



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

**Variação genética e filogenia de**  
***Pseudoplusia includens single nucleopolyhedrovirus***

**Saluana Rocha Craveiro**

Brasília-DF  
2012



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

# **Variação genética e filogenia de *Pseudoplusia includens single nucleopolyhedrovirus***

**Saluana Rocha Craveiro**

**Dissertação apresentada ao Programa de Pós-graduação em Biologia Molecular do Departamento de Biologia Celular do Instituto de Ciências Biológicas da Universidade de Brasília como requisito à obtenção do título de Mestre em Biologia Molecular**

**Orientadora:** Dra. Maria Elita Batista de Castro

**Co-orientador:** Dr. Peter Ward Inglis

Brasília-DF  
2012

Craveiro, Saluana Rocha  
Variação genética e filogenia de *Pseudoplusia includens single  
nucleopolyhedrovirus*

151p

Dissertação (Mestrado) – Universidade de Brasília. Instituto de  
Ciências Biológicas. Departamento de Biologia Celular.  
Pós-Graduação em Biologia Molecular. Brasília,  
2012.

Orientadora: Dra. Maria Elita Batista de Castro

Co-Orientador: Dr. Peter Ward Inglis

1.Filogenia 2.Variantes genotípicos 3.Baculovirus 4.PsinSNPV

Pesquisa realizada no Laboratório de Virologia de Insetos (LVI) da Embrapa Recursos Genéticos e Biotecnologia (CENARGEN), sob orientação da Dra. Maria Elita Batista de Castro e co-orientação do Dr. Peter Ward Inglis, com suporte financeiro da EMBRAPA e CNPq/UnB.

Brasília-DF, 17 de fevereiro de 2012

**Banca Examinadora:**

Dra. Maria Elita Batista de Castro (Orientadora) - Pesquisadora/CENARGEN

Dr. Bergmann Morais Ribeiro - Professor/UnB

Dra. Débora Pires Paula - Pesquisadora/CENARGEN

Suplente: Dra. Glaucia Salles Cortopassi Buso - Pesquisadora/CENARGEN

Aos meus queridos pais,  
Rosana e Saulo exemplos de  
garra e perseverança.

# Agradecimentos

A realização desta Dissertação só foi possível devido à ajuda de vários colegas e instituições aos quais presto meus sinceros agradecimentos.

À Embrapa Recursos Genéticos e Biotecnologia (CENARGEN) pelo suporte técnico e financeiro.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela concessão da bolsa de estudo.

Ao Decanato de Pós-Graduação da Universidade de Brasília (UnB) pelo suporte financeiro na participação de eventos científicos.

Ao Programa de Pós-Graduação em Biologia Molecular do Instituto de Ciências Biológicas da Universidade de Brasília pela oportunidade de realização do curso de mestrado.

---

A Deus por me dar forças e fazer de mim perseverante para poder realizar mais um de meus objetivos.

À minha orientadora Dra. Elita por ter me apoiado desde a graduação e acreditado no meu potencial, repassando conhecimento e me incentivando a buscá-lo cada dia mais. Obrigada por seu inspirador amor a Ciência e pela amizade, paciência e confiança dedicados a mim nos últimos quatro anos.

Ao Dr. Peter pela contínua colaboração e co-orientação. Por me acolher tão bem em seu laboratório e pelo inestimável treinamento técnico com os programas de análise e interpretações dos resultados em Filogenia Molecular. Muito obrigada pela paciência e dedicação!

À minha mãe querida Rosana, exemplo de mulher guerreira, que sempre me incentivou e deu força para buscar e conseguir realizar todos meus sonhos.

Ao meu pai José Saulo, exemplo de conduta, que me ensinou os caminhos da vida e confiou e apoiou as minhas decisões.

Ao meu irmão Allan Saul pelo amor e convívio amigo durante todos esses anos e, especialmente, por me presentear com uma sobrinha linda, Emanuella, fonte de alegria e inspiração.

À minha vovó Eutêmia e seu marido José Marques pelo amor e confiança depositados em mim.

Ao Fernando, meu namorado, pelo apoio na realização do trabalho, principalmente, durante a elaboração desta dissertação. Obrigada pelo amor e companheirismo de todos os dias. Sem dúvidas a vida ao seu lado é mais tranqüila e alegre.

A toda minha família, pais, irmãos, avós, tios e primos, pela preocupação demonstrada durante toda a realização do trabalho e pelo carinho, amor e dedicação proporcionados todos esses anos.

Aos membros da banca examinadora, Dr. Bergmann M. Ribeiro, Dra. Débora P. Paula e Dra. Glaucia S. C. Buso, por aceitarem aos convites de participarem da defesa e qualificação ajudando com perguntas e sugestões enriquecedoras ao projeto.

Aos técnicos do Laboratório de Virologia de Invertebrados (LVI), onde realizei maior parte dos experimentos, pelo valioso auxílio técnico prestado em diversas fases e aspectos deste trabalho. Zildinha, obrigada pelo apoio na realização dos experimentos e pelo cuidado e carinho de mãe sempre me ajudando muito. William, obrigada por sempre ajudar na realização dos experimentos e discussão dos resultados, sem contar dos maravilhosos momentos de conversas e risadas. Raimundinha, obrigada por prover todo o



suporte necessário para a realização dos experimentos e pelo carinho e alegria de todos os dias.

À Briana que me ensinou muito e foi uma parceira na realização do trabalho no laboratório e amiga fiel nos momentos de desespero, desabafos e descontração.

A Moema pela amizade e ajuda na realização dos experimentos, principalmente, com a amplificação e purificação dos produtos de PCR.

A todos os colegas de laboratório em especial Marina e Milene pela colaboração no dia-a-dia do trabalho.

À Dra. Marlinda pela sua atenção, acompanhamento e sugestões ao trabalho durante esses anos no dia-a-dia do laboratório, além da sua valiosa participação na banca examinadora de qualificação.

Ao Dr. Bergmann que muito ajudou com ensinamentos, colaboração e concessão do seu laboratório sempre quando necessário.

À Dra. Débora Paula pela amizade e generosidade sempre presentes. Pela ajuda com a discussão do trabalho e empréstimo dos equipamentos do Laboratório de Biologia Molecular – Embrapa/Cenargen. Obrigada pelos valiosos conselhos e é claro, sem esquecer, agradeço pelas divertidas baladas.

Ao Dr. Rogério e aos amigos do Laboratório de Micologia de Invertebrados – Embrapa/Cenargen, Daniela, Silas e Irene, pela paciência e boa disposição no empréstimo dos equipamentos.

À Dra. Rose Monnerat e aos colegas do Laboratório de Bactérias Entomopatogênicas - Embrapa/Cenargen, Érica e Lílian, pela colaboração prestada no decorrer do trabalho.

À Dra. Vânia Cristina por me receber tão bem no Laboratório de Genética Vegetal, principalmente, para a realização das reações de seqüenciamento e utilização do seqüenciador automático.

Ao Dr. Flávio Moscardi e Embrapa Soja pela doação de larvas infectadas com isolados virais de PsinSNPV utilizados neste estudo.

A todos, Dani, Silas, Cris, Diogo, Magno, Carla e Luciana, pela agradável convivência diária na Embrapa e pelos valiosos momentos de diversão.

À secretaria do Prédio de Controle Biológico da Embrapa/Cenargen, na pessoa da Rosângela, pela colaboração, impressões e cópias de artigos utilizadas no decorrer deste trabalho.

À secretaria da Pós-Graduação em Biologia Molecular da Universidade de Brasília, principalmente à Ana, pela disponibilidade em ajudar, principalmente a resolver as questões burocráticas.

As amigas Ana Flávia e Patrícia pelo apoio e confiança. Queridas, mesmo longe vocês tornaram esse momento mais alegre.

À Lauriana, Víctor, Laura, Vivian e Diogo pela torcida e vibração nas conquistas adquiridas durante o curso. Em especial agradeço a Vivian por me ajudar na criação dos desenhos esquemáticos (Figuras 3 e 4) presentes nesta Dissertação.

Ao casal Millane e Thiago e todo o grupo de amigos pelos vários momentos de diversão, principalmente, os memoráveis churrascos e forrós.

**Muito Obrigada!!!**

# Índice

Índice de Figuras .....	3
Índice de Tabelas .....	5
Abreviaturas e Siglas .....	6
Resumo .....	11
Abstract .....	12
1. Introdução .....	13
2. Revisão Bibliográfica .....	15
2.1. <i>Pseudoplusia includens single nucleopolyhedrovirus</i> (PsinSNPV): um vírus patogênico à lagarta falsa-medideira .....	15
2.2. Características gerais dos baculovirus .....	18
2.3. Aplicações biotecnológicas: baculovirus como agentes de controle biológico de insetos-praga .....	22
2.4. Genoma dos baculovirus .....	24
2.4.1. Grupo de genes conservados entre os baculovirus - <i>core genes</i> ...	24
2.4.2. Genes <i>lef-8</i> , <i>lef-9</i> , <i>pif-2</i> e <i>polh/gran</i> : uma ferramenta para estudos de identificação e filogenia dos baculovirus .....	26
2.4.3. Gene <i>DNA photolyase</i> ( <i>phr</i> ) .....	28
2.5. Evolução e taxonomia dos baculovirus .....	30
2.6. Filogenia molecular .....	34
2.7. Variação genética em populações de baculovirus .....	37
3. Objetivo Geral .....	39
3.1. Objetivos específicos .....	39
4. Materiais e Métodos .....	41
4.1. Vírus .....	41
4.2. Ferramentas de bioinformática .....	41
4.3. Purificação de corpos de oclusão (OBs) .....	42
4.4. Extração de DNA a partir de partículas OBs purificadas .....	43
4.5. Clivagem de DNA viral por endonucleases de restrição .....	44
4.6. Amplificação de DNA por reação em cadeia da polimerase - <i>Polymerase Chain Reaction</i> (PCR) .....	45
4.7. Purificação dos produtos de PCR .....	47
4.8. Sequenciamento dos produtos de PCR .....	48

4.9. Análise das sequências .....	49
4.10. Análise filogenética .....	50
4.10.1. Alinhamento das sequências utilizadas na análise filogenética ....	52
4.10.2. Análise da Máxima Parcimônia .....	52
4.10.3. Análise Bayesiana .....	53
4.10.4. Medidas de apoio estatístico: ILD - teste da incongruência da diferença de comprimento e PABA - teste da adição de partição alterando o <i>bootstrap</i> .....	55
4.11. Agrupamento dos isolados de PsinSNPV por UPGMA .....	57
5. Resultados .....	58
5.1. Identificação dos genes <i>lef-8</i> , <i>lef-9</i> , <i>phr</i> , <i>pif-2</i> e <i>polh</i> no genoma dos isolados de PsinSNPV (IA a IG) .....	58
5.2. Análise das sequências nucleotídicas parciais dos genes <i>lef-8</i> , <i>lef-9</i> , <i>phr</i> , <i>pif-2</i> e <i>polh</i> de PsinSNPV (IA a IG) .....	59
5.3. Relações filogenéticas entre isolados virais de PsinSNPV .....	70
5.4. Análise fenética dos RFLPs de isolados de PsinSNPV .....	73
5.5. Reconstrução das relações filogenéticas da família <i>Baculoviridae</i> incluindo o vírus PsinSNPV- IE .....	75
6. Discussão .....	88
7. Conclusões e Perspectivas .....	94
8. Referências Bibliográficas .....	97
Apêndices .....	121
Apêndice I .....	122
Apêndice II .....	127
Apêndice III .....	129
Glossário .....	137

## Índice de Figuras

<b>Figura 1.</b> Larvas <i>Pseudoplusia includens</i> sadia (A) e morta por infecção viral (B).	16
<b>Figura 2.</b> Micrografia eletrônica do corpo oclusão (OB) de PsinSNPV.	17
<b>Figura 3.</b> Desenho esquemático de corpos de oclusão dos baculovirus (Craveiro,2012).	19
<b>Figura 4.</b> Desenho esquemático de formas virais produzidas durante o ciclo de infecção dos baculovirus (Craveiro, 2012).	20
<b>Figura 5.</b> Desenho esquemático de fotoprodutos de DNA induzidos pela radiação UV.)	28
<b>Figura 6.</b> Taxonomia da família <i>Baculoviridae</i> .	32
<b>Figura 7.</b> Produtos de PCR dos genes <i>lef-8</i> , <i>lef-9</i> , <i>phr</i> , <i>pif-2</i> e <i>polh</i> de isolados de PsinSNPV (IA a IG).	58
<b>Figura 8.</b> Alinhamento das sequências nucleotídicas parciais do gene <i>lef-8</i> dos isolados de PsinSNPV (IA a IG).	60
<b>Figura 9.</b> Alinhamento das sequências nucleotídicas parciais do gene <i>lef-9</i> dos isolados de PsinSNPV (IA a IG).	61
<b>Figura 10.</b> Alinhamento das sequências nucleotídicas parciais do gene <i>polh</i> dos isolados de PsinSNPV (IA a IG).	63
<b>Figura 11.</b> Alinhamento das sequências nucleotídicas parciais do gene <i>phr</i> dos isolados de PsinSNPV (IA a IG).	65
<b>Figura 12.</b> Eletroferogramas do seqüenciamento do produto de PCR do gene <i>phr</i> .	66
<b>Figura 13.</b> Alinhamento das sequências nucleotídicas parciais do gene <i>pif-2</i> dos isolados de PsinSNPV (IA a IG).	68
<b>Figura 14.</b> Relações filogenéticas entre os isolados de PsinSNPV (IA a IG) baseadas nas sequências concatenadas dos genes <i>lef-8</i> , <i>lef-9</i> , <i>phr</i> , <i>pif-2</i> e <i>polh</i> .	72
<b>Figura 15.</b> Perfis de restrição de DNAs dos isolados de PsinSNPV (IA a IG).	73
<b>Figura 16.</b> Dendrograma de similaridade genotípica resultante da análise dos RFLPs de isolados de PsinSNPV (IA a IG) utilizando algoritmo UPGMA.	74
<b>Figura17.</b> Filogenia do vírus PsinSNPV com outros baculovirus inferida da concatenação das sequências nucleotídicas dos genes <i>lef-8</i> , <i>polh/gran</i> , <i>lef-9</i> e <i>pif-2</i> .	81
<b>Figura 18.</b> Filograma da concatenação das sequências nucleotídicas dos genes <i>lef-8</i> , <i>polh/gran</i> , <i>lef-9</i> e <i>pif-2</i> representando a relação filogenética de PsinSNPV-IE com outros membros da família <i>Baculoviridae</i> .	85
<b>Figura 19.</b> Reconstrução filogenética da família <i>Baculoviridae</i> incluindo o vírus PsinSNPV -IE.	87
<b>Figura 20.</b> Resultado do alinhamento BLAST da sequência deduzida de aminoácidos (A) LEF-8, (B) LEF-9, (C) PHR, (D) PIF-2 e (E) POLH de PsinSNPV-IE.	122

- Figura 21.** Árvores obtidas do teste PABA (Adição de partição alterando o *bootstrap*) da análise MP dos genes individuais (A, B, C, D) e da concatenação ordenada das sequências nucleotídicas dos genes *lef-8*, *polh/gran*, *lef-9* e *pif-2* (E, F e G)..... 127
- Figura 22.** Árvores filogenéticas representando as relações filogenéticas da família *Baculoviridae* obtidas da análise MP não ponderada baseada nas sequências nucleotídicas parciais dos genes individuais *lef-8*, *lef-9*, *pif-2* e *polh/gran*..... 129
- Figura 23.** Árvores filogenéticas representando as relações filogenéticas da família *Baculoviridae* obtidas da análise MP com reponderação progressiva baseada nas sequências nucleotídicas parciais dos genes individuais *lef-8*, *lef-9*, *pif-2* e *polh/gran*. .... 130
- Figura 24.** Árvores filogenéticas representando as relações filogenéticas da família *Baculoviridae* obtidas da análise MP não ponderada baseada nas sequências individuais LEF-8, LEF-9, PIF-2 e POLH/GRAN.. .... 131
- Figura 25.** Árvores filogenéticas representando as relações filogenéticas da família *Baculoviridae* obtidas da análise MP com reponderação progressiva baseada nas sequências peptídicas LEF-8, LEF-9, PIF-2 e POLH/GRAN. .... 132
- Figura 26.** Árvores filogenéticas representando as relações filogenéticas da família *Baculoviridae* obtidas da análise MP e IB baseada na concatenação das sequências nucleotídicas dos genes *lef-8*, *lef-9* e *pif-2*.. .... 133
- Figura 27.** Árvores filogenéticas representando as relações filogenéticas da família *Baculoviridae* obtidas da análise MP e IB baseada na concatenação das sequências peptídicas LEF-8, LEF-9 e PIF-2..... 134
- Figura 28.** Árvores filogenéticas representando as relações filogenéticas da família *Baculoviridae* obtidas da análise MP e IB baseada na concatenação das sequências nucleotídicas dos genes *lef-8*, *lef-9*, *pif-2* e *polh/gran*. .. 135
- Figura 29.** Árvores filogenéticas representando as relações filogenéticas da família *Baculoviridae* obtidas da análise MP e IB baseada na concatenação das sequências peptídicas LEF-8, POLH/GRAN, LEF-9 e PIF-2..... 136

## Índice de Tabelas

<b>Tabela 1.</b> Genes conservados entre os baculovirus ( <i>core genes</i> ) .....	25
<b>Tabela 2.</b> Genes conservados em baculovirus específicos de lepidópteros utilizados em estudos filogenéticos .....	27
<b>Tabela 3.</b> Classificação taxonômica da família <i>Baculoviridae</i> .....	33
<b>Tabela 4.</b> Isolados de PsinSNPV obtidos de larvas <i>P. includens</i> infectadas ..	41
<b>Tabela 5.</b> Oligonucleotídeos degenerados utilizados no PCR .....	47
<b>Tabela 6.</b> Baculovirus utilizados na construção das árvores filogenéticas.....	51
<b>Tabela 7.</b> Modelos evolutivos selecionados para as sequências nucleotídicas de cinco genes de isolados de PsinSNPV* .....	54
<b>Tabela 8.</b> Modelos evolutivos selecionados para as sequências nucleotídicas de quatro genes de baculovirus * .....	55
<b>Tabela 9.</b> Modelos evolutivos selecionados para quatro sequências peptídicas de baculovirus*.....	55
<b>Tabela 10.</b> Sequências de cinco genes de isolados de PsinSNPV depositadas no <i>GenBank</i> .....	59
<b>Tabela 11.</b> Posição nucleotídica (nt) dos polimorfismos detectados nas sequências parciais de genes de PsinSNPV .....	69
<b>Tabela 12.</b> Características das sequências concatenadas dos genes <i>lef-8</i> , <i>lef-9</i> , <i>phr</i> , <i>pif-2</i> e <i>polh</i> .....	70
<b>Tabela 13.</b> Estatísticas das árvores MP obtidas com a análise das sequências concatenadas dos genes <i>lef-8</i> , <i>lef-9</i> , <i>phr</i> , <i>pif-2</i> e <i>polh</i> .....	71
<b>Tabela 14.</b> Características das sequências nucleotídicas individual e concatenada dos genes <i>lef-8</i> , <i>polh/gran</i> , <i>lef-9</i> e <i>pif-2</i> .....	77
<b>Tabela 15.</b> Características das sequências deduzidas de aminoácidos, individual e concatenada, LEF-8, POLH/GRAN, LEF-9 e PIF-2.....	77
<b>Tabela 16.</b> Estatísticas das árvores obtidas da análise MP não ponderada das sequências nucleotídicas individual e concatenada dos genes <i>lef-8</i> , <i>polh/gran</i> , <i>lef-9</i> e <i>pif-2</i> .....	78
<b>Tabela 17.</b> Estatísticas das árvores obtidas da análise MP reponderada das sequências nucleotídicas individual e concatenada dos genes <i>lef-8</i> , <i>polh/gran</i> , <i>lef-9</i> e <i>pif-2</i> .....	78
<b>Tabela 18.</b> Estatísticas das árvores obtidas com a análise MP não ponderada das sequências deduzidas de aminoácidos, individual e concatenada, LEF-8, POLH/GRAN, LEF-9 e PIF-2.....	79
<b>Tabela 19.</b> Estatísticas das árvores obtidas com a análise MP reponderada das sequências deduzidas de aminoácidos, individual e concatenada, LEF-8, POLH/GRAN, LEF-9 e PIF-2.....	79
<b>Tabela 20.</b> Probabilidade de significância da incongruência das diferentes combinações das partições de dados - teste ILD.....	82
<b>Tabela 21.</b> Teste da adição de partição alterando o <i>bootstrap</i> (PABA) com base na árvore consenso estrito da análise de parcimônia não ponderada*.....	83

## Abreviaturas e Siglas

AcMNPV-C6 - *Autographa californica mutiple nucleopolyhedrovirus* -C6  
 AdhoNPV - *Adoxophyes honmai nucleopolyhedrovirus*  
 AdorGV - *Adoxophyes orana granulovirus*  
 AdorNPV - *Adoxophyes orana nucleopolyhedrovirus*  
 AgipMNPV - *Agrotis ipsilon nucleopolyhedrovirus* - Illions  
 AgMNPV- *Anticarsia gemmatalis multiple nucleopolyhedrovirus*  
 AgseGV - *Agrotis segetum granulovirus*  
 AgseNPV - *Agrotis segetum nucleopolyhedrovirus*  
 AnpeMN PV-L2 - *Antheraea pernyi multiple nucleopolyhedrovirus* -L2  
 AnpeMNPV- *Antheraea pernyi multiple nucleopolyhedrovirus*  
 ApciNPV - *Apocheima cinerarium nucleopolyhedrovirus*  
 BmNPV – T3 - *Bombyx mori nucleopolyhedrovirus* – T3  
 BomaNPV – S1 - *Bombyx mandarina nucleopolyhedrovirus* -S1  
 CfDEFMNPV-*Choristoneura fumiferana defective multiple nucleopolyhedrovirus*  
 CfMNPV - *Choristoneura fumiferana multiple nucleopolyhedrovirus*  
 ChchNPV - *Chrysodeixis chalcites nucleopolyhedrovirus*  
 ChocGV - *Choristoneura occidentalis granulovirus*  
 ClbiNPV-DZ1 - *Clanis bilineata nucleopolyhedrovirus* -DZ1  
 CpGV - *Cydia pomonella granulovirus*  
 CrleGV - *Cryptophlebia leucotreta granulovirus*-CV3  
 CuniNPV - *Culex nigripalpus nucleopolyhedrovirus*  
 EcobNPV-A1 - *Ecotropis obliqua nucleopolyhedrovirus* -A1  
 EpgrSNPV - *Epinotia granitalis single nucleopolyhedrovirus*  
 EppoNPV - *Epiphyas postvittana nucleopolyhedrovirus*  
 EupsNPV-Hangzhou - *Euproctis pseudoconsersa nucleopolyhedrovirus* -  
 Hangzhou  
 HearGV - *Helicoverpa armigera granulovirus*  
 HearMNPV - *Helicoverpa armigera multiple nucleopolyhedrovirus*  
 HearNPV-NNg1 - *Helicoverpa armigera nucleopolyhedrovirus* -NNg1  
 HearSNPV - *Helicoverpa armigera single nucleopolyhedrovirus*  
 HearSNPV-G4 - *Helicoverpa armigera single nucleopolyhedrovirus* -G4  
 HycuNPV - *Hyphantria cunea nucleopolyhedrovirus*



HzSNPV - *Helicoverpa zea* single nucleopolyhedrovirus  
 LdMNPV - *Lymantria dispar* multiple nucleopolyhedrovirus  
 LeseNPV-AH1 - *Leucania separata* nucleopolyhedrovirus -AH1  
 LyxyMNPV - *Lymantria xyli* multiple nucleopolyhedrovirus  
 MacoNPV-A90/2 - *Mamestra configurata* nucleopolyhedrovirus -A90/2  
 MacoNPV-A90/4 - *Mamestra configurata* nucleopolyhedrovirus -A90/4  
 MacoNPV-B - *Mamestra configurata* nucleopolyhedrovirus -B  
 MaviNPV - *Maruca vitrata* nucleopolyhedrovirus  
 NeabNPV - *Neodiprion abietis* nucleopolyhedrovirus  
 NeleNPV - *Neodiprion lecontei* nucleopolyhedrovirus  
 NeseNPV - *Neodiprion sertifer* nucleopolyhedrovirus  
 OpMNPV - *Orgyia pseudotsugata* multiple nucleopolyhedrovirus  
 OrleNPV-CFS77 - *Orgyia leucostigma* nucleopolyhedrovirus -CFS77  
 PhorGV - *Phtorimaea operculella* granulovirus  
 PlxyGV - *Plutella xylostella* granulovirus  
 PlxyNPV- CL3 - *Plutella xylostella* nucleopolyhedrovirus -CL3  
 PrGV - *Pieris rapae* granulovirus  
 PsinSNPV – *Pseudoplusia includens* single nucleopolyhedrovirus  
 PsunGV-Hawaiin - *Pseudaletia unipuncta* granulovirus- Hawaiin  
 RoNPV - *Rachiplusia ou* nucleopolyhedrovirus  
 SeMNPV - *Spodoptera exigua* multiple nucleopolyhedrovirus  
 SfMNPV-19 - *Spodoptera frugiperda* nucleopolyhedrovirus -19  
 SfMNPV-3AP2 - *Spodoptera frugiperda* nucleopolyhedrovirus -3AP2  
 SpltGV-K1 - *Spodoptera litura* granulovirus-K1  
 SpltMNPV-G2 - *Spodoptera litura* multiple nucleopolyhedrovirus -G2  
 SpltNPV-II - *Spodoptera litura* nucleopolyhedrovirus -II  
 ThorMNPV – *Thysanoplusia orichacea* multiple nucleopolyhedrovirus  
 TnSNPV - *Trichoplusia ni* single nucleopolyhedrovirus  
 XecnGV - *Xestia c-nigrum* granulovirus  
 %AT – porcentagem da composição nucleotídica composta por adeninas e timinas  
 8-HDF - 8-hidroxi 7,8-didemetil-5-deazaflavina  
 aa – aminoácido  
 AIC – Critério Informação de Akaike

BIC – Critério de Informação Bayesiana  
BLAST – *Basic Local Alignment Search Tool*  
*bro* – ORF repetitiva de baculovirus (*baculovirus repeated ORF*)  
BS - *bootstrap*  
BV – vírus extracelular (*budded virus*)  
CPD – dímero de pirimidina ciclobutano (*cyclobutane pyrimidine dimer*)  
*dnapol* – gene *DNA polimerase (DNA polymerase)*  
dNTP – 2'-desoxinucleotídeo 5'-fosfato  
dNTP - desorribonucleotídeo trifosfatado  
EDTA – ácido etilenodiamino tetra-acético  
EDTA - ácido etilenodiamino tetra-acético (*ethylenediamine tetraacetic acid*)  
*egt* – gene *ecdisteróide UDP-glicosil transferase (ecdysteroid UDP-glucosyltransferase)*  
Expasy - *Expert Protein Analysis System*  
FAD – flavina adenina dinucleotídeo (*flavine adenine dinucleotide*)  
*gp37* – gene *glicoproteína 37 (glycoprotein 37)*  
*gp41* – gene *glicoproteína 41 (glycoprotein 41)*  
*gta* – gene *transativação global (global transactivator)*  
GTR – *General Time Reversible*  
GV – *Granulovirus*  
h p.i. - horas pós-infecção  
*hrs* – região de repetição homóloga (*homologous repeat region*)  
*iap-1* – gene *inibidor de apoptose - 1 (inhibitor of apoptosis 1)*  
IB – Inferência Bayesiana  
IC – Índice de consistência  
*ie-2* – gene *imediatamente precoce ( immediate early 2)*  
IH – Índice de homoplasia  
ILD - teste da incongruência da diferença de comprimento (*Incongruence lenght difference test*)  
IPM – Manejo Integrado de Pragas (*Integrated Pest Mannagement*)  
IR – Índice de Retenção  
*lef* – gene *fator de expressão tardia (late expression factor)*  
MCMC – Monte Carlo via Cadeia de Markov  
ML – Máxima Verossimilhança (*Maximum Likelihood*)

MNPV – *multiple nucleopolyhedrovirus*

MP – Máxima Parcimônia

MTHF – 5,10-metiltetrahidrofolato

MUSCLE - *Multiple Sequence Alignment*

NCBI – *National Center for Biotechnology Information*

NJ – *Neighbour joining*

NPV – *nucleopolyhedrovirus*

nt – nucleotídeo

NTSYS - *Numerical Taxonomy and Multivariate System*

OB – corpo de oclusão (*occlusion body*)

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

ORF - *Open Reading Frame*

PABA - teste da adição de partição alterando o *bootstrap* (*Partition Addition Bootstrap Alteration*)

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

*phr* – gene *DNA photolyase*

*pif* – gene *fator de infectividade oral* (*per os infectivity factor*)

*polh/gran* – gene *poliedrina / granulina* (*polyhedrin/granulin*)

*ptp* – gene *fosfotirosina fosfatase* (*phosphotyrosine phosphatase*)

RAPD – DNA polimórfico amplificado ao acaso (*random amplification of polymorphic DNA*)

RC – Índice de Consistência Rescalonado

REN – endonucleases de restrição (*restriction endonucleases*)

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

SAHN - Agrupamento sequencial, aglomerativo, hierárquico e exclusivo (*Sequential Agglomerative, Hierarquic, Nonoverlapping, Clustering Methods*)

SAP - fosfatase alcalina de camarão (*shrimp alkaline phosphatase*)

SDS – dodecil sulfato de sódio (*sodium dodecyl sulfato*)

SNPs - polimorfismos de base única ou polimorfismos com base na substituição de nucleotídeos únicos (*Single Nucleotide Polymorphisms*)

SNPV – *single nucleopolyhedrovirus*

TAE - tampão Tris-Acetato-EDTA

TBR – *Tree Bisection Reconnection*

UV – ultravioleta

UPGMA – método da média aritmética não ponderada (*Unweighted pair group method with arithmetic mean*)

WAG - *Whelan And Goldman*

kb - quilobases

kDa - quiloDaltons M – molar

min – minuto(s)

pb – pares de bases

rpm – rotação por minuto

rpm – rotações por minuto

s – segundo

V – Volts

h – horas

$\delta$  - valor da alteração do suporte dos ramos

mm – mili ( $10^{-3}$ ) metro

$\mu$ L – micro ( $10^{-6}$ ) litro

$\mu$ m - micro ( $10^{-6}$ ) metro

$\eta$ g – nano ( $10^{-9}$ ) grama

$\eta$ m - nano ( $10^{-9}$ ) metro

## Resumo

A lagarta falsa-medideira (*Pseudoplusia includens*; Walker, 1857) tornou-se uma praga de grande importância nas lavouras de soja no Brasil. Sete isolados virais, caracterizados como variantes genotípicos do baculovirus *Pseudoplusia includens single nucleopolyhedrovirus* (PsinSNPV), foram obtidos de larvas *P. includens* coletadas em plantações de soja e algodão no Brasil e Guatemala. Visando determinar a diversidade genética e filogenia de sete isolados de PsinSNPV (IA a IG), sequências parciais dos genes *lef-8* e *lef-9* (*late expression factor*), *pif-2* (*per os infectivity factor 2*), *phr* (*DNA photolyase*) e *polh/gran* (*polyhedrin/granulin*) foram obtidas e árvores filogenéticas foram construídas usando análises da máxima parcimônia (MP) e Bayesiana (IB). DNA de sete isolados foi amplificado por PCR utilizando oligonucleotídeos degenerados específicos aos genes *lef-8*, *lef-9*, *pif-2*, *phr* e *polh/gran* e as sequências obtidas foram de 513, 260, 357, 780 e 510 pb de tamanho, respectivamente. A relação filogenética de PsinSNPV com outros 57 baculovirus sequenciados foi também determinada a partir de sequências nucleotídica e peptídica, individual e concatenada, dos genes *lef-8*, *lef-9*, *pif-2* e *polh/gran*. As árvores filogenéticas demonstraram que PsinSNPV agrupou-se com os baculovirus do gênero *Alphabaculovirus* Grupo II, onde PsinSNPV é mais proximamente relacionado ao *Chrysodeixis chalcites* NPV (ChchNPV) e *Trichoplusia ni* SNPV (TnSNPV). Árvores filogenéticas por máxima parcimônia e inferência Bayesiana foram obtidas da concatenação das sequências nucleotídicas dos cinco genes *lef-8*, *lef-9*, *phr*, *pif-2* e *polh*, onde *pif-2* foi o mais variável, contrariando dados da literatura. As árvores exibiram dois grupos monofiléticos estreitamente relacionados, um contendo os isolados IB, IC e ID e o outro contendo os isolados IA, IE, IF e IG.

## Abstract

The soybean looper (*Pseudoplusia includens*; Walker, 1857) has become a major pest in soybean crops in Brazil. Seven isolates, characterized as genotypic variants of the *Pseudoplusia includens* single nucleopolyhedrovirus (PsinSNPV) baculovirus were obtained from *P. includens* larvae collected in soybean and cotton crops in Brazil and Guatemala. In order to determine the genetic diversity and phylogeny of the seven isolates PsinSNPV (IA - IG), partial sequences of the genes *lef-8* and *lef-9* (late expression factor), *pif-2* (per os infectivity factor 2), *phr* (DNA photolyase) and *polh/gran* (polyhedrin/granulin) were obtained and phylogenetic trees constructed using maximum-parsimony (MP) and Bayesian analyses (IB). DNA from seven isolates was amplified by PCR using *lef-8*, *lef-9*, *pif-2*, *phr* and *polh/gran* specific degenerate primers and the sequences obtained were 513, 260, 357, 780 and 510 bp long, respectively. The phylogenetic relationship of PsinSNPV with 57 other baculoviruses was also determined from the individual and concatenated nucleotide and amino acid sequences of the genes *lef-8*, *lef-9*, *pif-2* and *polh*. The phylogenetic trees clearly place PsinSNPV with the group II *Alphabaculovirus*, where PsinSNPV is most closely related to ChchNPV and TnSNPV. Maximum parsimony and Bayesian phylogenetic trees were obtained from the concatenation of the nucleotide sequence of the five genes, *lef-8*, *lef-9*, *phr*, *pif-2* and *polh*, where *pif-2* was by far the most variable, in contrast to literature data. The trees exhibited two closely related monophyletic groups, one containing the isolates IB, IC and ID and another containing the isolates IA, IE, IF and IG.

# 1. Introdução

*Pseudoplusia includens single nucleopolyhedrovirus* (PsinSNPV) é um baculovirus patogênico à lagarta *Pseudoplusia includens* (Noctuidae: Lepidoptera) (Walker, 1857), praga emergente da cultura da soja, ocorrendo também em outras culturas como: algodão, feijão, tomate, batata, fumo, amendoim, girassol, hortaliças, entre outros (Souza *et al.*, 2009).

A lagarta *P. includens* apresentou, nas safras de 2003/2004 e 2007/2008, surtos em vários estados como Mato Grosso do Sul, São Paulo, Goiás e Paraná (Bueno *et al.*, 2009). Essa lagarta vem ganhando grande importância nos últimos anos, pois, além do aumento populacional, essas larvas apresentam uma maior capacidade de consumo de folhas quando comparada a lagarta da soja (*Anticarsia gemmatilis*) e se mostram resistentes a muitos inseticidas químicos necessitando de maiores doses para seu controle (Morales *et al.*, 1995).

Na natureza, os baculovirus possuem uma importância ecológica na regulação do tamanho das populações de insetos (Odindo, 1983; Evans, 1986). Características como alta especificidade e oclusão cristalina que protege as partículas virais das condições adversas do meio favorecem o uso do baculovirus para controle biológico (Payne, 1986; Moscardi e Sosa-Gómez, 1992a; 1992b).

Como parte preliminar deste trabalho, sete isolados virais (IA a IG), obtidos de larvas *P. includens* mortas por infecção viral em plantações de soja e algodão no Brasil e Guatemala, foram identificados e, quando avaliados, mostraram diferenças quanto a sua virulência. Dados referentes a análises de DNA mostraram diferentes perfis de restrição com presença de bandas

submolares demonstrando a existência de uma mistura de genótipos. Esses isolados virais foram então caracterizados como variantes genotípicos de PsinSNPV (Craveiro *et al.*, 2009; Alexandre *et al.*, 2010).

O baculovirus *Pseudoplusia includens* SNPV, objeto desta pesquisa, até então, não possui qualquer informação sobre sua sequência genômica e evolução. A utilização de ferramentas da bioinformática na biologia molecular e filogenia genética tem sido de fundamental importância em estudos de caracterização dos baculovirus. Com base nisso, neste estudo foram investigadas as variações genéticas observadas nas sequências parciais dos genes *late expression factor 8 e 9 (lef-8 e lef-9)*, *per os infectivity factor 2 (pif-2)*, *DNA photolyase (phr)* e *polyhedrin/granulin (polh/gran)* de PsinSNPV e determinadas as relações filogenéticas entre os isolados de PsinSNPV (IA a IG) e entre seu representante (PsinSNPV-IE) com outros baculovirus da família *Baculoviridae*.

A identificação de larvas *P. includens* mortas naturalmente pelo baculovirus PsinSNPV abre perspectivas para a utilização desse vírus no controle de larvas *P. includens*. A caracterização e o maior conhecimento sobre a virulência desse vírus auxiliará nos estudos que buscarão sua utilização como um agente de controle biológico.



## 2. Revisão Bibliográfica

### 2.1. *Pseudoplusia includens single nucleopolyhedrovirus* (PsinSNPV): um vírus patogênico à lagarta falsa-medideira

*Pseudoplusia includens single nucleopolyhedrovirus* (PsinSNPV) é um baculovirus patogênico à lagarta falsa-medideira, *Pseudoplusia includens*, também classificada como *Chrysodeixis includens* (Walker, 1857) (Lepidoptera: Noctuidae), uma praga emergente da cultura da soja, ocorrendo também em outras culturas como: algodão, feijão, tomate, batata, fumo, amendoim, girassol (Bueno *et al.*, 2007; Souza *et al.*, 2009).

A lagarta falsa-medideira, considerada como praga secundária da cultura da soja, apresentou, nas safras de 2003/2004 e 2007/2008, surtos em vários estados como Mato Grosso do Sul, São Paulo, Goiás e Paraná (Bueno *et al.*, 2009). O aumento da população dessa praga nas extensas plantações de soja tem sido atribuído às condições locais observadas como desfavoráveis à incidência de seus inimigos naturais (Bueno *et al.*, 2007). Essa espécie vem ganhando uma maior importância nos últimos anos, pois, além do aumento populacional, essas larvas necessitam de maiores doses de inseticida para seu controle e apresentam maior capacidade de consumo de folhas quando comparada à lagarta da soja *Anticarsia gemmatilis* (Morales *et al.*, 1995).

