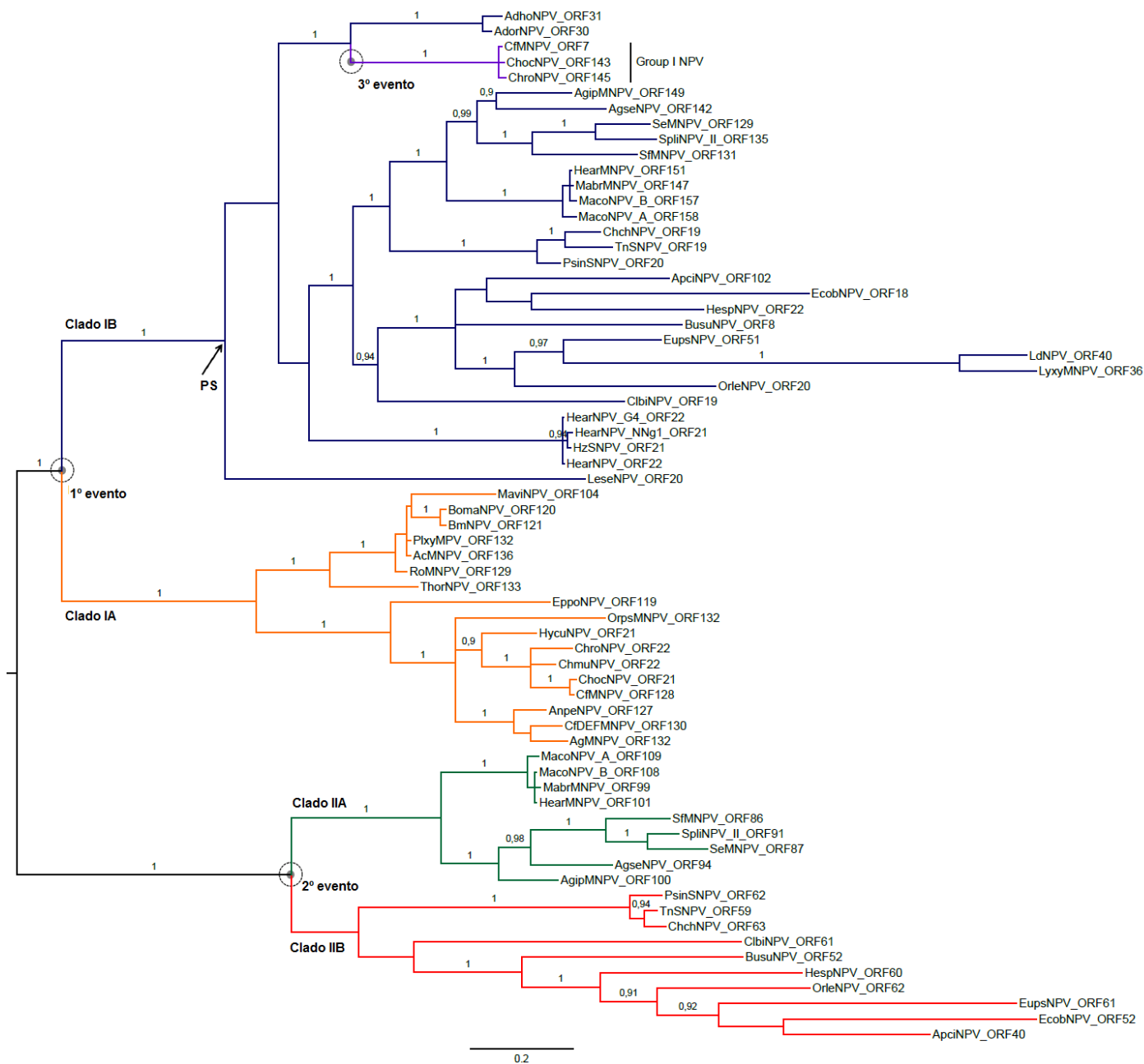
**Figura 15.** Média do ponto isoeletrico (pI) da proteína P26 nas posições adjacentes ao gene *p10* (P1) e ao gene *iap-2* (P2). Diferença significativa foi observada entre os pI das duas posições (Teste *t*-Student).

### 3.3.2. Eventos independentes de aquisição do gene *p26* dentro da família *Baculoviridae*

A árvore filogenética foi obtida por meio da inferência filogenética Bayesiana (IB) usando as sequências deduzidas de aminoácidos dos genes *p26* encontrados em todos os genomas de *Alphabaculovirus* até então sequenciados. O filograma apresentou quatro clados bem determinados (IA, IB, IIA e IIB) aninhados dentro dos clados maiores I e II (Figura 16). O clado IA contém as proteínas P26 presentes nos *Alphabaculovirus* Grupo I e o clado II presentes nos *Alphabaculovirus* Grupo II. As proteínas P26 presentes no clado IB pertencem também aos *Alphabaculovirus* do Grupo II, exceto no grupo monofilético com CfMNPV\_ORF31, ChocMNPV\_ORF143 e ChroNPV\_ORF145 que pertencem aos *Alphabaculovirus* Grupo I. Os clados I (IA e IB) e II estão correlacionados com a posição do gene *p26* no genoma: o clado I contém as

cópias *p26* adjacentes ao gene *p10* (posição 1), exceto para CfMNPV\_ORF31, ChocMNPV\_ORF143 e ChroNPV\_ORF145 que são adjacentes a *ptp1* e *ptp2* (posição 3). O clado II contém cópias *p26* adjacentes a *iap-2* (posição 2). Este padrão de agrupamento sugere a ocorrência de três eventos independentes de aquisição do gene *p26* pelos baculovírus. O primeiro evento de aquisição ocorreu na posição 1 do genoma ancestral comum de todos os baculovírus contendo este gene. O segundo evento de aquisição gerou a segunda cópia de *p26* na posição 2 do genoma dos *Alphabaculovirus* Grupo II com dois homólogos de *p26*. Finalmente, o terceiro evento de aquisição ocorreu no grupo monofilético CfMNPV, ChocNPV e ChroNPV em que a segunda cópia de *p26* na posição 3 foi adquirida pelos *Alphabaculovirus* Grupo I com dois homólogos de *p26*.

A aquisição de genes de baculovírus pode ser resultado de duplicação e transferência horizontal de genes por elementos transponíveis e/ou recombinação homóloga. O segundo evento de aquisição ocorreu provavelmente por transferência horizontal de genes. Neste caso, a hipótese de duplicação de genes pode ser refutada uma vez que a similaridade entre as cópias de *p26* de um mesmo vírus é baixa (menos de 30% de identidade). Além disso, clados I e II são claramente separados indicando que as cópias *p26* não se originaram a partir de um ancestral comum recente. No terceiro caso de aquisição, há uma alta similaridade entre a segunda cópia *p26* dos *Alphabaculovirus* Grupo I e a primeira cópia *p26* dos *Alphabaculovirus* Grupo II, agrupando esses genes no mesmo clado (clado IB). Portanto, a segunda cópia *p26* dos *Alphabaculovirus* Grupo I pode ter sido adquirida de um *Alphabaculovirus* Grupo II.



**Figura 16. Cladograma baseado nas seqüências deduzidas de aminoácidos de P26.** O padrão de agrupamento sugere a ocorrência de três eventos independentes de aquisição do gene *p26* que estão marcados por círculos. A seta indica o nó em que ocorreu a aquisição do peptídeo sinal. Valores nos ramos indicam as probabilidades posterior Bayesiana (valor > 0,9). A árvore foi enraizada por abordagem *midpoint*. A escala da barra indica o número de substituições por sítio.



### 3.3.3. Presença de peptídeo sinal nas proteínas P26 de *Alphabaculovirus* com dois homólogos

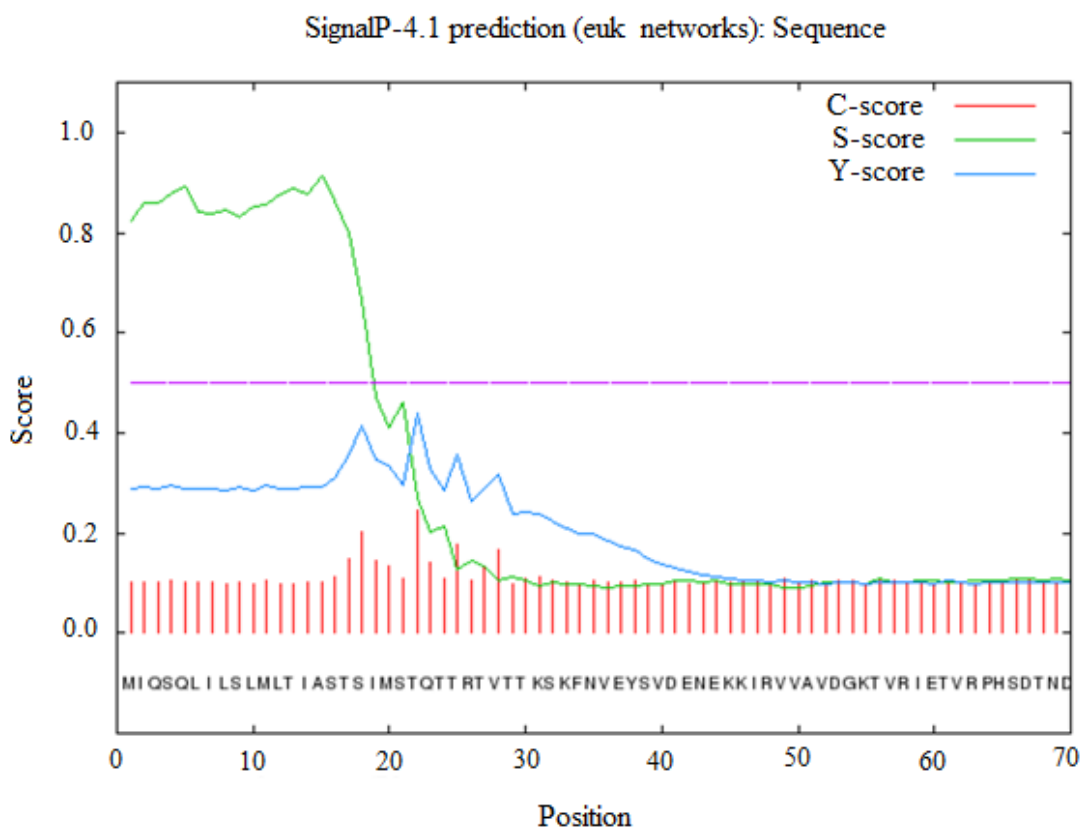
A presença e localização de sítios de clivagem de peptídeo sinal nas proteínas P26 dos *Alphabaculovirus* foram analisadas utilizando *SignalP v.4.1* e os resultados estão apresentados na Tabela 11 e Figura 17. A presença ou ausência de peptídeo sinal nas proteínas P26 está correlacionada com o agrupamento obtido na análise filogenética em que apenas as proteínas P26 do clado IB possuem peptídeo sinal. No entanto, quatro sequências pertencentes a esse clado, LeseNPV\_ORF20, EcobNPV\_ORF18, OrleNPV\_ORF20 e ChchNPV\_ORF19, não apresentaram peptídeo sinal.