*Pseudoplusia includens* pertence à ordem Lepidoptera, família Noctuidae e subfamília Plusiinae. Sua distribuição é restrita ao hemisfério ocidental, ocorrendo desde o norte dos Estados Unidos da América (EUA) até o sul da América do Sul (Alford e Hammond, 1982; Eichling e Cunningham, 1978). Essa lagarta apresenta coloração verde claro com linhas brancas longitudinais uma de cada lado do corpo e duas linhas finas dorsais (Silvie *et al.*, 2007). Ela

possui três pares de pernas abdominais, fazendo com que seu deslocamento ocorra intenso movimento do corpo, parecendo medir palmos, característica que lhe confere o nome popular de lagarta-mede-palmo ou lagarta falsa-medideira (Figura 1). Outra característica marcante dessa lagarta é que ela consome grandes áreas das folhas mantendo íntegras as nervuras principais, o que confere um aspecto rendilhado às folhas atacadas (Herzog, 1980).

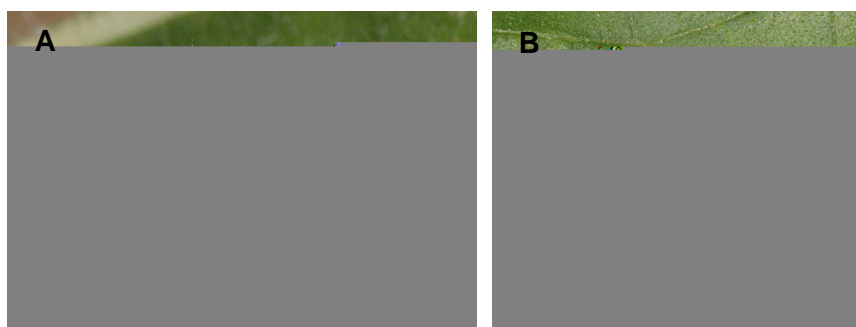


Figura 1. Larvas *Pseudoplusia includens* sadia (A) e morta por infecção viral (B). Possui coloração verde-clara com linhas longitudinais brancas no dorso. Presença de 3 pares de pernas abdominais que lhe confere movimento do corpo parecendo medir palmos. Foto cedida pela Embrapa Soja (Londrina-PR).

Na literatura são encontrados poucos estudos de caracterização de vírus patogênico à lagarta *P. includens*. Na década de 90, estudos com o vírus obtido de larvas *P. includens* provenientes de uma cultura de algodão na Guatemala, abordaram aspectos como dosagem e resposta da temperatura, efeitos de potenciadores virais, persistência no campo e eficácia de controle (Ali e Young, 1991; Livingston *et al.*, 1980; Kunimi *et al.*, 1997; Mcleod *et al.*, 1982; Zou e Young, 1996; Young, 2001; Young e Yearian, 1982).

Recentemente, isolados desse vírus, obtidos de larvas *P. includens* coletadas em plantações de soja no Brasil, foram identificados morfologicamente como *nucleopolyhedrovirus* (NPV), e estes exibiam uma ultraestrutura poliédrica contendo vírions com apenas um nucleocapsídeo por

envelope, sendo, portanto, do tipo “single”. A micrografia eletrônica de um representante do vírus *Pseudoplusia includens single nucleopolyhedrovirus* (PsinSNPV) é mostrada na Figura 2 (Alexandre *et al.*, 2010). Em continuidade, com base em análises de restrição de DNA e de infectividade viral, esses isolados virais juntamente com o vírus obtido na Guatemala foram caracterizados, inicialmente, como variantes genotípicos do vírus PsinSNPV (Alexandre *et al.*, 2010).

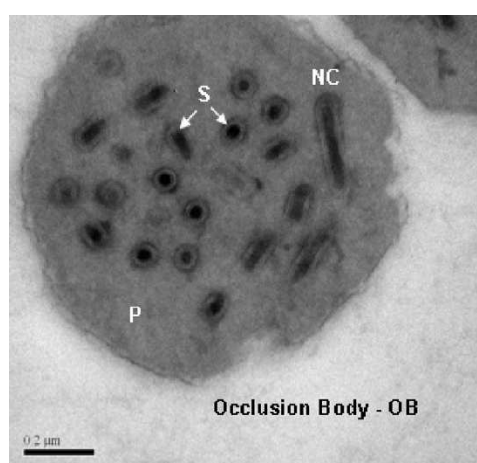


Figura 2. Micrografia eletrônica do corpo oclusão (OB) de PsinSNPV. OB contendo vírions, com um nucleocapsídeo por envelope, imersos na matriz protéica poliedrica. S: *single*, NC: nucleocapsídeo, P: poliedrina. Barra: 0,2 µm. (Alexandre *et al.*, 2010).

Em estudos de caracterização do gene *DNA photolyase* em genomas de baculovirus (NPV Grupo II) que infectam insetos da subfamília Plusiinae, outros dois isolados de PsinSNPV foram relatados, um proveniente de Los Angeles - EUA, um outro citado como proveniente da Guatemala (Xu *et al.*, 2008).

Devido o aumento da importância da lagarta falsa-medideira nas plantações de soja do Brasil, estudos vêm sendo realizados visando caracterizar o baculovirus PsinSNPV e avaliar seu potencial para controle biológico da praga *P. includens*.

## 2.2. Características gerais dos baculovirus

Os baculovirus constituem o maior grupo de vírus específicos de artrópodes que infectam e matam insetos da ordem Lepidoptera, Hymenoptera e Diptera (Theilmann *et al.*, 2005; Jehle *et al.*, 2006a).

Esses vírus possuem genoma de DNA circular, fita dupla, variando de 80 a 180 kb (Friesen, 2007). O material genético super condensado é circundado por uma proteína básica (p6.9) com massa molecular de 7 kDa que neutraliza a carga negativa do DNA. Esta estrutura está envolta por um capsídeo protéico em forma de bastonete, constituindo a unidade infectiva do vírus, o nucleocapsídeo de 30-60 x 250-300 nm (Jehle *et al.*, 2006a).

Os nucleocapsídeos envelopados formam os vírions e 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, 1998). Os vírus que apresentam corpos de oclusão de forma poliédrica, também chamados de poliedros, são os *nucleopolyhedrovirus* (NPVs) e os vírus que apresentam corpos de oclusão de forma ovocilíndrica, denominados grânulos, são os *granulovirus* (GVs) (Figura 3) (Theilmann *et al.*, 2005).

Os NPVs possuem vários vírions por cristal protéico contendo apenas um nucleocapsídeo por vírion (*Single* - SNPV) ou vários nucleocapsídeos por vírion (*Multiple* - MNPV), enquanto os GVs geralmente possuem um, ou raramente dois a três vírions por grânulo (Figura 3) (Friesen e Miller, 2001; Theilmann *et al.*, 2005).

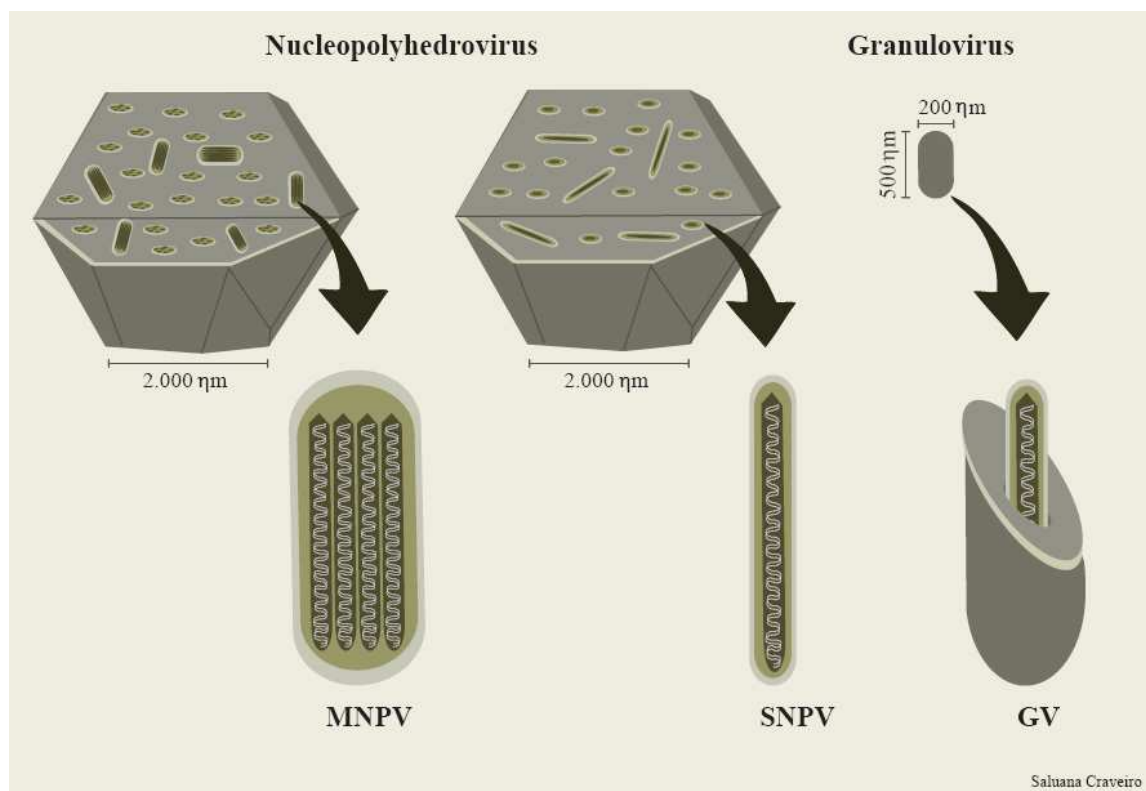


Figura 3. Desenho esquemático de corpos de oclusão dos baculovirus (Craveiro,2012). Os *nucleopolyhedrovirus* (NPVs) possuem vários vírions por poliedro apresentando um nucleocapsídeo (*single* - SNPV) ou vários nucleocapsídeos (*multiple* - MNPV) por envelope. Os *granulovirus* (GVs) apresentam um ou raramente dois a três vírions por grânulo.

As proteínas poliedrina e granulina, com massa molecular variando de 29 a 33 kDa, são os principais componentes protéicos dos corpos de oclusão dos NPVs e GVs, respectivamente. Essas proteínas correspondem cerca de 95% do conteúdo protéico total dos corpos de oclusão (Rohrmann, 2011).

Uma exceção entre os baculovirus é o vírus *Culex nigripalpus* NPV (CuniNPV), pelo fato de que a principal proteína do corpo de oclusão desse vírus não é homóloga à poliedrina e granulina dos vírus patogênicos aos lepidópteros e himenópteros (Jehle *et al.*, 2006a; Perera *et al.*, 2006).

A transmissão e replicação dos baculovirus ocorrem exclusivamente no estágio larval do inseto produzindo dois tipos de progênes infecciosas com

funções diferentes durante o ciclo de infecção. A forma ocluída do vírus – *occlusion derived virus* (ODV) - é responsável pela transmissão de inseto para inseto, enquanto a forma não ocluída, o vírus extracelular – *budded virus* (BV) - é responsável pela transmissão de célula para célula em um mesmo indivíduo (infecção sistêmica) (Figura 4). Essas duas progênes são essenciais para o processo de infecção e apesar de possuírem a mesma estrutura do nucleocapsídeo, diferem na composição e formação do seu envelope (Blissard, 1996; Funk *et al.*, 1997; Friesen, 2007).

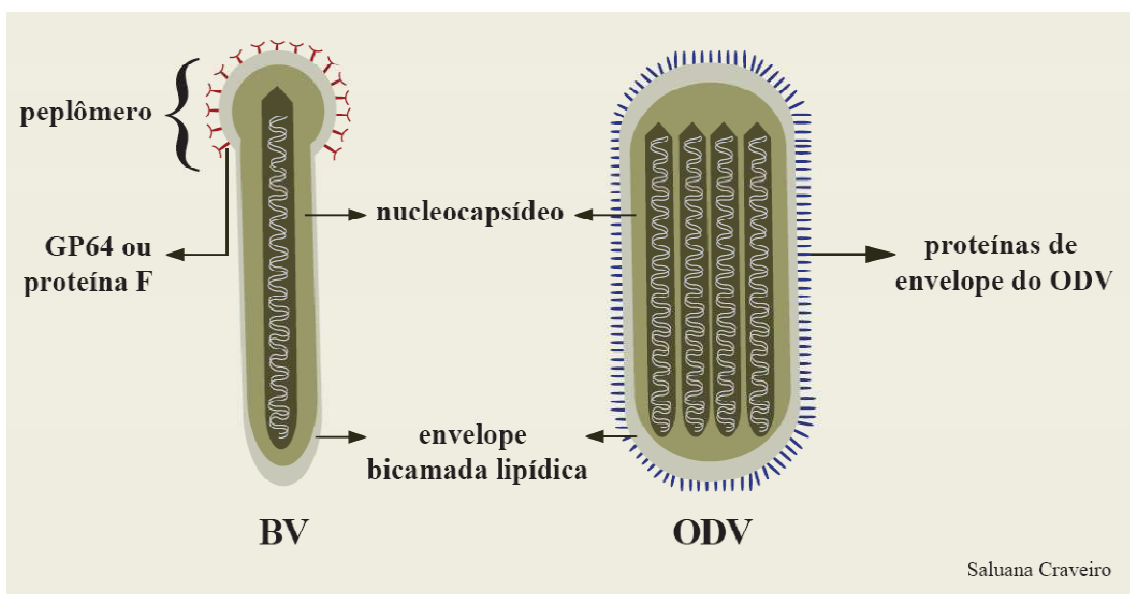


Figura 4. Desenho esquemático de formas virais produzidas durante o ciclo de infecção dos baculovirus (Craveiro, 2012). Vírus extracelular (BV) responsável pela infecção sistêmica e o vírus derivado de oclusão (ODV) responsável pela infecção de inseto para inseto.

No processo de infecção, a expressão gênica ocorre ordenadamente, em cascata, com a expressão dos genes de uma fase temporal regulando a expressão de genes das fases seguintes (Blissard e Rohrmann, 1990; O'Reilly *et al.*, 1992; van Oers e Vlak, 2007). A expressão temporal pode ser dividida em três fases: precoce (*early*) até 6 hp.i., tardia (*late*) de 6 a 18 hp.i. e muito tardia (*very late*) a partir de 18 hp.i.. Essas fases correspondem biologicamente

à reprogramação celular para replicação viral, produção de BVs e produção de OBs, respectivamente (Blissard e Rohrmann, 1990).

O ciclo de infecção dos baculovirus se inicia com a ingestão, pelo inseto, de partículas OBs presentes na superfície das folhas. Ao chegar ao intestino médio do inseto, o vírus é submetido a pH alcalino e ação de proteases que dissolvem as proteínas dos corpos de oclusão, liberando as partículas oclusas – ODVs (Horton e Burand, 1993). As partículas infectivas atravessam a membrana peritrófica e penetram nas células colunares do intestino médio da lagarta por fusão de membrana mediada por receptores específicos (Horton e Burand, 1993). Os nucleocapsídeos podem ser dirigidos ao núcleo para que ocorra a replicação viral ou podem atravessar a célula e atingir o sistema traqueal e outros tecidos. A replicação do DNA resulta na produção de BVs que brotam da membrana plasmática da célula hospedeira para infectar as células adjacentes (Engelhard *et al.*, 1994; Filipssen *et al.*, 1995). Os BVs penetram nas células por endocitose mediada por receptores (Volkman e Goldsmith, 1985; Volkman *et al.*, 1986). Em estágios avançados, 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 pela multiplicação das partículas OBs leva a uma série de mudanças comportamentais e morfológicas que iniciam com a redução na alimentação, retardo do crescimento do inseto e descoloração e rompimento do tegumento que culminam na morte da larva e liberação de grande quantidade de OBs no ambiente, servindo de inóculo para infecção de outras larvas do inseto hospedeiro (Granados e Williams, 1986; Volkman e Keddie, 1990).

### **2.3. Aplicações biotecnológicas: baculovirus como agentes de controle biológico de insetos-praga**

O controle de insetos-praga é realizado, principalmente, com uso de inseticidas químicos que muitas vezes é usado de forma inadequada favorecendo o aparecimento e a expansão de problemas de resistência de pragas, erupção de pragas secundárias, ressurgência e intoxicação humana e ambiental (Pedigo e Rice, 2005). Desta forma, a utilização de métodos alternativos como o uso do controle biológico torna-se essencial para a diminuição das conseqüências indesejáveis provenientes do uso de inseticidas químicos.

Os baculovirus são agentes de controle biológico utilizados com sucesso em programas de manejo integrado de pragas (IPM), sendo uma alternativa biológica aos inseticidas químicos (Moscardi *et al.*, 2011; Szewczyk *et al.*, 2011). Devido sua alta especificidade, os baculovirus 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, 2009; Souza *et al.*, 2007; Erlandson, 2008). A alta virulência dos baculovirus em um amplo espectro de hospedeiros e a estabilidade no meio ambiente são outras vantagens da utilização desses vírus como inseticidas microbianos.

O biocontrole baseado em baculovirus tem sido efetivo contra pragas como a lagarta-da-soja (*Anticarsia gemmatalis*) (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*) (Fritsch *et al.*, 2005; Sauphanor *et al.*, 2006; Vincent *et al.*, 2007), mariposa cigana (*Lymantria dispar*) (Cook *et al.*, 2003), mandarová da mandioca (*Erinnyis ello*) (Bellotti, 1999), entre outros.

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 (Moscardi *et al.*, 2011; Szewczyk *et al.*, 2011). No Brasil, esse programa chegou a atingir pico máximo de 2 milhões de hectares durante os anos 2003/2004. Posteriormente, o surgimento de pragas emergentes (exemplo: lagarta falsa-medideira), 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.

Uma das dificuldades da utilização dos baculovirus no campo é a resistência dos agricultores que, para matar mais rapidamente os insetos-praga, preferem utilizar inseticidas químicos. O fator limitante que mais afeta e compromete o desempenho dos baculovirus no campo são os efeitos da radiação solar que inativa grande parte da atividade viral dos corpos de oclusão em menos de 24 h (Jaques, 1985). Para aumentar a persistência no campo, formulações de baculovirus são acrescidas de branqueadores ópticos que aumentam a resistência do vírus à radiação UV.

Um maior conhecimento sobre a genômica dos baculovirus permitiu também a sua utilização como vetores de expressão para a produção de proteínas heterólogas, produção de vacinas e vetores na terapia gênica (Thiem e Cheng, 2009).

## **2.4. Genoma dos baculovirus**

O conteúdo total de genes dos baculovirus foi definido em estudos baseados em todos os genes dos 57 baculovirus seqüenciados revelando a existência de 895 ORFs diferentes presentes nos genomas virais (Miele *et al.*, 2011). Entretanto, o conteúdo de genes presente em cada membro da família *Baculoviridae* corresponde em torno de 90 a 181 genes dos quais 31 genes, denominados *core genes* - grupo de genes conservados entre os baculovirus, representam 3% do conteúdo genético viral (Rohrmann, 2011; Miele *et al.*, 2011).

### **2.4.1. Grupo de genes conservados entre os baculovirus - *core genes***

Os *core genes* são considerados genes mais ancestrais e altamente conservados entre todos baculovirus (Rohrmann, 2011; Miele *et al.*, 2011). Estão envolvidos nos diferentes estágios do ciclo viral como replicação do DNA, transcrição do RNA, composição protéica das partículas virais e infectividade oral (Tabela 1). Os genes conservados entre os baculovirus, quando comparados aos outros, possuem menor tolerância às mutações, por isso, alterações nesses genes podem implicar na perda da viabilidade viral (Herniou *et al.*, 2003; Miele *et al.*, 2011).

**Tabela 1. Genes conservados entre os baculovirus (core genes)**

<b>Genes</b>	<b>AcMNPV (ORF)</b>	<b>LdMNPV (ORF)</b>	<b>CpGV (ORF)</b>	<b>NeseNPV (ORF)</b>	<b>CuniNPV (ORF)</b>
<b>Replicação</b>					
<i>lef-1</i>	14	123	74	68	45
<i>lef-2</i>	6	137	41	57	25
<i>DNA polymerase</i>	65	83	111	28	91
<i>helicase</i>	95	97	90	61	89
<b>Transcrição</b>					
<i>lef-4</i>	90	93	9	62	96
<i>lef-5</i>	50	51	131	81	26
<i>lef-8</i>	62	64	117	40	59
<i>lef-9</i>	40	48	68	49	73
<i>p47</i>	99	100	87	58	88
<b>Ciclo celular: entrada, empacotamento e exportação</b>					
<i>p6.9</i>	100	101	86	36	23
<i>vp39</i>	89	92	96	89	24
<i>vlf-1</i>	77	86	106	45	18
<i>alk-exo</i>	133	157	125	31	53
<i>vp1054</i>	54	57	138	85	8
<i>vp91/p95</i>	83	91	101	84	35
<i>gp41</i>	80	88	104	47	33
<i>38 k</i>	98	99	88	59	87
<i>p33</i>	92	94	93	24	14
<i>odv-ec43</i>	109	107	55	70	69
<i>p49</i>	142	20	15	63	30
<i>odv-nc42</i>	68	80	114	41	58
<i>odv-e18</i>	143	19	14	65	31
<i>desmoplakin</i>	66	82	112	29	92
<i>odv-e27</i>	144	18	97	66	32
<i>ac81</i>	81	89	103	48	106
<b>Infectividade oral</b>					
<i>pif-0/p74</i>	138	27	60	50	74
<i>pif-1</i>	119	155	75	79	29
<i>pif-2</i>	22	119	48	55	38
<i>pif-3</i>	115	143	35	69	46
<i>pif-4/19k/odv-e28</i>	96	98	89	60	90
<i>pif-5/odv-e56</i>	148	14	18	38	102

#### **2.4.2. Genes *lef-8*, *lef-9*, *pif-2* e *polh/gran*: uma ferramenta para estudos de identificação e filogenia dos baculovirus**

Os genes *late expression factor 8 (lef-8)*, *late expression factor 9 (lef-9)* e *per os infectivity factor 2 (pif-2)* pertencem ao grupo de genes altamente conservados dentre os baculovirus. Esses genes, juntamente com o gene *polyhedrin/granulin (polh/gran)*, vêm sendo amplamente utilizados nos estudos de análise filogenética, diversidade genotípica e evolução de baculovirus.

Neste trabalho, estes quatro genes foram escolhidos por ser considerados uma ferramenta adequada e de comprovada eficiência para estudos de filogenia e evolução de NPVs que não possuem seus genomas seqüenciados (Herniou *et al.*, 2004; Lange *et al.*, 2004; Jehle *et al.*, 2006b).

Os genes *lef-8* e *lef-9* são expressos na fase precoce e são essenciais para expressão dos promotores dos genes tardios e muito tardios (Lu e Miller, 1994, 1995; Li *et al.*, 1999; Todd *et al.*, 1995). As proteínas LEF-8 e LEF-9 juntamente com as proteínas LEF-4 e P47 formam as quatro subunidades da RNA polimerase viral (Lu e Miller, 1994; Titterington *et al.*, 2003). Os genes *lef-8* e *lef-9* possuem os motivos conservados da subunidade  $\beta$  e  $\beta'$  que formam parte do sítio catalítico da enzima RNA polimerase viral (Tabela 2) (Lu e Miller, 1994; Passarelli *et al.*, 1994).

O gene *pif-2* pertence à família *pif* constituída por genes expressos na fase tardia que codificam fatores essenciais para a infectividade oral, principal rota da transmissão primária dos baculovirus (Tabela 2), mas não para replicação do vírus em cultura de células (Kikhno *et al.*, 2002; Braunagel *et al.*, 2003; Pijlman *et al.*, 2003; Fang *et al.*, 2006). A família PIF é constituída pelas proteínas P74 (PIF-0) (Faulkner *et al.*, 1997) e PIF-1 a PIF-5 (Kikhno *et al.*,

2002; Fang *et al.*, 2006; Pijlman *et al.*, 2003; Okhawa *et al.*, 2005; Fang *et al.*, 2009; Harrison *et al.*, 2010; Xiang *et al.*, 2011). Essas proteínas estão localizadas no envelope e são exclusivas do fenótipo ODV estando relacionadas aos eventos iniciais da infecção como o processo de ligação e entrada das partículas virais nas células do intestino médio da lagarta hospedeira (Okhawa *et al.*, 2005; Peng *et al.*, 2010). Recentemente, dois novos fatores de infectividade oral foram identificados, um presente apenas nos *Alphabaculovirus* do Grupo II, gene que codifica a proteína HA100 associada aos ODVs de *Helicoverpa armigera* NPV (Luo *et al.*, 2011), e outro presente nos baculovirus que infectam lepidópteros, gene *sf58* (Simon *et al.*, 2012).

O gene *polh/gran* é altamente conservados dentre os baculovirus específicos de lepidópteros. Esse gene possui altos níveis de expressão na fase muito tardia e codifica a principal proteína constituinte dos corpos de oclusão, a poliedrina nos NPVs e a granulina nos GVs (Tabela 2) (Hooft van Iddekinge *et al.*, 1983; Rohrmann, 1986). A matrix protéica constituída por poliedrina ou granulina tem como função, a proteção dos vírions das condições adversas do meio ambiente (Jarvis *et al.*, 1991).

**Tabela 2. Genes conservados em baculovirus específicos de lepidópteros utilizados em estudos filogenéticos**

	Gene	Função	ORF AcMNPV
<b>Transcrição</b>	<i>lef-8</i>	Subunidade da RNA polimerase viral.	50
	<i>lef-9</i>	Subunidade da RNA polimerase viral. Condensação do DNA.	62
<b>Estrutural</b>	<i>pif-2</i>	Proteína de ligação com células do intestino médio da lagarta	22
	<i>polh/gran</i>	Principal proteína constituinte dos corpos de oclusão	8

### 2.4.3. Gene *DNA photolyase (phr)*

Os organismos, quando expostos a radiação ultravioleta (UV) da luz solar (250-320 nm), podem sofrer lesões no DNA do tipo dímeros de pirimidina ciclobutano (CPDs) ou fotoproduto (6-4) pirimidina-pirimidona [(6-4)PPs] (Sancar, 1994; Sun *et al.*, 2004) (Figura 5). Esses fotoprodutos são citotóxicos e podem bloquear a transcrição, induzir mutações e disparar o processo apoptótico (Ljungman e Zhang, 1996; Nishigaki *et al.*, 1998; Ozturk *et al.*, 2008).

O fator limitante que mais afeta e compromete o desempenho dos baculovirus no campo são os efeitos da radiação solar que inativa grande parte da atividade viral dos corpos de oclusão (OBs) em menos de 24 h (Jaques, 1985; Sun *et al.*, 2004). A exposição dos baculovirus a radiação UV diminuiu drasticamente sua infectividade oral (Shapiro *et al.*, 2002).

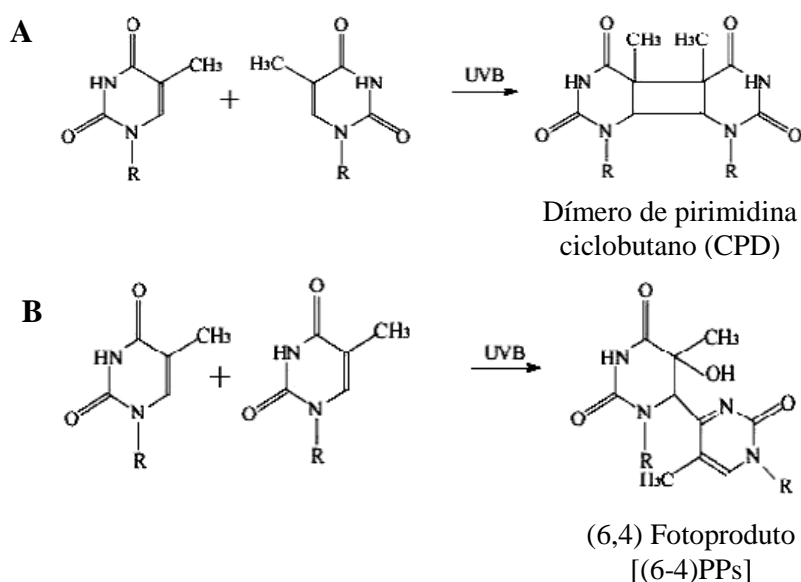


Figura 5. Desenho esquemático de fotoprodutos de DNA induzidos pela radiação UV. Lesões formadas por bases pirimidinas adjacentes em uma mesma fita de DNA exposta a luz UV-B, (A) dímeros de pirimidina ciclobutano (CPDs) e (B) fotoproduto (6,4) pirimidina-pirimidona [(6-4)PPs] (Adaptado de Montagner e Costa, 2009).

O gene *DNA photolyase (phr)* codifica uma proteína reparadora de dímeros de pirimidina ciclobutano, a enzima CPD-fotoliase que utiliza somente a luz como fonte de energia resultando em um caminho simples e rápido de reparo de DNA.

As CPD-fotoliasas são proteínas monoméricas com massa molecular de 44-66 kDa com 420-616 resíduos de aminoácidos (Sancar, 1990). Todas as CPD-fotoliasas conhecidas possuem como cofator catalítico o domínio flavina adenina dinucleotídeo (FAD) e um cofator que assimila a energia da luz, a 5,10-metiltetrahidrofolato – MTHF ou 8-hidroxi 7,8-didemetil-5-deazaflavina – 8-HDF (Deisenhofer, 2000; Eker *et al.*, 2009; Essen e Klar, 2006; Sancar, 2003).

As CPD-fotoliasas são divididas em dois grupos (classe I e II) em decorrência da divergência das sequências de aminoácidos dos diferentes organismos (Kanai *et al.*, 1997). A CPD-fotoliase de classe I está presente em vários microrganismos e a CPD-fotoliase de classe II em organismos eucariotos superiores (plantas e animais) e vários patógenos como poxvirus e microsporídios parasitas de insetos (Yasui *et al.*, 1994, Srinivasan *et al.*, 2001 Bennett *et al.*, 2003; Slamovits e Keeling, 2004).

Dois genes CPD-fotoliase de classe II, *phr 1* e *phr 2*, foram identificados no vírus *Chrysodeixis chalcites* NPV (ChchNPV) (van Oers *et al.*, 2004; 2005). Estudos posteriores mostram que o gene *phr* é conservado dentre os NPVs do Grupo II (*Alphabaculovirus*) que infectam insetos da subfamília Plusiinae (Xu *et al.*, 2008, Biernat *et al.*, 2011) estando presente em *Trichoplusia ni* SNPV (TnSNPV) (Willis *et al.*, 2005), *Clanis bilineata* NPV (ClbiNPV) (Zhu *et al.*, 2009), *Plusia acuta* NPV (PlacNPV) (Xu *et al.*, 2004), *Apocheima cinerarium* NPV (ApciNPV) (Biernat *et al.*, 2011) e *Thysanoplusia orichalcea* NPV isolado

B9 (ThorNPV-B9) (Xu *et al.*, 2004). Nos NPVs do Grupo I, que infectam insetos dessa mesma subfamília, o gene *phr* não foi encontrado e, entre os *Betabaculovirus*, esse gene foi identificado apenas em *Spodoptera litura* GV (Wang *et al.*, 2008).

## 2.5. Evolução e taxonomia dos baculovirus

A origem evolutiva dos baculovirus pode ser explicada por diferentes hipóteses. Uma das hipóteses propõe que os baculovirus originaram com os insetos da ordem Lepidoptera e posteriormente, por transferência horizontal, atingiram outras ordens de insetos (Rohrmann, 1986). Outra hipótese propõe que a origem dos baculovirus remonta a origem dos artrópodes com a co-cladogênese do vírus e seu hospedeiro (Federici, 1997). 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 e colaboradores (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 baculovirus pertencem à família *Baculoviridae* e são genética e morfologicamente distintos das outras famílias de vírus. A família *Baculoviridae*



possue uma complexidade de forma e função sugerindo uma longa linhagem evolutiva (Slack e Arif, 2007).

Até 2005, com o VIII Relatório do Comitê Internacional de Taxonomia de Vírus (Fauquet *et al.*, 2005), a família *Baculoviridae* foi taxonomicamente dividida 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. Os genes *dnapol*, *egt*, *gp41*, *chiA*, *cath*, *lef2*, *gp37* e, principalmente, o *polyhedrin/granulin* (*polh/gran*) foram os mais utilizados na análise filogenética que agrupavam os baculovirus em NPVs específicos de lepidópteros, subdivididos em Grupo I que inclui o vírus AcMNPV e OpMNPV e Grupo II que inclui HaSNPV e SeMNPV, e em GVs (Herniou *et al.*, 2003; Herniou e Jehle, 2007). Com o avanço do seqüenciamento do genoma completo, atualmente com 57 baculovirus seqüenciados (Miele *et al.*, 2011), a filogenia dos baculovirus pode ser melhor elucidada (Herniou e Jehle, 2007).

Herniou e colaboradores (2003), em um estudo sobre a evolução dos baculovirus baseado na análise filogenética de sequências genômicas completas de 13 baculovirus, apresentaram quatro grupos distintos: CuniNPV, GVs, NPVs Grupo I e NPVs Grupo II mostrando que o vírus CuniNPV não está agrupado aos outros NPVs de lepidópteros. Em continuidade, análise filogenética baseada em 29 genes comuns a todos os 29 baculovirus com o genoma completo seqüenciado mostrou que a classificação dos baculovirus deveria ser revisada (Jehle *et al.*, 2006a). Assim, a nova classificação reorganiza a família *Baculoviridae* em quatro gêneros: *Alphabaculovirus*, *Betabaculovirus*, *Gammabaculovirus* e *Deltabaculovirus* (Figura 6).

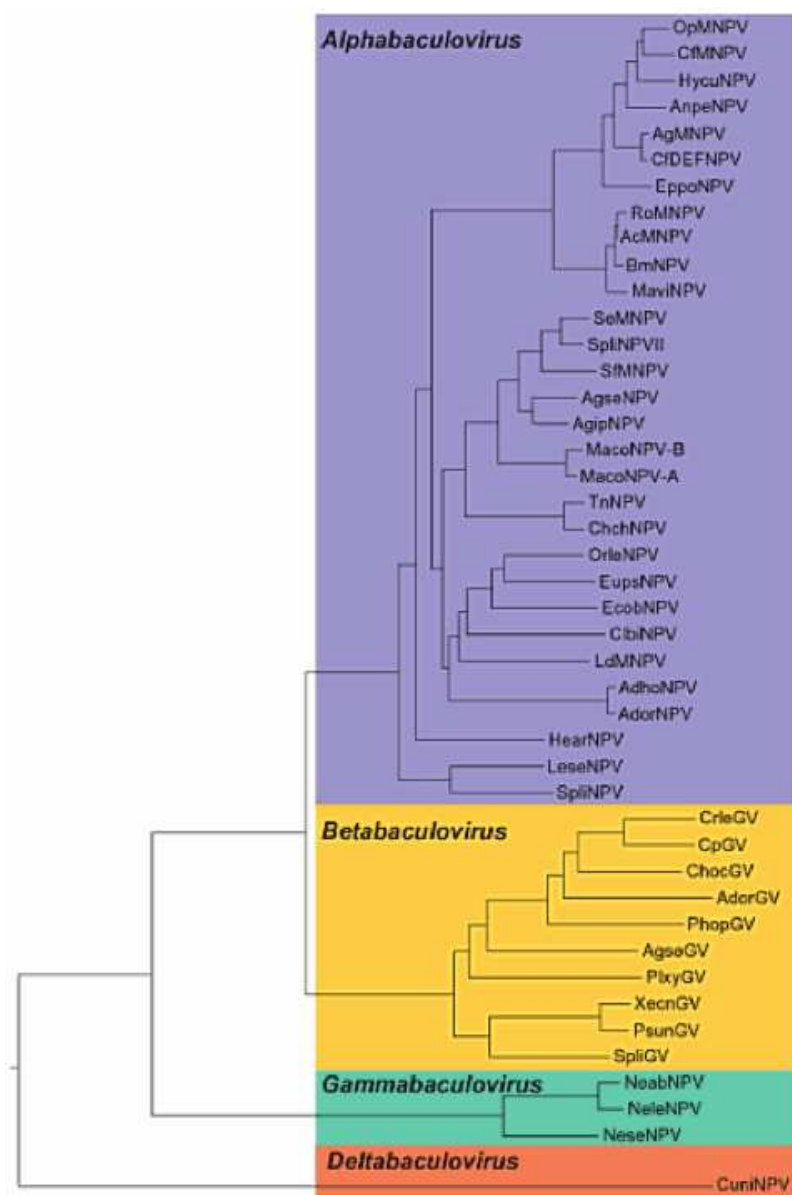


Figura 6. Taxonomia da família *Baculoviridae*. Árvore da Máxima Verossimilhança apresentando 44 baculovirus distribuídos em 4 gêneros *Alpha-*, *Beta-*, *Gamma-* e *Deltabaculovirus* (King *et al.*, 2011).

O gênero *Alphabaculovirus* inclui NPVs específicos de lepidópteros, com OBs de 0,4-15  $\mu\text{m}$  de tamanho e genoma de 100 a 180 kb. O *Betabaculovirus* inclui os GVs específicos de lepidópteros com OBs de 0,3-0,5 x 0,13-0,25  $\mu\text{m}$  e genoma de tamanho similar aos *Alphabaculovirus*. O *Gammabaculovirus* inclui os vírus específicos de himenópteros e atualmente é constituído pelo NeleSNPV, NeseSNPV e NeabNPV, com OBs de 0,4-1,1  $\mu\text{m}$  e genoma de 82 a 86 kb (Jehle *et al.*, 2006a). O *Deltabaculovirus* inclui os vírus específicos de

dípteros atualmente representado pelo CuniNPV com OBs de 0,4 µm de diâmetro e genoma de 108 kb (Afonso *et al.*, 2001) (Tabela 3).

**Tabela 3. Classificação taxonômica da família *Baculoviridae***

Gênero	Membros	Número de espécies*
<i>Alphabaculovirus</i>	NPV – Lepidópteros	456
<i>Betabaculovirus</i>	GV – Lepidópteros	148
<i>Gammabaculovirus</i>	NPV – Himenópteros	30
<i>Deltabaculovirus</i>	NPV – Dípteros	27

\* Número de espécies de insetos infectados por vírus (Rohrmann, 2011).

Os *Alphabaculovirus* diferem entre seus dois grupos: Grupo I e Grupo II (Zanotto *et al.*, 1993; Bulach *et al.*, 1999; Herniou *et al.*, 2001). 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 desempenhar esse papel (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, 2011). 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.*, 2006a).

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 vírus. 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).

## 2.6. Filogenia molecular

A filogenia molecular representa a hipótese das relações de parentesco dos organismos com base em eventos evolucionários presentes em dados moleculares como sequências nucleotídicas (DNA e RNA) e/ou sequências peptídicas (proteínas).

Partindo da premissa que os organismos sofrem contínuas modificações e que as espécies evoluem de um ancestral comum, surgiu grande interesse nos estudos das relações evolutivas e de parentesco entre as espécies.

Proposta pela primeira vez por Willi Hennig em 1950, a filogenia constitui a hipótese das relações de ancestralidade comum entre táxons ou grupos inferidas a partir de caracteres morfológicos, fisiológicos, moleculares e outros.

A utilização de dados moleculares para estudos filogenéticos surgiu com Zuckerkandl e Pauling (1962). A partir de estudos utilizando moléculas de hemoglobina, eles sugeriram que “a taxa de mudança evolutiva de qualquer proteína é aproximadamente constante ao longo do tempo”. Assim surgiu a idéia de relógio molecular que relaciona o tempo de divergência entre duas espécies com o número de modificações moleculares encontradas entre sequências de DNA ou proteínas.

As moléculas de DNA, RNA e proteínas são caracteres adequados a inferência filogenética, pois são dados de base genética e, uma vez que serão herdados, poderão refletir a relação de parentesco entre os organismos. Além

disso, os dados moleculares fornecem uma quantidade bem maior de informação do que os dados morfológicos, pois cada nucleotídeo ou aminoácido será um carácter independente na análise (Arriel *et al.*, 2006).

As árvores filogenéticas são representações gráficas das relações evolutivas entre os organismos. Existem dois tipos de análises utilizadas na construção das árvores: análise fenética e cladística.

Na análise fenética, as árvores filogenéticas são construídas por métodos que fazem a transformação da matriz de caracteres em uma matriz numérica de distância que contém as distâncias, tomadas em pares, do conjunto de dados. Exemplos desses métodos são: *Unweighted pair group method with arithmetic mean* (UPGMA) e *Neighbor joining* (NJ).

A análise cladística é baseada na matriz de caracteres, ou seja, a matriz é construída pela análise direta dos caracteres como é o método da Máxima Parcimônia (MP), Máxima Verossimilhança (ML) e Inferência Bayesiana (IB) (Schneider, 2007).

O método UPGMA é o método mais simples utilizado para a construção de fenogramas (árvores de similaridade). A árvore de similaridade é construída pelo método de agrupamento hierárquico dos pares de táxons não ponderados, baseado na média aritmética das distâncias. Esse método considera a árvore filogenética como aditiva e assume uma taxa constante de evolução, com todos os táxons igualmente distantes da raiz (Schneider, 2007).

O método da Máxima Parcimônia baseia-se no princípio da homologia, ou seja, se dois táxons compartilham a mesma característica é porque ela foi herdada do último ancestral comum. Na MP, a árvore que melhor representa a filogenia do grupo é a árvore que apresenta o menor número de mudanças

(passos) para explicar toda a variação presente na matriz de caracteres (Schneider, 2007).

A Análise Bayesiana é semelhante ao método da Máxima Verossimilhança (ML), mas difere na noção de probabilidade. A ML considera que todos os sítios evoluem independentemente e o cálculo da probabilidade para cada sítio é estimado baseando-se em modelos de substituição nucleotídica ou peptídica *a priori*. Na IB, a análise é baseada nas probabilidades *a posteriori*, ou seja, são estimadas as probabilidades da hipótese através da observação dos dados (Hall, 2001).

Com o intuito de desenvolver um método prático, rápido e com um alto poder de varredura para a identificação e classificação dos baculovirus, Lange *et al.* (2004) propuseram a construção de árvores filogenéticas baseada nas sequências parciais de genes altamente conservados a todos os baculovirus específicos de lepidópteros. Oligonucleotídeos degenerados para amplificação de sequências parciais de genes como *dnapol*, *pif2*, *lef8*, *lef9*, *polh/gran* foram utilizados para analisar a filogenia dos vírus *Thysanoplusia orichalcea* MNPV (ThorMNPV) (Cheng *et al.*, 2005), *Orgyia pseudotsugata* SNPV (OpSNPV) (Jakubowska *et al.*, 2007), *Epinotia granitalis* NPV (EpgrSNPV) (Takatsuka, 2007) e *Clanis bilineata* NPV (CibiNPV) (Zhu *et al.*, 2009) e analisar a diversidade genética e evolução de isolados de *Cydia pomonella* GV (CpGV) (Eberle *et al.*, 2009), *Spodoptera frugiperda* MNPV (SfMNPV) (Rowley *et al.*, 2010) e *Spodoptera exempta* NPV (SpexNPV) (Redman *et al.*, 2010).

## 2.7. Variação genética em populações de baculovirus

Apenas uma pequena proporção do genoma viral codifica genes que são comuns a todos os baculovirus (Jehle *et al.*, 2006a) e, 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 do genoma (Shackelton e Holmes, 2004).

Variantes genotípicos de um mesmo baculovirus são comumente encontrados em populações de campo de diferentes épocas (isolados temporais) ou localidades geográficas (isolados geográficos) como têm sido descrito para os vírus *Autographa californica* MNPV (Lee e Miller, 1978), *Cydia pomonella* GV (Eberle *et al.*, 2009), *Helicoverpa armigera* NPV (Figueiredo *et al.*, 2009), *Mamestra configurata* NPV (Li *et al.*, 2005), *Pseudoplusia includens* SNPV (Alexandre *et al.*, 2010) e *Spodoptera frugiperda* MNPV (Simón *et al.*, 2004; Rowley *et al.*, 2010). Além da variação em uma mesma população, variantes genotípicos de um mesmo vírus podem ser encontrados em uma única larva infectada (Cory *et al.*, 2005).

As variações são mutações pontuais, substituições, inserções e deleções que ocorrem por todo genoma, mas se concentram em regiões de alta variabilidade, as regiões de repetição homóloga - *homologous repeat regions* (*hrs*) e as ORFs repetidas dos baculovirus – *baculovirus repeated ORF* (*bro*) (de Jong *et al.*, 2005; Hayakawa *et al.*, 2000; Li *et al.*, 2002a). As *hrs* são regiões intergênicas formadas por sequências repetitivas que podem funcionar como ativadores em *cis* de *enhancers* e de origens de replicação viral do DNA

(Guarino *et al.*, 1986; Guarino e Summers, 1986; Hilton e Winstanley, 2007; Pearson *et al.*, 1992) 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).

Nas populações de campo, variantes genotípicos são facilmente detectados pela presença de bandas submolares e de polimorfismo de tamanho de fragmentos gerados nos perfis de restrição (Cory *et al.*, 2005).

A confirmação da mistura de genótipos presente em um único isolado ocorre com a purificação de clones virais utilizando métodos de diluição seriada do vírus, *in vitro* com células de inseto permissivas (*plaque assay*) e *in vivo* por passagens do vírus em lavas infectadas por injeção intra-hemocélica.

Ensaio de virulência podem apresentar diferenças de patogenicidade entre isolados virais de um mesmo baculovirus. Essas diferenças podem representar um importante fator para o desenvolvimento de novos bioinseticidas virais (Moscardi, 1999).



### 3. Objetivo Geral

Determinar as possíveis relações filogenéticas e evolucionárias entre sete isolados de PsinSNPV (IA a IG) e entre o vírus PsinSNPV com outros membros da família *Baculoviridae* para o estabelecimento de sua posição filogenética. Para este fim, variações genéticas entre esses isolados foram investigadas, caracterizadas e associadas às análises filogenéticas utilizando sequências parciais, individuais e concatenadas, de quatro genes conservados dentre os baculovirus específicos de lepidópteros (*lef-8*, *lef-9*, *pif-2* e *polh/gran*) e o gene *DNA photolyase (phr)* conservado nos *Alphabaculovirus* do Grupo II que infectam insetos da subfamília Plusiinae.

#### 3.1. Objetivos específicos

- Amplificar e identificar por PCR os genes conservados *lef-8*, *lef-9*, *pif-2* e *polh/gran* e o gene *phr* dos sete isolados de PsinSNPV (IA a IG);
- Purificar e seqüenciar os produtos de PCR;
- Confirmar a autenticidade das sequências obtidas por meio do programa BLAST e depositar as novas seqüências no *GenBank*;
- Detectar os polimorfismos das seqüências nucleotídicas parciais dos genes *lef-8*, *lef-9*, *pif-2*, *phr* e *polh/gran* dos isolados de PsinSNPV;
- Construir árvores filogenéticas com base na concatenação das sequências nucleotídicas parciais dos genes *lef-8*, *lef-9*, *pif-2*, *phr* e *polh/gran* dos isolados de PsinSNPV pela Máxima Parcimônia e Inferência Bayesiana;
- Reconstruir as relações filogenéticas dos baculovirus com base nas sequências nucleotídicas e peptídicas dos genes *lef-8*, *lef-9*, *pif-2* e

*polh/gran* de 58 baculovirus, incluindo o vírus PsinSNPV pela Máxima Parcimônia e Inferência Bayesiana.

- Analisar os polimorfismos de fragmentos de restrição de DNA (RFLP) dos isolados de PsinSNPV para a construção da matriz binária de presença/ausência de banda.
- Construir dendrograma de similaridade genotípica a partir da matriz binária dos dados de RFLP utilizando o algoritmo UPGMA.

## 4. Materiais e Métodos

### 4.1. Vírus

Isolados de *Pseudoplusia includens* SNPV, designados IA, IB, IC, ID, IE, IF e IG, foram obtidos a partir de larvas de *P. includens* mortas por infecção viral, cedidas pela Embrapa Soja, Londrina (PR). As larvas infectadas foram coletadas em plantações de soja e algodão no Brasil e Guatemala (Tabela 4).

**Tabela 4. Isolados de PsinSNPV obtidos de larvas *P. includens* infectadas**

Isolados Virais	Data de Coleta	Local de Coleta	Instituição	Cultura
IA	1972	Guatemala	University of Arkansas/EUA	Algodão
IB	Jan./2006	Londrina, PR	Embrapa Soja	Soja
IC	Jan./2006	Maringá, PR	Embrapa Soja	Soja
ID	Fev./2006	Iguaraçu, PR	Embrapa Soja	Soja
IE	Fev./2007	Iguaraçu, PR	Embrapa Soja	Soja
IF	Jan./2008	Dourados, MS	Embrapa Soja	Soja
IG	Jan./2008	Sertanópolis, PR	Embrapa Soja	Soja

### 4.2. Ferramentas de bioinformática

As ferramentas computacionais listadas a baixo foram utilizadas para a determinação das sequências e análises filogenéticas propostas neste estudo.

- Chromas *Pro* v1.5, 2003- 2009 (*Technelysium Pty*)
- *BioEdit Sequence Alignment Editor* v7.0.8.0, 1997-2007 (Hall, 1999)
- *FastGap* v1.2 (Borchsenius, 2009)
- *jModelTest* v0.1.1, 2005 (Posada, 2008)

- *ProtTest* v2.4, 2004-2011 (Abascal et al., 2005)  
(<http://darwin.uvigo.es/software/prottest.html>)
- *PAUP\** v4.0b10, 2001 (Swofford, 2003)
- *MrBayes* v3.1.2 (Huelsenbeck e Ronquist, 2001).
- *NTSYS - Numerical Taxonomy and Multivariate System* v2.21 (Rohlf, 1997)
- *BOOD* v3.0 (Coelho, 2002)
- *FigTree – Tree figure drawing toll* v1.3.1, 2006-2009 (Raumbaut, 2009)
- *Adobe Illustrator CS5* v15.0.0, 1987-2010
- *NCBI - National Center for Biotechnology Information:*
  - *BLAST - Basic Local Alignment Search Tool* (Astschul et al., 1997)  
(<http://blast.ncbi.nlm.nih.gov/Blast.cgi>)
  - *GenBank Database* (<http://www.ncbi.nlm.nih.gov/genbank/>)
  - *Sequin Application* v11.75 (<http://www.ncbi.nlm.nih.gov/Sequin/>)
- *MUSCLE - Multiple Sequence Alignment* v3.5 (Edgar, 2004)  
(<http://www.ebi.ac.uk/Tools/msa/muscle/>)
- *Expasy - Expert Protein Analysis System* (Gaspéiger et al., 2003):
  - *Translate Toll* (<http://web.expasy.org/translate/>)

### 4.3. Purificação de corpos de oclusão (OBs)

OBs de diferentes isolados de PsinSNPV (IA a IG) foram purificados essencialmente para estoque das amostras e para extração de DNA viral. Larvas de *P. includens* infectadas de cada local de coleta foram maceradas separadamente 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 15 min. O sedimento foi ressuspensão em tampão TE (Tris 0,01 M, pH 7,8 e EDTA 0,001 M) e SDS 0,5% e centrifugado a 12.000 rpm (Sorvall RC-5B, rotor SS-34) por 12 min. Essa centrifugação foi repetida com o sedimento ressuspensão em tampão TE e NaCl 0,5 M. Do material ressuspensão em água deionizada Milli-Q foram retirados 5 mL e aplicados em gradiente de sacarose contínuo de 40-65%, preparado em tampão TE a uma densidade de 1,17-1,31 g / mL. O material foi centrifugado a 24.000 rpm por 40 min, a 4°C (Sorvall OTD 75 U, rotor AH-627), e a banda correspondente aos OBs, posicionada no terço inferior do tubo, foi coletada com pipetas Pasteur, diluída 5 vezes em tampão TE e centrifugada a 12.000 rpm por 15 min, a 4°C (Sorvall RC-5B, rotor SS-34). A ressuspensão final foi feita em água deionizada Milli-Q autoclavada e os OBs foram armazenados a -20°C. A estimativa da concentração de partículas OBs em suspensão foi obtida a partir de contagem em Câmara de Neubauer com auxílio do microscópio de contraste de fase.

#### **4.4. Extração de DNA a partir de partículas OBs purificadas**

DNA dos isolados de PsinSNPV foram purificados, quantificados e verificados quanto a sua qualidade por análise em eletroforese de géis de agarose 0,8% para utilização nos diversos procedimentos que se seguem neste trabalho.

A extração de DNA viral foi baseada no protocolo descrito por O'Reilly *et al.* (1992). 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 incubação a 37°C por 30 min. Após a verificação da solubilização por microscópio óptico, foram adicionados SDS 1% (incubação por 10 min) e 500 µg / mL de proteinase K (incubação / 16 h). O material foi então centrifugado em microcentrifuga a 12.000 rpm (Eppendorf 5410 – rotor fixo), por 2 min, sendo o sedimento descartado. 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 3-5 min e o material centrifugado a 12.000 rpm (Eppendorf 5410 – rotor fixo), por 2 min. 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 / 16 h. 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 1 h. As amostras de DNA purificado foram armazenadas a 4 °C.

#### **4.5. Clivagem de DNA viral por endonucleases de restrição**

DNAs dos isolados de PsinSNPV foram clivados com diferentes enzimas de restrição para análise comparativa dos perfis eletroforéticos para estudo de polimorfismos dos fragmentos gerados (RFLP).

Os DNAs purificados foram clivados com as enzimas *Bam*HI, *Eco*RI, *Hind*III e *Pst*I. Os sistemas de digestão foram montados conforme as instruções

do fabricante, que consistiram da clivagem de 0,7-1,0 µg de DNA pela enzima de restrição de interesse com seu respectivo tampão 10x e incubação a 37°C / 16 h. Para a análise do DNA clivado, foi preparado gel de agarose 1% com tampão de corrida TAE 1x (TAE 50x – Tris-base 2M, ácido acético 2M e EDTA 0,05M, pH 8,0) e brometo de etídio (0,5 µg / mL). As amostras foram preparadas com tampão de amostra 1X ( tampão 5X: 0,25% de azul de bromofenol; 0,25% de xileno cianol; 15% de ficol), aplicadas no gel e submetidas a uma eletroforese com voltagem de 40V, 16 h. Após a corrida, os fragmentos de DNA foram visualizados em luz ultravioleta e fotografados em sistema de fotodocumentação (*Eagle Eye II* - Stratagene).

#### **4.6. Amplificação de DNA por reação em cadeia da polimerase - *Polymerase Chain Reaction* (PCR)**

A amplificação do DNA de PsinSNPV foi conduzida para detecção e sequenciamento de quatro genes conservados entre os baculovirus e do gene *DNA photolyase* (*phr*).

Os oligonucleotídeos degenerados para amplificação dos genes *lef-8*, *lef-9*, *polh/gran* e *pif-2* foram desenhados sobre regiões intragênicas altamente conservadas do genoma de AcMNPV e os oligonucleotídeos degenerados específicos ao gene *phr* foram desenhados com base em duas regiões altamente conservadas de CPD-fotolases de classe II (região protéica QVDAHN e YWAKKI). Os oligonucleotídeos degenerados para amplificação dos quatro genes conservados foram propostos em estudos anteriores de identificação molecular de baculovirus (Herniou *et al.*, 2004; Jehle *et al.*, 2006b; Lange *et al.*, 2004) e o oligonucleotídeo específico ao gene *phr* foi proposto

por Xu e colaboradores (2008) em estudos de identificação e análise desse gene na família *Baculoviridae*. Os DNAs purificados dos sete isolados de PsinSNPV foram utilizados como molde nas reações de PCR (volume final de 25  $\mu$ L) que consistiu de: 2% de tampão especial IVB 5x (composição indisponível – presença de cristais de  $MgCl_2$ ) (Phoneutria, Brasil), 100  $\mu$ M de cada dNTP, 0,2  $\mu$ M oligonucleotídeo senso, 0,2  $\mu$ M oligonucleotídeo anti-senso, 1 U (unit) de Taq DNA polimerase (Phoneutria, Brasil) e 40 ng de DNA. No programa utilizado para amplificação, as reações iniciaram com temperatura de 95  $^{\circ}C$  / 3 min passando por 36 ciclos com temperatura de desnaturação de 95  $^{\circ}C$  / 30 s, temperatura de anelamento do oligonucleotídeo / 1 min e temperatura de extensão de 72  $^{\circ}C$  / 30 s. A extensão final foi de 72  $^{\circ}C$  / 10 min. As sequências e a temperatura de anelamento específica para cada par de oligonucleotídeos constam na Tabela 5. Para análise do produto de PCR, amostras foram submetidas à eletroforese em gel de agarose 1%, sob voltagem de 60V por 3 h, em tampão de corrida TAE 1x e brometo de etídio (0,5 $\mu$ g / mL). Após a corrida, os produtos de PCR foram visualizados em luz ultravioleta e fotografados (*Eagle Eye II* - Stratagene). Foram realizadas pelo menos três reações de amplificação a partir de diferentes extrações de DNA para cada par de oligonucleotídeos amplificado por PCR.



**Tabela 5. Oligonucleotídeos degenerados utilizados no PCR**

Gene	Sequência do primer *	T (°C) anelamento	Banda (pb)**	Referências
<i>lef-8</i>	prL8F2 (senso): 5'-CTAAAAGCAGCCCGACTNNNACNRCNCAPCAAYCC-3'			Herniou <i>et al.</i> , 2004

\*B = C, G ou T; D = A, G ou T; H = A, C ou T; I = Inosina; N = C, A, T ou G; R = A ou G; S = C or G; Y = C ou T.

\*\*Tamanho de fragmento esperado para cada gene

#### 4.7. Purificação dos produtos de PCR

Os produtos de PCR dos genes *lef-8*, *lef-9*, *pif-2*, *phr* e *polh* foram purificados para a eliminação do excesso de oligonucleotídeos e dNTP's ou de possíveis fragmentos inespecíficos presentes na reação de PCR. Produtos obtidos das ampliações por PCR dos genes *lef-8*, *lef-9* e *phr* foram purificados com *ExoSAP-IT PCR Clean-up Kit* (GE Healthcare), de acordo com o fabricante com algumas adaptações. Foi adicionado aos produtos de PCR, 2 µL (para cada 5 µL da reação) do mix *ExoSAP-IT* que contém as enzimas hidrolíticas: exonuclease I e fosfatase alcalina de camarão (*shrimp alkaline phosphatase* - SAP). A atividade enzimática ocorreu a 37 °C / 45 min, seguida de incubação a 80 °C / 30 min para inativação das enzimas. Devido a presença de fragmentos inespecíficos presentes nas reações de PCR dos genes *pif-2* e *polh*, os produtos das ampliações desses genes foram purificados com o Kit *GFX™ PCR DNA and Gel Band Purification* (GE Healthcare) de acordo com as

instruções do fabricante descritas a seguir. Para eluição da banda, amostras foram aplicadas em gel de agarose de baixo ponto de fusão 1% (*low melting point*) e submetidas a uma eletroforese com voltagem de 60 V por 3 h. O gel foi corado com brometo de etídio (0,5µg / mL) e exposto a luz UV, em um menor intervalo de tempo possível. Para cada 10 mg de banda eluída de tamanho correspondente ao gene, foram acrescentados 10 µL de tampão de captura tipo 2 mantidos a 60 °C / 3 min. O produto foi então adicionado à coluna *GFX MicroSpin* permanecendo em temperatura ambiente por 5 min. A coluna montada com a amostra foi centrifugada a 12.000 rpm / 30 s (Eppendorf 5410 – rotor fixo) e, após adição de 500 µL de tampão de lavagem tipo 1, foi mantida a temperatura ambiente por mais 5 min. Em seguida, a coluna foi centrifugada a 12.000 rpm / 30 s (Eppendorf 5410 – rotor fixo), adicionados 30 µL de tampão de eluição tipo 6 e novamente centrifugada a 12.000 rpm / 5 min (Eppendorf 5410 – rotor fixo). O DNA purificado foi armazenado a 4 °C.

#### **4.8. Sequenciamento dos produtos de PCR**

Os produtos de PCR dos genes *lef-8*, *lef-9*, *pif-2*, *phr* e *polh* purificados foram sequenciados para determinação de suas sequências e detecção de polimorfismos de base única (*single nucleotide polymorphism* – SNPs). As reações de sequenciamento foram feitas em ambas as direções, senso (*forward*) e anti-senso (*reverse*), utilizando oligonucleotídeos específicos para cada gene (Tabela 5) e o kit *BigDye Terminator Cycle Sequencing Standard Version 3.1* (*Applied Biosystems*) de acordo com o fabricante. Como estratégia para o sequenciamento, a temperatura de anelamento foi elevada a 60°C, evitando que os produtos amplificados (*amplicons*) formassem estruturas

secundárias. Esses produtos foram purificados de acordo com o protocolo de precipitação por etanol / EDTA proposto pelo fabricante (*Applied Biosystems*). Foram adicionados à reação de seqüenciamento (volume final = 10 µL) 2,5 µL de EDTA 125 mM e 30 µL de etanol 100%. Para misturar a amostra, a placa de 96 poços foi invertida 4 a 5 vezes e mantida a temperatura ambiente por 20 min. A placa foi centrifugada a 2.250 x *g* / 30 min a 4 °C e o sobrenadante foi descartado. As amostras foram lavadas com etanol 70% e centrifugadas a 1.800 x *g* / 20 min a 4 °C. O sobrenadante foi novamente descartado e, para secar o precipitado, a placa foi incubada a 50 °C / 30 min. O precipitado foi ressuspensão em 9 µL de formamida Hi-Di (*Applied Biosystems*) e a amostra foi denaturada a 95 °C / 5 min. A corrida eletroforética foi realizada em sequenciador automático *ABI Prism 3730 DNA Analyser* (*Applied Biosystems*) e as seqüências foram analisadas com software *Sequencing Analysis 5.3.1* utilizando *KB™ Basecaller*.

#### **4.9. Análise das seqüências**

Os eletroferogramas resultantes do seqüenciamento automático foram analisados e as seqüências contíguas foram montadas com *ChromasPro v1.5*. Pesquisas para a verificação da autenticidade de todas as seqüências nucleotídicas foram realizadas usando o programa BLAST (Apêndice I). Para cada isolado de PsinSNPV, as seqüências deduzidas de aminoácidos para os cinco genes analisados (*lef-8*, *lef-9*, *pif-2*, *phr* e *polh*) foram obtidas utilizando o portal de bioinformática *Expasy*. Das seis matrizes de leitura obtidas pela ferramenta computacional *Translate tool* do portal *Expasy*, a determinação da seqüência deduzida de aminoácidos para cada gene de cada isolado de

PsinSNPV foi obtida pela busca de sequências similares no banco de dados *GenBank* utilizando-se a ferramenta BLAST. Todas as sequências foram depositadas no banco de dados *Genbank* do NCBI utilizando *Sequin*, ferramenta de submissão de sequências.

#### 4.10. Análise filogenética

As sequências dos genes *lef-8*, *lef-9*, *pif-2*, *phr* e *polh* foram analisadas pelos métodos da Máxima Parcimônia (MP), ponderada (reponderação progressiva - *reweighted*) e não ponderada, e da Inferência Bayesiana (IB) para o estabelecimento das relações filogenética entre os isolados de PsinSNPV e para a realização da reconstrução da filogenia da família *Baculoviridae* incluindo esse vírus. As relações filogenéticas entre os sete isolados de PsinSNPV (IA a IG) foram inferidas da concatenação da sequência nucleotídica parcial dos genes *lef-8*, *lef-9*, *pif-2*, *phr* e *polh* pelos métodos da máxima parcimônia e inferência Bayesiana. A relação filogenética do vírus PsinSNPV-IE com outros 57 baculovirus com genoma seqüenciado (Tabela 6) foi inferida com base no alinhamento dos genes *lef-8*, *lef-9*, *pif-2* e *polh/gran*. Foram construídas árvores baseadas nos quatro genes individuais, através da sequência nucleotídica e peptídica, utilizando análise MP. O mesmo foi realizado para a construção de árvores, agora incluindo também a inferência filogenética Bayesiana, utilizando duas diferentes concatenações: uma formada pelos quatro genes analisados e outra formada por três genes com exceção de *polh/gran* pelo fato deste gene não estar presente no vírus CuniNPV, representante do gênero *Deltabaculovirus*. A edição das árvores foi realizada com o auxílio dos programas *FigTree v1.3.1* e *Adobe Illustrator CS5 v15.0.0*.

**Tabela 6. Baculovirus utilizados na construção das árvores filogenéticas**

<b>NPV – Lepidoptera Grupo I</b>	<b>Abreviação</b>	<b>(kb)</b>	<b>Referência</b>	<b>GenBank*</b>
<i>Antheraea pernyi</i> MNPV-L2	AnpeMNPV-L2	126,246	Fan <i>et al.</i> , 2007	EF207986
<i>Antheraea pernyi</i> MNPV	AnpeMNPV	126,629	Nie <i>et al.</i> , 2007	DQ486030
<i>Anticarsia gemmatalis</i> MNPV	AgMNPV	132,239	Oliveira <i>et al.</i> , 2006	DQ813662
<i>Autographa californica</i> MNPV-C6	AcMNPV-C6	133,894	Ayres <i>et al.</i> , 1994	L22858
<i>Bombyx mori</i> NPV – T3	BmNPV – T3	128,413	Gomi <i>et al.</i> , 1999	L33180
<i>Bombyx mandarina</i> NPV-S1	BomaNPV – S1	126,770	Xu <i>et al.</i> , 2010	FJ882854
<i>Choristoneura fumiferana</i> MNPV	CfMNPV	129,609	De Jong <i>et al.</i> , 2005	AF512031
<i>C. fumiferana</i> DEFMNPV	CfDEFMNPV	131,160	Lauzon <i>et al.</i> , 2005	AY327402
<i>Epiphyas postvittana</i> NPV	EppoNPV	118,584	Hyink <i>et al.</i> , 2002	AY043265
<i>Hyphantria cunea</i> NPV	HycuNPV	132,959	Ikeda <i>et al.</i> , 2006	AP009046
<i>Maruca vitrata</i> NPV	MaviNPV	111,953	Chen <i>et al.</i> , 2008	EF125867
<i>Orgyia pseudotsugata</i> MNPV	OpMNPV	131,990	Ahrens <i>et al.</i> , 1997	U75930
<i>Plutella xylostella</i> NPV-CL3	PlxyNPV-CL3	134,417	Harrison e Lynn, 2007	DQ457003
<i>Rachiplusia ou</i> NPV	RoNPV	131,526	Harrison e Bonning, 2003	AY145471
<b>NPV – Lepidoptera Grupo II</b>				
<i>Adoxophyes honmai</i> NPV	AdhoNPV	113,220	Nakai <i>et al.</i> , 2003	AP006270
<i>Adoxophyes orana</i> NPV	AdorNPV	111,724	Hilton e Winstanley, 2008	EU591746
<i>Agrotis ipsilon</i> NPV- Illions	AgipMNPV	155,122	Harrison, não publ.	EU839994
<i>Agrotis segetum</i> NPV	AgseNPV	147,544	Jakubowska <i>et al.</i> , 2006	DQ123841
<i>Chrysodeixis chalcites</i> NPV	ChchNPV	149,622	Van Oers <i>et al.</i> , 2005	AY864330
<i>Clanis bilineata</i> NPV-DZ1	ClbiNPV-DZ1	135,545	Zhu <i>et al.</i> , 2009	DQ504428
<i>Ecotropis obliqua</i> NPV-A1	EcobNPV-A1	131,204	Ma <i>et al.</i> , 2007	DQ837165
<i>Euproctis pseudoconspersa</i> NPV- Hangzhou	EupsNPV-Hangzhou	141,291	Tang <i>et al.</i> , 2009	FJ227128
<i>Helicoverpa armigera</i> SNPV-G4	HearSNPV-G4	131,403	Zhang <i>et al.</i> , 2005	AF271059
<i>Helicoverpa armigera</i> SNPV	HearSNPV	130,759	Zhang <i>et al.</i> , 2005	AF303045
<i>Helicoverpa armigera</i> MNPV	HearMNPV	154,196	Tang <i>et al.</i> , não publ.	EU730893
<i>Helicoverpa armigera</i> NPV-NNg1	HearNPV-NNg1	132,425	Ogembo <i>et al.</i> , 2007	AP010907
<i>Helicoverpa zea</i> SNPV	HzSNPV	130,869	Chen <i>et al.</i> , 2002	AF334030
<i>Leucania separata</i> NPV-AH1	LeseNPV-AH1	168,041	Xiau e Qi, 2007	AY394490
<i>Lymantria dispar</i> MNPV	LdMNPV	161,046	Kuzio <i>et al.</i> , 1999	AF081810
<i>Mamestra configurata</i> NPV-A90/2	MacoNPV-A90/2	155,060	Li <i>et al.</i> , 1997	U59461
<i>Mamestra configurata</i> NPV-A90/4	MacoNPV-A90/4	153,656	Li <i>et al.</i> , 2005	AF539999
<i>Mamestra configurata</i> NPV-B	MacoNPV-B	158,482	Li <i>et al.</i> , 2002	AY126275
<i>Orgyia leucostigma</i> NPV-CFS77	OrleNPV-CFS77	156,179	Eveleigh <i>et al.</i> , não publ.	EU309041
<i>Spodoptera exigua</i> MNPV	SeMNPV	135,611	Ijkel <i>et al.</i> , 1999	AF169823
<i>Spodoptera frugiperda</i> NPV-19	SfMNPV-19	135,611	Wolff <i>et al.</i> , 2008	EU258200
<i>Spodoptera frugiperda</i> NPV-3AP2	SfMNPV-3AP2	131,331	Harrison <i>et al.</i> , 2008	EF035042
<i>Spodoptera litura</i> MNPV-G2	SpltMNPV-G2	139,342	Pang <i>et al.</i> , 2001	AF325155
<i>Spodoptera litura</i> NPV II	SpltNPV-II	148,634	Li <i>et al.</i> , não publ.	EU780426
<i>Trichoplusia ni</i> SNPV	TnSNPV	134,394	Willis <i>et al.</i> , 2005	DQ017380
<i>Lymantria xyliina</i> MNPV	LyxyMNPV	156,344	Nai <i>et al.</i> , 2010	GQ202541
<i>Apocheima cinerarium</i> NPV	ApciNPV	123,876	Zhang <i>et al.</i> , não publ.	FJ914221
<b>GV - Lepidoptera</b>				
<i>Adoxophyes orana</i> GV	AdorGV	99,657	WormLeaton <i>et al.</i> , 2003	AF547984
<i>Agrotis segetum</i> GV	AgseGV	131,680	Ai <i>et al.</i> , não publ.	AY522332
<i>Choristoneura occidentalis</i> GV	ChocGV	104,710	Escasa <i>et al.</i> , 2006	DQ333351
<i>Cryptophlebia leucotreta</i> GV-CV3	CrleGV	110,907	Lange e Jehle, 2003	AY229987
<i>Cydia pomonella</i> GV	CpGV	123,500	Luque <i>et al.</i> , 2001	U53466
<i>Helicoverpa armigera</i> GV	HearGV	169,794	Harrison e Poplam, 2008	EU255577
<i>Photorimaea operculella</i> GV	PhorGV	119,217	Taha <i>et al.</i> , 2000	AF499596
<i>Pieris rapae</i> GV	PrGV	108,592	Zhang <i>et al.</i> , não publ.	GQ884143
<i>Plutella xylostella</i> GV	PlxyGV	100,999	Hashimoto <i>et al.</i> , 2000	AF270937
<i>Pseudaletia unipuncta</i> GV- Hawaii	PsunGV-Hawaii	176,676	Li <i>et al.</i> , não publ.	EU678671
<i>Spodoptera litura</i> GV-K1	SpltGV-K1	124,121	Wang <i>et al.</i> , 2008	DQ288858
<i>Xestia c-nigrum</i> GV	XecnGV	178,733	Hayakawa <i>et al.</i> , 1999	AF162221
<b>NPV- Hymenoptera</b>				
<i>Neodiprion abietis</i> NPV	NeabNPV	84,264	Duffy <i>et al.</i> , 2006	DQ317692
<i>Neodiprion lecontei</i> NPV	NeleNPV	84,462	Lauzon <i>et al.</i> , 2004	AY349019
<i>Neodiprion sertifer</i> NPV	NeseNPV	81,755	Garcia-Maruniak <i>et al.</i> , 2004	AY430810
<b>NPV – Diptera</b>				
<i>Culex nigripalpus</i> NPV	CuniNPV	108,252	Afonso <i>et al.</i> , 2001	AF403738

\*Número de acesso no GenBank.

#### 4.10.1. Alinhamento das sequências utilizadas na análise filogenética

As sequências foram alinhadas individualmente utilizando o programa de *MUSCLE* v3.5 e subsequentemente concatenadas utilizando o programa *BioEdit* v7.0.9. Os *gaps* presentes no alinhamento original foram preservados e codificados separadamente usando o programa *FastGap* v1.0.8 por meio do método simples de codificação de *indels* (inserção / deleção) (Simmons e Ochoterena, 2000). O *FastGap* também foi utilizado para transformação do arquivo para formato nexus utilizado pelos programas PAUP\* v4.0b10 e MrBayes v3.1.2.

#### 4.10.2. Análise da Máxima Parcimônia

Na análise da Máxima Parcimônia o número de árvores possíveis aumenta rapidamente conforme é adicionado mais organismos a serem analisados (Felsenstein, 1978). Dessa forma, várias alternativas foram utilizadas com o intuito de buscar a árvore mais parcimoniosa sem a análise de todas as possibilidades evitando a estagnação da busca em um ponto local.

A análise da Máxima Parcimônia (MP) foi realizada utilizando o programa PAUP\* v4.0b10. A busca heurística consistiu de 1.000 repetições utilizando a adição randômica dos táxons (*Stepwise Addition*). A troca de ramos foi realizada pelo algoritmo TBR (*Tree bisection and reconnection*) salvando as primeiras 5.000 árvores mais parcimoniosas. Os comandos utilizados no PAUP para os métodos citados são: *hsearch*, *start=stepwise*, *addseq=random*, *nreps=1000*, *savereps=yes*, *swap=tbr* e *maxtrees=5000*. Os *indels* foram excluídos da análise. As árvores MP foram construídas pela análise de parcimônia não ponderada e pela reponderação progressiva e foi calculado o

número de caracteres informativos para parcimônia, o índice de consistência (IC), índice de homoplasia (IH), índice de retenção (IR) e índice de consistência rescalonado (RC). O valor de suporte para cada ramo foi estimado através da análise de *bootstrap* para 1.000 replicatas e o limiar de significância para os clados foram considerados significativos para valores acima de 95%, moderados para valores entre 70 - 94% e fracos para valores entre 51 - 69% (Hillis e Bull, 1993; Li, 1997; Schneider, 2007).

#### 4.10.3. Análise Bayesiana

A Inferência Bayesiana foi realizada utilizando o programa MrBayes v3.1.2. O programa jModelTest v0.1.1 foi utilizado para a determinação dos modelos de substituição nucleotídica mais apropriados ao alinhamento das sequências nucleotídicas individuais dos genes *lef-8*, *lef-9*, *pif-2*, *phr* e *polh* (Tabelas 7 e 8). Os modelos de evolução mais apropriados às sequências peptídicas também foram determinados utilizando ProtTest (Tabela 9). Os programas compararam 56 modelos de evolução alternativos utilizando para adoção do modelo o critério de informação Akaike (*Akaike Information Content*)  $AIC = -2\ln L + 2K$  (Akaike, 1974) e informação Bayesiana (*Bayesian Information Content*)  $BIC = 2\ln L + K\log N$  (Schwartz, 1978). Nas fórmulas, "L" corresponde ao valor de escore de verossimilhança, "K" o número de parâmetros estimáveis e "N" o tamanho da amostra. O algoritmo Monte Carlo via Cadeia de Markov (MCMC) foi iniciado a partir de uma árvore aleatória e seis Cadeias de Markov processadas para 1.000.000 a 5.000.000 gerações, dependendo da análise, com amostras coletadas 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. Os *indels* foram excluídos da análise.

A maioria dos modelos de evolução escolhidos pelo jModelTest (AIC) para as sequências nucleotídicas dos cinco genes, tanto na análise das relações filogenéticas entre os isolados quanto na filogenia dos baculovirus incluindo o vírus PsinSNPV (Tabelas 7 e 8), apresentou o mesmo número de parâmetros (nst=6) que o modelo GTR (*General Time Reversible*) para explicar as substituições nucleotídicas que ocorreram nas sequências analisadas. O modelo LG (Le e Gascuel, 2008), selecionado pelo programa ProtTest (Tabela 9), é uma variação do modelo WAG (*Whelan And Goldman*) (Whelan e Goldman, 2001), modelo empírico da evolução de proteínas globulares escolhido pelo programa MrBayes quando foi realizada a Inferência Filogenética Bayesiana das sequências peptídicas.