O peptídeo sinal direciona as proteínas para suas localizações celulares e extracelulares adequadas. As proteínas marcadas por uma sequência sinal N-terminal são exportadas por meio da via de secreção onde proteínas são transportadas através da membrana citoplasmática e, posteriormente, clivadas na região N-terminal do peptídeo sinal por uma peptidase sinal extracelular.

A presença ou ausência de peptídeo sinal pode explicar as diferenças de valores observados em relação ao peso molecular e aos pontos isoelétricos entre as proteínas P26 analisadas.

A presença de peptídeo sinal na primeira cópia *p26* nos *Alphabaculovirus* Grupo II e na segunda cópia *p26* nos *Alphabaculovirus* Grupo I sugere que este domínio protéico foi adquirido a partir de uma sequência ancestral comum destes vírus. E corrobora a hipótese da segunda cópia *p26* dos *Alphabaculovirus* Grupo I ter sido adquirida de um *Alphabaculovirus* Grupo II.

O genoma de ChinNPV possui duas cópias do gene *p26* e a análise filogenética das sequências de aminoácidos de P26 dos *Alphabaculovirus* apresentou três potenciais eventos de aquisição desses genes na família *Baculoviridae*. Embora a função de P26 não seja bem compreendida, apenas o clado IB apresentou proteínas P26 com peptídeo sinal, o que indica diferenças na atividade e atuação das duas cópias do gene *p26* encontradas no genoma da maioria dos *Alphabaculovirus*. No entanto, estudos da atividade funcional da proteína P26 deverão ser conduzidos para a melhor elucidação do modo de ação das diferentes proteínas P26 de baculovírus



**Figura 17. Peptídeo sinal predito na proteína P26 de ChinNPV-IE utilizando *SignalP v.4.1*.** A linha verde indica os motivos com alta probabilidade de conter o peptídeo sinal (posição 1-21, média de S-score = 0,791). O sítio de clivagem foi predito entre a Ser21 e Thy22 (IMS-TQ; D-score = 0,629).

### 3.4. Conclusões

- As cópias do gene *p26* de baculovírus se correlacionam de acordo com o seu posicionamento genômico estando adjacentes aos genes: *p10* nos *Alphabaculovirus* Grupos I e II; *iap-2* na segunda cópia dos *Alphabaculovirus* Grupo II; *ptp1* e *ptp2* na segunda cópia dos *Alphabaculovirus* Grupo I.
- A aquisição do gene *p26* ocorreu em três eventos independentes de captura dentro da família *Baculoviridae*.
- A presença de peptídeo sinal nas proteínas P26 dos *Alphabaculovirus* Grupo II (adjacente ao gene *iap-2*) indica função distinta dessas proteínas as outras classes de proteínas P26 dos baculovírus.

## DISCUSSÃO GERAL

O primeiro relato de variantes genotípicas ocorreu em 1976 quando mudanças morfológicas foram detectadas em placas de cultura de células infectadas com baculovírus TnSNPV (Potter *et al.*, 1976). Posteriormente, a análise de perfil de restrição foi conduzida para o estudo de variantes de AcMNPV com placas morfológicas diferenciadas (Lee e Miller, 1978). Desde então, a técnica de polimorfismo de comprimento de fragmentos de restrição (RFLP) vem sendo amplamente utilizada para estudos de variações genéticas em amostras de baculovírus. Concomitantemente, polimorfismos pontuais eram estudados por meio do sequenciamento de genes específicos, dentre esses, os genes conservados pertencentes ao grupo dos *core genes* de baculovírus (Khan *et al.*, 2003; Rowley *et al.*, 2010).

Variantes genotípicas de ChinNPV foram previamente identificados em amostras virais de lagartas coletadas em plantações de soja e algodão do Brasil e Guatemala. Por meio da análise de RFLP foram observados polimorfismos entre sete isolados de ChinNPV. Além disso, foram observados bandas submolares caracterizando a presença de mistura de genótipos dentro de um mesmo isolado (Alexandre *et al.*, 2010).

Neste trabalho, as variações genéticas em ChinNPV foram inicialmente investigadas pela análise dos genes *lef-9*, *pif-2* e *helicase*, genes altamente conservados entre os baculovírus. Juntamente com o gene *lef-9*, *pif-2* é considerado um dos genes ancestrais mais conservados e, do lado oposto, encontramos o gene *helicase* considerado um dos mais variáveis entre os *core genes* de baculovírus (Miele *et al.*, 2011). Em contraste com o esperado, o gene *pif-2* apresentou uma alta variação entre os isolados de ChinNPV. O mesmo não foi observado para os genes *lef-9* e *helicase*.

A busca de sinais de seleção diversificadora pode ser realizada em nível de sequências gênicas pois modelos de substituição de códons podem demonstrar como a seleção natural, exercida sobre uma proteína, atua na estrutura gênica. Para identificação das forças evolutivas que estão sob o gene *pif-2* de ChinNPV foi realizado uma análise de seleção com base na relação dN/dS utilizando modelos de substituição de códons. Entre os 158 polimorfismos observados no gene *pif-2* de ChinNPV, 3 mutações não sinônimas apresentaram

sítios sob seleção diversificadora. Mutações não sinônimas causam alterações na sequência de aminoácidos da proteína podendo gerar alterações fenotípicas entre isolados virais. Essas mutações podem ser alvos certos de seleção natural pois influenciam a estrutura da proteína podendo inclusive alterar sua função. Estudos relataram que uma substituição individual na sequência de aminoácidos da proteína Helicase foi capaz de ampliar o espectro de hospedeiro de AcMNPV (Kamita e Maeda, 1997; Argaud *et al.*, 1998). Dentre os 3 sítios sob seleção diversificadora, os sítios nas posições dos códons 275 e 373 apresentaram alterações radicais de aminoácidos que, conseqüentemente, podem causar mudança na estrutura secundária da proteína refletindo em variação fenotípica entre os isolados analisados. As mutações não sinônimas em sua maioria são delétera e rapidamente eliminadas por seleção negativa. Entretanto, essas mutações podem ser fixadas na população por conferir um aumento na eficiência ou uma nova função à proteína. A fixação de mutações comumente ocorre em famílias de multigenes em que há múltiplas cópias de um mesmo gene e a perda funcional de uma das cópias pode ser suprida por outro genótipo (Coleman *et al.*, 1998). Múltiplas cópias são encontradas em isolados de baculovírus podendo, até mesmo, ser observado diferentes genótipos em uma mesma partícula viral o que propicia que mutações não sinônimas sejam mais facilmente fixadas no genoma dos baculovírus.

A relação dN / dS é uma das medidas mais comuns e confiáveis para análise de pressões evolutivas em proteínas. Entretanto, o pequeno número de táxons e conjunto de dados e a baixa divergência são limitações presentes nos modelos utilizados para a análise de seleção. No entanto, sítios sob seleção diversificadora foram identificados por diferentes métodos e abordagens o que habilita a confiabilidade dos resultados obtidos nesse estudo. Além disso, a falta de ferramentas para estudos das pressões evolutivas sob genes inviabiliza outros tipos de análises além das análises aqui apresentadas.

O avanço das técnicas de sequenciamento proporcionou uma maior facilidade e um baixo custo para a determinação da sequência genômica de baculovírus. Assim, o número de estudos em que as variações genéticas eram investigadas ao longo de todo genoma aumentou. Baillie e Bouwer (2012) publicaram um dos primeiros estudos em que foi investigado em nível genômico a diversidade genética presente em uma amostra contendo um mix de diferentes

isolados de HearNPV. A partir de então, esse tipo de análise foi investigada em isolados de diferentes espécies virais como EeGV, HearNPV, AgMNPV, entre outros (Ardisson-Araújo *et al.*, 2014; Nouné e Hauxwell, 2015; Brito *et al.*, 2016).

Para avaliação da diversidade genética no contexto genômico, o genoma completo dos sete isolados de ChinNPV foram sequenciados. Após a montagem e anotação dos genes, o genoma de ChinNPV-IE, isolado utilizado como referência, foi publicado sendo este trabalho o primeiro relato da descrição genômica do baculovírus patogênico a larvas *C. includens*.

Mesmo apresentando uma alta similaridade com o mínimo de 96,4% de identidade, foi observada uma heterogeneidade genética nos genomas dos sete isolados de ChinNPV sendo contabilizados 1.995 variações pontuais ao longo de todo o genoma. Considerando que os genomas dos isolados de ChinNPV possuem tamanho médio de 140 kpb, foi identificado um número pequeno de mutações correspondendo menos que 2% do número de bases nucleotídicas presentes nos genomas. Além disso, erros de sequenciamento, em especial, inserções de adeninas em regiões homopoliméricas, são inerentes à tecnologia 454 Roche e podem estar causando uma superestimativa do número de variações genéticas entre o genoma dos sete isolados. De qualquer forma, essas poucas mutações foram capazes de causar variações fenotípicas sendo os isolados ChinNPV-IA, -IE e -IF os mais patogênicos entre os sete isolados analisados.