**Tabela 7. Modelos evolutivos selecionados para as sequências nucleotídicas de cinco genes de isolados de PsinSNPV\***

Gene	Modelo de evolução (nº critérios = 7)	
	Informação Akaike (AIC)	Informação Bayesiana (BIC)
<i>lef-8</i>	HKY (nst=2; rates=equal)	HKY (nst=2; rates=equal)
<i>lef-9</i>	TIM3ef (nst=6; rates=equal)	K80 (nst=2; rates=equal)
<i>pif-2</i>	TrNef + G (nst=6; rates =gamma)	K80 + G (nst=2; rates=gamma)
<i>phr</i>	TPM2uf (nst=6; rates=equal)	HKY (nst=2; rates=equal)
<i>polh</i>	TIM3ef (nst=6; rates=equal)	TPM3 (nst=6; rates=equal)

\*Busca a partir do alinhamento de sequências de isolados de PsinSNPV (IA a IG) e ChchNPV.



**Tabela 8. Modelos evolutivos selecionados para as sequências nucleotídicas de quatro genes de baculovirus \***

Gene	Modelo de evolução (nº critérios = 7)	
	Informação Akaike (AIC)	Informação Bayesiana (BIC)
<i>lef-8</i>	TIM1+I+G (nst=6; rates=invgamma)	TrN+G (nst=6; rates=gamma)
<i>lef-9</i>	TrN+I+G (nst=6; rates=invgamma)	TrN+I+G (nst=6; rates=invgamma)
<i>pif-2</i>	GTR+I+G (nst=6; rates=invgamma)	TrN+I+G (nst=6; rates=invgamma)
<i>polh/gran</i>	TIM1ef+I+G (nst=6; rates=invgamma)	TIM1ef+I+G (nst=6; rates=invgamma)

\*Busca a partir do alinhamento de sequências de 57 baculovirus seqüenciados (Tabela 6) e PsinSNPV-IE

**Tabela 9. Modelos evolutivos selecionados para quatro sequências peptídicas de baculovirus\***

Sequência peptídica	Modelo de evolução	
	Informação Akaike (AIC)	Informação Bayesiana (BIC)
LEF-8	LG+I+G	LG+I+G+F
LEF-9	-	LG+G
PIF-2	LG+I+G	-
POLH/GRAN	LG+G	LG+G

\*Busca a partir do alinhamento de sequências de 57 baculovirus seqüenciados (Tabela 6) e PsinSNPV-IE

#### 4.10.4. Medidas de apoio estatístico: ILD - teste da incongruência da diferença de comprimento e PABA - teste da adição de partição alterando o *bootstrap*

A congruência filogenética entre as partições dos dados (genes *lef-8*, *lef-9*, *pif-2* e *polh/gran*) foi analisada pelo teste ILD (*Incongruence length difference test*) (Farris *et al.*, 1995a; 1995b) utilizando o método da Máxima Parcimônia (PAUP\* v4.0b10). A busca heurística consistiu de 100 repetições, adição randômica dos táxons e troca dos ramos pelo algoritmo TBR salvando as

primeiras 100 árvores mais parcimoniosas. Neste teste, foram avaliadas 10 combinações entre 4 genes: *lef-8*, *polh/gran*, *lef-9*, *pif-2* e *gaps*; *lef-8*, *polh/gran*, *lef-9* e *pif-2*; *lef-8*, *polh/gran* e *lef-9*; *polh/gran*, *lef-9* e *pif2*; *lef-8* e *polh/gran*; *lef-9* e *pif-2*; *polh/gran* e *lef-9*; *lef-8* e *lef-9*; *lef-8* e *pif-2*; *polh/gran* e *pif-2*. A probabilidade de significância (valor *P*) da incongruência foi obtida pela análise da diferença do número de passos das árvores filogenéticas resultantes das combinações avaliadas.

O conflito entre as partições de dados também foi avaliado pelo teste PABA (*Partition Addition Bootstrap Alteration*). Este teste foi realizado utilizando o método da Máxima Parcimônia (PAUP\* 4.0b10) por meio da busca heurística com 1.000 repetições utilizando a adição randômica dos táxons. A troca de ramos foi realizada pelo algoritmo TBR salvando as primeiras 5.000 árvores mais parcimoniosas. O valor de suporte dos ramos foi estimado através da análise de *bootstrap* para 1.000 replicatas. A análise MP foi baseada nas sequências nucleotídicas individuais dos genes *lef-8*, *polh/gran*, *lef-9* e *pif-2* e na concatenação de *lef-8* com os outros genes adicionados um a um em ordem decrescente do valor de caracteres informativos. Os *gaps* foram codificados e inseridos nas análises. O valor de *bootstrap* para cada análise foi sobreposto à árvore filogenética de referência, árvore mais parcimoniosa obtida da concatenação das sequências nucleotídicas dos quatro genes (Apêndice II). Na árvore de referência, foram selecionados 20 ramos para o cálculo do *bootstrap* (BS) e o valor da alteração do suporte dos ramos ( $\delta$ ) à medida que cada partição dos dados fosse adicionada.

#### 4.11. Agrupamento dos isolados de PsinSNPV por UPGMA

A variação genética entre populações do vírus PsinSNPV foi avaliada pela análise de agrupamento hierárquico com base no polimorfismo dos fragmentos de restrição (RFLP) dos DNAs clivados com as enzimas *Bam*HI, *Eco*RI, *Hind*III e *Pst*I. A matriz binária gerada pela visualização da presença (1) e ausência (0) de fragmentos de restrição do DNA de PsinSNPV (IA a IG) foi utilizada para a construção da matriz de similaridade empregando o cálculo do coeficiente de similaridade de Sorensen-Dice (Dice, 1945; Sorensen, 1948) e Jaccard (Jaccard, 1901). Os dendrogramas baseados nas matrizes de similaridade foram construídos pelo método da média aritmética não ponderada (*Unweighted Pair Group Method Analysis* - UPGMA) e do agrupamento sequencial, aglomerativo, hierárquico e exclusivo (*Sequencial, Agglomerative, Hierarquic, Nonoverlapping, Clustering Methods* - SAHN). Os dendrogramas, os valores de correlação cofenética e teste de Mantel (Sokal e Rohlf, 1962) foram gerados no programa computacional estatístico NTSYS v2.21. O intervalo de confiança do dendrograma foi estimado utilizando o programa BOOD v3.0 com 10000 replicatas de *bootstrap*. Esta análise inclui todas as bandas submolares observadas nos perfis de restrição dos isolados.

## 5. Resultados

### 5.1. Identificação dos genes *lef-8*, *lef-9*, *phr*, *pif-2* e *polh* no genoma dos isolados de PsinSNPV (IA a IG)

A amplificação por PCR revelou a presença dos genes *lef-8*, *lef-9*, *phr*, *pif-2* e *polh* no genoma dos sete isolados de PsinSNPV (IA a IG). Os produtos obtidos pela amplificação foram detectados por eletroforese em gel de agarose apresentando bandas correspondentes aos tamanhos esperados de 513, 270, 357, 780 e 510 pb para os genes *lef-8*, *lef-9*, *pif-2*, *phr* e *polh*, respectivamente (Figura 7). Estes resultados estão de acordo com o tamanho de banda correspondente aos produtos de PCR desses genes em outros baculovirus descritos na literatura (Tabela 5).

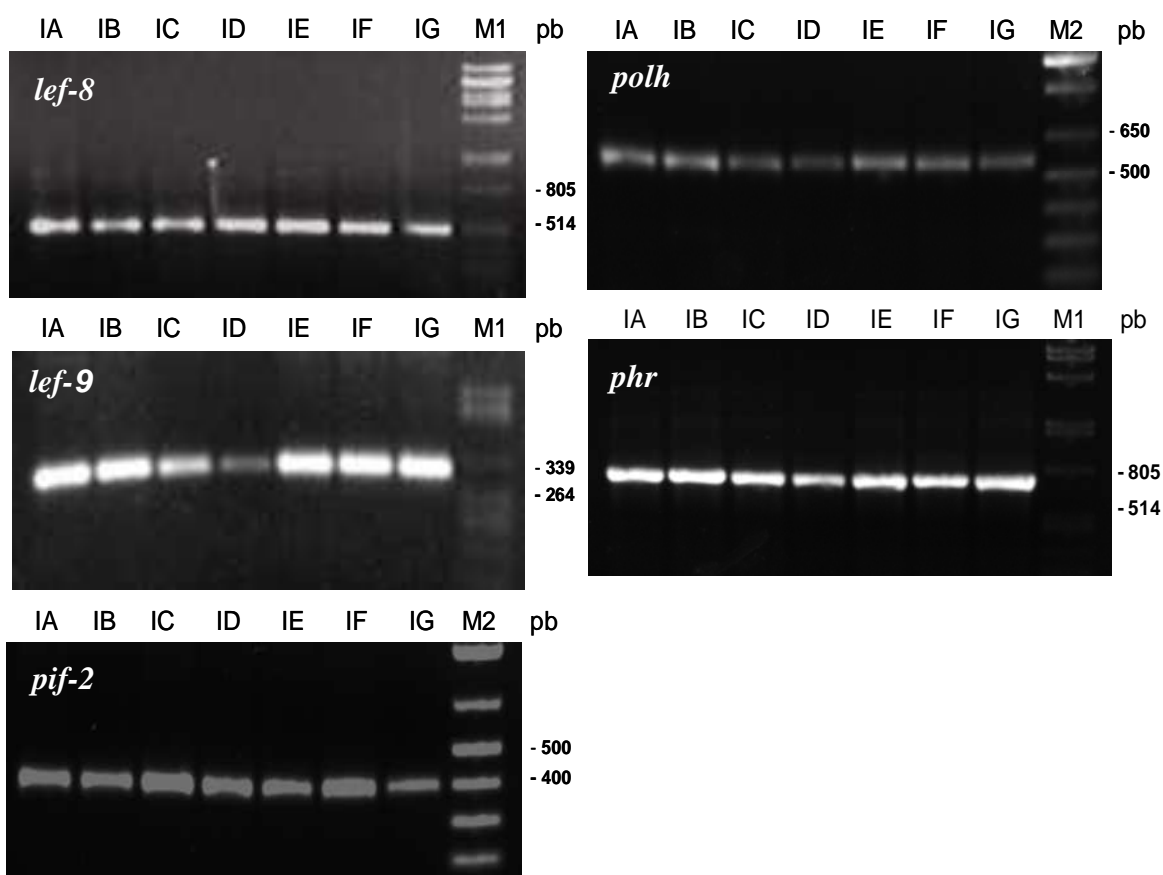


Figura 7. Produtos de PCR dos genes *lef-8*, *lef-9*, *phr*, *pif-2* e *polh* de isolados de PsinSNPV (IA a IG). Eletroforese em gel de agarose 0,8%. M1 (marcador): DNA  $\lambda$ /PstI. M2 (marcador): 1 kb plus DNA Ladder (Invitrogen).

## 5.2. Análise das sequências nucleotídicas parciais dos genes *lef-8*, *lef-9*, *phr*, *pif-2* e *polh* de PsinSNPV (IA a IG)

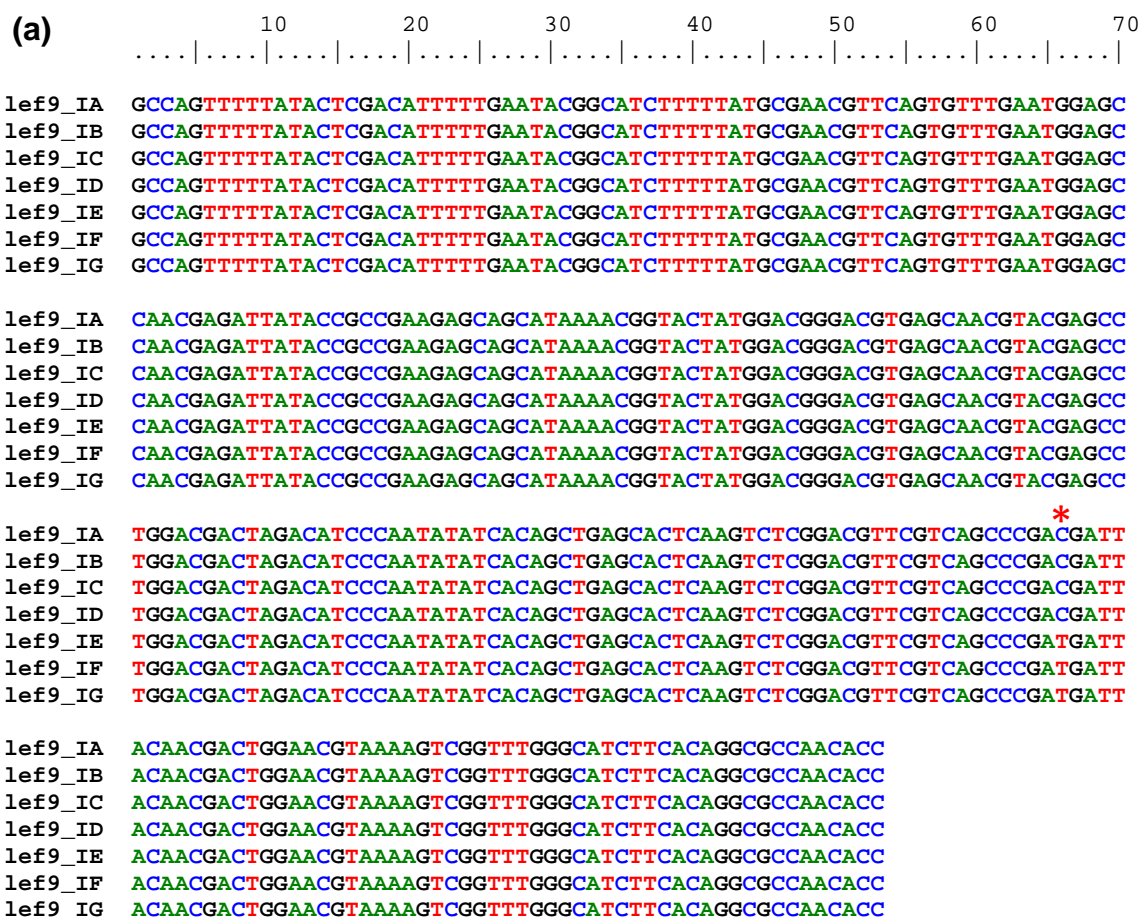
Os produtos de PCR correspondentes aos genes *lef-8*, *lef-9*, *phr*, *pif-2* e *polh* dos isolados de PsinSNPV foram sequenciados e suas sequências depositadas no *GenBank* (Tabela 10).

**Tabela 10. Sequências de cinco genes de isolados de PsinSNPV depositadas no *GenBank***

PsinSNPV Isolados	Nº de acesso no <i>GenBank</i>				
	<i>lef-8</i>	<i>lef-9</i>	<i>pif-2</i>	<i>phr</i>	<i>polh</i>
<b>IA</b>	JQ260765	JQ260772	JQ260786	JQ260779	JQ260793
<b>IB</b>	JQ260766	JQ260773	JQ260787	JQ260780	JQ260794
<b>IC</b>	JQ260767	JQ260774	JQ260788	JQ260781	JQ260794
<b>ID</b>	JQ260768	JQ260775	JQ260789	JQ260782	JQ260796
<b>IE</b>	JQ260769	JQ260776	JQ260790	JQ260783	JQ260797
<b>IF</b>	JQ260770	JQ260777	JQ260791	JQ260784	JQ260798
<b>IG</b>	JQ260771	JQ260778	JQ260792	JQ260785	JQ260799

O alinhamento individual das sequências nucleotídicas parciais do gene *lef-8* dos isolados de PsinSNPV apresentaram tamanho de 399 pb com média de 63,77% da composição nucleotídica composta por adeninas e timinas (%AT) (Figura 8a). As sequências nucleotídicas parciais do gene *lef-9* apresentaram tamanho de 263 pb e %AT média de 50,43% (Figura 9a). No alinhamento, os isolados IE, IF e IG apresentaram homogeneidade para dois polimorfismos de base única (SNP) com transições de citosina para timina na posição 160 nucleotídeos (nt) para o gene *lef-8* e 206 nt para o gene *lef-9*. Nos dois genes, os polimorfismos não causaram diferenças nas sequências deduzidas de aminoácidos entre os isolados analisados (Figuras 8b e 9b).





## (b) Sequência deduzida de aminoácidos LEF-9

Lef-9 QFLYSTFLNTASFYANVQCLNGANEIIPPKSSIKRYGRDVSNVRAWTRHPNISQLSTQVSDVRQPDDY  
NDWNVKVGLGIFTGANT

Figura 9. Alinhamento das sequências nucleotídicas parciais do gene *lef-9* dos isolados de PsinSNPV (IA a IG). Os asteriscos revelam as posições em que foram detectados polimorfismos (a). Sequência deduzida de aminoácidos obtida a partir das sequências nucleotídicas parciais do gene *lef-9* dos isolados de PsinSNPV (IA a IG) (b).

As sequências nucleotídicas parciais do gene *polh* dos isolados de PsinSNPV apresentaram tamanho de 511 pb e %AT média de 49,54% (Figura 10a). No alinhamento, os isolados IE e IF apresentaram duas bases polimórficas com transições de guanina para adenina na posição 292 nt e transição inversa na posição 342 nt. Diferentemente dos outros isolados, as sequências deduzidas de aminoácidos dos isolados IE e IF apresentaram arginina (R) ao invés de lisina (K) na posição 114 aminoácidos (aa) (Figura

10b). Esses dois aminoácidos são carregados positivamente sendo ambos hidrofílicos.

(a)

```

      10      20      30      40      50      60      70
...|...|...|...|...|...|...|...|...|...|...|...|...|...|
polh_IA CCTGGGACCCGGAAAGAACCACAAAACTCACTTTGTTC AAGGAGATTTCGTAATGTTAAGCCCGATACCATG
polh_IB CCTGGGACCCGGAAAGAACCACAAAACTCACTTTGTTC AAGGAGATTTCGTAATGTTAAGCCCGATACCATG
polh_IC CCTGGGACCCGGAAAGAACCACAAAACTCACTTTGTTC AAGGAGATTTCGTAATGTTAAGCCCGATACCATG
polh_ID CCTGGGACCCGGAAAGAACCACAAAACTCACTTTGTTC AAGGAGATTTCGTAATGTTAAGCCCGATACCATG
polh_IE CCTGGGACCCGGAAAGAACCACAAAACTCACTTTGTTC AAGGAGATTTCGTAATGTTAAGCCCGATACCATG
polh_IF CCTGGGACCCGGAAAGAACCACAAAACTCACTTTGTTC AAGGAGATTTCGTAATGTTAAGCCCGATACCATG
polh_IG CCTGGGACCCGGAAAGAACCACAAAACTCACTTTGTTC AAGGAGATTTCGTAATGTTAAGCCCGATACCATG

polh_IA AAGCTCGTCGTTAACTGGAGCGGCAAAGAGTTTCTTAGGGAAACTTGGACTCGTTTCATGGAGGACAGCT
polh_IB AAGCTCGTCGTTAACTGGAGCGGCAAAGAGTTTCTTAGGGAAACTTGGACTCGTTTCATGGAGGACAGCT
polh_IC AAGCTCGTCGTTAACTGGAGCGGCAAAGAGTTTCTTAGGGAAACTTGGACTCGTTTCATGGAGGACAGCT
polh_ID AAGCTCGTCGTTAACTGGAGCGGCAAAGAGTTTCTTAGGGAAACTTGGACTCGTTTCATGGAGGACAGCT
polh_IE AAGCTCGTCGTTAACTGGAGCGGCAAAGAGTTTCTTAGGGAAACTTGGACTCGTTTCATGGAGGACAGCT
polh_IF AAGCTCGTCGTTAACTGGAGCGGCAAAGAGTTTCTTAGGGAAACTTGGACTCGTTTCATGGAGGACAGCT
polh_IG AAGCTCGTCGTTAACTGGAGCGGCAAAGAGTTTCTTAGGGAAACTTGGACTCGTTTCATGGAGGACAGCT

polh_IA TTCCCATCGTAAACGACCAAGAAATCATGGACGTGTTCCCTTGTGGTCAACATGCGCCCGACGAGACCTAA
polh_IB TTCCCATCGTAAACGACCAAGAAATCATGGACGTGTTCCCTTGTGGTCAACATGCGCCCGACGAGACCTAA
polh_IC TTCCCATCGTAAACGACCAAGAAATCATGGACGTGTTCCCTTGTGGTCAACATGCGCCCGACGAGACCTAA
polh_ID TTCCCATCGTAAACGACCAAGAAATCATGGACGTGTTCCCTTGTGGTCAACATGCGCCCGACGAGACCTAA
polh_IE TTCCCATCGTAAACGACCAAGAAATCATGGACGTGTTCCCTTGTGGTCAACATGCGCCCGACGAGACCTAA
polh_IF TTCCCATCGTAAACGACCAAGAAATCATGGACGTGTTCCCTTGTGGTCAACATGCGCCCGACGAGACCTAA
polh_IG TTCCCATCGTAAACGACCAAGAAATCATGGACGTGTTCCCTTGTGGTCAACATGCGCCCGACGAGACCTAA

polh_IA CAGATGCTTTAAATTCCTAGCTCAACACGCTCTGCGTTGCGACCCCGACTACGTTCCCTCATGAGGTGATC
polh_IB CAGATGCTTTAAATTCCTAGCTCAACACGCTCTGCGTTGCGACCCCGACTACGTTCCCTCATGAGGTGATC
polh_IC CAGATGCTTTAAATTCCTAGCTCAACACGCTCTGCGTTGCGACCCCGACTACGTTCCCTCATGAGGTGATC
polh_ID CAGATGCTTTAAATTCCTAGCTCAACACGCTCTGCGTTGCGACCCCGACTACGTTCCCTCATGAGGTGATC
polh_IE CAGATGCTTTAAATTCCTAGCTCAACACGCTCTGCGTTGCGACCCCGACTACGTTCCCTCATGAGGTGATC
polh_IF CAGATGCTTTAAATTCCTAGCTCAACACGCTCTGCGTTGCGACCCCGACTACGTTCCCTCATGAGGTGATC
polh_IG CAGATGCTTTAAATTCCTAGCTCAACACGCTCTGCGTTGCGACCCCGACTACGTTCCCTCATGAGGTGATC

polh_IA AGAATCGTAGAGCCTTCTGGGTGGGCAGCAACAACGAGTACAGAATCAGTCTGGCTAAAAAGGGCGGCC
polh_IB AGAATCGTAGAGCCTTCTGGGTGGGCAGCAACAACGAGTACAGAATCAGTCTGGCTAAAAAGGGCGGCC
polh_IC AGAATCGTAGAGCCTTCTGGGTGGGCAGCAACAACGAGTACAGAATCAGTCTGGCTAAAAAGGGCGGCC
polh_ID AGAATCGTAGAGCCTTCTGGGTGGGCAGCAACAACGAGTACAGAATCAGTCTGGCTAAAAAGGGCGGCC
polh_IE AGAATCGTAGAGCCTTCTGGGTGGGCAGCAACAACGAGTACAGAATCAGTCTGGCTAAAAAGGGCGGCC
polh_IF AGAATCGTAGAGCCTTCTGGGTGGGCAGCAACAACGAGTACAGAATCAGTCTGGCTAAAAAGGGCGGCC
polh_IG AGAATCGTAGAGCCTTCTGGGTGGGCAGCAACAACGAGTACAGAATCAGTCTGGCTAAAAAGGGCGGCC

polh_IA GCTGCCCCATCATGAACCTCCACTCTGAATACACCAACTCGTTTGAAGAGTTTATTGCTCGCGTGATCTG
polh_IB GCTGCCCCATCATGAACCTCCACTCTGAATACACCAACTCGTTTGAAGAGTTTATTGCTCGCGTGATCTG
polh_IC GCTGCCCCATCATGAACCTCCACTCTGAATACACCAACTCGTTTGAAGAGTTTATTGCTCGCGTGATCTG
polh_ID GCTGCCCCATCATGAACCTCCACTCTGAATACACCAACTCGTTTGAAGAGTTTATTGCTCGCGTGATCTG
polh_IE GCTGCCCCATCATGAACCTCCACTCTGAATACACCAACTCGTTTGAAGAGTTTATTGCTCGCGTGATCTG
polh_IF GCTGCCCCATCATGAACCTCCACTCTGAATACACCAACTCGTTTGAAGAGTTTATTGCTCGCGTGATCTG
polh_IG GCTGCCCCATCATGAACCTCCACTCTGAATACACCAACTCGTTTGAAGAGTTTATTGCTCGCGTGATCTG

polh_IA GGAGAACTTCTACAAGCCCATCGTTTACGTAGGAACCGATTCCGCCGAGGAAGAGGAGATTCTTCTTGAA
polh_IB GGAGAACTTCTACAAGCCCATCGTTTACGTAGGAACCGATTCCGCCGAGGAAGAGGAGATTCTTCTTGAA
polh_IC GGAGAACTTCTACAAGCCCATCGTTTACGTAGGAACCGATTCCGCCGAGGAAGAGGAGATTCTTCTTGAA
polh_ID GGAGAACTTCTACAAGCCCATCGTTTACGTAGGAACCGATTCCGCCGAGGAAGAGGAGATTCTTCTTGAA
polh_IE GGAGAACTTCTACAAGCCCATCGTTTACGTAGGAACCGATTCCGCCGAGGAAGAGGAGATTCTTCTTGAA
polh_IF GGAGAACTTCTACAAGCCCATCGTTTACGTAGGAACCGATTCCGCCGAGGAAGAGGAGATTCTTCTTGAA
polh_IG GGAGAACTTCTACAAGCCCATCGTTTACGTAGGAACCGATTCCGCCGAGGAAGAGGAGATTCTTCTTGAA

polh_IA GTGTCTTTAGTCTTCAAGATT
polh_IB GTGTCTTTAGTCTTCAAGATT
polh_IC GTGTCTTTAGTCTTCAAGATT
polh_ID GTGTCTTTAGTCTTCAAGATT
polh_IE GTGTCTTTAGTCTTCAAGATT
polh_IF GTGTCTTTAGTCTTCAAGATT
polh_IG GTGTCTTTAGTCTTCAAGATT

```



(b)

**Alinhamento das sequências deduzidas de aminoácidos POLH**

```

          10      20      30      40      50      60      70
    .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
polh_IA  LGPGKNQKLTTLFKEIRNVKPD*TMKLVVNW*SGKEFLRETWTRFMEDSFPIVNDQEIMDVFLV*VNMRP*TRPN
polh_IE  .....
polh_IA  RCFKFLAQHALRCDPDYVPHEVIRIVEPSWVGSNNEYRISLAKKGGGCPIMNLHSEY*TNSFE*EFIAR*VIW
polh_IE  .....R.....
polh_IA  ENFYKPIVYVGTDSAEEEEILLEVSLVFKI
polh_IE  .....

```

Figura 10. Alinhamento das sequências nucleotídicas parciais do gene *polh* dos isolados de PsinSNPV (IA a IG). Os asteriscos revelam as posições em que foram detectados polimorfismos (a). Alinhamento das sequências deduzidas de aminoácidos do gene *polh* indicando a variação dos aminoácidos entre os isolados (b).

As sequências nucleotídicas parciais do gene *phr* dos isolados de PsinSNPV apresentaram tamanho de 719 pb e %AT média de 66,40% (Figura 11a). No alinhamento, os isolados IE, IF e IG apresentaram duas bases polimórficas: uma transição de adenina para guanina na posição 135 nt e uma transversão de citosina para adenina na posição 212 nt. Nesse alinhamento pode ser observado sobreposição de bases em uma mesma posição nucleotídica. Os dois picos para uma mesma posição nucleotídica presentes no sentido senso e anti-senso do eletroferograma indicam a mistura de genótipos em um mesmo isolado (Figura 12). Na posição 219 nt os isolados IE e IG apresentaram uma transição de citosina para timina enquanto que o isolado IF apresenta duas bases para essa mesma posição, citosina e timina. O mesmo ocorre na posição 406 nt em que o isolado IE apresenta essa mesma transição e os isolados IF e IG apresentam citosina e timina. Esses polimorfismos causaram diferenças nas sequências deduzidas de aminoácidos entre os isolados apresentando na posição 71 aa tirosina (Y) no lugar de serina (S) para os isolados IE, IF e IG. Os dois aminoácidos são polares apresentando

comportamento hidrofílico. Na posição 136 aa, o isolado IE apresentou uma tirosina ao invés de histidina (H) enquanto que, nos isolados IF e IG, estes dois aminoácidos estão presentes (Figura 11b). A histidina é um aminoácido com carga positiva e a tirosina possui na cadeia lateral um anel aromático, ambos portanto hidrofílicos.

(a)

```

      . . . . | . . . . | . . . . | . . . . | . . . . | . . . . | . . . . | . . . . | . . . . | . . . . | . . . . | . . . . | . . . . |
      10      20      30      40      50      60      70

phr_IA  GATGCGCATAAATGTTGTGCCTTGTGGTTAGTGTCTAATCAACCTATATTTGATACGTTTAGAGATAAAA
phr_IB  GATGCGCATAAATGTTGTGCCTTGTGGTTAGTGTCTAATCAACCTATATTTGATACGTTTAGAGATAAAA
phr_IC  GATGCGCATAAATGTTGTGCCTTGTGGTTAGTGTCTAATCAACCTATATTTGATACGTTTAGAGATAAAA
phr_ID  GATGCGCATAAATGTTGTGCCTTGTGGTTAGTGTCTAATCAACCTATATTTGATACGTTTAGAGATAAAA
phr_IE  GATGCGCATAAATGTTGTGCCTTGTGGTTAGTGTCTAATCAACCTATATTTGATACGTTTAGAGATAAAA
phr_IF  GATGCGCATAAATGTTGTGCCTTGTGGTTAGTGTCTAATCAACCTATATTTGATACGTTTAGAGATAAAA
phr_IG  GATGCGCATAAATGTTGTGCCTTGTGGTTAGTGTCTAATCAACCTATATTTGATACGTTTAGAGATAAAA

phr_IA  TTAACGAACAACCTCCAAAACTATTTAAAAACCGTTTCCATTGTTGTAAAAATCCCCATAAAATCAGTTGT
phr_IB  TTAACGAACAACCTCCAAAACTATTTAAAAACCGTTTCCATTGTTGTAAAAATCCCCATAAAATCAGTTGT
phr_IC  TTAACGAACAACCTCCAAAACTATTTAAAAACCGTTTCCATTGTTGTAAAAATCCCCATAAAATCAGTTGT
phr_ID  TTAACGAACAACCTCCAAAACTATTTAAAAACCGTTTCCATTGTTGTAAAAATCCCCATAAAATCAGTTGT
phr_IE  TTAACGAACAACCTCCAAAACTATTTAAAAACCGTTTCCATTGTTGTAAAAATCCCCATAAAATCAGTTGT
phr_IF  TTAACGAACAACCTCCAAAACTATTTAAAAACCGTTTCCATTGTTGTAAAAATCCCCATAAAATCAGTTGT
phr_IG  TTAACGAACAACCTCCAAAACTATTTAAAAACCGTTTCCATTGTTGTAAAAATCCCCATAAAATCAGTTGT

phr_IA  GCCGTGAGAAATCTTCGACCAACACTTATATCGATTGGTACGCTTTAATAACGTACGAAACATTGATCAA
phr_IB  GCCGTGAGAAATCTTCGACCAACACTTATATCGATTGGTACGCTTTAATAACGTACGAAACATTGATCAA
phr_IC  GCCGTGAGAAATCTTCGACCAACACTTATATCGATTGGTACGCTTTAATAACGTACGAAACATTGATCAA
phr_ID  GCCGTGAGAAATCTTCGACCAACACTTATATCGATTGGTACGCTTTAATAACGTACGAAACATTGATCAA
phr_IE  GCCGTGAGAAATCTTCGACCAACACTTATATCGATTGGTACGCTTTAATAACGTACGAAACATTGATCAA
phr_IF  GCCGTGAGAAATCTTCGACCAACACTTATATCGATTGGTACGCTTTAATAACGTACGAAACATTGATCAA
phr_IG  GCCGTGAGAAATCTTCGACCAACACTTATATCGATTGGTACGCTTTAATAACGTACGAAACATTGATCAA

phr_IA  * TCTATAGACACGGTTAAGTGGACCAAAACAGGCTATAGACCGGCAATGCTGAGACTAGCCTCGTTTATCG
phr_IB  TCTATAGACACGGTTAAGTGGACCAAAACAGGCTATAGACCGGCAATGCTGAGACTAGCCTCGTTTATCG
phr_IC  TCTATAGACACGGTTAAGTGGACCAAAACAGGCTATAGACCGGCAATGCTGAGACTAGCCTCGTTTATCG
phr_ID  TCTATAGACACGGTTAAGTGGACCAAAACAGGCTATAGACCGGCAATGCTGAGACTAGCCTCGTTTATCG
phr_IE  TATATAGATACGGTTAAGTGGACCAAAACAGGCTATAGACCGGCAATGCTGAGACTAGCCTCGTTTATCG
phr_IF  TATATAGATACGGTTAAGTGGACCAAAACAGGCTATAGACCGGCAATGCTGAGACTAGCCTCGTTTATCG
phr_IG  TATATAGATACGGTTAAGTGGACCAAAACAGGCTATAGACCGGCAATGCTGAGACTAGCCTCGTTTATCG

phr_IA  AGTTTTCTCTTTATATGTATAATAATAGATTAGAAAATGCATTTTCTATCAATCAAAGTGATTTGTCGCC
phr_IB  AGTTTTCTCTTTATATGTATAATAATAGATTAGAAAATGCATTTTCTATCAATCAAAGTGATTTGTCGCC
phr_IC  AGTTTTCTCTTTATATGTATAATAATAGATTAGAAAATGCATTTTCTATCAATCAAAGTGATTTGTCGCC
phr_ID  AGTTTTCTCTTTATATGTATAATAATAGATTAGAAAATGCATTTTCTATCAATCAAAGTGATTTGTCGCC
phr_IE  AGTTTTCTCTTTATATGTATAATAATAGATTAGAAAATGCATTTTCTATCAATCAAAGTGATTTGTCGCC
phr_IF  AGTTTTCTCTTTATATGTATAATAATAGATTAGAAAATGCATTTTCTATCAATCAAAGTGATTTGTCGCC
phr_IG  AGTTTTCTCTTTATATGTATAATAATAGATTAGAAAATGCATTTTCTATCAATCAAAGTGATTTGTCGCC

phr_IA  ATTTTTTCATTTCGGATTCTTATCGACACAACGAGTTATCTATTATTTAACTTTTCATATATTTAAAAACC
phr_IB  ATTTTTTCATTTCGGATTCTTATCGACACAACGAGTTATCTATTATTTAACTTTTCATATATTTAAAAACC
phr_IC  ATTTTTTCATTTCGGATTCTTATCGACACAACGAGTTATCTATTATTTAACTTTTCATATATTTAAAAACC
phr_ID  ATTTTTTCATTTCGGATTCTTATCGACACAACGAGTTATCTATTATTTAACTTTTCATATATTTAAAAACC
phr_IE  ATTTTTTCATTTCGGATTCTTATCGACACAACGAGTTATCTATTATTTAACTTTTATATATTTAAAAACC
phr_IF  ATTTTTTCATTTCGGATTCTTATCGACACAACGAGTTATCTATTATTTAACTTTTATATATTTAAAAACC
phr_IG  ATTTTTTCATTTCGGATTCTTATCGACACAACGAGTTATCTATTATTTAACTTTTATATATTTAAAAACC

```

Figura 11a - Continuação

```

phr_IA GAAGAGAATATGCAGACAATTATTGCTTATTCAATGCTAATTACAATAAAATTTAAAAATTTGGTGGAAACA
phr_IB GAAGAGAATATGCAGACAATTATTGCTTATTCAATGCTAATTACAATAAAATTTAAAAATTTGGTGGAAACA
phr_IC GAAGAGAATATGCAGACAATTATTGCTTATTCAATGCTAATTACAATAAAATTTAAAAATTTGGTGGAAACA
phr_ID GAAGAGAATATGCAGACAATTATTGCTTATTCAATGCTAATTACAATAAAATTTAAAAATTTGGTGGAAACA
phr_IE GAAGAGAATATGCAGACAATTATTGCTTATTCAATGCTAATTACAATAAAATTTAAAAATTTGGTGGAAACA
phr_IF GAAGAGAATATGCAGACAATTATTGCTTATTCAATGCTAATTACAATAAAATTTAAAAATTTGGTGGAAACA
phr_IG GAAGAGAATATGCAGACAATTATTGCTTATTCAATGCTAATTACAATAAAATTTAAAAATTTGGTGGAAACA

phr_IA CAATCGATATTATTTGGAACATCACACAACAGTATATCTTTACAGTTTGAAGAAGCTCGAGTTTTACAA
phr_IB CAATCGATATTATTTGGAACATCACACAACAGTATATCTTTACAGTTTGAAGAAGCTCGAGTTTTACAA
phr_IC CAATCGATATTATTTGGAACATCACACAACAGTATATCTTTACAGTTTGAAGAAGCTCGAGTTTTACAA
phr_ID CAATCGATATTATTTGGAACATCACACAACAGTATATCTTTACAGTTTGAAGAAGCTCGAGTTTTACAA
phr_IE CAATCGATATTATTTGGAACATCACACAACAGTATATCTTTACAGTTTGAAGAAGCTCGAGTTTTACAA
phr_IF CAATCGATATTATTTGGAACATCACACAACAGTATATCTTTACAGTTTGAAGAAGCTCGAGTTTTACAA
phr_IG CAATCGATATTATTTGGAACATCACACAACAGTATATCTTTACAGTTTGAAGAAGCTCGAGTTTTACAA

phr_IA ACTTACGACGTGCTGTGGAATAAACTCAAAATGATTTGCGAGAACATGGCAAAATATATCCAACATTAC
phr_IB ACTTACGACGTGCTGTGGAATAAACTCAAAATGATTTGCGAGAACATGGCAAAATATATCCAACATTAC
phr_IC ACTTACGACGTGCTGTGGAATAAACTCAAAATGATTTGCGAGAACATGGCAAAATATATCCAACATTAC
phr_ID ACTTACGACGTGCTGTGGAATAAACTCAAAATGATTTGCGAGAACATGGCAAAATATATCCAACATTAC
phr_IE ACTTACGACGTGCTGTGGAATAAACTCAAAATGATTTGCGAGAACATGGCAAAATATATCCAACATTAC
phr_IF ACTTACGACGTGCTGTGGAATAAACTCAAAATGATTTGCGAGAACATGGCAAAATATATCCAACATTAC
phr_IG ACTTACGACGTGCTGTGGAATAAACTCAAAATGATTTGCGAGAACATGGCAAAATATATCCAACATTAC

phr_IA GCGTATACTGGGCCAAAAA
phr_IB GCGTATACTGGGCCAAAAA
phr_IC GCGTATACTGGGCCAAAAA
phr_ID GCGTATACTGGGCCAAAAA
phr_IE GCGTATACTGGGCCAAAAA
phr_IF GCGTATACTGGGCCAAAAA
phr_IG GCGTATACTGGGCCAAAAA

```

## (b) Alinhamento das sequências deduzidas de aminoácidos Photolyase

```

          10      20      30      40      50      60      70
...|...|...|...|...|...|...|...|...|...|...|...|...|...|...|
phr_IA DAHNVVPCWLVSNQPIFDTRDKINEQLQNYLKPFPPIVVKHPHKSVPVPESSNTYIDWYALITSRNIDQ
phr_IE .....
phr_IF .....
phr_IG .....

phr_IA SIDTVKWKTKAGYRPAMRLASFIEFSLYMYNNRLENAFSINQSDLSPFFHFGFLSTQRVIYYLTFHILKT
phr_IE Y.....Y....
phr_IF Y.....X....
phr_IG Y.....X....

phr_IA RGRYGSKGKLNKVERFLENLLYRREYADNYCLFNANYNKFKILVEHNRYFFEHTTVYLYSLEELEFSQ
phr_IE .....
phr_IF .....
phr_IG .....

phr_IA TYDVLWNKTQNDLREHGKIYPTLRVYWAK
phr_IE .....
phr_IF .....
phr_IG .....