Nos genomas dos isolados ChinNPV-IA, -IC e -ID foi observada uma inserção de ~1.700 pb adicionando um gene *bro* a mais no genoma desses isolados. As proteínas BRO possuem um alto nível de plasticidade e variabilidade e são considerados sítios potencialmente favoráveis para hipervariabilidade (Muñoz *et al.*, 1999; Bideshi *et al.*, 2003; Cory *et al.*, 2005). Estudos descrevem algumas cópias dos genes *bro* envolvidas com a replicação de DNA e/ou regulação da transcrição do hospedeiro ou atuando como fator de regulação viral na fase tardia mas, até o momento, a função da família de proteínas BRO não está bem elucidada (Kang *et al.*, 1999; Zemskov *et al.*, 2000; Iyer *et al.*, 2002; Bideshi *et al.*, 2003). Em populações de AgMNPV, BmNPV e HearNPV também foram relatados diferenças no número de cópias de genes *bro* presentes no genoma de diferentes isolados virais (Zhou *et al.*, 2012a; Brito *et al.*, 2016; Ogembo *et al.*, 2009). Entretanto, não há estudos relatando como os

diferentes números de cópias de genes *bro* podem influenciar na variação fenotípica encontrada entre isolados virais. Em ChinNPV, o número de cópias de genes *bro* não está relacionado com a variação fenotípica como a diferença de virulência observada entre os sete isolados, pois ChinNPV-IA, -IE e -IF, considerados os mais patogênicos, apresentaram diferentes número de cópias de genes *bro* como o isolado ChinNPV-IA contendo uma cópia a mais.

Duas cópias do gene *p26* foram identificadas no genoma de ChinNPV o que despertou a curiosidade de se entender quais os eventos de aquisição e qual a ação desses genes durante o ciclo de infecção dos baculovírus.

A hipótese de aquisição do gene *p26* por meio da duplicação de genes foi refutada, mediante a pouca similaridade observada entre as cópias de *p26* e por não terem sua origem de um ancestral comum recente. Assim, os estudos descritos neste trabalho indicam que a aquisição das cópias do gene *p26* no genoma dos baculovírus ocorreu por transferência horizontal de genes. A hipótese proposta por esse estudo é que as cópias do gene *p26* foram inseridas no genoma dos baculovírus por 3 eventos de aquisição independentes. Os primeiro e segundo eventos de aquisição do gene *p26* na família *Baculoviridae* ocorreram, independentemente, com a transferência de *p26* de organismos de outras famílias virais ou até mesmo outros microorganismos para o genoma do baculovírus ancestral. Cópias do gene *p26* foram identificadas no genoma de entomopoxvirus sendo eles uma das possíveis origens desse gene. Entretanto, mais estudos são necessários para a identificação de qual organismo foi o “doador” do gene *p26* para os baculovírus. E, para concluirmos nossa hipótese, o terceiro evento de aquisição pode ter ocorrido pela transferência da segunda cópia *p26* dos *Alphabaculovirus* Grupo I para os *Alphabaculovirus* Grupo II durante a co-infecção de representantes virais desses dois clados em um mesmo inseto hospedeiro.

Apesar de não ser clara a ação da proteína P26 durante o ciclo de infecção dos baculovírus, nossos estudos indicam que as duas cópias de *p26* exercem funções distintas e não uma função com ação conjunta, como foi inicialmente pensado. Estudos relatam que a deleção do gene *p26* não foi capaz de influenciar na infectividade, virulência e produção de partículas OBs (Chaabihi *et al.*, 1993; Rodems e Friesen, 1993; Simón *et al.*, 2008; Wang *et al.*, 2009). Entretanto, esses estudos prévios descrevem a deleção de apenas uma das

cópias do gene *p26*, a cópia localizada adjacente ao gene *p10*. Recentemente, estudos descrevem a deleção individual das duas cópias do gene *p26* do vírus de *Spodoptera exigua* NPV que, na cópia correspondente a ORF-129, resultou na diminuição da patogenicidade das partículas OBs enquanto que, na cópia correspondente a ORF-87, não foi observado efeito significativo (Serrano *et al.*, 2015). Esses últimos dados corroboram com nossos resultados indicando que as duas cópias de *p26* possuem função distinta durante o ciclo de infecção dos baculovírus. Entretanto, mais estudos deverão ser conduzidos para elucidar com mais detalhes como as cópias *p26* atuam na infecção de células do hospedeiro.

A heterogeneidade fenotípica ocorre naturalmente e é mantida em populações de baculovírus no campo sugerindo que a diversidade genética é uma vantagem para a adaptação ao hospedeiro, a evolução e a sobrevivência dos baculovírus. Uma das maiores fontes de diversidade genética em populações virais é o fluxo contínuo de material genético do hospedeiro para as populações virais que ocorre durante a interação do vírus com o inseto-hospedeiro (Gilbert *et al.*, 2016). A colonização do inseto *C. includens* é recente no Brasil e a baixa heterogeneidade genética observada em diferentes populações brasileiras desse inseto pode ser um dos fatores que propiciam a alta similaridade do genoma de isolados de ChinNPV coletados em plantações de soja do país (Palma *et al.*, 2015).

## CONCLUSÕES GERAIS E PERSPECTIVAS

As informações geradas pelos dados obtidos deste estudo dão suporte para concluirmos que:

- ✓ O gene *pif-2* recebeu destaque em análises de diversidade genética presentes nos *core genes pif-2, lef-9* e *helicase* de isolados de ChinNPV pois, em contraste com o esperado, esse gene apresentou um grande número de polimorfismos (158 SNPs).
- ✓ Dentre os polimorfismos observados em *pif-2* de ChinNPV, 23 são variações não-sinônimas dentre os quais 2 sítios estão sob pressão diversificadora sendo variações potencialmente causadoras de alterações significativas na estrutura secundária da proteína PIF2.
- ✓ Este é o primeiro estudo detalhado sobre a descrição genômica de ChinNPV que apresenta DNA fita-dupla circular em torno de 140 kpb e contém 142 ORFs putativas: 37 *core genes* de baculovírus; 100 genes encontrados em outros baculovírus; 3 genes *bro, bro-a* (ORF-33), *bro-b* (ORF-69) e *bro-c* (ORF-110); duas ORFs únicas de ChinNPV (Psin5 e Psin8).
- ✓ Os genomas de sete isolados de ChinNPV possuem alta similaridade com identidade mínima de 96,4%. Entretanto, foi observado uma heterogeneidade genética entre os isolados que apresentaram polimorfismos, *indels* e inserções de pequenos fragmentos ao longo de todo o genoma.
- ✓ As cópias do gene *p26* de baculovírus se correlacionam de acordo com o seu posicionamento genômico: adjacente a *p10* (*Alphabaculovirus* com uma única cópia), *iap-2* (*Alphabaculovirus* Grupo II) e *ptp1* e *ptp2* (segunda cópia dos *Alphabaculovirus* Grupo I). Provavelmente, esses genes foram adquiridos pelos baculovírus em três eventos independentes de captura e a presença de diferentes domínios protéicos indica função distinta entre as duas cópias encontradas nos *Alphabaculovirus* Grupo II.



Os dados e informações obtidas neste estudo abrem importantes possibilidades de investigação:

- ✓ Prospecção de novos isolados do vírus patogênico a larvas *C. includens* para uso no controle da praga.
- ✓ Investigação de mecanismos de ação distintos das proteínas P26 encontradas na família *Baculoviridae*.
- ✓ Caracterização das ORFs únicas de ChinNPV como possíveis novas proteínas virais ainda não descritas na literatura.
- ✓ Estudo detalhado da diversidade genética com base no genoma completo dos isolados de ChinNPV investigando o efeito da heterogeneidade genética na variação fenotípica observada entre os isolados.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Abascal, F., Zardoya, R., Posada, D. (2005). **ProtTest: Selection of Best-fit Models of Protein Evolution**. *Bioinformatics* 21(9): 2104-2105.
- Adams, J. R., McClintock, T. J. (1991). **Baculoviridae: nuclear polyhedrosis virus**. In: *Atlas of invertebrate viruses*. Edited by J. R. Adams and J. R. Bonani. Boca Raton, CRC: 87-214.
- Afonso, C., Tulman, E., Lu, Z., Balinsky, C., Moser, B., Becnel, J., Rock, D., Kutish, G. (2001). **Genome sequence of a baculovirus pathogenic for *Culex nigripalpus***. *Journal of Virology* 75(22): 11157-11165.
- Afonso, C. L., Tulman, E. R., Lu, Z., Oma, E., Kutish, G. F., Rock, D. L. (1999). **The Genome of *Melanoplus sanguinipes* Entomopoxvirus**. *Journal of Virology* 73(1): 533-552.
- Alexandre, T. M., Ribeiro, Z. M. A., Craveiro, S. R., Cunha, F., Fonseca, I. C., Moscardi, F., Castro, M. E. B. (2010). **Evaluation of seven viral isolates as potential biocontrol agents against *Pseudoplusia includens* (Lepidoptera: Noctuidae) caterpillars**. *Journal of Invertebrate Pathology* 105(1): 98-104.
- Alford, R. A., Hammond, A. (1982). **Plusiinae (Lepidoptera: Noctuidae) populations in Louisiana soybean ecosystems as determined with loop lure-baited traps**. *Journal of Economic Entomology* 75(4): 647-650.
- Ali, A., Young, S. Y. (1991). **Influence of larval age and temperature on effectiveness of a nuclear polyhedrosis virus in the soybean looper, *Pseudoplusia includens* (Lepidoptera: Noctuidae) on soybean**. *Biological Control* 1(4): 334-338.
- 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., Bielawski, J. P., Yang, Z. (2002). **Accuracy and power of Bayes prediction of amino acid sites under positive selection**. *Molecular Biology and Evolution* 19(6): 950-958.
- 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): 1.
- Argaud, O., Croizier, L., López Ferber, M., Croizier, G. (1998). **Two key mutations in the host-range specificity domain of the p143 gene of *Autographa californica* nucleopolyhedrovirus are required to kill *Bombyx mori* larvae**. *Journal of General Virology* 79(4): 931-935.