```

Figura 11. Alinhamento das sequências nucleotídicas parciais do gene *phr* dos isolados de PsinSNPV (IA a IG). Os asteriscos revelam as posições em que foram detectados polimorfismos. O “Y” na posição 219 e 406 nt indica a presença de duas bases: citosina e timina (a). Alinhamento das sequências deduzidas de aminoácidos do gene *phr* indicando a variação dos aminoácidos entre os isolados. O “X” na posição 136 aa indica a presença dos dois aminoácidos: tirosina (Y) e histidina (H) (b).

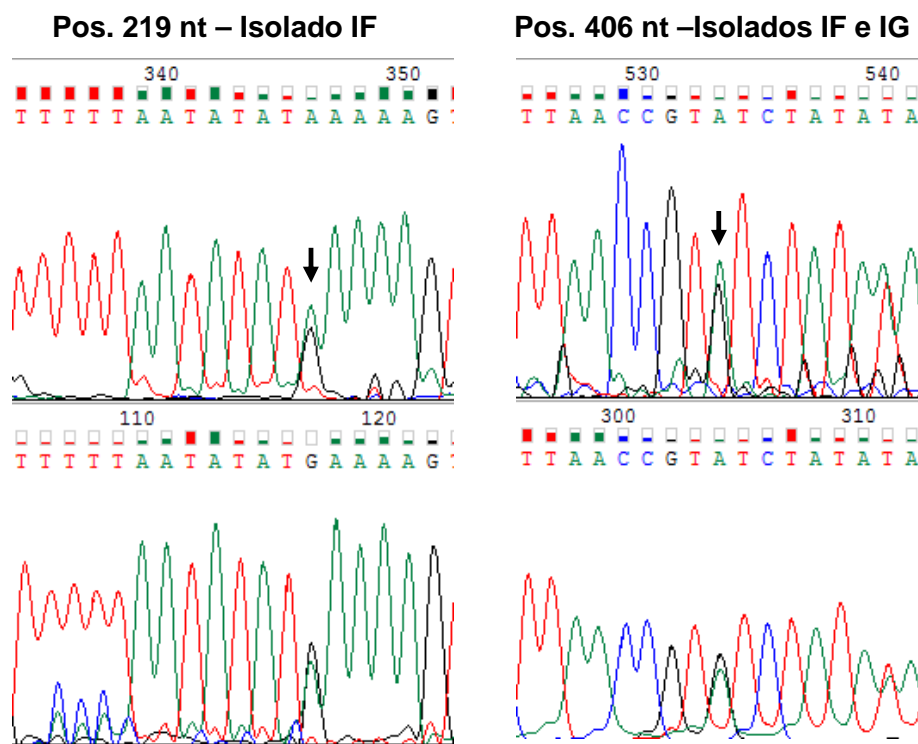
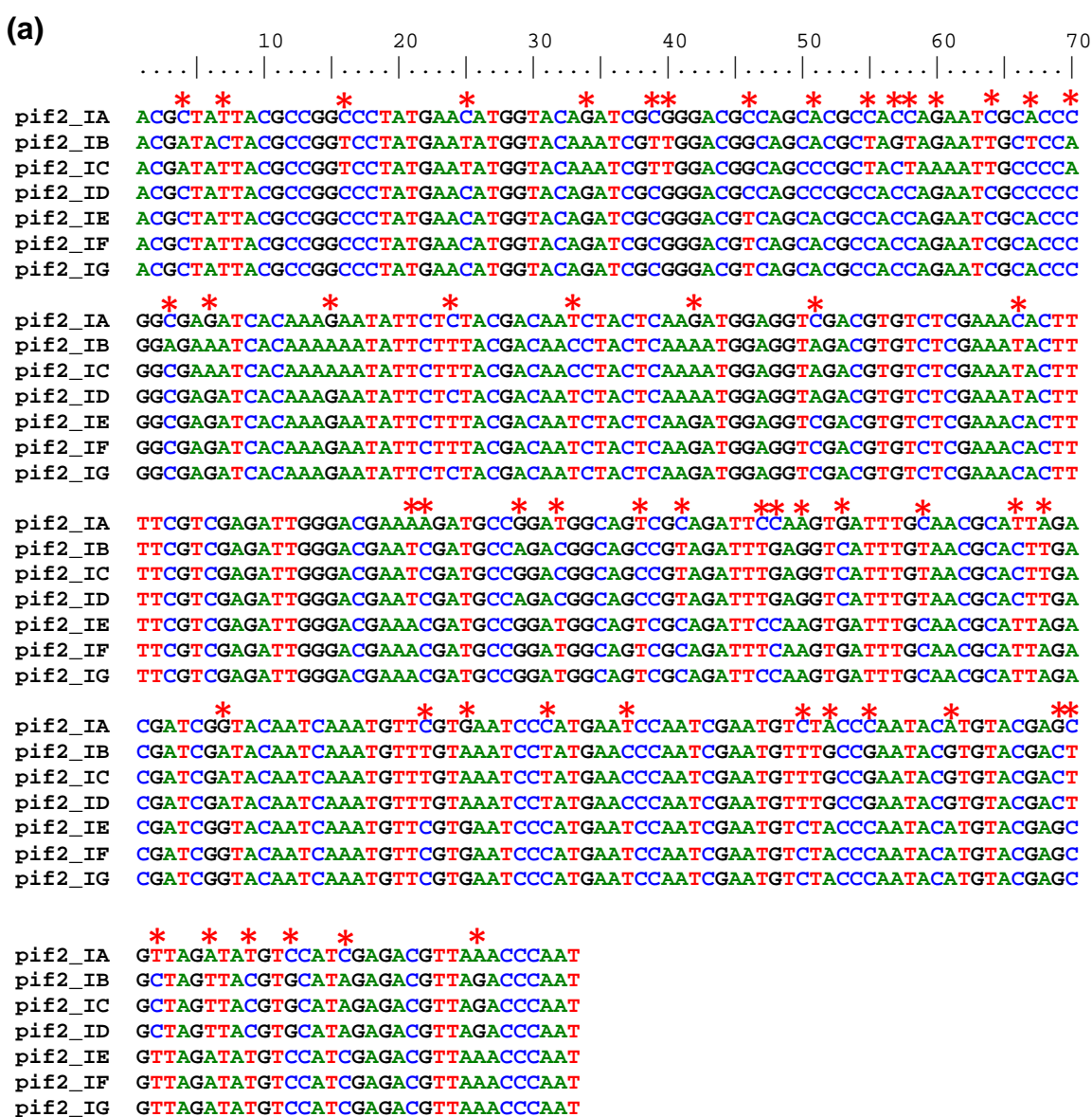


Figura 12. Eletroferogramas do seqüenciamento do produto de PCR do gene *phr*. As setas indicam as posições em que há sobreposição de bases nucleotídicas indicando mistura de genótipos.

As seqüências nucleotídicas parciais do gene *pif-2* dos isolados de PsinSNPV apresentaram tamanho de 313 pb e %AT média de 52,61% (Figura 13a). O alinhamento das seqüências parciais do gene *pif-2* dos isolados de PsinSNPV mostra a presença de 55 bases polimórficas. Os polimorfismos detectados na seqüência nucleotídica parcial deste gene causaram variações nas seqüências deduzidas de aminoácidos entre os isolados (Figura 13b) apresentando nas posições: 13 aa isolados IB e IC valina (V) ao invés de alanina (A); 17 aa isolados IC e ID prolina (P) – histidina(H); 19 aa isolado IB serina (S) – treonina (T); 20 aa: isolado IC lisina (R) – arginina (K); 54 aa isolado IA lisina (R), IB, IC e ID serina (S) e IE, IF e IG treonina (T); 63 aa: isolados IB, IC e ID ácido glutâmico (E) – glutamina (Q); 93 aa isolados IB, IC e ID treonina (T) – serina (S); 94 aa isolados IB, IC e ID alanina (A) – valina (V);

95 aa isolados IB, IC e ID serina (S) – arginina (K) e 102 aa isolados IB, IC e ID arginina (K) – lisina (R).

As alterações nas posições 17 e 19 aa apresentam variações de aminoácidos com diferentes graus de hidrofobicidade. Os aminoácidos treonina e prolina são hidrofóbicos enquanto serina e histidina possuem carga positiva e são hidrofílicos. Esse tipo de variação na estrutura primária pode ocasionar modificações na estrutura tridimensional das proteínas.



**(b) Alinhamento das sequências deduzidas de aminoácidos PIF-2**

```

          10      20      30      40      50      60      70
    ....|....|....|....|....|....|....|....|....|....|....|....|....|
pif2_IA RYYAGPMNMVQIAGRQHATRIAPGEITKNILYDNLKMEVDVSRNTFRRDWDEKMPDGSRRRFQVICNALD
pif2_IB .....V.....S.....S.....S.....E.....
pif2_IC .....V..P..K.....S.....E.....
pif2_ID .....P.....S.....E.....
pif2_IE .....T.....
pif2_IF .....T.....
pif2_IG .....T.....

pif2_IA DRYNQMFVNPMNPIECLPNTCTSVRYVHRDVKPNFET
pif2_IB .....TAS.....R.....
pif2_IC .....TAS.....R.....
pif2_ID .....TAS.....R.....
pif2_IE .....
pif2_IF .....
pif2_IG .....

```

Figura 13. Alinhamento das sequências nucleotídicas parciais do gene *pif-2* dos isolados de PsinSNPV (IA a IG). Os asteriscos revelam as posições em que foram detectados polimorfismos (a). Alinhamento das sequências deduzidas de aminoácidos dos genes *pif-2* indicando a variação dos aminoácidos entre os isolados (b). Entre os isolados de PsinSNPV (IA a IG), o gene *pif-2*

apresentou maior número de variações totalizando a presença de 55 bases polimórficas (Tabela 11), diferentemente dos das sequências dos genes *lef-8*, *lef-9*, *phr* e *polh* que apresentaram poucos polimorfismos na sua composição nucleotídica.

Considerando que o conteúdo de guanina e citosina influencia a expressão de genes, pode se inferir que os eventos de transcrição e replicação dos genes aqui estudados não estariam sendo afetados pela sua composição nucleotídica, uma vez que as sequências analisadas apresentaram a mesma quantidade relativa de guanina/citosina e adenina/timina.

As sequências deduzidas de aminoácidos obtidas das sequências nucleotídicas parciais dos genes *lef-8*, *lef-9*, *pif-2*, *phr* e *polh* foram analisadas por BLAST e apresentaram alta identidade (> 80%) sendo seqüências homólogas as de outros baculovirus (Anexo I). Os vírus ChchNPV e TnSNPV foram os que apresentaram maior identidade (> 90%, *E value* < 1e<sup>-50</sup>).

**Tabela 11. Posição nucleotídica (nt) dos polimorfismos detectados nas sequências parciais de genes de PsinSNPV**

Genes / Posição (nt)	Isolados de PsinSNPV						
	IA	IB	IC	ID	IE	IF	IG
<i>lef-8</i> 160	C	C	C	C	T	T	T
<i>lef-9</i> 206	C	C	C	C	T	T	T
<i>pif-2</i> 4	C	A	A	C	C	C	C
7	T	C	T	T	T	T	T
16	C	T	T	C	C	C	C
25	C	T	T	C	C	C	C
34	G	A	A	G	G	G	G
39	C	T	T	C	C	C	C
40	G	T	T	G	G	G	G
46	C	G	G	C	T	T	T
51	A	C	C	A	A	A	A
55	C	T	T	C	C	C	C
57	C	G	C	C	C	C	C
58	C	T	T	C	C	C	C
60	G	G	A	G	G	G	G
64	C	T	T	C	C	C	C
67	A	T	C	C	A	A	A
70	C	A	A	C	C	C	C
73	C	A	C	C	C	C	C
76	G	A	A	G	G	G	G
85	G	A	A	G	G	G	G
94	C	T	T	C	T	T	C
103	T	C	C	T	T	T	T
112	G	A	A	A	G	G	G
121	C	A	A	A	C	C	C
136	C	T	T	T	C	C	C
161	A	T	T	T	A	A	A
162	A	C	C	C	C	C	C
169	G	A	G	A	G	G	G
172	T	C	C	C	T	T	T
178	T	C	C	C	T	T	T
181	C	T	T	T	C	C	C
187	C	T	T	T	C	T	C
188	C	G	G	G	C	C	C
190	A	G	G	G	A	A	A
193	G	C	C	C	G	G	G
199	C	T	T	T	C	C	C
206	T	C	C	C	T	T	T
208	A	T	T	T	A	A	A
217	G	A	A	A	G	G	G
232	C	T	T	T	C	C	C
235	G	A	A	A	G	G	G
241	C	T	T	T	C	C	C
247	T	C	C	C	T	T	T
260	C	T	T	T	C	C	C
262	A	G	G	G	A	A	A
265	C	G	G	G	C	C	C
271	A	G	G	G	A	A	A
279	G	C	C	C	G	G	G
280	C	T	T	T	C	C	C
282	T	C	C	C	T	T	T
286	A	T	T	T	A	A	A
289	T	C	C	C	T	T	T
292	C	G	G	G	C	C	C
296	C	A	A	A	C	C	C
306	A	G	G	G	A	A	A
319	A	G	G	G	A	A	A
<i>phr</i> 135	A	A	A	A	G	G	G
212	C	C	C	C	A	A	A
219	C	C	C	C	T	C, T	T
406	C	C	C	C	T	C, T	C, T
<i>polh</i> 292	G	G	G	G	A	A	G
342	A	A	A	A	G	G	A

### 5.3. Relações filogenéticas entre isolados virais de PsinSNPV

A relação filogenética entre os isolados de PsinSNPV (IA a IG) foi identificada pela análise da Máxima Parcimônia (MP) e Inferência Bayesiana (IB) da sequência de 2.205 nucleotídeos derivados da concatenação das sequências parciais dos genes *lef-8*, *lef-9*, *phr*, *pif-2* e *polh* dos isolados de PsinSNPV. O vírus ChchNPV foi escolhido como grupo externo. Estudos preliminares mostraram que este vírus está proxicamente relacionado a PsinSNPV (Craveiro *et al.*, 2010). O vírus ChchNPV possui dois genes *DNA photolyase* em seu genoma e apenas o gene caracterizado como *phr1* foi utilizado na análise.

As sequências concatenadas dos genes *lef-8*, *lef-9*, *phr*, *pif-2* e *polh* foram analisadas usando o programa PAUP\*, sem a codificação das inserções e deleções (*indels*) para a estimativa do número de caracteres constantes, informativos e não informativos para a análise MP (Tabela 12).

**Tabela 12. Características das sequências concatenadas dos genes *lef-8*, *lef-9*, *phr*, *pif-2* e *polh***

Características	Total
Táxons analisados	8
Caracteres analisados	2.229
Caracteres constantes	1.848
Caracteres não informativos para parcimônia	323
Caracteres informativos para parcimônia	58
*Caracteres informativos peso = 1	2217
*Caracteres informativos peso ≠ 1	12**

\*Informações obtidas pela análise MP com reponderação progressiva

\*\*Caract. Inform. peso ≠ 1: 4 caract. = 0.333333, 4 caract.= 0.250000 e 4 caract.= 0

Na busca heurística inicial da análise MP da sequências concatenadas dos cinco genes analisados, uma árvore mais parcimoniosa foi conservada



com comprimento de 398 passos para a MP não ponderada e 378.66667 passos para MP com reponderação progressiva. Nas duas árvores, o índice de consistência foi superior a 0,95 (Tabela 13). Durante a reponderação progressiva, 2217 caracteres tiveram pesos iguais a 1 e 12 caracteres tiveram peso menor que 1 (4 caracteres = 0.333333, 4 caracteres = 0.250000 e 4 caracteres = 0). Um ciclo de reponderação progressiva foi suficiente para manter o índice de consistência rescalonado (RC) estável.

**Tabela 13. Estatísticas das árvores MP obtidas com a análise das sequências concatenadas dos genes *lef-8*, *lef-9*, *phr*, *pif-2* e *polh***

<b>Estatísticas das árvores</b>	<b>Análise não ponderada</b>	<b>Reponderação progressiva</b>
Número de árvores mais parcimoniosas	1	1
Comprimento da árvore (nº de passos)	398	378,66667
Índice de Consistência (IC)	0,9698	0,9938
Índice de Homoplasia (IH)	0,0302	0,0062
Índice de Retenção (IR)	0,8983	0,9767
Índice de Consistência Rescalonado (RC)	0,8712	0,9706

Para a Inferência Bayesiana, o modelo de substituição nucleotídica que mais apropriado à sequência dos cinco genes analisados é o modelo GTR (*General Time Reversible*) com taxa de variação entre os sítios igual (comando MrBayes: nst=6 rates=equal). As Cadeias de Markov processadas para 1.000.000 gerações foram suficientes para que a diferença das frequências obtivesse desvio padrão menor que 0,01 ( $p = 0.009757$ ).

As árvores mais parcimoniosas obtidas da análise não ponderada e da reponderação progressiva apresentaram mesma topologia com valores de *bootstrap* aproximados (1000 replicatas) (Figura 14). Nas árvores, os sete isolados estão distribuídos em dois grupos monofiléticos, um contendo os isolados IA, IE, IF e IG e o outro contendo os isolados IB, IC e ID. Os isolados

virais IE e IF formam um grupo robusto (*bootstrap* > 90%) o que corrobora com a hipótese proposta por Alexandre *et al.* (2010) de que esses isolados apresentam genomas similares. Os outros isolados também estão distribuídos em grupos fortemente sustentados (valores de *bootstrap* acima de 90%).

O mesmo agrupamento ocorre na árvore obtida pela Inferência Bayesiana que também apresentou grupos fortemente sustentados com probabilidade posterior acima de 0,9 (Figura 14).

Nestes resultados, observa-se que os isolados IA, IE e IF, mais virulentos entre os sete isolados analisados (Alexandre *et al.*, 2010), e os isolados IB, IC e ID, coletados na mesma safra (2006) em plantações de soja da região do Paraná, estão agrupados em diferentes grupos monofiléticos (Figura 14). Assim, não pode ser descartada a hipótese de que os isolados de PsinSNPV podem estar relacionados com a virulência e local de coleta.

Os isolados se apresentam divididos em diferentes grupos com altos valores de suporte dos ramos, *bootstrap* > 90% e probabilidade posterior > 0,9, indicando a presença de diversidade genética entre os isolados estudados.

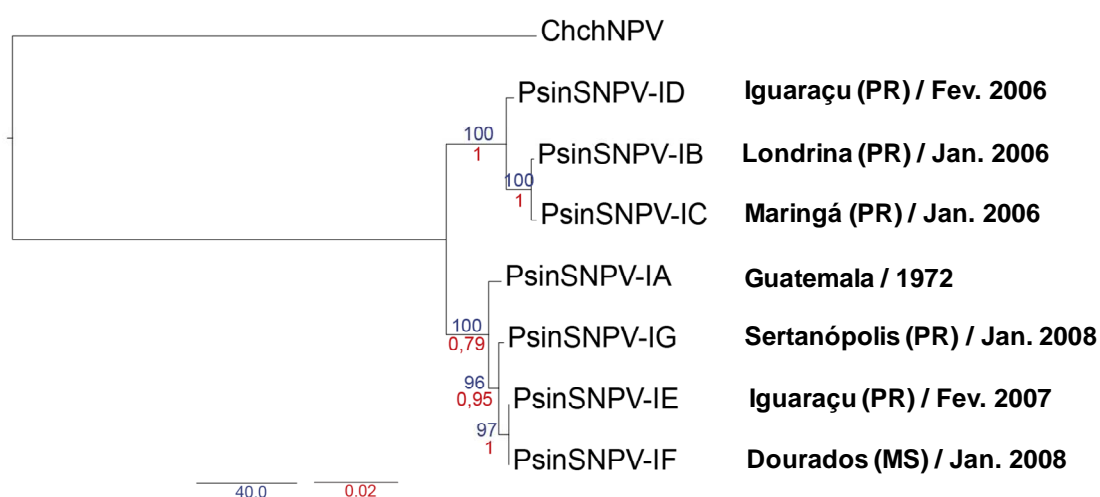


Figura 14. Relações filogenéticas entre os isolados de PsinSNPV (IA a IG) baseadas nas sequências concatenadas dos genes *lef-8*, *lef-9*, *phr*, *pif-2* e *polh*. Árvore filogenética mais parcimoniosa obtida da busca heurística. Os números nos nós indicam o valor de suporte dos ramos, em azul *bootstrap* (1000 replicatas) e em vermelho probabilidade posterior obtida da análise IB. O vírus ChchNPV foi utilizado como grupo externo. A escala em barra representa o número de substituição para cada sítio.

#### 5.4. Análise fenética dos RFLPs de isolados de PsinSNPV

Poucos polimorfismos foram detectados nos fragmentos de restrição obtidos pela clivagem de DNA dos isolados de PsinSNPV (IA a IG) com quatro enzimas de restrição (Figura 15). A análise dos RFLPs por algoritmo UPGMA apresentou um alto nível de similaridade entre os isolados de PsinSNPV com coeficiente de similaridade de Dice igual a 0,92.

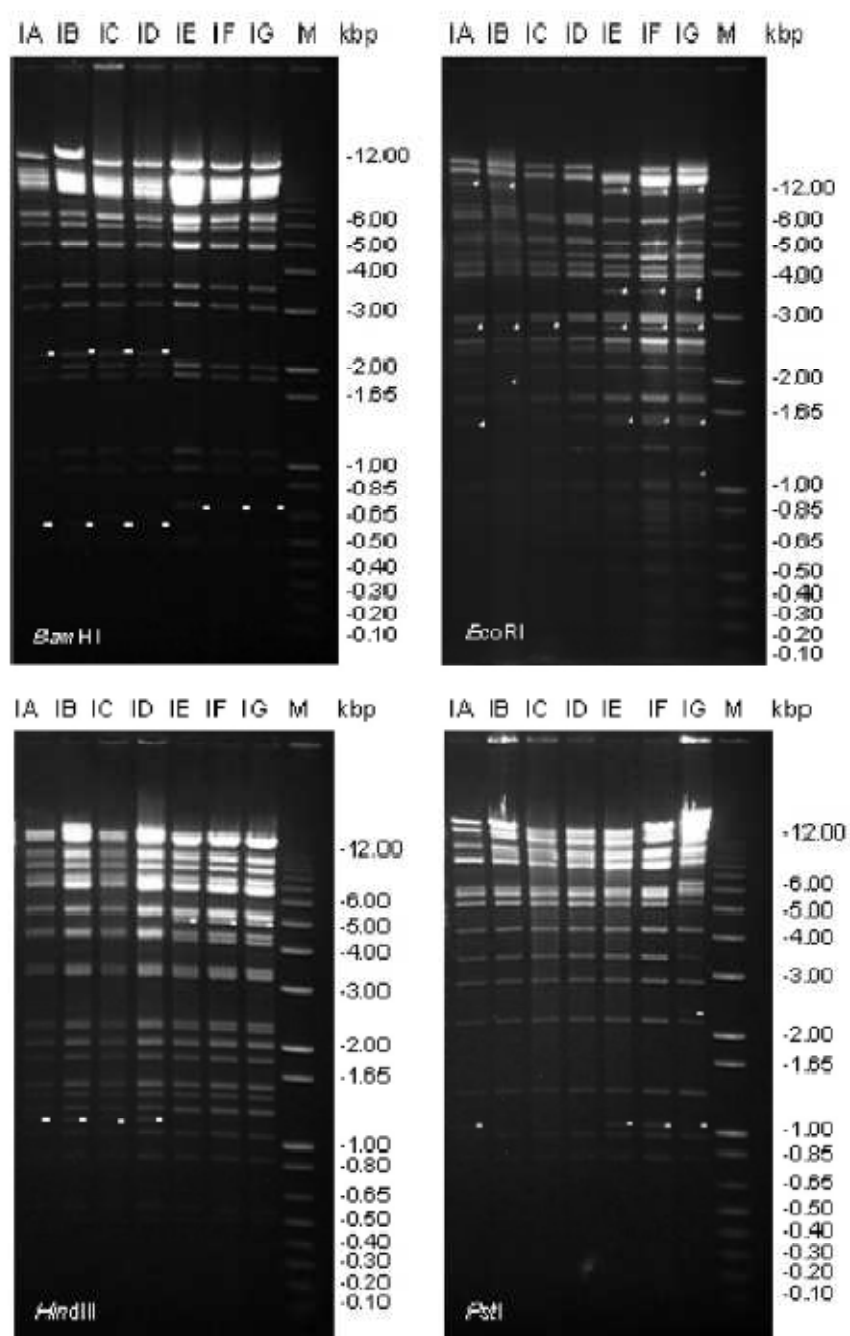


Figura 15. Perfis de restrição de DNAs dos isolados de PsinSNPV (IA a IG). DNAs clivados com as enzimas *Bam*HI, *Eco*RI, *Hind*III e *Pst*I. Marcador molecular (M): 1 kb plus DNA Ladder (Invitrogen). Bandas submolares marcadas por asterisco. Eletroforese em gel de agarose 1% (Craveiro *et al.*, 2009; Alexandre *et al.*, 2010).

Os dendrogramas construídos utilizando os coeficientes de similaridade de Dice e Jaccard, quando comparados, não apresentaram diferenças relevantes nos valores do coeficiente de similaridade e *bootstrap*. Os coeficientes de Sorensen-Dice e Jaccard são idênticos em suas expressões, exceto pelo fato de Dice atribuir peso dobrado a co-ocorrência positiva, ou seja, ocorrência de bandas de um mesmo peso molecular em dois perfis (*matching*).

A correlação cofenética foi 0,95 (Mantel t-test = 3,85) revelando um bom ajuste entre a matriz de similaridade e o dendrograma (Sokal e Rohlf, 1962). Como podemos observar no dendrograma, os isolados virais estão distribuídos em dois grupos suportados por fortes valores de *bootstrap* (> 85%) e subdivididos em cinco grupos de similaridade: (1) PsinSNPV-IA, (2) PsinSNPV-IB, (3) PsinSNPV-IC e ID, (4) PsinSNPV-IE e IF e (5) PsinSNPV-IG. Esses dados evidenciam a presença de diversidade entre os genótipos avaliados (Figura 16).

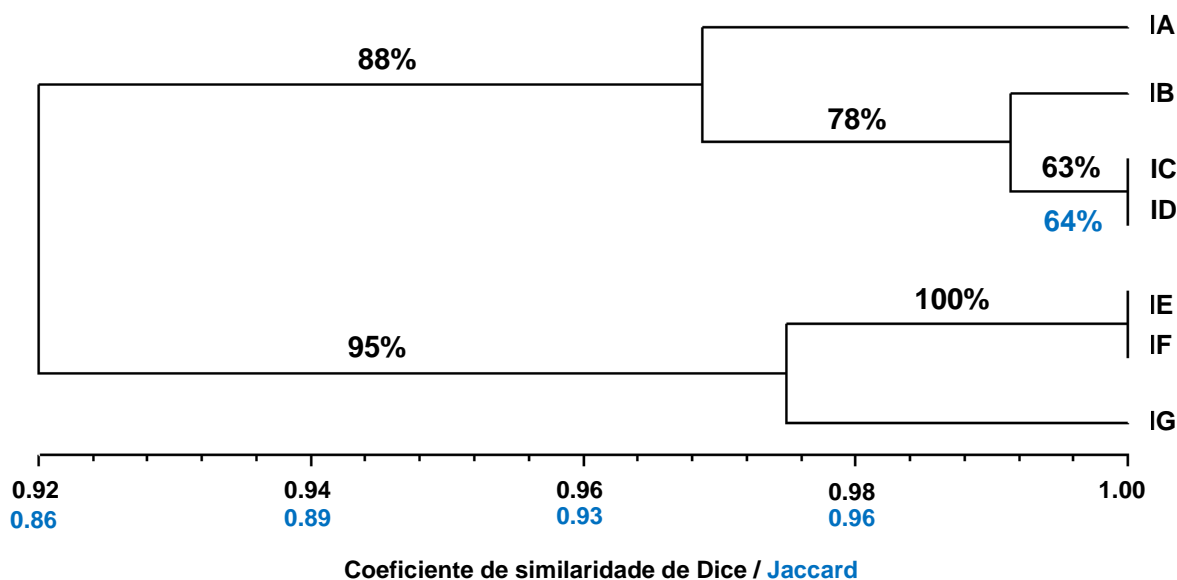


Figura 16. Dendrograma de similaridade genotípica resultante da análise dos RFLPs de isolados de PsinSNPV (IA a IG) utilizando algoritmo UPGMA. Os coeficientes de similaridade de Dice e Jaccard foram calculados com base na matriz da presença/ausência de fragmentos gerada pela análise dos perfis de restrição de DNAs clivados com *Bam*HI, *Eco*RI, *Hind*III e *Pst*I. O número nos ramos indicam o valor de *bootstrap* (%) para 1000 replicatas. Em azul estão os valores calculados pelo coeficiente de Jaccard.

As relações filogenéticas dos isolados virais mostraram diferenças entre o dendograma obtido pela análise fenética (Figura 16) e as árvores filogenéticas obtidas pelas análises cladísticas (Figura 14). Diferente da árvore filogenética, o dendograma apresentou o isolado IA agrupado aos isolados IB, IC e ID. A incongruência observada entre as análises e a detecção de poucos caracteres variáveis (polimorfismos) são fatores que podem explicar a inconsistência do agrupamento entre os isolados obtido pela análise fenética.

### **5.5. Reconstrução das relações filogenéticas da família *Baculoviridae* incluindo o vírus PsinSNPV- IE**

As relações filogenéticas da família *Baculoviridae* foram reconstruídas utilizando as sequências nucleotídicas e peptídicas dos genes *lef-8*, *lef-9*, *pif-2* e *polh/gran* de PsinSNPV e outros 57 baculovirus com genoma seqüenciado. Como representante do vírus PsinSNPV foi escolhido o isolado PsinSNPV-IE por se tratar de um isolado brasileiro e ser um dos mais virulentos dentre os sete isolados analisados (Alexandre *et al.*, 2010). Para tanto, um total de 28 árvores filogenéticas foram construídas (Apêndice III):

- 16 árvores a partir da sequência individual dos quatro genes utilizando máxima parcimônia ponderada e não ponderada;
- 6 árvores obtidas da concatenação dos genes *lef-8*, *lef-9* e *pif-2* utilizando máxima parcimônia e inferência Bayesiana, sendo 2 MP ponderada, 2 MP não ponderada e 2 IB;
- 6 árvores obtidas da concatenação dos genes *lef-8*, *polh/gran*, *lef-9* e *pif-2* utilizando máxima parcimônia e inferência Bayesiana, sendo 2 MP ponderada, 2 MP não ponderada e 2 IB.

Nas análises filogenéticas, o vírus *Culex nigripalpus* NPV (CuniNPV) foi usado como grupo externo, exceto nas inferências filogenéticas em que foi utilizado o gene *polh/gran*, pelo fato da proteína dos OBs de CuniNPV não ser homóloga à proteína dos OBs dos outros baculovirus. Nas árvores que incluem o gene *polh/gran*, foi utilizado como grupo externo o vírus *Neodiprion sertifer* NPV (NeseNPV).

Características das diferentes sequências utilizadas nas análises filogenéticas, como o número de caracteres constantes, informativos e não informativos para a parcimônia, foram estimadas através do programa PAUP, sem a codificação das inserções e deleções e estão apresentadas nas Tabelas 14 e 15.

As estatísticas das árvores como número de árvores mais parcimoniosas, índice de homoplasia e consistência estão especificados nas Tabelas 16, 17, 18 e 19. Destaca-se o índice de consistência, que apresentou baixos valores ( $IC < 0,8$ ) para todas as árvores mais parcimoniosas com exceção da árvore mais parcimoniosa reponderada obtida da sequência peptídica POLH/GRAN. Isto mostra uma grande quantidade de homoplasias presentes nas matrizes de dados.

**Tabela 14. Características das sequências nucleotídicas individual e concatenada dos genes *lef-8*, *polh/gran*, *lef-9* e *pif-2***

Características	Genes				Concatenação	
	<i>lef-8</i>	<i>polh/gran</i>	<i>lef-9</i>	<i>pif-2</i>	4 genes	3 genes**
Total de táxons analisados	58	57 <sup>a</sup>	58	58	57 <sup>a</sup>	58
Total de caracteres	503	566	424	348	1.786	1.333
Número de caracteres constantes	95	169	120	82	418	294
Número de caracteres não informativos para parcimônia	26	20	52	23	121	143
Número de caracteres informativos para parcimônia	382	377	252	243	1.247	896
*Número de caracteres informativos peso = 1	158	285	208	118	742	521
*Número de caracteres informativos peso ≠ 1	345	281	216	230	1.044	812

\*Informações obtidas pela análise MP com reponderação progressiva. \*\*Concatenação dos 3 genes: *lef-8*, *lef-9* e *pif-2*.

<sup>a</sup>Não está incluso o vírus CuniNPV na análise.

**Tabela 15. Características das sequências deduzidas de aminoácidos, individual e concatenada, LEF-8, POLH/GRAN, LEF-9 e PIF-2**

Características	Sequências deduzidas de aminoácidos				Concatenação	
	LEF-8	POLH/GRAN	LEF-9	PIF-2	4 seq. pept.	3 seq.pept.**
Total de táxons analisados	58	57 <sup>a</sup>	58	58	57 <sup>a</sup>	58
Total de caracteres	185	183	126	117	584	427
Número de caracteres constantes	30	55	19	35	141	85
Número de caracteres não informativos para parcimônia	17	8	28	12	39	53
Número de caracteres informativos para parcimônia	138	120	79	70	404	289
*Número de caracteres informativos peso = 1	72	136	76	51	301	189
*Número de caracteres informativos peso ≠ 1	113	47	50	66	283	238

\*Informações obtidas pela análise MP com reponderação progressiva. \*\*Concatenação dos 3 genes: *lef-8*, *lef-9* e *pif-2*.

<sup>a</sup>Não está incluso o vírus CuniNPV na análise.

**Tabela 16. Estatísticas das árvores obtidas da análise MP não ponderada das sequências nucleotídicas individual e concatenada dos genes *lef-8*, *polh/gran*, *lef-9* e *pif-2***

Estatísticas das árvores não ponderadas	Genes				Concatenação	
	<i>lef-8</i>	<i>polh/gran</i>	<i>lef-9</i>	<i>pif-2</i>	4 genes	3 genes*
Número de árvores mais parcimoniosas	1	33	16	2	6	9
Comprimento da árvore (nº de passos)	3587	2861	2113	2403	10673	8205
Índice de Consistência (IC)	0,2623	0,2628	0,2721	0,2351	0,2581	0,2605
Índice de Homoplasia (IH)	0,7377	0,7372	0,7279	0,7649	0,7419	0,7395
Índice de Retenção (IR)	0,6066	0,6387	0,6116	0,6089	0,6125	0,5988
Índice de Consistência Rescalonado (RC)	0,1519	0,1679	0,1664	0,1432	0,1581	0,1560

\*Concatenação dos 3 genes: *lef-8*, *lef-9* e *pif-2*.

**Tabela 17. Estatísticas das árvores obtidas da análise MP reponderada das sequências nucleotídicas individual e concatenada dos genes *lef-8*, *polh/gran*, *lef-9* e *pif-2***

Estatísticas das árvores reponderadas	Genes				Concatenação	
	<i>lef-8</i>	<i>polh/gran</i>	<i>lef-9</i>	<i>pif-2</i>	4 genes	3 genes*
Número de árvores mais parcimoniosas	2	5	1	1	1	1
Comprimento da árvore (nº de passos)	579,79	505,34	371,34	345,69	1766,53	1336,34
Índice de Consistência (IC)	0,4203	0,5249	0,5156	0,3710	0,4628	0,4533
Índice de Homoplasia (IH)	0,5797	0,4751	0,4844	0,6290	0,5372	0,5467
Índice de Retenção (IR)	0,7225	0,8324	0,7536	0,7065	0,7587	0,7119
Índice de Consistência Rescalonado (RC)	0,3037	0,4370	0,3886	0,2622	0,3511	0,3227

\*Concatenação dos 3 genes: *lef-8*, *lef-9* e *pif-2*.