- Assefa, S., Keane, T. M., Otto, T. D., Newbold, C., Berriman, M. (2009). **ABACAS: algorithm-based automatic contiguation of assembled sequences.** *Bioinformatics* 25(15): 1968-1969.
- Ayres, M. D., Howard, S. C., Kuzio, J., Lopez-Ferber, M., Possee, R. D. (1994). **The complete DNA sequence of Autographa californica nuclear polyhedrosis virus.** *Virology* 202(2): 586-605.
- Baillie, V., Bouwer, G. (2012). **High levels of genetic variation within Helicoverpa armigera nucleopolyhedrovirus populations in individual host insects.** *Archives of Virology* 157(12): 2281-2289.
- Bawden, A. L., Glassberg, K. J., Diggans, J., Shaw, R., Farmerie, W., Moyer, R. W. (2000). **Complete genomic sequence of the Amsacta moorei Entomopoxvirus: Analysis and comparison with other poxviruses.** *Virology* 274(1): 120-139.
- Becnel, J. J., White, S. E., Moser, B. A., Fukuda, T., Rotstein, M. J., Undeen, A. H., Cockburn, A. (2001). **Epizootiology and transmission of a newly discovered baculovirus from the mosquitoes Culex nigripalpus and C. quinquefasciatus.** *Journal of General Virology* 82(2): 275-282.
- Bellotti, A. C., Smith, L., Lapointe, S. L. (1999). **Recent advances in cassava pest management.** *Annual Review of Entomology* 44(1): 343-370.
- Bernal, A., Williams, T., Muñoz, D., Caballero, P., Simón, O. (2013). **Complete genome sequences of five Chrysodeixis chalcites nucleopolyhedrovirus genotypes from a Canary Islands isolate.** *Genome Announcements* 1(5): e00873-00813.
- Bernardi, O., Malvestiti, G. S., Dourado, P. M., Oliveira, W. S., Martinelli, S., Berger, G. U., Head, G. P., Omoto, C. (2012). **Assessment of the high-dose concept and level of control provided by MON 87701 X MON 89788 soybean against Anticarsia gemmatalis and Pseudoplusia includens (Lepidoptera: Noctuidae) in Brazil.** *Pest Management Science*.
- Bideshi, D. K., Renault, S., Stasiak, K., Federici, B. A., Bigot, Y. (2003). **Phylogenetic analysis and possible function of bro-like genes, a multigene family widespread among large double-stranded DNA viruses of invertebrates and bacteria.** *Journal of General Virology* 84(9): 2531-2544.
- Bilimoria, S. L. (1991). **The biology of nuclear polyhedrosis viruses.** In: *Viruses of Invertebrates*. Edited by E. Kurstak. New York, Marcel Dekker: 1-72.
- Bottimer, L. J. (1926). **Notes on Some Lepidoptera from Eastern Texas.** *Journal of Agricultural Research* 33: 797-819.
- Boucias, D., Pendland, J. C. (2012). **Principles of insect pathology,** Springer Science & Business Media.

- Braunagel, S., Russell, W., Rosas-Acosta, G., Russell, D., Summers, M. (2003). **Determination of the protein composition of the occlusion-derived virus of *Autographa californica* nucleopolyhedrovirus.** *Proceedings of the National Academy of Sciences* 100(17): 9797-9802.
- Braunagel, S. C., Summers, M. D. (1994). **Autographa californica nuclear polyhedrosis virus, PDV, and ECV viral envelopes and nucleocapsids: structural proteins, antigens, lipid and fatty acid profiles.** *Virology* 202(1): 315-328.
- Braunagel, S. C., Summers, M. D. (2007). **Molecular biology of the baculovirus occlusion-derived virus envelope.** *Current Drug Targets* 8(10): 1084-1095.
- Brito, A. F., Braconi, C. T., Weidmann, M., Dilcher, M., Alves, J. M. P., Gruber, A., De Andrade Zanotto, P. M. (2016). **The pangenome of the *Anticarsia gemmatalis* multiple nucleopolyhedrovirus (AgMNPV).** *Genome Biology and Evolution* 8(1): 94-108.
- Bueno, R. C. O. D. F., Bueno, A. D. F., Moscardi, F., Postali Parra, J. R., Hoffmann-Campo, C. B. (2011). **Lepidopteran larva consumption of soybean foliage: basis for developing multiple-species economic thresholds for pest management decisions.** *Pest Management Science* 67(2): 170-174.
- Bueno, R. C. O. F., Parra, J. R. P., Bueno, A. F., Haddad, M. L. (2009). **Desempenho de Tricogramatídeos como Potenciais Agentes de Controle de *Pseudoplusia includens* Walker (Lepidoptera: Noctuidae).** *Neotropical Entomology* 38: 389-394.
- Bueno, R. C. O. F., Parra, J. R. P., Bueno, A. F., Moscardi, F., Oliveira, J. R. G., Camillo, M. F. (2007). **Sem barreira.** *Revista Cultivar* 93: 12-15.
- Carstens, E. B., Ball, L. A. (2009). **Ratification vote on taxonomic proposals to the International Committee on Taxonomy of Viruses (2008).** *Archives of Virology* 154(7): 1181-1188.
- Chaabihi, H., Ogliastro, M., Martin, M., Giraud, C., Devauchelle, G., Cerutti, M. (1993). **Competition between baculovirus *polyhedrin* and *p10* gene expression during infection of insect cells.** *Journal of Virology* 67(5): 2664-2671.
- Chen, X., Ijkel, W. F. J., Tarchini, R., Sun, X., Sandbrink, H., Wang, H., Peters, S., Zuidema, D., Lankhorst, R. K., Vlak, J. M., Hu, Z. (2001). **The sequence of the *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus genome.** *Journal of General Virology* 82(1): 241-257.
- Chevreux, B., Wetter, T. and Suhai, S. (1999). **Genome sequence assembly using trace signals and additional sequence information.** In:

Computer Science and Biology: Proceedings of the German Conference on Bioinformatics (GCB), Vol. 99 (1999), pp. 45-56.

- Chevreux, B., Pfisterer, T., Drescher, B., Driesel, A. J., Müller, W. E. G., Wetter, T., Suhai, S. (2004). **Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs.** *Genome Research* 14(6): 1147-1159.
- Clem, R. J., Passarelli, A. L. (2013). **Baculoviruses: sophisticated pathogens of insects.** *PLoS Pathogens* 9(11).
- Coleman, M., Freeman, S., Herron, J. (1998). **Evolutionary Analysis.** *Animal Behaviour* 55(6): 1736-1737.
- Cook, S. P., Webb, R. E., Podgwaite, J. D., Reardon, R. C. (2003). **Increased mortality of gypsy moth *Lymantria dispar* (L.)(Lepidoptera: Lymantriidae) exposed to gypsy moth nuclear polyhedrosis virus in combination with the phenolic glycoside salicin.** *Journal of Economic Entomology* 96(6): 1662-1667.
- Copping, L. G., Menn, J. J. (2000). **Biopesticides: a review of their action, applications and efficacy.** *Pest Management Science* 56(8): 651-676.
- Cory, J. S., Green, B. M., Paul, R. K., Hunter-Fujita, F. (2005). **Genotypic and phenotypic diversity of a baculovirus population within an individual insect host.** *Journal of Invertebrate Pathology* 89(2): 101-111.
- Craveiro, S. R. (2012). **Varição genética e filogenia de *Pseudoplusia includens* single nucleopolyhedrovirus.** Dissertação de Mestrado. Departamento de Biologia Celular, UnB, Brasília-DF. pp.141.
- Craveiro, S. R., Melo, F. L., Ribeiro, Z. M. A., Ribeiro, B. M., Bão, S. N., Inglis, P. W., Castro, M. E. B. (2013). ***Pseudoplusia includens* single nucleopolyhedrovirus: genetic diversity, phylogeny and hypervariability of the *pif-2* gene.** *Journal of Invertebrate Pathology* 114(3): 258-267.
- Croizier, G., Croizier, L., Argaud, O., Poudevigne, D. (1994). **Extension of *Autographa californica* nuclear polyhedrosis virus host range by interspecific replacement of a short DNA sequence in the *p143* helicase gene.** *Proceedings of the National Academy of Sciences* 91(1): 48-52.
- De Jong, J. G., Lauzon, H. A., Dominy, C., Poloumienko, A., Carstens, E. B., Arif, B. M., Krell, P. J. (2005). **Analysis of the *Choristoneura fumiferana* nucleopolyhedrovirus genome.** *Journal of General Virology* 86(4): 929-943.
- Delport, W., Poon, A. F. Y., Frost, S. D. W., Pond, S. L. K. (2010). **Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology.** *Bioinformatics* 26(19): 2455-2457.

- Drummond, A., Ashton, B., Buxton, S., Cheung, M., Cooper, A., Duran, C., Field, M., Heled, J., Kearse, M., Markowitz, S. (2014). Geneious Pro v6. 1.7.
- Edgar, R. C. (2004). **MUSCLE: multiple sequence alignment with high accuracy and high throughput.** *Nucleic Acids Research* 32(5): 1792-1797.
- Eichlin, T. D., Cunningham, H. B. (1978). **The Plusiinae (Lepidoptera: Noctuidae) of America north of Mexico, emphasizing genitalic and larval morphology,** Department of Agriculture, Agricultural Research Service.
- Embrapa (2011). **Tecnologias de produção de soja—região central do Brasil 2012 e 2013.** *Embrapa Soja, Londrina.*
- Engelhard, E., Kam-Morgan, L., Washburn, J., Volkman, L. (1994). **The insect tracheal system: a conduit for the systemic spread of *Autographa californica* nuclear polyhedrosis virus.** *Proceedings of the National Academy of Sciences* 91(8): 3224-3227.
- Erlanson, M. (2008). **Insect pest control by viruses.** In: *Encyclopedia of Virology.* Edited by B. W. J. Mahy and M. H. V. van Regenmortel. Oxford, Elsevier. 3: 125-133.
- Fang, M., Nie, Y., Harris, S., Erlanson, M. A., Theilmann, D. A. (2009). ***Autographa californica* multiple nucleopolyhedrovirus core gene ac96 encodes a per os infectivity factor (*pif-4*).** *Journal of Virology* 83(23): 12569-12578.
- Fang, M., Nie, Y., Wang, Q., Deng, F., Wang, R., Wang, H., Wang, H., Vlak, J. M., Chen, X., Hu, Z. (2006). **Open reading frame 132 of *Helioverpa armigera* nucleopolyhedrovirus encodes a functional per os infectivity factor (PIF-2).** *Journal of General Virology* 87(9): 2563-2569.
- Faulkner, P., Kuzio, J., Williams, G. V., Wilson, J. A. (1997). **Analysis of p74, a PDV envelope protein of *Autographa californica* nucleopolyhedrovirus required for occlusion body infectivity in vivo.** *Journal of General Virology* 78(12): 3091-3100.
- Fauquet, C. M., Mayo, M. A., Maniloff, J., Desselberger, U., Ball, L. A. (2005). **Virus taxonomy: VIIIth report of the International Committee on Taxonomy of Viruses,** Academic Press.
- Federici, B. A. (1997). **Baculovirus pathogenesis.** In: *The baculoviruses.* Springer: 33-59.
- Felberbaum, R. S. (2015). **The baculovirus expression vector system: a commercial manufacturing platform for viral vaccines and gene therapy vectors.** *Biotechnology Journal* 10(5): 702-714.
- Ferrelli, M. L., Berretta, M. F., Belaich, M. N., Ghiringhelli, P. D., Sciocco-Cap, A., Romanowski, V. (2012). **The Baculoviral Genome.** In: *Viral Genomes -*

*Molecular Structure, Diversity, Gene Expression Mechanisms and Host-Virus Interactions*. Edited by M. Garcia, InTech.