**Tabela 18. Estatísticas das árvores obtidas com a análise MP não ponderada das sequências deduzidas de aminoácidos, individual e concatenada, LEF-8, POLH/GRAN, LEF-9 e PIF-2**

Estatísticas das árvores não ponderadas	Sequências deduzidas de aminoácidos				Concatenação	
	LEF-8	POLH/GRAN	LEF-9	PIF-2	4 seq. pept.	3 seq. pept.*
Número de árvores mais parcimoniosas	458	1592	500	281	22	120
Comprimento da árvore (nº de passos)	1318	378	478	595	2651	2442
Índice de Consistência (IC)	0,5531	0,7063	0,6381	0,5261	0,5571	0,5467
Índice de Homoplasia (IH)	0,4469	0,2937	0,3619	0,4739	0,4429	0,4533
Índice de Retenção (IR)	0,7435	0,9127	0,8188	0,7563	0,7834	0,7433
Índice de Consistência Rescalonado (RC)	0,4112	0,6447	0,5225	0,3978	0,4365	0,4063

\*Concatenação de 3 sequências peptídicas: LEF-8, LEF-9 e PIF-2.

**Tabela 19. Estatísticas das árvores obtidas com a análise MP reponderada das sequências deduzidas de aminoácidos, individual e concatenada, LEF-8, POLH/GRAN, LEF-9 e PIF-2**

Estatísticas das árvores reponderadas	Sequências deduzidas de aminoácidos				Concatenação	
	LEF-8	POLH/GRAN	LEF-9	PIF-2	4 seq. pept.	3 seq. pept.*
Número de árvores mais parcimoniosas	2	120	18	4	1	2
Comprimento da árvore (nº de passos)	532,32	243,92	256,30	226,84	1133,86	982,81
Índice de Consistência (IC)	0,6359	0,8440	0,7630	0,6126	0,6669	0,6410
Índice de Homoplasia (IH)	0,3641	0,1560	0,2370	0,3874	0,3331	0,3590
Índice de Retenção (IR)	0,7997	0,9585	0,8805	0,8124	0,8506	0,7997
Índice de Consistência Rescalonado (RC)	0,5085	0,8090	0,6718	0,4977	0,5672	0,5126

\*Concatenação de 3 sequências peptídicas: LEF-8, LEF-9 e PIF-2.

A árvore de consenso estrito (Figura 17), resultante das seis árvores mais parcimoniosas obtidas da análise MP não ponderada da sequência nucleotídica concatenada dos genes *lef-8*, *polh/gran*, *lef-9* e *pif-2*, apresentou três grupos: grupo externo formado pelo vírus NeseNPV que juntamente com NeabNPV e NeleNPV formam o gênero *Gammabaculovirus* (NPVs específicos de himenópteros), grupo formado pelos *Betabaculovirus* (GVs específicos de lepidópteros) e grupo formado pelos *Alphabaculovirus* (NPVs específicos de lepidópteros). Estes grupos estão bem definidos e fortemente sustentados com *bootstrap* acima de 95%.

No gênero *Alphabaculovirus*, estão presentes o Grupo I, grupo bem definido (*bootstrap* = 100%) formado pelos representantes AcMNPV, BmNPV, AgMNPV e CfMNPV e o Grupo II pouco conservado, com alta politomia, formado por HearNPV, SfMNPV, TnSNPV entre outros. O Grupo I é claramente subdividido em dois clados: clado a (AcMNPV, PlxyNPV, BomaNPV, BmNPV, RoNPV e MaviNPV) e clado b (AgMNPV, EppoNPV, AnpeMNPV, CfMNPV, CfDEFMNPV, HycuNPV e OpMNPV). No Grupo II, a falta de agrupamentos bem resolvidos não permite a formação de subdivisões (Miele *et al.*, 2011).

Os *Alphabaculovirus* do Grupo I e *Gammabaculovirus* em todas as árvores mostradas neste trabalho apresentaram grupos bem resolvidos suportados por *bootstrap* acima de 95%. Esses grupos formam os clados mais recentes da família *Baculoviridae* e possuem uma menor diversidade no conteúdo genético, pois, desde o aparecimento de seu ancestral comum, esses grupos tiveram menor tempo para sofrer mutações e incorporar novas sequências (Miele *et al.*, 2011).

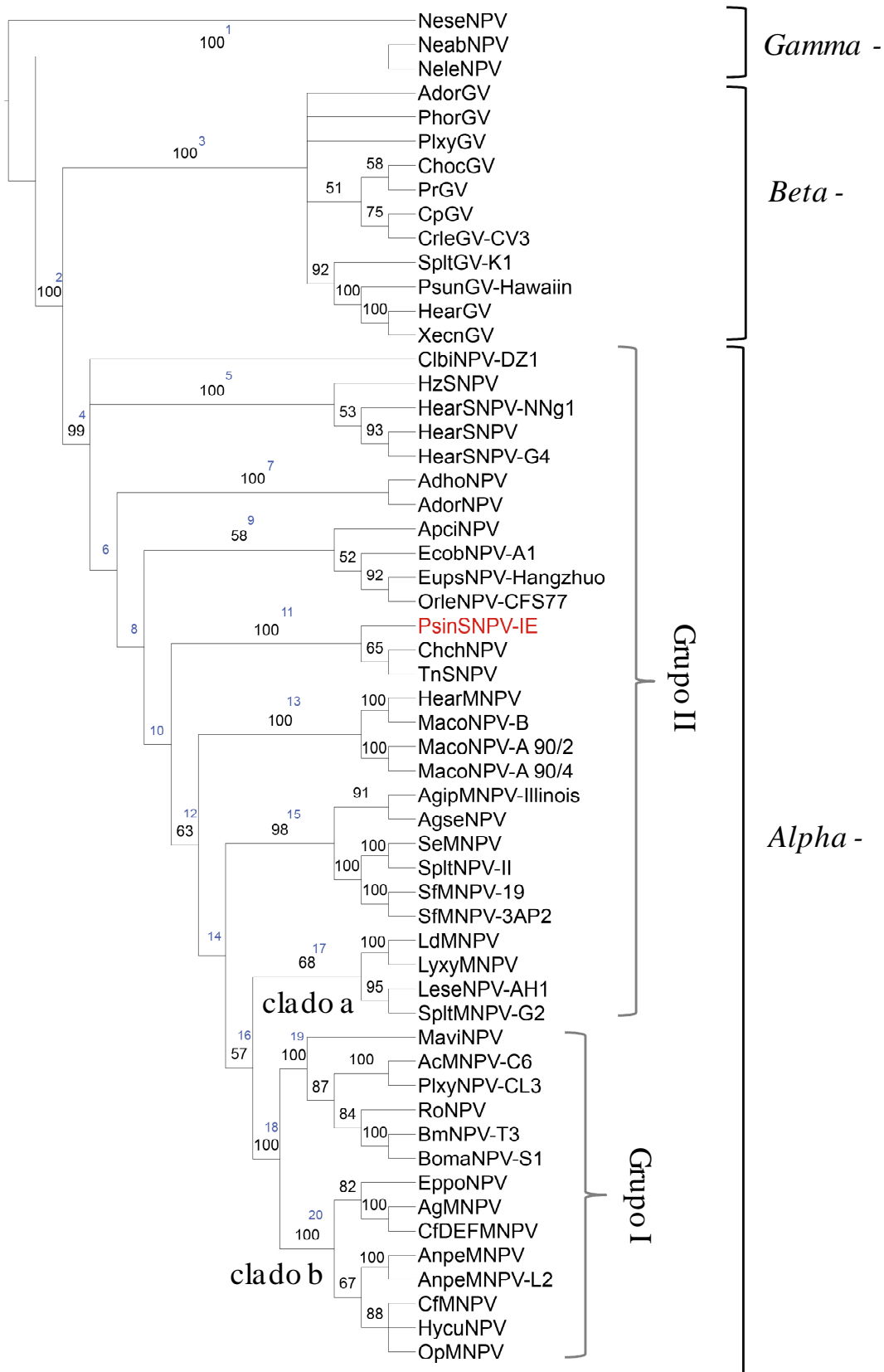


Figura17. Filogenia do vírus PsinSNPV com outros 57 baculovirus inferida da concatenação das sequências nucleotídicas dos genes *lef-8*, *polh/gran*, *lef-9* e *pif-2*. Árvore consenso estrito das seis árvores mais parcimoniosas obtidas da análise MP não ponderada. Os números acima dos ramos em preto representam os valores de *bootstrap* (> 50%) para 1000 replicatas e em azul indicam os ramos utilizados no teste PABA (Tabela 20). O vírus PsinSNPV-IE está assinalado em vermelho e NeseNPV foi utilizado como grupo externo.

No cladograma, o vírus PsinSNPV-IE se apresenta agrupado aos outros baculovirus do Grupo II dos *Alphabaculovirus* estando proximamente relacionado com os vírus ChchNPV e TnSNPV.

Para avaliar o conflito e a incongruência entre as partições utilizadas nesta análise, foram realizados os testes ILD (incongruência da diferença de comprimento e PABA (adição de partição alterando o *bootstrap*).

O teste ILD foi realizado para as diferentes combinações das partições do conjunto de dados formados pelos genes *lef-8 polh/gran*, *lef-9* e *pif-2*. Cada combinação das partições do conjunto de dados apresentou resultados significantes estatisticamente (Tabela 20), *P* valor < 0,05, indicando a incongruência entre os sinais filogenéticos dos genes analisados.

**Tabela 20. Probabilidade de significância da incongruência das diferentes combinações das partições de dados - teste ILD**

<b>Concatenação dos genes</b>	<b>Valor <i>P</i></b>
<i>lef-8, polh/gran, lef-9 pif-2 e gaps</i>	0,01
<i>lef-8, polh/gran, lef-9 e pif-2</i>	0,01
<i>lef-8, polh/gran e lef-9</i>	0,01
<i>polh/gran, lef-9 e pif-2</i>	0,01
<i>lef-8 e polh/gran</i>	0,01
<i>lef-9 e pif-2</i>	0,01
<i>polh/gran e lef-9</i>	0,01
<i>lef-8 e lef-9</i>	0,02
<i>lef-8 e pif-2</i>	0,03
<i>polh/gran e pif-2</i>	0,01

No teste PABA, a adição das partições do conjunto de dados favoreceu a predominância dos valores positivos da taxa de alteração do suporte dos ramos ( $\delta > 0$ ) e o aumento da média do valor de *bootstrap* dos 20 ramos selecionados e do somatório de todos os ramos da árvore (Tabela 21). Diferente do teste ILD, o teste PABA apresentou concordância entre as partições dos dados utilizadas na análise.

Tabela 21. Teste da adição de partição alterando o *bootstrap* (PABA) com base na árvore consenso estrito da análise de parcimônia não ponderada\*

Nós	Valor de <i>bootstrap</i> - análise do gene					Valor da alteração do suporte dos ramos ( $\delta$ )							
	<i>lef-8</i>	<i>polh/gran</i>	<i>lef-9</i>	<i>pif-2</i>	<i>gaps</i>	<i>polh/gran</i> (2°)		<i>lef-9</i> (3°)		<i>pif-2</i> (4°)		<i>gaps</i> (5°)	
	BS	BS	BS	BS	BS	BS	$\delta$	BS	$\delta$	BS	$\delta$	BS	$\delta$
1	99	100	100	99	96	100	+1	100	0	100	0	100	0
2	100	100	100	100	100	100	0	100	0	100	0	100	0
3	53	100	12	33	0	100	+47	100	0	100	0	100	0
4	99	65	99	49	0	95	-4	98	+3	99	+1	99	0
5	100	100	100	100	0	100	0	100	0	100	0	100	0
6	0	31	5	1	0	8	+8	37	+29	25	-12	24	-1
7	100	100	100	100	10	100	0	100	0	100	0	100	0
8	0	2	1	1	0	2	+2	10	+8	17	+7	17	0
9	5	1	69	1	0	8	+3	57	+49	59	+2	58	-1
10	2	19	2	0	0	27	+25	45	+18	36	-9	35	-1
11	100	99	100	100	0	100	0	100	0	100	0	100	0
12	3	0	1	3	0	25	+22	36	+11	62	+26	63	+1
13	100	95	100	100	0	100	0	100	0	100	0	100	0
14	0	0	0	6	0	14	+14	17	+3	36	+19	36	0
15	19	60	28	96	25	82	+63	86	+4	97	+11	98	+1
16	9	0	1	2	0	30	+21	45	+15	56	+11	57	+1
17	33	18	2	9	0	61	+28	50	-11	68	+18	68	0
18	90	2	99	89	0	97	+7	100	+3	100	0	100	0
19	100	0	97	100	0	88	-12	99	+11	100	+1	100	0
20	100	3	75	61	0	100	0	100	0	100	0	100	0
<b>Média</b>	55,6	44,75	54,55	52,5	11,55	66,85	+11,25	74	+11,35	77,75	+3,7	77,8	0
<b>Total nós</b>	3054	2595	2756	2505	535	3436	+225	3770	+227	3975	+74	3980	0

\* Árvore consenso estrito das 6 árvores mais parcimoniosas (Figura 18).

BS = valor de *bootstrap* para cada combinação das partições.  $\delta$  = alteração do valor de suporte *bootstrap* para cada adição sequencial da partição na ordem decrescente do número de caracteres informativos. Total nós = somatório do *bootstrap* de todos os ramos da árvore.

Para a Inferência Filogenética Bayesiana, o modelo de substituição nucleotídica que mais apropriado à sequência dos quatro genes analisados é o modelo GTR + I +G (*General Time Reversible* - GTR), modelo geral de reversão ao longo do tempo com taxa de variação entre os sítios *gamma* – distribuída e proporção de sítios invariáveis (comando MrBayes: nst=6 rates=invgamma). As Cadeias de Markov processadas para 2.260.000 gerações foram suficientes para que a diferença das frequências obtivesse desvio padrão menor que 0,01 ( $p = 0,009558$ ).

A árvore obtida da Inferência Bayesiana baseada na concatenação das sequências nucleotídicas dos quatro genes analisados apresentou o mesmo agrupamento presente na árvore MP (Figura 18) com ramos fortemente sustentados por probabilidade posterior acima de 0,9. A diferença encontrada entre os dois métodos é o valor de probabilidade posterior obtida da inferência Bayesiana superior ao valor de suporte dos ramos *bootstrap* obtida da análise da máxima parcimônia.

No gênero *Betabaculovirus*, a subdivisão em clado a (XnGV, HearGV, PsunGV, SpliGV e PlxyGV) e clado b (AdorGV, PhorGV, CpGV, CrleGV e ChocGV), proposta por Miele *et al.* (2011), foi observada apenas na árvore obtida pela Inferência Bayesiana. Entretanto, nessa árvore, o vírus PlxyGV forma um grupo parafilético a esses dois cladros.

Na árvore obtida pela Inferência Bayesiana, o vírus PsinSNPV – IE também está agrupado aos *Alphabaculovirus* – Grupo II.

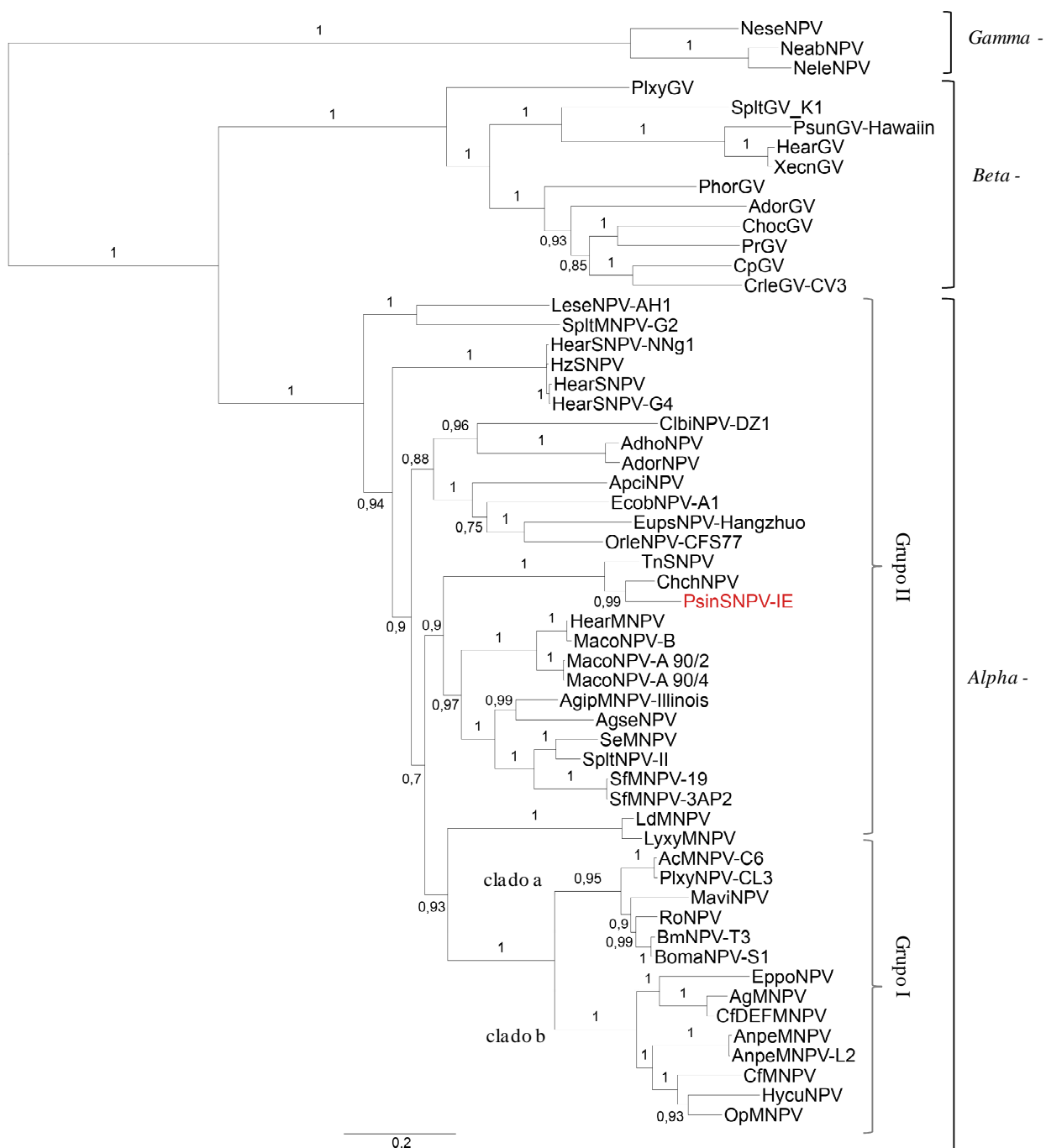


Figura 18. Filograma da concatenação das seqüências nucleotídicas dos genes *lef-8*, *polh/gran*, *lef-9* e *pif-2* representando a relação filogenética de *PsinSNPV-IE* com outros membros da família *Baculoviridae*. Árvore filogenética inferida da análise Bayesiana utilizando modelo GTR+G+I com os números nos ramos representando os valores de probabilidade posterior. O vírus *PsinSNPV-IE* está assinalado em vermelho e *NeseNPV* foi utilizado como grupo externo. Na árvore estão representados os gêneros *Alphabaculovirus* Grupo I e II (*Alpha-*), *Betabaculovirus* (*Beta-*) e *Gammabaculovirus* (*Gamma-*). A escala em barra representa o número de substituição nucleotídica para cada base.

As árvores filogenéticas obtidas da análise da Máxima Parcimônia baseadas nas sequências individuais, nucleotídica e peptídica, dos quatro genes *lef-8*, *lef-9*, *pif-2* e *polh/gran* (Apêndice III) apresentaram ramos com baixos valores de suporte (*bootstrap* < 75%), alta politomia e inconsistência no agrupamento dos baculovirus.

Entre os quatro genes analisados individualmente, apenas os genes *lef-8* e *polh/gran* apresentaram árvores com a clara presença dos quatro gêneros da família *Baculoviridae*. Entretanto, inconsistências nos agrupamentos foram encontradas: na árvore baseada no gene *lef-8*, foi observado um grupo monofilético formado por *Beta-* e *Gammabaculovirus* e, na árvore baseada no gene *polh/gran*, os vírus *Autographa californica* MNPV (AcMNPV) e *Plutella xylostella* MNPV (PlxyMNPV) estão presentes no Grupo II. As análises de genoma completo apresentam o vírus AcMNPV pertencendo ao Grupo I dos *Alphabaculovirus*. Essa discordância pode ser explicada pelo fato de que o gene da poliedrina de AcMNPV é um gene quimera que consiste do mosaico de sequências específicas tanto do Grupo I quanto do Grupo II (Jehle, 2004). Essa explicação pode ser correlacionada ao vírus PlxyMNPV por ser um vírus similar ao AcMNPV (Harrison, 2009).

Para a reconstrução da família *Baculoviridae* com a representação de seus quatro gêneros, árvores filogenéticas foram construídas com base na concatenação das sequências nucleotídicas e peptídicas dos genes *lef-8*, *lef-9* e *pif-2* (Apêndice III). A árvore MP obtida da concatenação da sequência peptídica desses três genes (Figura 20) apresentou cinco grupos: *Alphabaculovirus* Grupo I e II, *Betabaculovirus*, *Gammabaculovirus* e



*Deltabaculovirus*. Estes grupos estão bem definidos e fortemente sustentados com *bootstrap* acima de 95%.

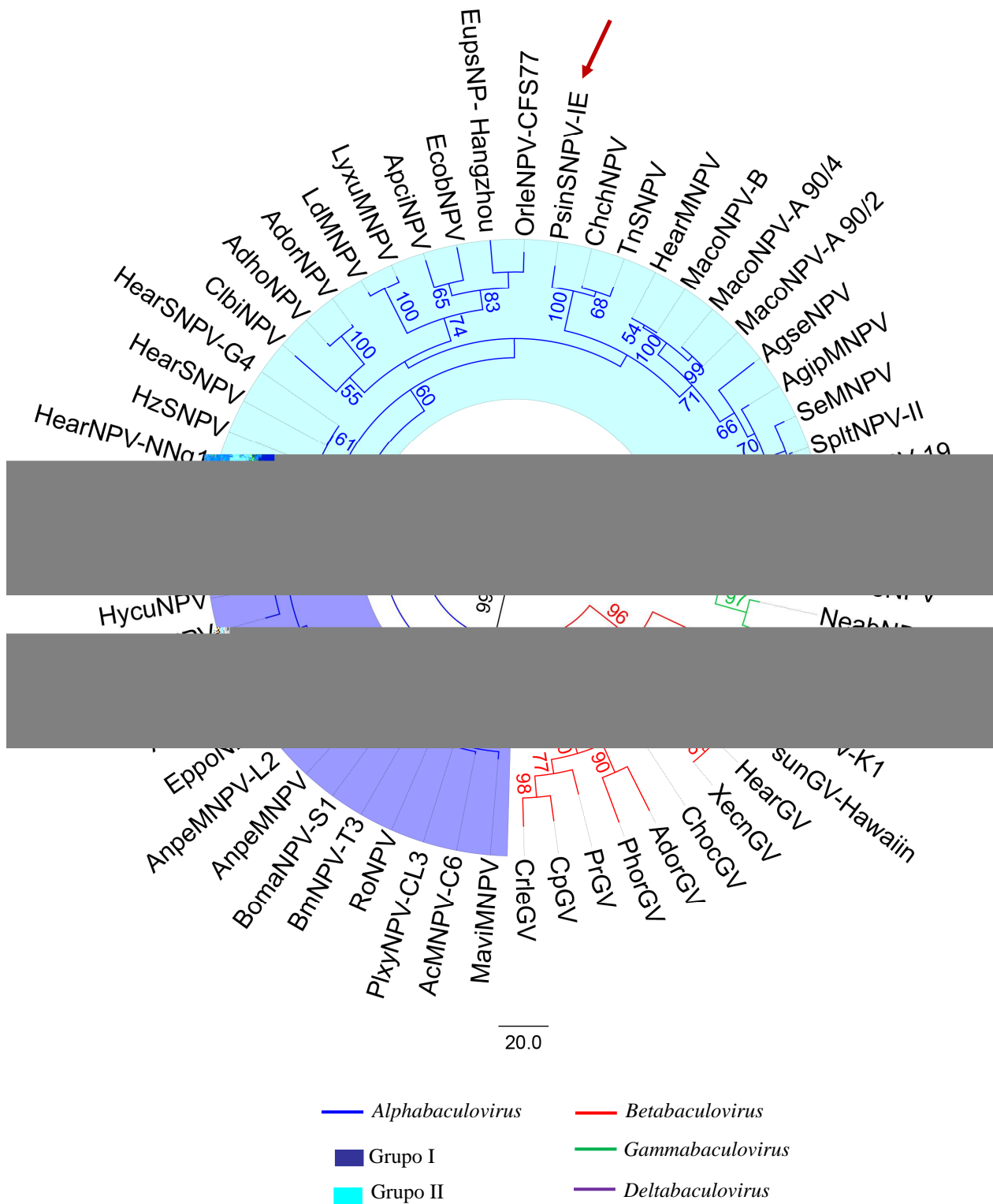


Figura 19. Reconstrução filogenética da família *Baculoviridae* incluindo o vírus *PsinSNPV-IE*. Árvore filogenética mais parcimoniosa baseada nas sequências peptídicas LEF-8, LEF-9 e PIF-2. Os números nos ramos representam os valores de *bootstrap* (< 50%) para 1000 replicatas. O vírus *PsinSNPV-IE* está indicado pela seta vermelha. O vírus *CuniNPV* foi utilizado como grupo externo. A escala em barra representa o número de substituição para cada sítio.

## 6. Discussão

O presente trabalho teve como principal foco de interesse estudar o vírus *Pseudoplusia includens* SNPV (PsinSNPV), investigando as variações genéticas entre sete isolados desse vírus (IA a IG) e a sua relação filogenética com os outros membros da família *Baculoviridae*.

Em um estudo de parceria entre a Embrapa Recursos Genéticos e Biotecnologia e Embrapa Soja, esses mesmos isolados (PsinSNPV IA a IG) se apresentaram diferenciados com relação aos perfis de restrição de DNA, com a presença de bandas submolares, e quanto à atividade biológica, destacando os isolados IA, IE e IF como os mais virulentos (Alexandre *et al.*, 2010).

Variações genéticas como essas são comumente encontradas entre isolados da mesma espécie de baculovirus coletados em diferentes safras e/ou localizações geográficas. Polimorfismos dos fragmentos de restrição (RFLPs), como os detectados nos isolados de PsinSNPV (Alexandre *et al.*, 2010), são utilizados em análises de variações genéticas presentes em diferentes populações de baculovirus (Cory *et al.*, 2005; Hitchman *et al.*, 2007; Figueiredo *et al.*, 2009).

Com base no exposto, foi realizado um estudo detalhado e aprofundado para um maior conhecimento e caracterização do vírus PsinSNPV. Inicialmente, foi realizada a análise, por algoritmo UPGMA, dos polimorfismos encontrados nos perfis de restrição de DNA dos isolados de PsinSNPV (IA a IG) digeridos com as enzimas *Bam*HI, *Eco*RI, *Hind*III e *Pst*I. Mesmo com a presença de alta similaridade (CD=0,92), os isolados estão distribuídos em grupos distintos sugerindo a existência de heterogeneidade entre eles.

Na busca de outra estratégia para uma melhor elucidação das variações genéticas, genes altamente conservados entre os baculovirus foram utilizados como alvo da amplificação por PCR e do estudo filogenético. Essa metodologia foi utilizada com grande sucesso em estudos de identificação e distinção de isolados de *Cydia pomonella* GV (CpGV) e *Spodoptera frugiperda* MNPV (SfMNPV) (Eberle *et al.*, 2009; Rowley *et al.*, 2010).

A relação filogenética entre os isolados de PsinSNPV (IA a IG) foi realizada por meio da análise dos genes conservados *lef-8*, *lef-9*, *pif-2* e *polh*. Mas, considerando que o gene *phr* é conservado no subgrupo dos *Alphabaculovirus* do Grupo II que infectam insetos da subfamília Plusiine (Xu *et al.*, 2008; Biernat *et al.*, 2011) e que o baculovirus PsinSNPV pertence a esse subgrupo (Craveiro *et al.*, 2010), estudos foram conduzidos para identificação desse gene no genoma dos sete isolados de PsinSNPV para sua posterior utilização na análise das relações filogenéticas entre os isolados de PsinSNPV (IA a IG) . Esse gene é pouco conservado entre os baculovirus podendo funcionar como um forte sinal filogenético para a obtenção de uma melhor resolução das relações entre indivíduos de uma mesma espécie, como é o caso dos isolados de PsinSNPV.

Dessa forma, a análise das sequências parciais dos cinco genes (*lef-8*, *lef-9*, *phr*, *pif-2* e *polh*) dos sete isolados de PsinSNPV (IA a IG) foi realizada e poucas variações entre os isolados foram encontradas. Dessas variações, o gene *pif-2* foi o que apresentou maior número de polimorfismos, embora a expectativa fosse para o gene *phr*. Dentre os 31 genes conservados, o *pif-2* possui a sequência ancestral mais conservada entre os baculovirus (Miele *et al.*, 2011). Em contraste, este gene, entre os cinco genes analisados,

apresentou maior variação entre diferentes isolados de uma mesma espécie de vírus, o baculovirus PsinSNPV. A presença de polimorfismos nas sequências parciais dos cinco genes analisados corrobora com a hipótese de Alexandre *et al.*, (2010) da existência de variantes genotípicos entre os isolados. A mistura de genótipos ocorrendo naturalmente no campo pode constituir uma vantagem no desenvolvimento de bioinseticida com base em baculovirus (Simón *et al.*, 2005).

A análise filogenética desses cinco genes apresentou árvores filogenéticas robustas com os isolados de PsinSNPV distribuídos em dois grupos monofiléticos. As análises, fenética e cladística, aqui realizadas, mostram que os sete isolados são diferentes linhagens do mesmo baculovirus (PsinSNPV) e indicam que ocorre uma diversidade genética intraespecífica entre os isolados estudados. Entretanto, foi observada uma discordância entre o agrupamento baseado nos RFLPs e o agrupamento obtido da análise filogenética com base na sequência parcial dos quatro genes conservados e no gene *phr*. A análise hierárquica de grupo utilizando RFLPs é uma técnica que percorre o genoma completo do vírus, mas torna-se limitada por identificar apenas as variações presentes nos sítios de ligação das enzimas (Baillie e Bower, 2011), por isso, a discordância observada pode ter ocorrido por haver poucos marcadores moleculares utilizados na análise por UPGMA. Utilizando esse método, seria necessário maiores esforços para avaliar a variação genética entre os isolados. Uma forma, seria a utilização de um maior número de enzimas de restrição possibilitando um aumento do número de caracteres variáveis que auxiliarão a análise em questão.

A utilização de oligonucleotídeos degenerados específicos aos genes *lef-8*, *lef-9* e *polh/gran* é uma ferramenta de comprovada eficiência para a identificação e determinação da inferência filogenética de baculovirus não caracterizados (Lange *et al.*, 2004; Jehle *et al.*, 2006b; Jakubowska *et al.*, 2007; Eberle *et al.*, 2009; Rowley *et al.*, 2010). Estudos utilizando como alvo os genes *lef-8* e *pif-2* também foram eficientes na classificação e identificação de baculovirus (Herniou *et al.*, 2004; Jakubowska *et al.*, 2007).

As árvores filogenéticas obtidas da concatenação dos genes *lef-8*, *lef-9*, *pif-2* e *polh/gran* (Figuras 18 e 19), apresentaram topologia compatível com as árvores produzidas pela concatenação dos genes conservados de 58 baculovirus (Miele *et al.*, 2011). Esta observação comprova a eficiência da utilização desses quatro genes conservados que representam uma ferramenta confiável e robusta para a identificação e classificação dos baculovirus.

Nas árvores obtidas da análise individual dos genes *lef-8*, *lef-9*, *pif-2* e *polh/gran* podemos observar diferentes topologias indicando a heterogeneidade do sinal filogenético contido entre o conjunto de dados formado pelos quatro genes. A inconsistência e discordância entre as árvores refletem uma inferência filogenética imprecisa ou a falta de um sinal filogenético robusto (Herniou *et al.*, 2003), sendo esta última hipótese a mais provável para o presente estudo. A utilização de um único gene, é insuficiente para a definição de uma espécie (Herniou *et al.*, 2004). Herniou *et al.* (2001) ressaltam a importância da utilização de dois ou mais genes para a inferência filogenética dos baculovirus. Assim a forte discordância observada entre as partições dos dados podem indicar que os quatro genes analisados possuem diferentes histórias evolutivas.

Entretanto, a árvore construída com base na sequência concatenada dos quatro genes apresentou os quatro grupos da família *Baculoviridae* bem definidos e suportados por altos valores de *bootstrap* e probabilidade posterior.

Foram realizados testes para avaliação da congruência entre as partições do conjunto de dados (genes *lef-8*, *lef-9*, *pif-2* e *polh/gran*) utilizados na inferência filogenética. O teste incongruência da diferença de comprimento (ILD) apresentou valores significantes estatisticamente indicando incongruência entre as partições.

O teste ILD é o mais utilizado para indicar a significância estatística entre as partições de dados, porém ele não mede a contribuição de cada partição, não avalia se uma dada partição pode ser ignorada ou receber um peso menor na análise, não revela qual taxa ou nó causa a incongruência e não indica se outros fatores que não a incongruência entre as partições podem gerar valores significativos para ILD (Gatesy *et al.*, 1999; Salle e Brower, 1997; Hipp *et al.*, 2004; Struck *et al.*, 2006).

Embora seja amplamente utilizado, foi questionada a validade do teste ILD como critério de congruência e permutabilidade (Cunningham, 1997; Barker e Lutzoni, 2002). Críticas ao teste ILD basearam-se na observação que, em muitas vezes, as partições dos dados foram identificadas como incongruentes quando na verdade essas partições se combinavam para produzir uma estimativa precisa da filogenia dos organismos (Yoder *et al.*, 2001). Um exemplo é o resultado deste estudo que, apesar de ter encontrado valor *P* significativo para a partição dos dados, apresentou uma inferência filogenética robusta para a concatenação dos quatro genes.

Estudos rejeitam a hipótese do teste ILD como medida imparcial sobre congruência filogenética sugerindo que o valor  $P$  significativo não deve ser tomado como uma demonstração conclusiva de que a análise da partição dos dados irá produzir filogenias enganosas (Dolphin *et al.*, 2000; Yoder *et al.*, 2001; Barker e Lutzoni, 2002; Darlu e Lecointre, 2002; Dowton e Austin, 2002).

Assim, o teste PABA (adição de partição alterando o bootstrap ) foi realizado para avaliar se as partições possuem sinais filogenéticos congruentes ou conflitantes. O teste PABA apresentou valores crescentes de *bootstrap* a medida que foi adicionado mais um gene à análise ( $\delta$  positivo) demonstrando congruência entre as sequências concatenadas.

As árvores robusta inferidas da sequência concatenada dos quatro genes analisados e o valor positivo da alteração do suporte dos ramos (*bootstrap*) obtido pelo teste PABA confirmam a necessidade da utilização de dois ou mais genes para realização da análise filogenética. A análise filogenética baseada no genoma completo ou na concatenação de dois ou mais genes apresenta resultados mais consistentes, pois cada gene contribuirá para o sinal filogenético geral e um efeito sinérgico será produzido pela combinação dos sinais correspondentes a cada gene (Herniou *et al.*, 2003).

## 7. Conclusões e Perspectivas

A utilização das sequências parciais dos genes *lef-8*, *lef-9*, *pif-2*, *phr* e *polh/gran* permitiram de forma prática a caracterização, identificação e o estabelecimento das relações filogenéticas entre os isolados virais e entre o vírus PsinSNPV e outros membros da família *Baculoviridae*.

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

- ✓ Os isolados de PsinSNPV distribuídos em grupos filogenéticos distintos, mostram a existência de diversidade genética intraespecífica nesse vírus, constituindo como uma vantagem para o uso de PsinSNPV no controle biológico.
- ✓ O gene *pif-2* de PsinSNPV exibe alta variabilidade genética, contrariamente ao descrito até então na literatura.
- ✓ O vírus PsinSNPV ocupa a posição filogenética de *Alphabaculovirus* do Grupo II, proximalmente relacionado com os vírus TnSNPV e ChchNPV.
- ✓ A presença do gene *phr* no genoma de PsinSNPV confirma a proposta de Xu *et al.* (2008) da conservação deste gene dentro os *Alphabaculovirus* do Grupo II que infectam insetos da subfamília Plusiinae.



Em relação aos resultados e metodologias utilizadas também podemos concluir que:

- ✓ Os resultados obtidos da análise conjunta dos genes foram confirmados com dados presentes na literatura apresentando uma inferência filogenética precisa para os 58 baculovirus analisados.
- ✓ A árvore obtida da concatenação dos genes *lef-8*, *lef-9* e *pif-2* mostra que, neste caso, apenas três dos quatro genes estudados foram suficientes para a obtenção de uma análise filogenética robusta dos baculovirus.
- ✓ A análise individual dos genes foi insuficiente para resolver as relações filogenéticas da família *Baculoviridae*, o que reforça a observação de que, análises filogenéticas baseadas em genomas completos ou na concatenação de dois ou mais genes apresentam resultados mais consistentes.
- ✓ A técnica RFLP, da forma que foi utilizada no presente estudo, apresentou limitações, pois foram obtidos poucos caracteres variáveis o que pode ter comprometido a análise das relações filogenéticas entre os sete isolados de PsinSNPV (IA a IG).
- ✓ Os métodos da máxima parcimônia e inferência Bayesiana foram complementares no estudo da filogenia de baculovirus, pois o conjunto dos dados obtidos permitiram resultados mais conclusivos.

Este trabalho tem como perspectivas de estudo imediatas:

- ✓ A investigação da alta variabilidade genética inesperada obtida da análise do gene *pif-2* de PsinSNPV.
- ✓ A realização do sequenciamento e análise do genoma completo dos sete isolados de PsinSNPV.

Finalizando, os dados apresentados neste trabalho sobre a filogenia e diversidade do baculovirus PsinSNPV, aliados a outras informações biológicas e moleculares, contribuirão para maior entendimento das relações desse vírus com o inseto hospedeiro e com os outros integrantes da família *Baculoviridae*, além de auxiliar em futuros trabalhos que tem como perspectivas o desenvolvimento e a produção de um bioinseticida eficaz para o controle biológico da lagarta falsa-medideira, atualmente uma das principais pragas da cultura da soja no Brasil.