- Flipsen, J., Martens, J., Van Oers, M., Vlak, J., Van Lent, J. (1995). **Passage of *Autographa californica* nuclear polyhedrosis virus through the midgut epithelium of *Spodoptera exigua* larvae.** *Virology* 208(1): 328-335.
- Folsom, J. W. (1936). **Notes on little-known insects.** *Journal of Economic Entomology* 29: 282-285.
- Friesen, P. D., Miller, L. K. (2001). **Fundamental Virology.** In: *Insect viruses*. Edited by D. M. Knipe, P. M. Howley, D. E. Griffin, M. A. Martin, R. A. Lamb, B. Roizman and S. E. Straus. Philadelphia, Lippincott Williams & Wilkins Press. 4: 443–472.
- Fritsch, E., Undorf-Spahn, K., Kienzle, J., Zebitz, C. P., Huber, J. (2007). **Codling moth granulovirus: First indication of variations in the susceptibility of local codling moth populations.** *IOBC WPRS BULLETIN* 30(1): 181.
- Funk, C. J., Braunagel, S.C., Rohrmann, G.F., (1997). **Baculovirus structure.** In: *The Baculovirus*. Edited by L. K. Miller. New York, Plenum Press: 7 - 32.
- Gilbert, C., Peccoud, J., Chateigner, A., Moumen, B., Cordaux, R., Herniou, E. A. (2016). **Continuous influx of genetic material from host to virus populations.** *PLoS Genetics* 12(2): e1005838.
- Goater, B., Ronkay, L., Fibiger, M. (2003). *Noctuidae Europaeae, Catocalinae and Plusiinae*, 10: 452pp, Entomological Press, Sorø.
- Granados, R. R., Williams, K. A. (1986). **In vivo infection and replication of baculoviruses.**
- Guarino, L. A., Gonzalez, M. A., Summers, M. D. (1986). **Complete sequence and enhancer function of the homologous DNA regions of *Autographa californica* nuclear polyhedrosis virus.** *Journal of Virology* 60(1): 224-229.
- Guarino, L. A., Summers, M. D. (1986). **Interspersed homologous DNA of *Autographa californica* nuclear polyhedrosis virus enhances delayed-early gene expression.** *Journal of Virology* 60(1): 215-223.
- Guarino, L. A., Xu, B., Jin, J., Dong, W. (1998). **A virus-encoded RNA polymerase purified from baculovirus-infected cells.** *Journal of Virology* 72(10): 7985-7991.
- Haas-Stapleton, E. J., Washburn, J. O., Volkman, L. E. (2004). **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* 78(13): 6786-6791.

- Hanada, K., Shiu, S.-H., Li, W. H. (2007). **The nonsynonymous/synonymous substitution rate ratio versus the radical/conservative replacement rate ratio in the evolution of mammalian genes.** *Molecular Biology and Evolution* 24(10): 2235-2241.
- Harrison, R. L., Puttler, B., Popham, H. J. R. (2008). **Genomic sequence analysis of a fast-killing isolate of *Spodoptera frugiperda* multiple nucleopolyhedrovirus.** *Journal of General Virology* 89(3): 775-790.
- Harrison, R. L., Sparks, W. O., Bonning, B. C. (2010). ***Autographa californica* multiple nucleopolyhedrovirus ODV-E56 envelope protein is required for oral infectivity and can be substituted functionally by *Rachiplusia ou* multiple nucleopolyhedrovirus ODV-E56.** *Journal of General Virology* 91(Pt 5): 1173-1182.
- Hasegawa M., Kishino H., T., Y. (1985). **Dating of the human-ape splitting by a molecular clock of mitochondrial DNA.** *Journal of Molecular Biology* 22(2): 160-174.
- Hayakawa, T., Ko, R., Okano, K., Seong, S.-I., Goto, C., Maeda, S. (1999). **Sequence analysis of the *Xestia c-nigrum* granulovirus genome.** *Virology* 262(2): 277-297.
- Hayakawa, T., Rohrmann, G. F., Hashimoto, Y. (2000). **Patterns of genome organization and content in lepidopteran baculoviruses.** *Virology* 278(1): 1-12.
- Hefferon, K. L., Oomens, A. G. P., Monsma, S. A., Finnerty, C. M., Blissard, G. W. (1999). **Host cell receptor binding by baculovirus GP64 and kinetics of virion entry.** *Virology* 258(2): 455-468.
- Hegedus, D., Erlandson, M., Gillott, C., Toprak, U. (2009). **New insights into peritrophic matrix synthesis, architecture, and function.** *Annual Review of Entomology* 54: 285-302.
- Hensley, S. D., Newson, L. D., Chapin, J. (1964). **Observations on the looper complex of the Noctuid subfamily Plusiinae.** *Journal of Economic Entomology* 57: 1006-1007.
- Herniou, E. A., Arif, B. M., Becnel, J. J., Blissard, G. W., Bonning, B., Harrison, R., Jehle, J. A., Theilmann, D. A., Vlask, J. M. (2012). **Baculoviridae.** In: *Virus Taxonomy*. Edited by A. M. Q. King, M. J. Adams, E. B. Carstens and E. J. Lefkowitz. Oxford, Elsevier: 163-174.
- Herniou, E. A., Jehle, J. A. (2007). **Baculovirus phylogeny and evolution.** *Current Drug Targets* 8(10): 1043-1050.



- 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.
- Herniou, E. A., Olszewski, J. A., O'reilly, D. R., Cory, J. S. (2004). **Ancient coevolution of baculoviruses and their insect hosts.** *Journal of Virology* 78(7): 3244-3251.
- Herzog, D. C. (1980). **Sampling soybean looper on soybean.** In: *Sampling Methods in Soybean Entomology*. Edited by M. Kogan and D. C. Herzog. New York, NY, Springer New York: 141-168.
- Herzog, D. C., Todd, J. H. (1980). **Sampling velvetbean carterpillar on soybean.** In: *Sampling Methods in Soybean Entomology*. Edited by M. Kogan and D. C. Herzog. New York, Springer-Verlag: 107-140.
- Hilton, S., Winstanley, D. (2007). **Identification and functional analysis of the origins of DNA replication in the *Cydia pomonella* granulovirus genome.** *Journal of General Virology* 88(5): 1496-1504.
- 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.
- Horton, H. M., Burand, J. (1993). **Saturable attachment sites for polyhedron-derived baculovirus on insect cells and evidence for entry via direct membrane fusion.** *Journal of Virology* 67(4): 1860-1868.
- Huang, Y., Zhang, L. (2004). **Rapid and sensitive dot-matrix methods for genome analysis.** *Bioinformatics* 20(4): 460-466.
- Huelsenbeck, J. P., Ronquist, F. (2001). **MRBAYES: Bayesian inference of phylogenetic trees.** *Bioinformatics* 17(8): 754-755.
- Ijkel, W. F., Westenberg, M., Goldbach, R. W., Blissard, G. W., Vlak, J. M., Zuidema, D. (2000). **A novel baculovirus envelope fusion protein with a proprotein convertase cleavage site.** *Virology* 275(1): 30-41.
- Iyer, L. M., Koonin, E. V., Aravind, L. (2002). **Extensive domain shuffling in transcription regulators of DNA Viruses and implications for the origins of fungal APSES transcription factors.** *Genome Biology* 3: 1-11.
- Jakob, N. J., Müller, K., Bahr, U., Darai, G. (2001). **Analysis of the first complete DNA sequence of an invertebrate Iridovirus: Coding strategy of the genome of *Chilo iridescent virus*.** *Virology* 286(1): 182-196.

- Jehle, J. A., Blissard, G. W., Bonning, B. C., Cory, J. S., Herniou, E. A., Rohrmann, G. F., Theilmann, D. A., Thiem, S. M., Vlak, J. M. (2006). **On the classification and nomenclature of baculoviruses: A proposal for revision.** *Archives of Virology* 151(7): 1257-1266.
- Jost, D. J., Pitre, H. N. (2002). **Soybean looper (Lepidoptera: Noctuidae) oviposition on cotton and soybean of different growth stages: influence of olfactory stimuli.** *Journal of Economic Entomology* 95(2): 286-293.
- Kamita, S. G., Maeda, S. (1997). **Sequencing of the putative DNA helicase-encoding gene of the Bombyx mori nuclear polyhedrosis virus and fine-mapping of a region involved in host range expansion.** *Gene* 190(1): 173-179.
- Kamiya, K., Zhu, J., Murata, M., Laviña-Caoili, B. A., Ikeda, M., Kobayashi, M., Kawamura, S. (2004). **Cloning and comparative characterization of three distinct nucleopolyhedroviruses isolated from the common cutworm, *Spodoptera litura* (Lepidoptera: Noctuidae) in Japan.** *Biological Control* 31(1): 38-48.
- 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.
- Katsuma, S. (2015). **Baculovirus controls host caterpillars by manipulating host physiology and behavior.** *AGRI-Bioscience Monographs* 5(1): 1-27.
- Katsuma, S., Kobayashi, J., Koyano, Y., Matsuda-Imai, N., Kang, W., Shimada, T. (2012). **Baculovirus-encoded protein BV/ODV-E26 determines tissue tropism and virulence in lepidopteran insects.** *Journal of Virology* 86(5): 2545-2555.
- Khan, S., Sneddon, K., Fielding, B., Ward, V., Davison, S. (2003). **Functional characterization of the ecdysteroid UDP-glucosyl transferase gene of Helicoverpa armigera Single-enveloped nucleopolyhedrovirus isolated in South Africa.** *Virus Genes* 27(1): 17-27.
- Kikhno, I., Gutiérrez, S., Croizier, L., Croizier, G., Ferber, M. L. (2002). **Characterization of pif, a gene required for the per os infectivity of Spodoptera littoralis nucleopolyhedrovirus.** *Journal of General Virology* 83(12): 3013-3022.
- King, A. M., Adams, M. J., Lefkowitz, E. J. (2011). **Virus taxonomy: classification and nomenclature of viruses: Ninth Report of the International Committee on Taxonomy of Viruses,** Elsevier.
- Krogh, A., Larsson, B., Von Heijne, G., Sonnhammer, E. L. L. (2001). **Predicting transmembrane protein topology with a hidden markov model:**

- application to complete genomes.** *Journal of Molecular Biology* 305(3): 567-580.
- Kuzio, J., Jaques, R., Faulkner, P. (1989). **Identification of p74, a gene essential for virulence of baculovirus occlusion bodies.** *Virology* 173(2): 759-763.
- Kuzio, J., Pearson, M. N., Harwood, S. H., Funk, C. J., Evans, J. T., Slavicek, J. M., Rohrmann, G. F. (1999). **Sequence and analysis of the genome of a baculovirus pathogenic for *Lymantria dispar*.** *Virology* 253(1): 17-34.
- Kwang, T. W., Zeng, X., Wang, S. (2016). **Manufacturing of AcMNPV baculovirus vectors to enable gene therapy trials.** *Molecular therapy. Methods & Clinical Development* 3: 15050.
- Lacey, L., Grzywacz, D., Shapiro-Ilan, D., Frutos, R., Brownbridge, M., Goettel, M. (2015). **Insect pathogens as biological control agents: back to the future.** *Journal of Invertebrate Pathology* 132: 1-41.
- Lacey, L. A., Unruh, T. R. (2005). **Biological control of codling moth (*Cydia pomonella*, Lepidoptera: Tortricidae) and its role in integrated pest management, with emphasis on entomopathogens.** *Vedalia* 12(1): 33-60.
- Lange, M., Wang, H., Zhihong, H., Jehle, J. A. (2004). **Towards a molecular identification and classification system of lepidopteran-specific baculoviruses.** *Virology* 325(1): 36-47.
- Le, S. Q., Gascuel, O. (2008). **An improved general amino acid replacement matrix.** *Molecular Biology and Evolution* 25(7): 1307-1320.
- Lee, H. H., Miller, L. K. (1978). **Isolation of genotypic variants of *Autographa californica* nuclear polyhedrosis virus.** *Journal of Virology* 27(3): 754-767.
- Li, L., Harwood, S. H., Rohrmann, G. F. (1999). **Identification of additional genes that influence baculovirus late gene expression.** *Virology* 255(1): 9-19.
- Li, Q., Donly, C., Li, L., Willis, L. G., Theilmann, D. A., Erlandson, M. (2002). **Sequence and organization of the *Mamestra configurata* nucleopolyhedrovirus genome.** *Virology* 294(1): 106-121.
- Liu, A., Qin, J., Rankin, C., Weaver, R. R. (1986). **Nucleotide sequence of a portion of the *Autographa californica* nuclear polyhedrosis virus genome containing the EcoRI site-rich region (hr5) and an open reading frame just 5' of the p10 gene.** *Journal of General Virology* 67: 2565-2570.
- Liu, H., Gao, H., Guo, G., Li, Y., Li, Y., Wang, J., Zhang, Z., Yu, Z. (2015). **The homologous regions of *Spodoptera litura* (Lepidoptera: Noctuidae)**

**nucleopolyhedrovirus II have both the function as origin of DNA replication and enhancer.** *Journal of Insect Science* 15(1): 89.

- Livingston, J. T., Yearian, W. (1972). **A nuclear polyhedrosis virus of *Pseudoplusia includens* (Lepidoptera: Noctuidae).** *Journal of Invertebrate Pathology* 19(1): 107-112.
- Lu, A., Miller, L. K. (1995). **The roles of eighteen baculovirus late expression factor genes in transcription and DNA replication.** *Journal of Virology* 69(2): 975-982.
- Machado, E. B. (2006). **Controle de *Condylorrhiza vestigialis* (Guenée, 1854)(Lepidoptera: Crambidae), a mariposa do álamo, com o uso de *C. vestigialis* multiplenucleopolyhedrovirus em condições de laboratório e campo.** Departamento de Ciências Agrárias, UFPR, Paraná. pp.124
- Maeda, S., Kamita, S. G., Kondo, A. (1993). **Host range expansion of *Autographa californica* nuclear polyhedrosis virus (NPV) following recombination of a 0.6-kilobase-pair DNA fragment originating from *Bombyx mori* NPV.** *Journal of Virology* 67(10): 6234-6238.
- Makkonen, K.-E., Airene, K., Ylä-Herttulala, S. (2015). **Baculovirus-mediated gene delivery and RNAi applications.** *Viruses* 7(4): 2099-2125.
- Marsaro Júnior, A. L., Pereira, P. R. V. S., Silva, W. R., Griffel, S. C. P. (2010). **Flutuação populacional de insetos-praga na cultura da soja no Estado de Roraima.** *Revista Acadêmica de Ciências Agrárias e Ambientais* 8(1): 71-76.
- Martin, P., Lingren, P., Greene, G. (1976). **Relative abundance and host preferences of cabbage looper, soybean looper, tobacco budworm, and corn earworm on crops grown in northern Florida.** *Environmental Entomology* 5(5): 878-882.
- Maruniak, J. E. (1986). **Baculovirus structural proteins and protein synthesis.**
- Mcdougal, V. V., Guarino, L. A. (2000). **The *Autographa californica* nuclear polyhedrosis virus p143 gene encodes a DNA helicase.** *Journal of Virology* 74(11): 5273-5279.
- Mcdougal, V. V., Guarino, L. A. (2001). **DNA and ATP Binding activities of the baculovirus DNA helicase P143.** *Journal of Virology* 75(15): 7206-7209.
- Mcleod, P. J., Young, S. Y., Yearian, W. C. (1982). **Application of a baculovirus of *Pseudoplusia includens* to soybean: Efficacy and seasonal persistence.** *Environmental Entomology* 11(2): 412-416.
- Miele, S. A., Garavaglia, M. J., Belaich, M. N., Ghiringhelli, P. D. (2011). **Baculovirus: molecular insights on their diversity and conservation.** *International Journal of Evolutionary Biology* 2011: 379-424.

- Mikhailov, V. S., Okano, K., Rohrmann, G. F. (2003). **Baculovirus alkaline nuclease possesses a 5'→3' exonuclease activity and associates with the DNA-binding protein LEF-3.** *Journal of Virology* 77(4): 2436-2444.
- Mitsuhashi, W., Miyamoto, K., Wada, S. (2014). **The complete genome sequence of the Alphaentomopoxvirus *Anomala cuprea entomopoxvirus*, including its terminal hairpin loop sequences, suggests a potentially unique mode of apoptosis inhibition and mode of DNA replication.** *Virology* 452–453(0): 95-116.
- Miwa, M., Tanaka, M., Matsushima, T., Sugimura, T. (1974). **Purification and properties of glycohydrolase from calf thymus splitting ribose-ribose linkages of poly(adenosine diphosphate ribose).** *Journal of Biological Chemistry* 11: 3475-3482.
- Monsma, S. A., Oomens, A. G., Blissard, G. W. (1996). **The GP64 envelope fusion protein is an essential baculovirus protein required for cell-to-cell transmission of infection.** *Journal of Virology* 70(7): 4607-4616.
- Moscardi, F. (1999). **Assessment of the application of baculoviruses for control of Lepidoptera.** *Annual Review of Entomology* 44(1): 257-289.
- Moscardi, F. (2007). **A nucleopolyhedrovirus for control of the velvetbean caterpillar in brazilian soybeans.** *Biological Control: A Global Perspective*: 344-352.
- Moscardi, F., Bueno, A., Sosa-Gómez, D. R., Roggia, S., Hoffmann-Campo, C. B., Pomari, A. F., Corso, I. C., Yano, S. a. C. (2012). **Artrópodes que atacam as folhas da soja.** *Soja. Manejo integrado de insetos e outros artrópodes-praga. Brasília, DF: Embrapa*: 214-334.
- Moscardi, F., Souza, M. L., Castro, M. E. B., Lara Moscardi, M., Szewczyk, B. (2011). **Baculovirus pesticides: Present state and future perspectives.** In: *Microbes and Microbial Technology*. Edited by I. Ahmad, F. Ahmad and J. Pichtel, Springer New York: 415-445.
- Moser, B. A., Becnel, J. J., White, S. E., Afonso, C., Kutish, G., Shanker, S., Almira, E. (2001). **Morphological and molecular evidence that *Culex nigripalpus* baculovirus is an unusual member of the family *Baculoviridae*.** *Journal of General Virology* 82(2): 283-297.
- Mu, J., Van Lent, J. W. M., Smagghe, G., Wang, Y., Chen, X., Vlak, J. M., Van Oers, M. M. (2014). **Live imaging of baculovirus infection of midgut epithelium cells: a functional assay of per os infectivity factors.** *Journal of General Virology* 95(Pt 11): 2531-2539.
- Muñoz, D., Murillo, R., Krell, P. J., Vlak, J. M., Caballero, P. (1999). **Four genotypic variants of a *Spodoptera exigua* nucleopolyhedrovirus (Se-SP2) are distinguishable by a hypervariable genomic region.** *Virus Research* 59(1): 61-74.