## 8. Referências Bibliográficas

- Abascal, F., Zardoya, R., Posada, D., 2005. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* 21: 2104–2105.
- Adams, J.R., Mcclintock, T.J., 1991. Baculoviridae: Nuclear Polyhedrosis Virus. In: Adams, J.R.; Bonani, J.R. (Eds.), Atlas of Invertebrate Viruses. Boca Raton: Crc, pp. 87-204.
- Afonso C.L., Tulman E.R., Lu Z., Balinsky C.A., Moser B.A., Becnel J.J., Rock D.L., Kutish G.F., 2001. Genome sequence of a baculovirus pathogenic for *Culex nigripalpus*. *Journal of Virology* 75: 11157–11165.
- Ahrens C.H., Russell R.L.Q., Funk C.J., Evans J.T., Harwood S.H., Rohrmann G.F., 1997. The sequence of the *Orgyia pseudotsugata* multinucleocapsid nuclear polyhedrosis virus genome. *Virology* 229: 381–399.
- Akaike, H., 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19: 716–723.
- Alexandre, T.M., Ribeiro, Z.M.A., Craveiro, S.R., Cunha, F., Fonseca, I.C.B., 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: 98-104.
- Alford, A.R., Hammond Junior, A.M., 1982. Plusiinae (Lepidoptera: Noctuidae) populations in Louisiana soybean ecosystems as determined with looplure-baited traps [*Pseudoplusia includens*, *Rachiplusia ou*, *Trichoplusia spp.*]. *Journal of Economic Entomology* 75: 647-650.
- Ali, A., Young, S., 1991. Influence of larval age and temperature on effectiveness of a nuclear ployhedrosis virus in the soybean looper, *Pseudoplusia includens* (Lepidoptera: Noctuidae) on soybean. *Biological Control* 1: 334-338.
- Altschul, S.F., Madden, T.L., Schaffer, A.A., Altschul, S.F., 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: 3389-402.
- 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: 586–605.

- Baillie, V.L., Bouwer, G., 2011. High levels of genetic variation within core *Helicoverpa armigera nucleopolyhedrovirus* genes. *Virus Genes* 44: 149-162.
- Barker, F.K., Lutzoni, F.M., 2002. The Utility of the Incongruence Length Difference Test. *Systematic Biology* 51: 625–637.
- Bellotti, A. C. 1999. Recent advances in cassava pest management. *Annual Review of Entomology* 44: 345-370.
- Bennett, C.J., Webb, M., Willer, D.O., Evans, D.H., 2003. Genetic and phylogenetic characterization of the type II cyclobutane pyrimidine dimer photolyases encoded by Leporipoxviruses. *Virology* 315: 10-19.
- 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: 2531–2544.
- Biernat, M.A., Ros, V.I., Vlak, J.M., van Oers, M.M., 2011. Baculovirus cyclobutane pyrimidine dimer photolyases show a close relationship with lepidopteran host homologues. *Insect Molecular Biology* 20: 457-64.
- Bilimoria, S.L., 1991. The biology of nuclear polyhedrosis viruses. In: Kurstak, E. (Ed.), *Viruses of Invertebrates*. New York: Marcel Dekker, pp. 1-72.
- Black, B.C., Brennan, L.A., Dierks, P.M., Gard, I.E., 1997. Commercialization of baculoviral insecticides. In: Miller, L.K. (Ed.), *The Baculoviruses*. Plenum Press, New York, pp. 341–387
- Blissard G.W., 1996. Baculovirus-insect cell interactions. *Cytotechnology* 20: 73-93.
- Blissard, G.W., Rohrmann, G.F., 1990. Baculovirus Diversity and Molecular Biology. *Annual Review of Entomology* 35: 127-155.
- Borchsenius F., 2007. FastGap 1.0.8. Software distributed by the author ([http://192.38.46.42/aubot/fb/FastGap\\_home.htm](http://192.38.46.42/aubot/fb/FastGap_home.htm)).
- Boucias, D., Pendland, J., 1998. Principles of insect pathology. Norwell-Massachusetts : Kluwer Academics Publishers, 550p.
- Braunagel, S.C., Russell, W.K., Rosas-Acosta, G., Russell, D.H., Summers, M.D., 2003. Determination of the protein composition of the occlusion-derived virus of *Autographa californica nucleopolyhedrovirus*. *Proceedings of the National Academy of Sciences* 100: 9797–9802.

- 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). *Biological Control* 38: 389-394.
- Bueno, R.C.O.F., Parra, J.R.P., Bueno, A.F., Moscardi, F., Di Oliveira, J.R.G., Camillo, M.F., 2007. Sem Barreira. *Cultivar Grandes Culturas* 93: 12-15.
- Bulach D.M., Kumar C.A., Zaia A., Liang, B., Tribe, D.E., 1999. Group II nucleopolyhedrovirus subgroups revealed by phylogenetic analysis of polyhedrin and DNA polymerase gene sequences. *Journal Invertebrate Pathology* 73: 59-73.
- Chen Y.R., Wu C.Y., Lee S.T., Wu Y.J., Lo C.F., Tsai M.F., Wang C.H., 2008. Genomic and host range studies of *Maruca vitrata nucleopolyhedrovirus*. *Journal of General Virology* 89: 2315–2330.
- Chen, X., WF, I.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: 241–257.
- Chen, X., Zhang, W.J., Wong, J., Chun, G., Lu, A., McCutchen, B.F., Presnail, J.K., Herrmann, R., Dolan, M., Tingey, S., Hu, Z.H., Vlak, J.M., 2002. Comparative analysis of the complete genome sequences of *Helicoverpa zea* and *Helicoverpa armigera* single-nucleocapsid nucleopolyhedroviruses. *Journal of General Virology* 83: 673– 684.
- Cheng, X.W., Carner, G.R., Lange, M., Jehle, J.A., Arif, B.M., 2005. Biological and molecular characterization of a multicapsid nucleopolyhedrovirus from *Thysanoplusia orichalcea* (L.) (Lepdoptera: Noctuidae). *Journal of Invertebrate Pathology* 88: 126-135.
- Coelho, A.S.G., 2002. BOOD: avaliação de dendrogramas baseados em estimativa de distâncias/similaridades genéticas através do procedimento de bootstrap. Goiânia: UFG, 1 disquete.
- 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: 1662–1667.

- Copping, L.G., Menn, J.J., 2000. Biopesticides: A review of their action, applications and efficacy. *Pest Management Science* 56: 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: 101-111.
- Craveiro, S.R., Ribeiro, Z.M.A., Alexandre, T.M., Moscardi, F., Castro, M.E.B., 2009. Infectividade e análise molecular de sete isolados de PsinSNPV. Brasília: Embrapa Recursos Genéticos e Biotecnologia, Boletim de Pesquisa e Desenvolvimento 230.
- Craveiro, S.R., Ribeiro, Z.M.A., Inglis, P.W., Castro, M.E.B., 2010. Relações filogenéticas do *Pseudoplusia includens single nucleopolyhedrovirus* isolado IE (PsinSNPV-IE). Brasília: Embrapa Recursos Genéticos e Biotecnologia, Boletim de Pesquisa e Desenvolvimento 230.
- Craveiro, S.R.; Alexandre, T.M.; Ribeiro, Z.M.A.; Moscardi, F.; Castro, M.E.B., 2008. Caracterização de isolados virais patogênicos à lagarta falsa-medideira (*Pseudoplusia includens*). Brasília: Embrapa Recursos Genéticos e Biotecnologia, Boletim de Pesquisa e Desenvolvimento, 230 15 p.
- Cunningham, C.W., 1997. Can three incongruence tests predict when data should be combined? *Molecular Biology Evolution* 14: 733–740.
- Dalhus, B., Laerdahl, J.K., Backe, P.H., Bjoras, M., 2009. DNA base repair and checkpoint signaling. *Nature Review. Molecular Cell Biology* 3(5): 317-327.
- Darlu, P., Lecointre, G., 2002. When does the incongruence length difference test fail? *Molecular Biology Evolution* 19: 432–437.
- 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: 929–943.
- de Salle, R., Brower, A.V.Z., 1997. Process partition, congruence, and the independence of characters: Inferring relationships among closely related Hawaiian *Drosophila* from multiple gene regions. *Systematic Biology* 46: 751-764.
- de Souza, M.L., de Castro, M.E.B., Moscardi, F., 2007. Virus for biological control of insects. In: Pimentel, D. (Ed.), *Encyclopedia of Pest Management*. CRC Press, Boca Raton, 2: 689–692.

- Deisenhofer, J., 2000. DNA photolyases and cryptochromes. *Mutation Resistance* 460: 143–149.
- Dice, L.R., 1945. Measures of the amount of ecologic association between species. *Ecology* 26: 297-302.
- Dolphin, K., Belshaw R., Orme C.D.L., Quicke, D.L.J., 2000. Noise and incongruence: Interpreting results of the incongruence length difference test. *Molecular Phylogenetic Evolution* 17: 401–406.
- Dowton, M., Austin, A.D., 2002. Increased congruence does not necessarily indicate increased phylogenetic accuracy – the behavior of the incongruence length difference test in mixed-model analyses. *Systematic Biology* 51: 19-31.
- Duffy S.P., Young A.M., Morin B., Lucarotti C.J., Koop B.F., Levin D.B., 2006. Sequence analysis and organization of the *Neodiprion abietis* nucleopolyhedrovirus genome. *Journal of Virology* 80: 6952–6963.
- Eberle, K.E., Sayed, S., Rezapanah, M., Shojai-Estrabagh, S., Jehle, J.A., 2009. Diversity and evolution of the *Cydia pomonella granulovirus*. *Journal of General Virology* 90: 662-671.
- Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32:1792-1797.
- Eichlin, T.D., Cunningham, H.B., 1978. The Plusiinae (Lepidoptera:Noctuidae) of America North of México, emphasizing genitalic and larval morphology. USDA Agricultural Research Service, Washington, 222p. [Technical Bulletin n.1567].
- Eker, A.P., Quayle, C., Chaves, I., Van Der Horst, G.T., 2009. DNA repairs in mammalian cells: Direct DNA damage reversal: elegant solution for nasty problems. *Cell Molecular Life Science* 66: 968-980.
- Engelhard, E.K., Kammorgan, L.N.W., Washburn, J.O., Volkman, L.E., 1994. The Insect Traqueal System - A conduit for the systemic spread of *Autographa californica* M. Nuclear Polyhedrosis. *Proceedings of the National Academy of Science* 91: 3224-3227.
- Erlandson, M., 2008. Insect pest control by viruses. In: Mahy, B.W.J.; van Regenmortel, M.H.V. (Eds.), *Encyclopedia of Virology*. Oxford, Elsevier 3:125-133.

- Escasa S.R., Lauzon H.A.M, Mathur A.C., Krell P.J., Arif B.M., 2006. Sequence analysis of the *Choristoneura occidentalis granulovirus* genome. *Journal of General Virology* 87: 1917–1933.
- Essen L.O., Klar T., 2006. Light-driven DNA repair by photolyases. *Cell Molecular Life Science* 63: 1266–1277.
- Evans, H.F., 1986. Ecology & epizootiology of baculoviruses. In: Granados, R.R.; Federici, B.A. (Eds.), *The biology of baculoviruses*. Boca Raton: CRC, 2: 89-132.
- Fan, Q., Li, S., Wang, L., Zhang, B., Ye, B., Zhao, Z., Cui, L., 2007. The genome sequence of the multinucleocapsid nucleopolyhedrovirus of the Chinese oak silkworm *Antheraea pernyi*. *Virology* 366: 304-315.
- Fang, M. G., Nie, Y. C., Wang, Q., Deng, F., Wang, R. R., Wang, H. Z., Wang, H. L., Vlak, J. M., Chen, X. W., Hu, Z. H., 2006. Open reading frame 132 of *Helicoverpa armigera nucleopolyhedrovirus* encodes a functional per os infectivity factor (PIF-2). *Journal General Virology* 87: 2563–2569.
- Fang, M., Nie, Y., Harris, S., Erlandson, 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: 12569–12578.
- Farris, J.S., Källersjö, M., Kluge, A.G., Bult, C., 1995a. Constructing a significance test for incongruence. *Systematic Biology* 44: 570–572.
- Farris, J.S., Källersjö, M., Kluge, A.G., Bult, C., 1995b. Testing significance of incongruence. *Cladistics* 10: 315–319.
- 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: 3091–3100.
- Fauquet, C.M., Mayo, M.A., Maniloff, J., Desselberger, U., Ball, L.A., 2005. *Virus Taxonomy - 8th Reports of the International Committee on Taxonomy of Viruses*. Academic Press, pp. Granados, R.R., Williams, K.A., 1986. *Vivo Infection and Replication of Baculoviruses*. In: Granados, R.R.; Federici, B.A. (Eds.), *The Biology of Baculoviruses*. Boca Raton 1: 89-108.
- Federici, B.A. 1997. Baculovirus pathogenesis. In Miller, L.K., (Ed.), "The Baculoviruses". Plenum Press, New York.



- Felsenstein, J., 1978. The number of evolutionary trees. *Systematic Zoology* 27: 27-33.
- Figueiredo, E., Munõz, D., Murillo, R., Mexia, A., Caballero, P., 2009. Diversity of Iberian nucleopolyhedrovirus wild-type isolates infecting *Helicoverpa armigera* (Lepidoptera: Noctuidae). *Biological Control* 50: 43-49.
- Flipsen J.T.M., Martens J.W.M., van Oers M.M., et al. 1995. Passage of *Autographa californica* nuclear polyhedrosis-virus through the midgut epithelium of *Spodoptera-exigua* larvae. *Virology* 208: 328-335.
- Friesen P.D., 2007. Insect Viruses. In: Knipe D.M.; Howley P.M.; Griffin, D.E.; Lamb, R.A.; Martin, M.A.; Roizman, B.; Straus S.E. (Eds.), *Fields Virology*. Philadelphia, Lippincott Williams e Wilkins 5th ed, 1: 707-736.
- Friesen, P.D., Miller, L.K., 2001. Insect viruses. In: Knipe, D.M.; Howley, P.M.; Griffin, D.E.; Martin, M.A.; Lamb, R.A.; Roizman, B.; Straus, S.E. (Eds.), *Fundamental Virology*. Philadelphia, Lippincott Williams e Wilkins Press, 4: 443–472.
- Fritsch, E., Undorf-Spahn, K., Kienzle, J., Zebitz, C. P. W., Huber, J., 2005. Codling moth granulovirus: First indications of variations in the susceptibility of local codling moth populations. *Nachrichtenbl. Deut. Pflanzenschutz* 57: 29-34.
- Funk, C.J., Braunagel, S.C., Rohrmann, G.F., 1997. Baculovirus structure. In: Miller L.K. (Ed.), *The Baculovirus*. New York: Plenum Press, pp.7-32.
- Garcia-Maruniak, A., Maruniak J. E., Zanutto P. M. A., Doumbouya A.E., Liu J.C., Merritt T.M., Lanoie J.S., 2004. Sequence analysis of the genome of the *Neodiprion sertifer* nucleopolyhedrovirus. *Journal of Virology* 78: 7036–7051.
- Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., Bairoch, A., 2003. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research* 31: 3784–3788.
- Gatesy, J., O'Grady, P., Baker, R.H., 1999. Corroboration among data sets in simultaneous analysis: Hidden support for phylogenetic relationships among higher level Artiodactyla taxa. *Cladistics* 15: 271-313.
- Gauvreau, K., Pagano, M., 1994. Why 5%? *Nutrition* 10: 93-4.

- Gomi S., Majima K., Maeda S., 1999. Sequence analysis of the genome of *Bombyx mori nucleopolyhedrovirus*. *Journal of General Virology* 80: 1323–1337.
- Granados, R.R., Williams, K.A., 1986. Vivo Infection and Replication of Baculoviruses. In: Granados, R.R., Federici, B.A. (Eds.), *The Biology of Baculoviruses*. Boca Raton 1: 89-108.
- Gross, C.H., Shuman, S., 1998. Characterization of a baculovirus-encoded RNA 5'-triphosphatase. *Journal of Virology* 72: 7057-7063.
- Guarino, L.A. Xu, B., Jin, J., Dong, W., 1998. A virus-encoded RNA polymerase purified from baculovirus-infected cells. *Journal of Virology* 72: 7985-7991.
- 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* 60:224–229.
- Guarino, L.A., Summers, M.D., 1986. Interspersed homologous DNA of *Autographa californica* nuclear polyhedrosis virus enhances delayed-early gene expression. *Journal of Virology* 60:215–223.
- Hall, B.G., 2001. *Phylogenetics trees made easy. A how to manual for molecular biologists*. Massachusetts: Sinauer Associates, pp. 179.
- Hall, T.A., 1999. Bioedit: A user-friendly biological sequence alignment editor and analysis program for windows 95/98/nt. *Nucleic Acids Symposium Series* 41: 95-98.
- Harrison R.L., Popham H.J.R., 2008. Genomic sequence analysis of a granulovirus isolated from the Old World bollworm, *Helicoverpa armigera*. *Virus Genes* 36: 565–581.
- 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: 775–790.
- Harrison, R.L., 2009. Structural divergence among genomes of closely related baculoviruses and its implications for baculovirus evolution. *Journal of Invertebrate Pathology* 101: 181-186.
- Harrison, R.L., Bonning, B.C., 2003. Comparative analysis of the genomes of *Rachiplusia ou* and *Autographa californica* multiple nucleopolyhedroviruses. *Journal of General Virology* 84: 1827-1842.

- Harrison, R.L., Lynn, D.E., 2007. Genomic sequence analysis of a nucleopolyhedrovirus isolated from the diamondback moth, *Plutella xylostella*. *Virus Genes* 35: 857-873.
- 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 General Virology* 91: 1173–1182.
- Hashimoto Y., Hayakawa T., Ueno Y., Fujita T., Sano Y., Matsumoto T., 2000. Sequence analysis of the *Plutella xylostella granulovirus* genome. *Virology* 275: 358–372.
- 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: 277–297.
- Hayakawa, T., Rohrmann, G.F., Hashimoto, Y., 2000. Patterns of genome organization and content in lepidopteran baculoviruses. *Virology* 278: 1–12.
- Herniou E.A., Luque T., Chen X., Vlak, J.M., Winstanley, D., Cory, J.S., O'Reilly D.R., 2001. Use of whole genome sequence data to infer baculovirus phylogeny. *Journal of Virology* 75: 8117-8126.
- 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: 3244-3251.
- Herniou, E.A., Jehle, J.A., 2007. Baculovirus phylogeny and evolution. *Current Drug Targets* 8: 1043-1050
- Herniou, E.A., Olszewski, J.A., Cory, J.S., O'Reilly, D., 2003. The genome sequence and evolution of baculoviruses. *Annual Reviews Entomology* 48: 211–34.
- Herzog, D.C., 1980. Sampling soybean looper on soybean. In: Kogan, M. and Herzog, D.C. (Eds.). *Sampling methods in soybean entomology*. Springer-Verlag, New York, pp. 140-168.
- Hillis, D.M., Bull, J.J., 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Systematic Biology* 42: 182-192.

- Hilton S., Winstanley D., 2008. Genomic sequence and biological characterization of a nucleopolyhedrovirus isolates from the summer fruit tortrix, *Adoxophyes orana*. *Journal of General Virology* 89: 2898–2908.
- 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: 1496–1504.
- Hipp, A.L., Hall, J.C., Systma, K.J., 2004. Congruence versus phylogenetic accuracy: Revisiting the incongruence length difference test. *Systematic Biology* 53 :81-89.
- Hitchman, R.B., Hodgson, D.J., King, L.A., Hails, R.S., Cory, J.S., Possee, R.D., 2007. Host mediated selection of pathogen genotypes as a mechanism for the maintenance of baculovirus diversity in the field. *Journal Invertebrate Pathology* 94: 153–162.
- Hooft van Iddekinge B.J.L., Smith, G.E., Summers, M.D., 1983. Nucleotide sequence of the polyhedrin gene of *Autographa californica* nuclear polyhedrosis virus. *Virology* 131: 561-565
- Horton, M.H., Burand, J.P., 1993. Saturable attachment sites for polyhedron-derived baculovirus on insect cells and evidence for entry via direct membrane fusion. *Journal of Virology* 67:1860-1868.
- Huelsenbeck, J.P., Ronquist, F., 2001. Bayesian inference of phylogenetic tree. *Bioinformatics Application Note* 17: 754-755.
- Hyink O., Dellow R.A., Olsen M.J., Caradoc-Davies K.M.B., Drake K., Herniou E.A., Cory J.S., O'Reilly D.R., Ward V.K., 2002. Whole genome analysis of the *Epiphyas postvittana nucleopolyhedrovirus*. *Journal of General Virology* 83: 957–971.
- IJkel W.F., Westenberg M., Goldbach R.W., Blissard, G.W., Vlak, J.M., Zuidema, D., 2000. A novel baculovirus envelope protein with a proprotein convertase cleavage site. *Virology* 275: 30-41.
- IJkel, W.F.J., van Strien, E.A., Heldens, J.G., Broer, R., Zuidema, D., Goldbach, R.W., Vlak, J.M., 1999. Sequence and organization of the *Spodoptera exigua* multicapsid nucleopolyhedrovirus genome. *Journal of General Virology* 80: 3289-3304

- Ikeda M., Shikata M., Shirata N., Chaeychomsri S., Kobayashi M., 2006. Gene organization and complete sequence of the *Hyphantria cunea* nucleopolyhedrovirus genome. *Journal of General Virology* 87: 2549–2562.
- Iorio, C., Vialard, J. E., McCracken, S., Lagacé, M., Richardson, C. D., 1998. The late expression factors 8 and 9 and possibly the phosphoprotein p78/83 of *Autographa californica* multicapsid nucleopolyhedrovirus are components of the virus-induced RNA polymerase. *Intervirology* 41: 35–46.
- Jaccard, P., 1901. Étude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bulletin de la Société Vaudoise des Sciences Naturelles* 37: 547-79.
- Jakubowska, A.K., 2010. Genomic support for speciation and specificity of baculoviruses. 138f. Tese - Wageningen University, Wageningen. 2010.
- Jakubowska, A.K., Peters, S.A., Ziemnicka, J., Vlak, J.M., van Oers M.M., 2006. Genome sequence of an enhancer gene-rich nucleopolyhedrovirus (NPV) from *Agrotis segetum*: collinearity with *Spodoptera exigua* multiple NPV. *Journal of General Virology* 87: 537–551.
- Jakubowska, A.K., van Oers, M.M., Otvos, I.S., Vlak, J.M., 2007. Phylogenetic analysis of *Orgyia pseudotsugata* Single-nucleocapsid nucleopolyhedrovirus. *Virologica Sinica* 22: 257-265.
- Jaques, R.P., 1985. Stability of insect viruses in the environment. In: Maramorosch, K.; Sherma, K.E. (Eds), *Viral Insecticides for Biological Control*. Academic Press, San Diego, pp. 285-360.
- Jarvis, D.L., Bohlmeier, D.A., Garcia, Jr A., 1991. Requirements for nuclear localization and supramolecular assembly of a baculovirus polyhedrin protein. *Virology* 185: 795-810.
- Jehle, J.A., 2004. The mosaic structure of the polyhedrin gene of the *Autographa californica* nucleopolyhedrovirus (AcMNPV). *Virus Genes* 29: 5-8.
- 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., 2006a. On the classification and nomenclature of baculoviruses: A proposal for revision. *Archives of Virology* 151:1257-1266.

- Jehle, J.A., Lange, M., Wang, H., Hu, Z., Wang, Y., Hauschild, R., 2006b. Molecular identification and phylogenetic analysis of baculoviruses from Lepidoptera. *Virology* 346: 180-193.
- Kanai, S., Kikuno, R., Toh, H., Ryo, H., Todo, T., 1997. Molecular evolution of the photolyase-blue-light photoreceptor family. *Journal of Molecular Evolution* 45: 535-548.
- Kikhno, I., Gutiérrez, S., Croizier, L., Croizier, G., López-Ferber, M., 2002. Characterization of pif, a gene required for the per os infectivity of *Spodoptera littoralis nucleopolyhedrovirus*. *Journal General Virology* 83: 3013-3022.
- Kimura, S., Tahira, Y., Ishibashi, T., Mori, Y., Mori, T., Hashimoto, J., Sakaguchi, K., 2004. DNA repair in higher plants; photoreactivation is the major DNA repair pathway in non proliferating cells while excision repair (nucleotide excision repair and base excision repair) is active in proliferating cells. *Nucleotide Acids Research* 32: 2760-2767.
- King, A.M.Q., Adams, M.J., Carstens, E.B., Lefkowitz, E.J., 2011. Virus taxonomy: classification and nomenclature of viruses: Ninth Report of the International Committee on Taxonomy of Viruses. San Diego. Elsevier Academic Press.
- Kunimi, Y., Fuxa, J.R., Richter, A.R., 1997. Survival time and lethal doses for wild and recombination *Autographa californica* nuclear polyhedrosis viruses in different instar of *Pseudaletia includens*. *Biological Control* 9: 129-135.
- 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: 17-34.
- Lange M., Jehle J.A., 2003. The genome of the *Cryptophlebia leucotreta granulovirus*. *Virology* 317: 220-236.
- Lange, M., Wang, H., Zhihong, H., Jehle, J.A., 2004. Towards a molecular identification and classification system of lepidopteran-specific baculoviruses. *Virology* 325: 36-47.
- Lauzon H.A.M., Jamieson P.B., Krell P.J., Arif B.M., 2005. Gene organization and sequencing of the *Choristoneura fumiferana* defective nucleopolyhedrovirus genome. *Journal of General Virology* 86: 945-961.

- Lauzon H.A.M., Lucarotti C.J., Krell P.J., Feng Q., Retnakaran A., Arif B.M., 2004. Sequence and organization of the *Neodiprion lecontei* nucleopolyhedrovirus genome. *Journal of Virology* 78: 7023–7035.
- Le, S.Q., Gascuel, O., 2008. LG: An Improved, General Amino-Acid Replacement Matrix. *Molecular Biology and Evolution* 25: 1307-20.
- Lee H.H., Miller L.K., 1978. Isolation of genotypic variants of *Autographa californica* nuclear polyhedrosis virus. *Journal of Virology* 27: 54– 67.
- Li J., Heinz K.M., Flexner J.L., McCutchen B.F., 1999. Effects of recombinant baculoviruses on three nontarget heliothine predators. *Biological Control* 15: 293– 302.
- Li L., Donly C., Li Q., Willis L.G., Keddie B.A., Erlandson M.A., Theilmann D.A., 2002b. Identification and genomic analysis of a second species of nucleopolyhedrovirus isolated from *Mamestra configurata*. *Virology* 297: 226–244.
- Li L., Li Q., Willis L.G., Erlandson M., Theilmann D.A., Donly C., 2005. Complete comparative genomic analysis of two field isolates of *Mamestra configurata* nucleopolyhedrovirus-A. *Journal of General Virology* 86: 91–105.
- Li S., Erlandson M., Moody D., Gillott C., 1997. A physical map of the *Mamestra configurata* nucleopolyhedrovirus genome and sequence analysis of the polyhedrin gene. *Journal of General Virology* 78: 265–271.
- Li W.H., 1997. Molecular evolution. Sinauer Associates, Inc. Sunderland, MA.
- Li, Q., Donly, C., Li, L., Willis, L.G., Theilmann, D.A., Erlandson, M., 2002a. Sequence and organization of the *Mamestra configurata* nucleopolyhedrovirus genome. *Virology* 294: 106–121.
- Livingston, J.M., Mcleod, P.J., Yearia, W.C., Young, S.Y., 1980. Laboratory and field evaluation of nuclear polyhedrosis virus of the soybean Looper, *Pseudoplusia includens*. *Journal of the Georgia Entomological Society* 15: 94-199.
- Ljungman, M., Zhang, F., 1996. Blockage of RNA polymerase as a possible trigger for u.v. light-induced apoptosis. *Oncogene* 13: 823–831.
- Lu A., Miller LK., 1994. Identification of three late expression factor genes within the 33.8 to 43.4 map unit region of *Autographa californica* nuclear polyhedrosis virus. *Journal of Virology* 68: 6710–6718.

- Lu, A., Miller, L.K., 1995. Differential requirements for baculovirus late expression factor genes in two cell lines. *Journal of Virology* 69: 6265–6272.
- Luo, S., Zhang, Y., Xu, X., Westenberg, M., Vlak, J.M., Wang, H., Hu, Z., Deng, F., 2011. *Helicoverpa armigera* nucleopolyhedrovirus occlusion-derived virus-associated protein, HA100, affects oral infectivity in vivo but not virus replication in vitro. *Journal of General Virology* 92: 1324–1331.
- Luque T., Finch R., Crook N., O'Reilly D.R., Winstanley D., 2001. The complete sequence of the *Cydia pomonella granulovirus* genome. *Journal of General Virology* 82: 2531–2547.
- Ma X.C., Shang J.Y., Yang Z.N., Bao Y.Y., Xiao Q., Zhang C.X., 2007. Genome sequence and organization of a nucleopolyhedrovirus that infects the tea looper caterpillar, *Ectropis obliqua*. *Virology* 360: 235–246.
- Mcleod, P.J., Young, S.Y., Yearian, W.C., 1982. Application of a baculovirus of *Pseudoplusia includens* to soybean: efficacy and seasonal persistence. *Environment Entomology* 11: 412-416.
- Miele, S.A.B., Garavaglia, M.J., Belaich, M.N., Ghiringhelli, P.D., 2011. Baculovirus: molecular insights on their diversity and conservation. *International Journal of Evolutionary Biology* 15p.
- Miller, L.K., Lu, A., 1997. The molecular basis of baculovirus host range. In: Miller, L.K. (Ed.), *The Baculoviruses*. Plenum Press, New York. pp. 217-235.
- Montagner S., Costa A., 2009. Bases biomoleculares do fotoenvelhecimento. *Anais Brasileiros de Dermatologia* 84: 263-269.
- Morales, L., Moscardi, F., Kastelic, J.G., Sosa-Gomez, D.R., Paro, F.R., Soldorio, I.L., 1995. Suscetibilidade de *Anticarsia gemmatalis* Hubner e *Chrysodeixis includens* (Walker) (Lepidoptera: Noctuidae), a *Bacillus thuringiensis* (Berliner). *Annais da Sociedade Entomológica do Brasil* 24: 593–598.
- Moscardi, F. 2007. A Nucleopolyhedrovirus for control of the velvetbean caterpillar in Brazilian Soybeans. In: Vincent, C. M., Goethel, S., Lazarovits, G., (Eds.), *Biological Control: A Global Perspective*. Oxfordshire, UK, Cambridge, USA. CAB International pp. 344-352.
- Moscardi, F., 1999. Assessment of the application of baculoviruses for control of Lepidoptera. *Annual Review of Entomology* 44: 257-289.



- Moscardi, F., Morales, L., Santos, B., 2002. The successful use of AgMNPV for the control of velvetbean caterpillar, *Anticarsia gemmatalis*, in soybean in Brazil. In: 35th Annual Meeting of the SIP. Foz do Iguaçu, Brasil, pp. 86-91.
- Moscardi, F., Sosa-Gómez, D.R., 1992a. A case study in biological control: soybean defoliating caterpillars in Brazil. In: International Crop Science Congress, 1. Ames, Iowa. International Crop Science I. Madison: Crop Science Society of America, pp.115-119, 1993.
- Moscardi, F., Sosa-Gómez, D.R., 1992b. Use of viruses against soybean caterpillars in Brazil. In: Copping, I.G.; Green, M.B.; Reeds, R.T. (Eds.), Pest Management in Soybean. London: Elsevier Applied Science, pp. 98-109.
- Moscardi, F., Souza, M.L., Castro, M.E.B., Moscardi, M.L., Szewczyk, B., 2011. Baculovirus pesticides: present state and future perspectives. In: Ahmad *et al.* (Org.), Microbes and Microbial Technology. Springer, pp.415-445.
- Nai Y.S., Wu C.Y., Wang T.C., Chen Y.R., Lau W.H., Lo C.F., Tsai M.F., Wang C.H., 2010. Genomic sequencing and analyses of *Lymantria xyli* multiple nucleopolyhedrovirus. *BioMed Central Genomics*.
- Nakai M., Goto C., Kang W., Shikata M., Luque T., Kunimi Y., 2003. Genome sequence and organization of a nucleopolyhedrovirus isolated from the smaller tea tortrix, *Adoxophyes honmai*. *Virology* 316: 171–183.
- Nie, Z. M., Zhang, Z. F., Wang, D., He, P. A., Jiang, C. Y., Song, L., Chen, F., Xu, J., Yang, L., Yu, L.L., Chen, J., Lv, Z.B., Lu, J.J., Wu, X.F., Zhang, Y.Z., 2007. Complete sequence and organization of *Antheraea pernyi* nucleopolyhedrovirus, a dr-rich baculovirus. *BioMed Central Genomics*.
- Nishigaki, R., Mitani, H., Shima, A., 1998. Evasion of UVC-induced apoptosis by photorepair of cyclobutane pyrimidine dimers. *Experimental Cell Research* 244: 43–53.
- O'Reilly, D., Miller, L.K., Luckow, V.A., 1992. Baculovirus Expression Vectors: A Laboratory Manual. W.H. Freeman and Company 347, New York.
- Odindo, M. O., 1983. Epizootiological observations on a nuclear polyhedrosis of the African armyworm *Spodoptera exempta* (Walker). *Insect Science and its Application* 4: 291-298.

- 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 Sericology* 76: 39–49.
- Okhawa, 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: 15258–15264.
- Okuno S., Takatsuka J., Nakai M., Ototake S., Masui A., Kunimi Y., 2003. Viral-enhancing activity of various stilbene-derived brighteners for a *Spodoptera litura* (Lepidoptera: Noctuidae) nucleopolyhedrovirus. *Biological Control* 26: 146–152.
- Oliveira J.V.C., Wolff J.L.C., Garcia-Maruniak A., Ribeiro B.M., Castro M.E.B., Souza M.L., Moscardi F., Maruniak J.E., Zanotto P.M.A., 2006. Genome of the most widely used viral biopesticide: *Anticarsia gemmatalis* multiple nucleopolyhedrovirus. *Journal of General Virology* 87: 3233–3250.
- Öztürk, N., Kao, Y.T., Selby, C.P., Kavakli, I.H., Partch, C.L., Zhong, D., Sancar, A., 2008. Purification and characterization of a type III photolyase from *Caulobacter crescentus*. *Biochemistry* 47: 10255-10261.
- Pang Y., Yu J., Wang L., Hua X., Baob W., Lib G., Chenb C., Hanb H., Hub S., Yangb H., 2001. Sequence analysis of the *Spodoptera litura* multicapsid nucleopolyhedrovirus genome. *Virology* 287: 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 RNAPolymerases. *Journal of Virology* 68:4673–78.
- Payne, C.C., 1986. Insect pathogenic viruses as pest control agents. *Fortschritte der Zoologie* 32:183-200.
- Pearson M., Groten, C., Rohrmann, G.F., 2000. Identification of the *Lymantria dispar* nucleopolyhedrovirus envelope fusion protein provides evidence for a phylogenetic division of Baculoviridae. *Journal of Virology* 74: 6126-6131.

- Pearson, M., Bjornson, R., Pearson, G., Rohrmann, G.F., 1992. The *Autographa californica* baculovirus genome: evidence for multiple replication origins. *Science* 257:1382–1384.
- Pedigo, L.P., Rice, M.E., 2005. Entomology and Pest Management, Prentice Hall, New Jersey, 784p.
- 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: 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: 35–42.
- 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 General Virology* 84:2041–2049.
- Posada, D., 2008. jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution* 25: 1253–1256.
- Rambaut, A., 2009. FigTree v1.3.1. [<http://tree.bio.ed.ac.uk/software/figtree/>] webcite Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, United Kingdom.
- Redman, E.M., Wilson, K., Grzywacz, D., Cory, J.S., 2010. High levels of genetic diversity in *Spodoptera exempta* NPV from Tanzania. *Journal of Invertebrate Pathology* 105: 190-193.
- Rodriguez, F., Oliver, J.F., Marin, A., Medina, J.R., 1990. The general stochastic model of nucleotide substitution. *Journal of Molecular Evolution* 142: 485-501.
- Rohlf, F.J., 1997. NTSYSpc: numerical taxonomy and multivariate analysis system. Setauket, NY: Exeter Software, 1 CD-ROM.
- Rohrmann, G.F., 1986. Polyhedrin structure. *Journal of General Virology* 67: 1499-1513.
- Rohrmann, G.F., 2011. Baculovirus Molecular Biology: Second Edition [Internet]. National Library of Medicine (US), National Center for Biotechnology Information.
- Rokas, A., Holland, P.W.H., 2000. Rare genomic changes as a tool for phylogenetics. *Trends in Ecology and Evolution* 15: 454-459.

- 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.
- Rycyna, R.E., Alderfer, J.L., 1985. UV irradiation of nucleic acids: formation, purification and solution conformational analysis of the '6-4 lesion' of dTpdT. *Nucleic Acids Research*. 13: 5949-5963.
- Sancar, A., 1994. Structure and function of DNA photolyase. *Biochemistry* 33: 2-9.
- Sancar, A., 2003. Structure and function of DNA photolyase and cryptochrome blue-light photoreceptors. *Chemical Reviews* 103: 2203-2237.
- Sancar, G.B., 1990. DNA photolyases: Physical properties, action mechanism, and roles in dark repair. *DNA Repair* 236: 147-160.
- Sauphanor, B., Berling, M., Toubon, J.-F., Reyes, M., Delnatte, J., Allemoz, P., 2006. Carpcapsc des pommes. Cas de résistance au virus de la granulose en vergers biologique. *Phytoma – La Défense des Végétaux* 590: 24-27.
- Schneider, H., 2007. Método de análise filogenética: um guia prático. 2ª edição. Ribeirão Preto: Holos, Editora e Sociedade Brasileira de Genética, 200p.
- Schwarz, G.E., 1978. Estimating the dimension of a model. *Annals of Statistics* 6: 461–464.
- Setlow, R.B., Carrier, W.L., 1966. Pyrimidine dimers in ultraviolet-irradiated DNA's. *Journal Molecular Biology* 17: 237-254.
- 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: 458–465.
- Shapiro, M., Farrar Jr, R.R., Domek, J., Javaid, I., 2002. Effects of virus concentration and ultraviolet irradiation on the activity of corn earworm and beet armyworm (Lepidoptera: Noctuidae) nucleopolyhedroviruses. *Journal of Economic Entomology* 95: 243–249.
- Silvie, P., Bélot, J. L., Michel, B., 2007. Manual de identificação das pragas e seus danos no cultivo de algodão. Boletim Técnico 34, 2. ed. Cascavel: COODETEC/CIRAD-CA, 120p.