- Murrell, B., Wertheim, J. O., Moola, S., Weighill, T., Scheffler, K., Pond, S. L. K. (2012). **Detecting individual sites subject to episodic diversifying selection.** *PLoS Genetics* 8(7): e1002764.
- Myers E.W., Sutton G.G., Delcher A.L., Dew I.M., Fasulo D.P., Flanigan M.J., Kravitz S.A., Mobarry C.M., Reinert K.H., Remington K.A. (2000). **A whole-genome assembly of *Drosophila*.** *Science* 287:2196-2204.
- Nguyen, Q., Nielsen, L., Reid, S. (2013). **Genome scale transcriptomics of baculovirus-insect interactions.** *Viruses* 5(11): 2721-2747.
- Nie, Y., Fang, M., Erlandson, M. A., Theilmann, D. A. (2012). **Analysis of the *Autographa californica* Multiple Nucleopolyhedrovirus overlapping gene pair *lef3* and *ac68* reveals that AC68 is a per os infectivity factor and that LEF3 is critical, but not essential, for virus replication.** *Journal of Virology* 86(7): 3985-3994.
- Noune, C., Hauxwell, C. (2015). **Complete genome sequences of *Helicoverpa armigera* single nucleopolyhedrovirus strains AC53 and H25EA1 from Australia.** *Genome Announcements* 3(5): e01083-01015.
- O'reilly, D. R., Miller, L. K., Luckow, V. A. (1994). **Baculovirus expression vectors: a laboratory manual,** Oxford University Press.
- Ogembo, J. G., Chaeychomsri, S., Kamiya, K., Ishikawa, H., Katou, Y., Ikeda, M., Kobayashi, M. (2007). **Cloning and comparative characterization of nucleopolyhedroviruses isolated from African bollworm, *Helicoverpa armigera* (Lepidoptera: Noctuidae) in different geographic regions.** *Journal of Insect Biotechnology and Sericulture* 76(1): 39-49.
- Ogembo, J., Caoili, B., Shikata, M., Chaeychomsri, S., Kobayashi, M., Ikeda, M., (2009). **Comparative genomic sequence analysis of novel *Helicoverpa armigera* nucleopolyhedrovirus (NPV) isolated from Kenya and three other previously sequenced *Helicoverpa* spp. NPVs.** *Virus Genes* 39: 261-272.
- Ohkawa, T., Washburn, J. O., Sitapara, R., Sid, E., Volkman, L. E. (2005). **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* 79(24): 15258-15264.
- Palma, J., Maebe, K., Guedes, J. V. C., Smagghe, G. (2015). **Molecular variability and genetic structure of *Chrysodeixis includens* (Lepidoptera: Noctuidae), an important soybean defoliator in Brazil.** *PLoS One* 10(3): e0121260.
- Pang, Y., Yu, J., Wang, L., Hu, X., Bao, W., Li, G., Chen, C., Han, H., Hu, S., Yang, H. (2001). **Sequence analysis of the *Spodoptera litura* multicapsid nucleopolyhedrovirus genome.** *Virology* 287(2): 391-404.

- Passarelli, A. L., Todd, J. W., Miller, L. K. (1994). **A baculovirus gene involved in late gene expression predicts a large polypeptide with a conserved motif of RNA polymerases.** *Journal of Virology* 68(7): 4673-4678.
- Pearson, M., Bjornson, R., Pearson, G., Rohrmann, G. (1992). **The *Autographa californica* baculovirus genome: evidence for multiple replication origins.** *Science* 257(5075): 1382-1384.
- Pearson, M. N., Groten, C., Rohrmann, G. F. (2000). **Identification of the *Lymantria dispar* nucleopolyhedrovirus envelope fusion protein provides evidence for a phylogenetic division of the *Baculoviridae*.** *Journal of Virology* 74(13): 6126-6131.
- Pearson, M. N., Rohrmann, G. F. (1995). ***Lymantria dispar* nuclear polyhedrosis virus homologous regions: characterization of their ability to function as replication origins.** *Journal of Virology* 69(1): 213-221.
- Pedigo, L. P., Rice, M. E. (2014). **Entomology and pest management,** Waveland Press.
- Peng, K., Van Oers, M. M., Hu, Z., Van Lent, J. W. M., Vlak, J. M. (2010). **Baculovirus per os infectivity factors form a complex on the surface of occlusion-derived virus.** *Journal of Virology* 84(18): 9497-9504.
- Perera, O. P., Valles, S. M., Green, T. B., White, S., Strong, C. A., Becnel, J. J. (2006). **Molecular analysis of an occlusion body protein from *Culex nigripalpus* nucleopolyhedrovirus (CuniNPV).** *Journal of Invertebrate Pathology* 91(1): 35-42.
- Petersen, T. N., Brunak, S., Von Heijne, G., Nielsen, H. (2011). **SignalP 4.0: discriminating signal peptides from transmembrane regions.** *Nature Methods* 8(10): 785-786.
- Pijlman, G. P., Pruijssers, A. J., Vlak, J. M. (2003). **Identification of *pif-2*, a third conserved baculovirus gene required for per os infection of insects.** *Journal of General Virology* 84(Pt 8): 2041-2049.
- Pond, S. L., Frost, S. D. W., Grossman, Z., Gravenor, M. B., Richman, D. D., Brown, A. J. L. (2006). **Adaptation to different human populations by HIV-1 revealed by codon-based analyses.** *PLoS Computational Biology* 2(6): e62.
- Pond, S. L. K., Frost, S. D. W. (2005a). **Datamonkey: rapid detection of selective pressure on individual sites of codon alignments.** *Bioinformatics* 21(10): 2531-2533.
- Pond, S. L. K., Frost, S. D. W. (2005b). **Not so different after all: A comparison of methods for detecting amino acid sites under selection.** *Molecular Biology and Evolution* 22(5): 1208-1222.

- Pond, S. L. K., Sergei L. Kosakovsky, Frost, S. D. W., Muse, S. V. (2005). **HyPhy: hypothesis testing using phylogenies**. *Bioinformatics* 21(5): 676-679.
- Posada, D. (2008). **jModelTest: Phylogenetic model averaging**. *Molecular Biology and Evolution* 25(7): 1253-1256.
- Potter, J., Faulkner, P., Mackinnon, E. (1976). **Strain selection during serial passage of Trichoplusia ni nuclear polyhedrosis virus**. *Journal of Virology* 18(3): 1040-1050.
- Raykin, C., Ladin, B. F., Weaver, R. F. (1986). **Physical mapping of temporally regulated overlapping transcripts in the region of the 10K protein gene in Autographa californica nuclear polyhedrosis virus**. *Journal of Virology* 57: 18-27.
- Rodems, S. M., Friesen, P. D. (1993). **The hr5 transcriptional enhancer stimulates early expression from the Autographa californica nuclear polyhedrosis virus genome but is not required for virus replication**. *Journal of Virology* 67(10): 5776-5785.
- Rohrmann, G. F. (2013). **Baculovirus Molecular Biology**. Bethesda, National Center for Biotechnology Information
- Rohrmann, G. (1986). **Polyhedrin structure**. *Journal of General Virology* 67(8): 1499-1513.
- Rowley, D. L., Farrar Jr., R. R., Blackburn, M. B., Harrison, R. L. (2010). **Genetic and biological variation among nucleopolyhedrovirus isolates from the fall armyworms, Spodoptera frugiperda (Lepdoptera;Noctuidae)**. *Virus Genes* 40: 458-468.
- Rowley, D. L., Popham, H. J. R., Harrison, R. L. (2011). **Genetic variation and virulence of nucleopolyhedroviruses isolated worldwide from the heliothine pests Helicoverpa armigera, Helicoverpa zea, and Heliiothis virescens**. *Journal of Invertebrate Pathology* 107(2): 112-126.
- Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M. A., Barrell, B. (2000). **Artemis: sequence visualization and annotation**. *Bioinformatics* 16(10): 944-945.
- Salmela, L., Schröder, J. (2011). **Correcting errors in short reads by multiple alignments**. *Bioinformatics* 27(11): 1455-1461.
- Schmieder, R., Edwards, R. (2011). **Quality control and preprocessing of metagenomic datasets**. *Bioinformatics*.
- Serrano, A., Pijlman, G. P., Vlak, J. M., Muñoz, D., Williams, T., Caballero, P. (2015). **Identification of Spodoptera exigua nucleopolyhedrovirus genes involved in pathogenicity and virulence**. *Journal of Invertebrate Pathology* 126: 43-50.



- Shackelton, L. A., Holmes, E. C. (2004). **The evolution of large DNA viruses: combining genomic information of viruses and their hosts.** *Trends in microbiology* 12(10): 458-465.
- Silvie, P., Belot, J., Michel, B., Takizawa, E., Busarello, G., Thomazoni, D. (2007). **Manual de identificação das pragas e seus danos no cultivo do algodão.** Cascavel, COODETEC, *Boletim Técnico* 34.
- Simon, O., Palma, L., Beperet, I., Munoz, D., Lopez-Ferber, M., Caballero, P., Williams, T. (2011). **Sequence comparison between three geographically distinct Spodoptera frugiperda multiple nucleopolyhedrovirus isolates: Detecting positively selected genes.** *Journal of Invertebrate Pathology* 107(1): 33-42.
- Simón, O., Williams, T., Caballero, P., Possee, R. D. (2008). **Effects of Acp26 on in vitro and in vivo productivity, pathogenesis and virulence of Autographa californica multiple nucleopolyhedrovirus.** *Virus Research* 136(1–2): 202-205.
- Slack, J., Arif, B. M. (2006). **The Baculoviruses occlusion derived virus: virion structure and function.** In: *Advances in Virus Research*. Edited by M. Karl and J. S. Aaron, Academic Press. Volume 69: 99-165.
- Slack, J., Arif, B. M. (2007). **The baculoviruses occlusion-derived virus: virion structure and function.** *Advances in virus research* 69: 100.
- Slack, J., Arif, B. M., Karl, M., Aaron, J. S. (2007). **The baculoviruses occlusion-derived virus: Virion structure and function.** In: *Advances in Virus Research*. Academic Press. Volume 69: 99-165.
- Slack, J. M., Lawrence, S. D., Krell, P. J., Arif, B. M. (2010). **A soluble form of P74 can act as a per os infectivity factor to the Autographa californica multiple nucleopolyhedrovirus.** *Journal of General Virology* 91(4): 915-918.
- Song, J., Wang, R., Deng, F., Wang, H., Hu, Z. (2008). **Functional studies of per os infectivity factors of Helicoverpa armigera single nucleocapsid nucleopolyhedrovirus.** *Journal of General Virology* 89(9): 2331-2338.
- Sonnhammer, E. L., Von Heijne, G., Krogh, A. (1998). A hidden Markov model for predicting transmembrane helices in protein sequences. *Ismb*.
- Sosa-Gómez, D., Moscardi, F., Santos, B., Alves, L., Alves, S. (2008). **Produção e uso de vírus para o controle de pragas na América Latina.** *Controle Microbiano de Pragas na America Latina. FEALQ. Piracicaba, Brazil.*
- Sosa-Gómez, D. R., Bassoi, M. C. (2010). **Manual de identificação de insetos e outros invertebrados da cultura da soja,** Embrapa Soja.

- Souza, M. L., Castro, M. E. B., Moscardi, F. (2007). **Virus for biological control of insects.** In: *Encyclopedia of Pest Management*. Edited by D. Pimentel. Boca Raton, CRC Press. 2: 689–692.
- Sparks, W. O., Harrison, R. L., Bonning, B. C. (2011). **Autographa californica multiple nucleopolyhedrovirus ODV-E56 is a per os infectivity factor, but is not essential for binding and fusion of occlusion-derived virus to the host midgut.** *Virology* 409(1): 69-76.
- Specht, A., Paula-Moraes, S. V. D., Sosa-Gómez, D. R. (2015). **Host plants of *Chrysodeixis includens* (Walker) (Lepidoptera, Noctuidae, Plusiinae).** *Revista Brasileira de Entomologia* 59: 343-345.
- Srinivasa, M., Jagadeesh Babu, C., Anitha, C., Girish, G. (2008). **Laboratory evaluation of available commercial formulations of HaNPV against *Helicoverpa armigera* (Hub.).** *Journal of Biopesticides* 1(2): 138-139.
- Stiles, B., Himmerich, S. (1998). **Autographa californica NPV isolates: Restriction endonuclease analysis and comparative biological activity.** *Journal of Invertebrate Pathology* 72(2): 174-177.
- 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., Lobo De Souza, M. (2006). **Baculoviruses re-emerging biopesticides.** *Biotechnology Advances* 24(2): 143-160.
- Szewczyk, B., Rabalski, L., Krol, E., Sihler, W., Lobo De Souza, M. (2009). **Baculovirus biopesticides - a safe alternative to chemical protection of plants.** *Journal of Biopesticides* 2: 209-216.
- Szewczyk, B., Souza, M. L., Castro, M. E. B., Moscardi, M. L., Moscardi, F. (2011). **Baculovirus biopesticides.** In: *Pesticides - Formulations, Effects, Fate*. Edited by M. Stoytcheva, InTech: 25-36.
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S. (2011). **MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods.** *Molecular Biology and Evolution*.
- Theilmann, D. A., Blissard, G.W., Bonning, B., Jehle, J., O'Reilly, D.R., Rohrmann, G.F., Thiem, S., Vlak, J.M (2005). **Baculoviridae.** In: *Virus taxonomy: classification and nomenclature of viruses. Eighth Report of The International Committee on Taxonomy of Viruses*. Edited by C. M. Fauquet, Mayo, M.A., Maniloff, J., Desselberger, U., Ball, L.A. Amsterdam, Elsevier Academic Press.
- Thézé, J., Bézier, A., Periquet, G., Drezen, J.-M., Herniou, E. A. (2011). **Paleozoic origin of insect large dsDNA viruses.** *Proceedings of the National Academy of Sciences* 108(38): 15931-15935.

- Thiem, S., Cheng, X.-W. (2009). **Baculovirus host-range**. *Virologica Sinica* 24(5): 436-457.
- Todd, J. W., Passarelli, A. L., Miller, L. K. (1995). **Eighteen baculovirus genes, including *lef-11*, *p35*, *39K*, and *p47*, support late gene expression**. *Journal of Virology* 69(2): 968-974.
- Trichilo, P. J., Mack, T. (1989). **Soybean leaf consumption by the soybean looper (Lepidoptera: Noctuidae) as a function of temperature, instar and larval weight**. *Journal of Economic Entomology* 82(2): 633-638.
- Valicente, F., Cruz, I. (1991) "**Controle biológico da lagarta do cartucho, *Spodoptera frugiperda*, com Baculovirus Spodoptera**." 15, 1-23.
- Van Oers, M. M., Pijlman, G. P., Vlak, J. M. (2015). **Thirty years of baculovirus–insect cell protein expression: from dark horse to mainstream technology**. *Journal of General Virology* 96(1): 6-23.
- Van Oers, M. M., Vlak, J. M. (2007). **Baculovirus Genomics**. *Current Drug Targets* 8(10): 1051-1068.
- Vincent, C., Andermatt, M., Valéro, J. (2007). **Madex® and VirosoftCP4®, viral biopesticides for codling moth control**. *Biological control: a global perspective*. CABI, Wallingford, UK: 336-343.
- Volkman, L., Keddie, B. (1990). **Nuclear polyhedrosis virus pathogenesis**. *Seminary of Virology* 1: 249-256.
- Wang, L., Salem, T. Z., Campbell, D. J., Turney, C. M., Kumar, C. S., Cheng, X.-W. (2009). **Characterization of a virion occlusion-defective *Autographa californica* multiple nucleopolyhedrovirus mutant lacking the *p26*, *p10* and *p74* genes**. *Journal of General Virology* 90(7): 1641-1648.
- Wang, Q., Bosch, B. J., Vlak, J. M., Van Oers, M. M., Rottier, P. J., Van Lent, J. W. M. (2016). **Budded baculovirus particle structure revisited**. *Journal of Invertebrate Pathology* 134: 15-22.
- Westenberg, M., Uijtdewilligen, P., Vlak, J. M. (2007). **Baculovirus envelope fusion proteins F and GP64 exploit distinct receptors to gain entry into cultured insect cells**. *Journal of General Virology* 88(12): 3302-3306.
- Willis, L. G., Siepp, R., Stewart, T. M., Erlandson, M. A., Theilmann, D. A. (2005). **Sequence analysis of the complete genome of *Trichoplusia ni* single Nnucleopolyhedrovirus and the identification of a baculoviral photolyase gene**. *Virology* 338(2): 209-226.
- Wolcott, G. N. (1936). **Insectae Borinquenses**. *Journal Agriculture University* 20: 1-627.

- Xiang, X., Chen, L., Guo, A., Yu, S., Yang, R., Wu, X. (2011). **The *Bombyx mori* nucleopolyhedrovirus (BmNPV) ODV-E56 envelope protein is also a per os infectivity factor.** *Virus Research* 155(1): 69-75.
- Xu, F., Vlak, J. M., Van Oers, M. M. (2008). **Conservation of DNA photolyase genes in group II nucleopolyhedroviruses infecting Plusiinae insects.** *Virus Research* 136(1–2): 58-64.
- Yang, Z. (2001). **Adaptive Molecular Evolution.** In: *Handbook of Statistical Genetics*. Edited by D. Balding, M. Bishop and C. Cannings, John Wiley & Sons, Ltd: 327-350.
- Yang, Z. (2007). **PAML 4: Phylogenetic Analysis by Maximum Likelihood.** *Molecular Biology and Evolution* 24(8): 1586-1591.
- Yang, Z., Bielawski, J. P. (2000). **Statistical methods for detecting molecular adaptation.** *Trends in Ecology & Evolution* 15(12): 496-503.
- Yao, L., Zhou, W., Xu, H., Zheng, Y., Qi, Y. (2004). **The *Heliothis armigera* single nucleocapsid nucleopolyhedrovirus envelope protein P74 is required for infection of the host midgut.** *Virus Research* 104(2): 111-121.
- Young, S. (2001). **Enhancement of nucleopolyhedrovirus activity in *Helicoverpa zea* (Boddie) and *Pseudoplusia includens* (Walker) larvae with a fluorescent brightener.** *Journal of Entomological Science* 36(2): 162-168.
- Young, S., Yearian, W. (1979). **Soil application of *Pseudoplusia* NPV: persistence and incidence of infection in soybean looper caged on soybean.** *Environmental Entomology* 8(5): 860-864.
- Young, S., Yearian, W. (1982). **Nuclear polyhedrosis virus infection of *Pseudoplusia includens* [Lep.: Noctuidae] larvae: Effect on post larval stages and transmission.** *Entomophaga* 27(1): 61-66.
- Young, S., Yearian, W. (1988). **Secondary transmission of nuclear polyhedrosis virus by *Pseudoplusia includens* and *Anticarsia gemmatilis* larvae on semisynthetic diet.** *Journal of Invertebrate Pathology* 51(2): 133-138.
- Yu, M., Carstens, E. B. (2010). **Identification of a domain of the baculovirus *Autographa californica* multiple nucleopolyhedrovirus single-strand DNA-binding protein LEF-3 essential for viral DNA replication.** *Journal of Virology* 84(12): 6153-6162.
- Zemskov, E. A., Kang, W. K., 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.

- Zhang, C.-X., Ma, X.-C., Guo, Z.-J. (2005). **Comparison of the complete genome sequence between C1 and G4 isolates of the *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus.** *Virology* 333(1): 190-199.
- Zhou, J.-B., Li, X.-Q., De-Eknamkul, W., Suraporn, S., Xu, J.-P. (2012a). **Identification of a new *Bombyx mori* nucleopolyhedrovirus and analysis of its *bro* gene family.** *Virus Genes* 44(3): 539-547.
- Zhou, J. B., Li, X. Q., De-Eknamkul, W., Suraporn, S., Xu, J. P. (2012b). **Identification of a new *Bombyx mori* nucleopolyhedrovirus and analysis of its *bro* gene family.** *Virus Genes* 44(3): 539-547.
- Zhu, S., Wang, W., Wang, Y., Yuan, M., Yang, K. (2013). **The baculovirus core gene ac83 is required for nucleocapsid assembly and per os infectivity of *Autographa californica* nucleopolyhedrovirus.** *Journal of Virology* 87(19): 10573-10586.
- Zhu, Z., Yin, F., Liu, X., Hou, D., Wang, J., Zhang, L., Arif, B. M., Wang, H., Deng, F., Hu, Z. (2014). **Genome sequence and analysis of *Buzura suppressaria* nucleopolyhedrovirus: A Group II *Alphabaculovirus*.** *PLoS ONE* 9(1): e86450.
- Zou, Y., Young, S. Y. (1996). **Use of a fluorescent brightener to improve *Pseudoplusia includens* (Lepidoptera: Noctuidae) nuclear polyhedrosis virus activity in the laboratory and field.** *Journal of Economic Entomology* 89(1): 92-96.