- Simmons, M.P., Ochoterena, H., 2000. Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology* 49: 369–381.
- Simón O., Palma L., Williams T., López-Ferber M., Caballero P., 2012. Analysis of a naturally-occurring deletion mutant of *Spodoptera frugiperda* multiple nucleopolyhedrovirus reveals *sf58* as a new per os infectivity factor of lepidopteran-infecting baculoviruses. *Journal of Invertebrate Pathology* 109: 117–126.
- Simón, O., Gutiérrez, S., Williams, T., Caballero, P., López-Ferber, M., 2005. Nucleotide sequence and transcriptional analysis of the *pif* gene of *Spodoptera frugiperda* nucleopolyhedrovirus (SfMNPV). *Virus Research* 108: 213-220.
- Simón, O., Williams, T., López-Ferber, M., Caballero, P., 2004. Genetic structure of a *Spodoptera frugiperda* nucleopolyhedrovirus population: high prevalence of deletion genotypes. *Applied and Environmental Microbiology* 70: 5579-5588.
- Slack, J.M., Arif, B.M., 2007. The baculoviruses occlusion-derived virus; virion structure and function. *Virus Research* 69: 99-165.
- Slamovits, C.H., Keeling, P.J., 2004. Class II photolyase in a microsporidian intracellular parasite. *Journal of Molecular Biology* 341: 713-721.
- Sokal, R. R., Rohlf, F. J., 1962. The comparison of dendrograms by objective methods. *Taxon* 11: 33-40.
- Sorensen, T., 1948. A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons. *Kongelige Danske Videnskabernes Selskabs. Biologiske Skrifter* 5: 1-34.
- Sosa-Gomez, D.R., Moscardi, F., Santos, B., Alves, L.F., Alves, S.B., 2008. Produção e uso de vírus para o controle de pragas na América Latina. In: Alves, Sérgio Batista; Rogério Biaggioni Lopes. (Org.). Controle Microbiano de pragas na América Latina: Avanços e Desafios. 1 ed. Piracicaba: FEALQ, 14: 49-68.
- Souza, J.C, Reis, P.R., Silva, R.A., Santa-Cecília, L.V.C., 2009. Lagarta falsa-medideira da soja: provavelmente será uma nova praga na cultura da batata. *Revista Batata Show* 23.

- Srinivasa, M., Jagadeesh Babu, C.S., Anitha, C.N., Girish, G., 2008. Laboratory evaluation of available commercial formulations of HaNPV against *Helicoverpa armigera* (Hub.). *Journal of Biopesticides* 1: 138-139.
- Srinivasan, V., Schnitzlein, W.M., and Tripathy, D.N., 2001. Fowlpox virus encodes a novel DNA repair enzyme, CPD-photolyase, that restores infectivity of UV light-damaged virus. *Journal of Virology* 75: 1681-1688.
- Struck, T.H., Purschken, G., Halanych, K.M., 2006. Phylogeny of Eunicidae (Annelida) and exploring data congruence using a Partition Addition Bootstrap Alteration (PABA) approach. *Systematic Biology* 55: 1-20.
- Sun X., Peng H., 2007. Recent advances in biological control of pest insects by using viruses in China. *Virologica Sinica* 22: 158–162.
- Sun, X., Sun, X., Van der Werf, W., Vlak, J.M., 2004. Field inactivation of wild-type and genetically modified *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus in cotton, *Biocontrol Science and Technology* 14: 185–192.
- Swofford, D.L., 2003. Paup\*: Phylogenetic analysis using parsimony (\*and other methods) version 4.0b10. Sunderland, Massachusetts: Sinauer Associates.
- Szewczyk, B., Hoyos-Carvajal, L., Paluszek, M., Skrzecz, I., Souza, M.L., 2006. Baculoviruses reemerging biopesticides. *Biotechnology Advances* 24: 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: M. Stoytcheva (Org.), Pesticides – Formulations. Effects, Fate, Rijeka: InTech, pp. 25-36.
- Taha A., Nour-el-Din A., Croizier L., Ferber M.L., Croizier G., 2000. Comparative analysis of the granulins regions of the *Phthorimaea operculella* and *Spodoptera littoralis granuloviruses*. *Virus Genes* 21: 147–155.
- Takatsuga, J., 2007. Characterization of a nucleopolyhedrovirus of *Epinotia granitalis* (Lepdoptera: Tortricidae). *Journal of Invertebrate Pathology* 96: 265-269.

- Tang X.D., Xiao Q., Ma X.C., Zhu Z.R., Zhang C.X., 2009. Morphology and genome of *Euproctis pseudoconpers nucleopolyhedrovirus*. *Virus Genes* 38: 38495–506.
- 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: Fauquet, C.M., Mayo, M.A., Maniloff, J., Desselberger, U., Ball, L.A. (Eds.). *Virus taxonomy: classification and nomenclature of viruses*. Eighth Report of The International Committee on Taxonomy of Viruses. Elsevier Academic Press: Amsterdam, pp. 177-185.
- 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: 15931-15935.
- Thiem, S.M., Cheng, X.W., 2009. Baculovirus host-range. *Virologica Sinica* 24: 436-457.
- Titterton, J.S., Nun, T.K., Passarelli, A.L., 2003. Functional dissection of the baculovirus late expression factor-8 gene: sequence requirements for late gene promoter activation. *Journal of General Virology* 84: 1817-1826.
- 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: 968-974.
- Valicente, F.H., Cruz, I., 1991. Controle biológico da lagarta-do-cartucho, *Spodoptera frugiperda*, com o baculovirus. Sete Lagoas: Embrapa-CNPMS, Circular técnica, 15, 23p.
- van Oers M.M., Abma-Henkens M.H.C., Herniou E.A., de Groot J.C.W., Peters S., Vlak J. M., 2005. Genome sequence of *Chrysodeixis chalcites nucleopolyhedrovirus*, a baculovirus with two DNA photolyase genes. *Journal of General Virology* 86: 2069–2080.
- van Oers, M.M., Herniou, E.A., Usmany, M., Messelink, G.J., Vlak, J.M., 2004. Identification and characterization of a DNA photolyase-containing baculovirus from *Chrysodeixis chalcites*. *Virology* 330: 460-470.
- van Oers, M.M., Vlak, J.M., 2007. Baculovirus genomics. *Current Drug Targets* 8: 1051-1068.
- Vande Berg, B.J., Sancar, G.B., 1998. Evidence for dinucleotide flipping by DNA photolyase. *Journal of Biological Chemistry* 273: 20276-20284.

- Vincent, C., Andermatt, M., Valero, J., 2007. Madex and virosoft, viral biopesticides for codling moth control. *Biological control: a global perspective – case studies around the world*. Wallingford: Cabi Publishing, pp. 336-343.
- Volkman, L.E., Goldsmith, P.A., 1985. Mechanism of neutralization of budded *Autographa californica* nuclear polyhedrosis virus by a monoclonal antibody: Inhibition of entry by adsorptive endocytosis. *Virology* 143: 185-195.
- Volkman, L.E., Goldsmith, P.A., Hess, R.T., 1986. Alternate pathway of entry of budded *Autographa californica* nuclear polyhedrosis virus: fusion at the plasma membrane. *Virology* 148: 288-297.
- Volkman, L.E., Keddie, B.A., 1990. Nuclear polyhedrosis virus pathogenesis. *Seminars in Virology* 1: 249-256.
- Wang Y., Choi J.Y., Roh J.Y., Woo S.D., Jin B.R., Je Y.H., 2008. Molecular and phylogenetic characterization of *Spodoptera litura granulovirus*. *Journal of Microbiology* 46: 704–708.
- Whelan, S., Goldman N., 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum likelihood approach. *Molecular Biology and Evolution* 18: 691-699.
- 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 nucleopolyhedrovirus* and the identification of a baculoviral photolyase gene. *Virology* 338: 209–226.
- Wolff J.L.C., Valicente F.H., Martins R., Oliveira J.V., Zanotto P.M., 2008. Analysis of the genome of *Spodoptera frugiperda* nucleopolyhedrovirus (SfMNPV-19) and of the high genomic heterogeneity in group II nucleopolyhedroviruses. *Journal of General Virology* 89: 1202–1211.
- Wormleaton S., Kuzio J., Winstanley D., 2003. The complete sequence of the *Adoxophyes orana granulovirus* genome. *Virology* 311: 350–365.
- 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: 69–75.
- Xiao H., Qi, Y., 2007. Genome sequence of *Leucania seperata nucleopolyhedrovirus*. *Virus Genes* 35: 845–856.



- Xu Y.P., Ye Z.P., Niu C.Y., Bao Y.Y., Wang W.B, Shen W.D, Zhang C.X., 2010. Comparative analysis of the genomes of *Bombyx mandarina* and *Bombyx mori* nucleopolyhedroviruses. *Journal of Microbiology* 48: 102–110.
- Xu, F., Vlak, J. M., van Oers, M. M., 2008. Conservation of DNA photolyase genes in group II nucleopolyhedroviruses infecting plusiine insects. *Virus Research* 136: 58–64.
- Yasui, A., Eker, A.P.M., Yasuhira, S., Yajima, H., Kobayashi, T., Takao, M., Oikawa, A., 1994. A new class of DNA photolyases present in various organisms including aplacental mammals. *European Molecular Biology Organization (EMBO) Journal* 13: 6143–6151.
- Yoder, A.D., Irwin, J.A., Payseur, B.A., 2001. Failure of the ILD to determine data combinability for slow loris phylogeny. *Systematic Biology* 50:408–424.
- Young, S.Y., 2001. Enhancement of nucleopolyhedrovirus activity in *Helicoverpa zea* (Boddie) and *Pseudoplusia includens* (Walker) larvae with a fluorescent brightener. *Journal Entomology Society* 36:162-168.
- Young, S.Y., Yearian, W.C., 1982. Nuclear polyhedrosis virus infection of *Pseudoplusia includes* (Lepidoptera: Noctuidae) larvae: effect on post larval stages and transmission. *Entomophaga* 27: 61-66.
- Zanotto P.M., Kessing B.D., Maruniak J.E., 1993. Phylogenetic interrelationships among baculoviruses: evolutionary rates and host associations. *Journal of Invertebrate Pathology* 62: 147-164.
- 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: 190–199.
- Zhang, G., 1994. Research, development and application of *Heliothis* viral pesticide in China. *Resource Environment* 3: 1–6.
- Zhu, S.Y., Yi, J.P., Shen, W.D., Wang, L.Q., He, H.G., Wang, Y., Li, B., Wang, W.B., 2009. Genomic sequence, organization and characteristics of a new nucleopolyhedrovirus isolated from *Clanis bilineata* larva. *BioMed Center Genomics* 10: 91-100.
- Zou Y., Young S.Y., 1996. Use of fluorescent brightener to improve *Pseudoplusia includens* (Lepidoptera: Noctuidae) nuclear polyhedrosis virus activity in the laboratory and field. *Journal of Economy Entomology* 89: 92–96.

Zuckermandl, E., Pauling, L.B., 1962. Molecular disease, evolution, and genetic heterogeneity. In: Kasha, M.; Pullman, B (Eds), *Horizons in Biochemistry*. Academic Press, New York, pp. 189–225.

# APÊNDICES

# Apêndice I

## (A) BLAST – Sequência deduzida de aminoácidos LEF-8

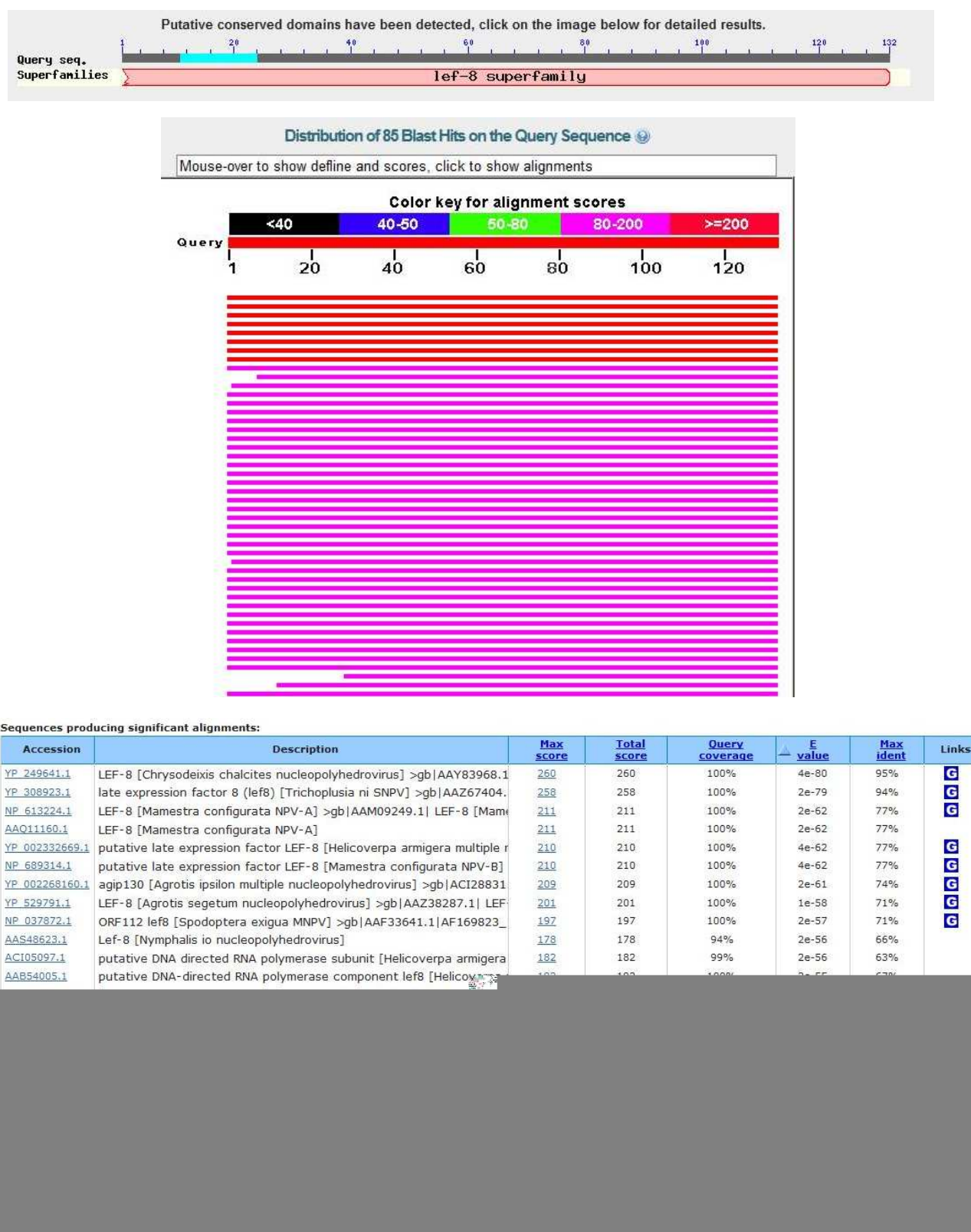
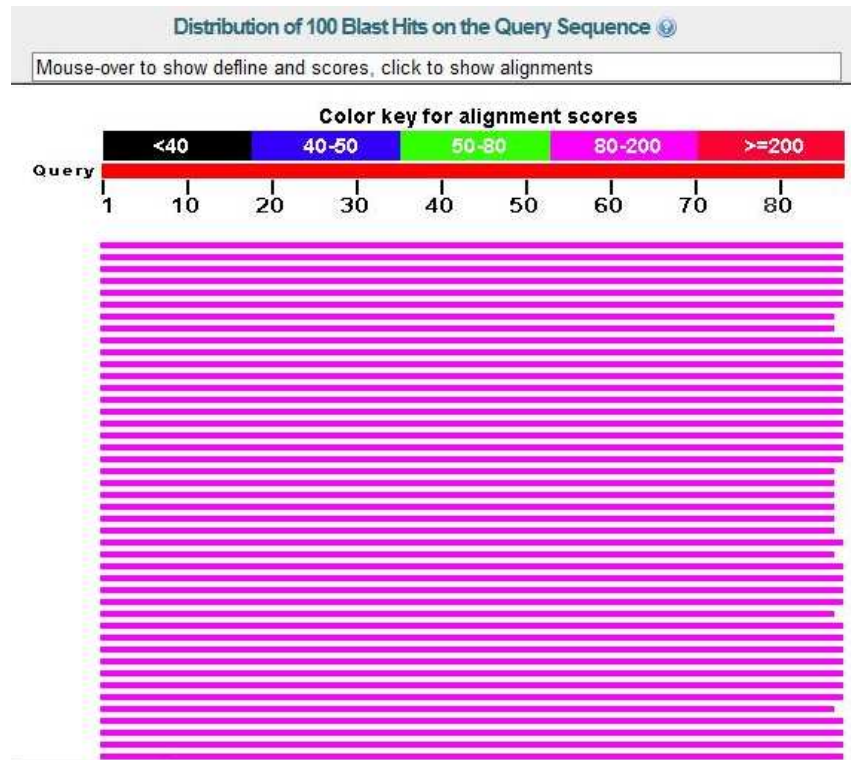
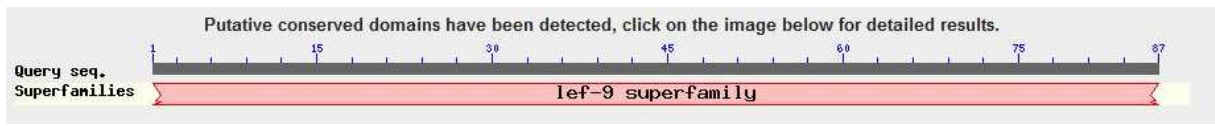


Figura 20. Resultado do alinhamento *Blast* da sequência deduzida de aminoácidos (A) LEF-8, (B) LEF-9, (C) PHR, (D) PIF-2 e (E) POLH de PsinSNPV-IE.

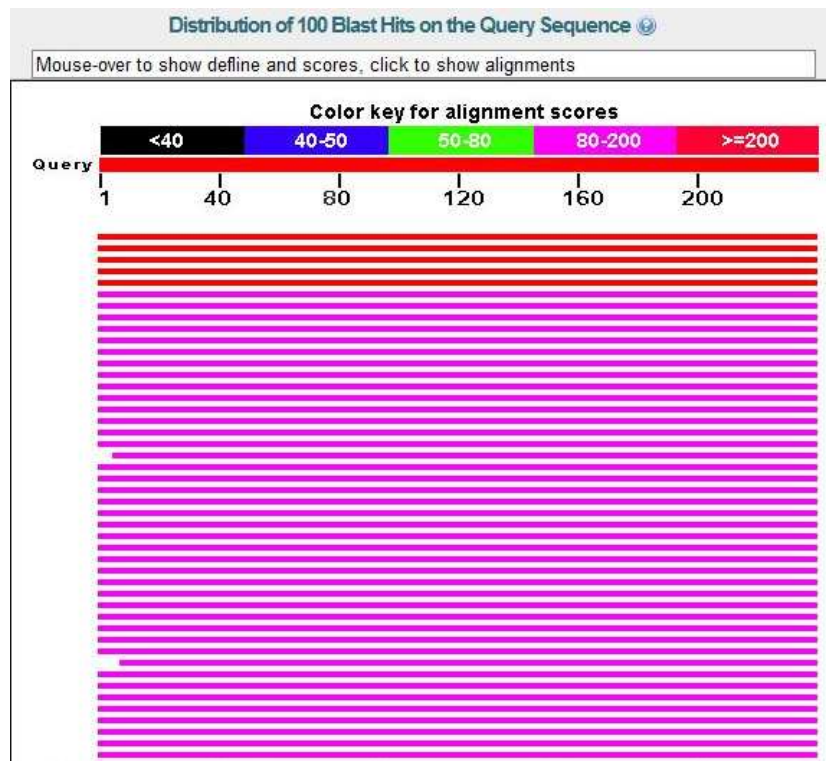
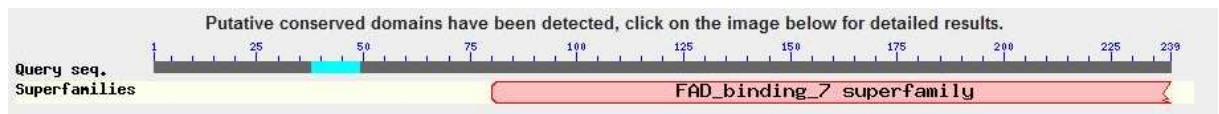
## (B) BLAST – Sequência deduzida de aminoácidos LEF-9



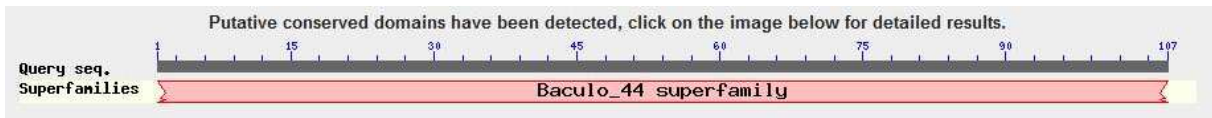
Sequences producing significant alignments:

Accession	Description	Max score	Total score	Query coverage	E value	Max ident	Links
<a href="#">AAW49135.1</a>	lef-9 [Plusia acuta nucleopolyhedrovirus] >gb ADY88160.1  late expr	182	182	100%	3e-59	100%	<a href="#">G</a> <a href="#">G</a>
<a href="#">YP_308939.1</a>	late expression factor 9 (lef9) [Trichoplusia ni SNPV] >gb AAZ67420.	186	186	100%	7e-56	100%	
<a href="#">YP_249656.1</a>	LEF-9 [Chrysodeixis chalcites nucleopolyhedrovirus] >gb AAZ83983.1	186	186	100%	8e-56	100%	
<a href="#">AAW49124.1</a>	lef-9 [Mamestra brassicae MNPV] >gb AAW49125.1  lef-9 [Mamestra	168	168	100%	1e-53	91%	
<a href="#">AAW49130.1</a>	lef-9 [Peridroma margaritosa nucleopolyhedrovirus]	167	167	100%	1e-53	90%	
<a href="#">BAF81898.1</a>	late expression factor 9 [Epinotia granitalis nucleopolyhedrovirus]	167	167	100%	2e-53	90%	
<a href="#">ADW96153.1</a>	lef-9 [Cerapteryx graminis nucleopolyhedrovirus]	166	166	98%	4e-53	88%	
<a href="#">AAT10238.1</a>	lef-9 [Busseola fusca nucleopolyhedrovirus] >gb ADY88246.1  late ex	166	166	98%	6e-53	87%	
<a href="#">ADY88176.1</a>	late expression factor-9 [Helicoverpa armigera multiple nucleopolyhed	166	166	100%	7e-53	90%	
<a href="#">AAW49093.1</a>	lef-9 [Agrotis segetum nucleopolyhedrovirus]	161	161	100%	3e-51	87%	
<a href="#">AAT10220.1</a>	lef-9 [Agrotis insilon multiple nucleopolyhedrovirus]	161	161	100%	4e-51	87%	

## (C) BLAST – Sequência deduzida de aminoácidos CPD-fotoliase



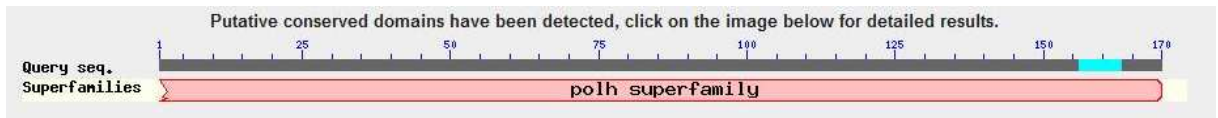
## (D) BLAST – Sequência deduzida de aminoácidos PIF-2



### Sequences producing significant alignments:

Accession	Description	Max score	Total score	Query coverage	E value	Max ident	Links
<a href="#">YP_249752.1</a>	PIF-2 [Chrysoideixis chalcites nucleopolyhedrovirus] >gb AAZ84079.1	222	222	100%	7e-71	96%	<a href="#">G</a>
<a href="#">YP_309030.1</a>	orf141 [Trichoplusia ni SNPV] >gb AAZ67511.1  orf141 [Trichoplusia	209	209	100%	7e-66	90%	<a href="#">G</a>
<a href="#">YP_001651022.1</a>	per os infectivity factor 2 [Orgyia leucostigma NPV] >gb ABY65838.1	172	172	100%	3e-51	73%	<a href="#">G</a>
<a href="#">ADB84445.1</a>	per os infectivity factor 2 [Apocheima cinerarium nucleopolyhedrovirus]	168	168	100%	2e-50	70%	
<a href="#">NP_047756.1</a>	LdOrf-119 peptide [Lymantria dispar MNPV] >gb AAC70305.1  LdOrf-	169	169	100%	6e-50	69%	<a href="#">G</a>
<a href="#">YP_002854731.1</a>	pif-2 [Euproctis pseudoconsersa nucleopolyhedrovirus] >gb ACO535	166	166	100%	8e-49	68%	<a href="#">G</a>
<a href="#">YP_003517878.1</a>	PIF-2 [Lymantria xyliana MNPV] >gb ADD73847.1  PIF-2 [Lymantria xy	167	167	100%	3e-48	68%	<a href="#">G</a>
<a href="#">YP_529706.1</a>	PIF-2 [Agrotis segetum nucleopolyhedrovirus] >gb AAZ38202.1  PIF-	164	164	100%	6e-48	69%	<a href="#">G</a>
<a href="#">AAS13393.1</a>	ac22-like protein [Lymantria monacha nucleopolyhedrovirus]	155	155	99%	9e-48	65%	
<a href="#">NP_689218.1</a>	hypothetical protein [Mamestra configurata NPV-B] >gb AAM95030.1	164	164	100%	1e-47	68%	<a href="#">G</a>
<a href="#">YP_717645.1</a>	PIF-2 [Clanis bilineata nucleopolyhedrosis virus] >gb ABF47448.1  PIF	162	162	100%	1e-47	69%	<a href="#">G</a>
<a href="#">NP_613131.1</a>	hypothetical protein Mamestra configurata NPV-A1 >gb AAM95030.1	163	163	100%	2e-47	66%	<a href="#">G</a>

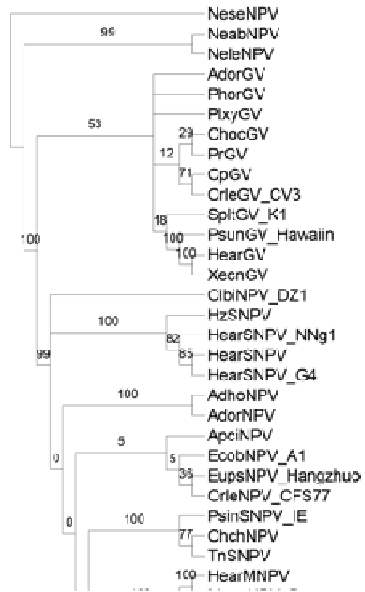
## (E) BLAST – Sequência deduzida de aminoácidos POLH

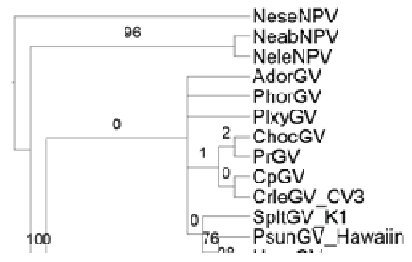


Sequences producing significant alignments:

Accession	Description	Max score	Total score	Query coverage	E value	Max ident	Links
<a href="#">ADY89161.1</a>	polyhedrin [Pseudoplusia includens nucleopolyhedrovirus]	352	352	100%	3e-124	100%	
<a href="#">YP_249605.1</a>	polyhedrin [Chrysodeixis chalcites nucleopolyhedrovirus] >gb AAZ839	355	355	100%	5e-124	100%	<a href="#">G</a>
<a href="#">AAW49202.1</a>	polyhedrin [Plusia acuta nucleopolyhedrovirus]	351	351	100%	1e-123	99%	
<a href="#">ACD01440.1</a>	polyhedrin [Pseudoplusia includens nucleopolyhedrovirus]	351	351	100%	1e-123	99%	
<a href="#">ACD01438.1</a>	polyhedrin [Chrysodeixis chalcites nucleopolyhedrovirus]	351	351	99%	2e-123	100%	
<a href="#">AAC64160.1</a>	polyhedrin [Trichoplusia ni SNPV]	353	353	100%	3e-123	99%	
<a href="#">YP_308889.1</a>	polyhedrin [Trichoplusia ni SNPV] >gb AAZ67370.1  polyhedrin [Trich	353	353	100%	4e-123	99%	<a href="#">G</a>
<a href="#">AAC64234.1</a>	polyhedrin protein [Plusia orichalcea nuclear polyhedrosis virus]	352	352	100%	7e-123	99%	
<a href="#">ACD01439.1</a>	polyhedrin [Pseudoplusia includens nucleopolyhedrovirus]	348	348	99%	2e-122	99%	
<a href="#">ADB84363.1</a>	polyhedrin [Apocheima cinerarium nucleopolyhedrovirus]	348	348	100%	2e-121	96%	
<a href="#">NP_818648.1</a>	polyhedrin [Adoxophyes honmai NPV] >dbj BAC67252.1  polyhedrin [A	348	348	100%	3e-121	95%	<a href="#">G</a>
<a href="#">AAW49195.1</a>	polyhedrin [Mamestra brassicae MNPV] >gb AAW49197.1  polyhedrin	345	345	100%	4e-121	95%	







# Apêndice III

## Análise MP não ponderada – Sequência nucleotídica

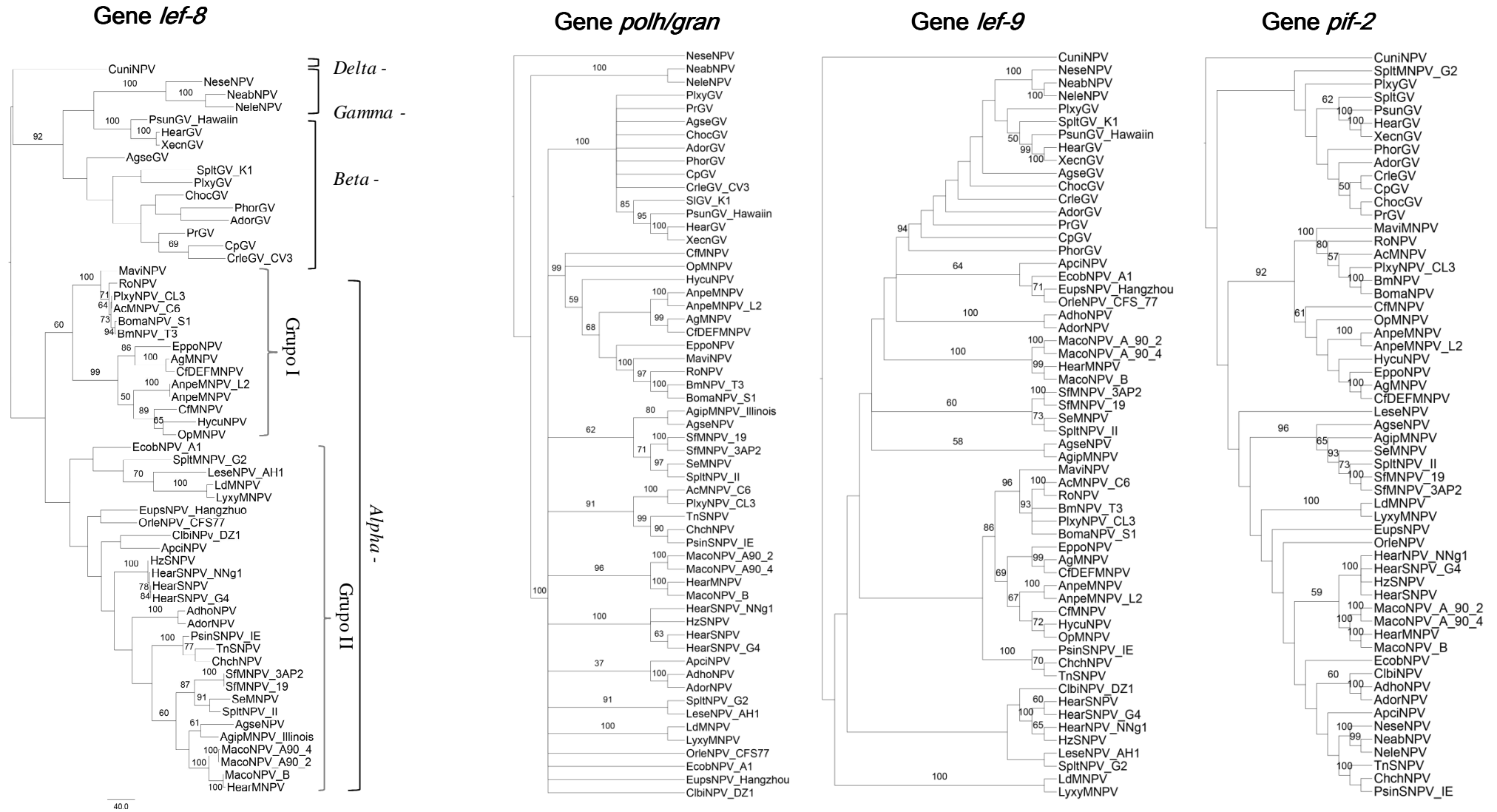


Figura 22. Árvores filogenéticas representando as relações filogenéticas da família *Baculoviridae* obtidas da análise MP não ponderada baseada nas sequências nucleotídicas parciais dos genes individuais *lef-8*, *lef-9*, *pif-2* e *polh/gran*. (A) Filograma da sequência nucleotídica do gene *lef-8*. (B) Cladograma do consenso de 33 árvores MP obtidas da análise da sequência nucleotídica do gene *polh/gran*. (C) Cladograma do consenso de 16 árvores MP obtidas da análise da sequência nucleotídica do gene *lef-9*. (D) Cladograma do consenso de 2 árvores MP obtidas da análise da sequência nucleotídica do gene *pif-2*. Os números nos ramos representam os valores de *bootstrap* (> 50%) para 1.000 replicatas. A barra de escala representa o número de substituições nucleotídicas por base.

## Árvores MP com reponderação progressiva – Sequência nucleotídica

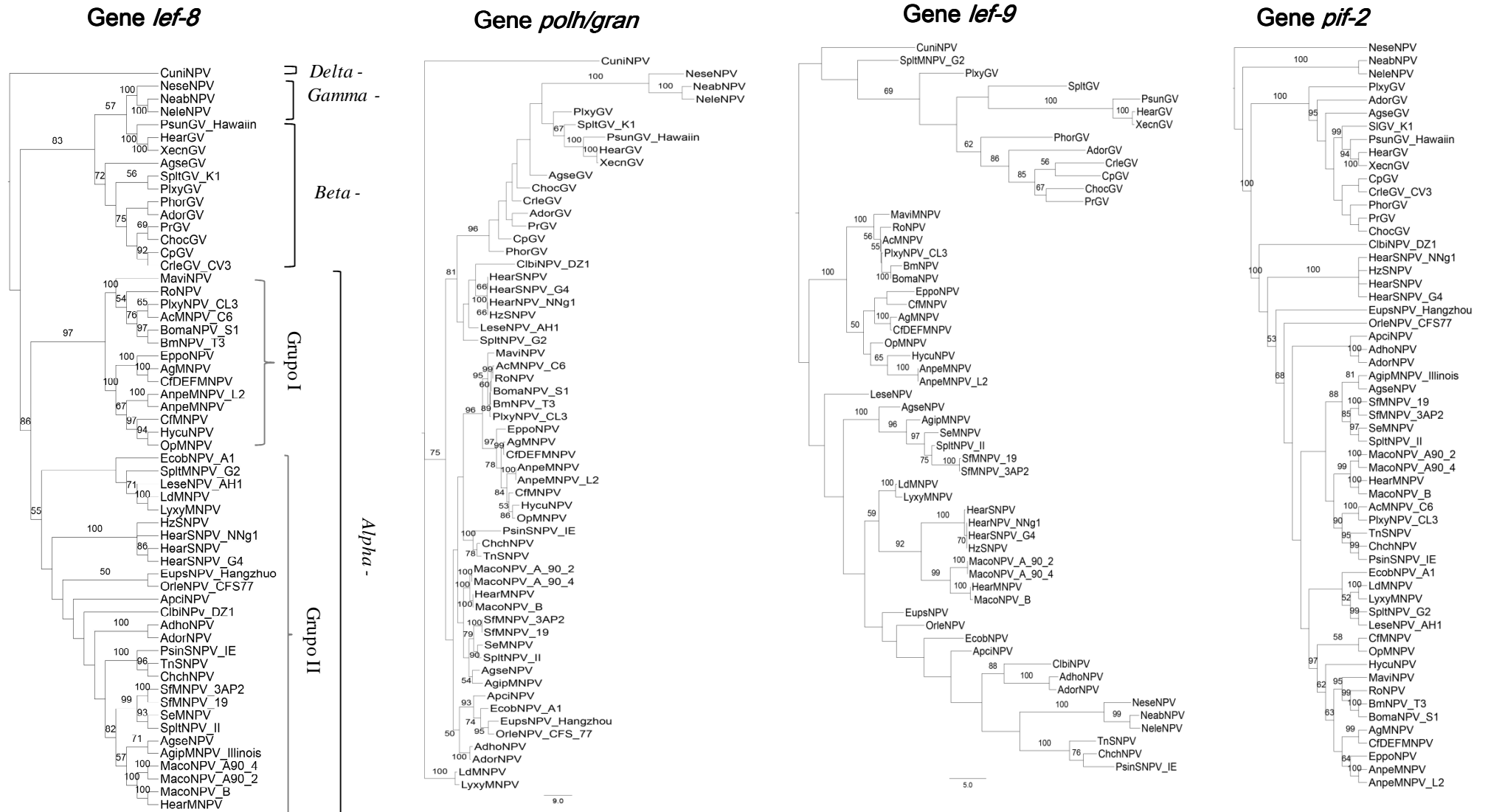


Figura 23. Árvores filogenéticas representando as relações filogenéticas da família *Baculoviridae* obtidas da análise MP com reponderação progressiva baseada nas sequências nucleotídicas parciais dos genes individuais *lef-8*, *lef-9*, *pif-2* e *polh/gran*. (A) Cladograma do consenso de 2 árvores MP obtidas da análise da sequência nucleotídica do gene *lef-8*. (B) Cladograma do consenso de 5 árvores MP obtidas da análise da sequência nucleotídica do gene *polh/gran*. (C) Filograma da sequência nucleotídica do gene *lef-9*. (D) Filograma da sequência nucleotídica do gene *pif-2*. Os números nos ramos representam os valores de *bootstrap* (> 50%) para 1.000 replicatas. A barra de escala representa o número de substituições nucleotídicas por base.

## Análise MP não ponderada – Sequência peptídica

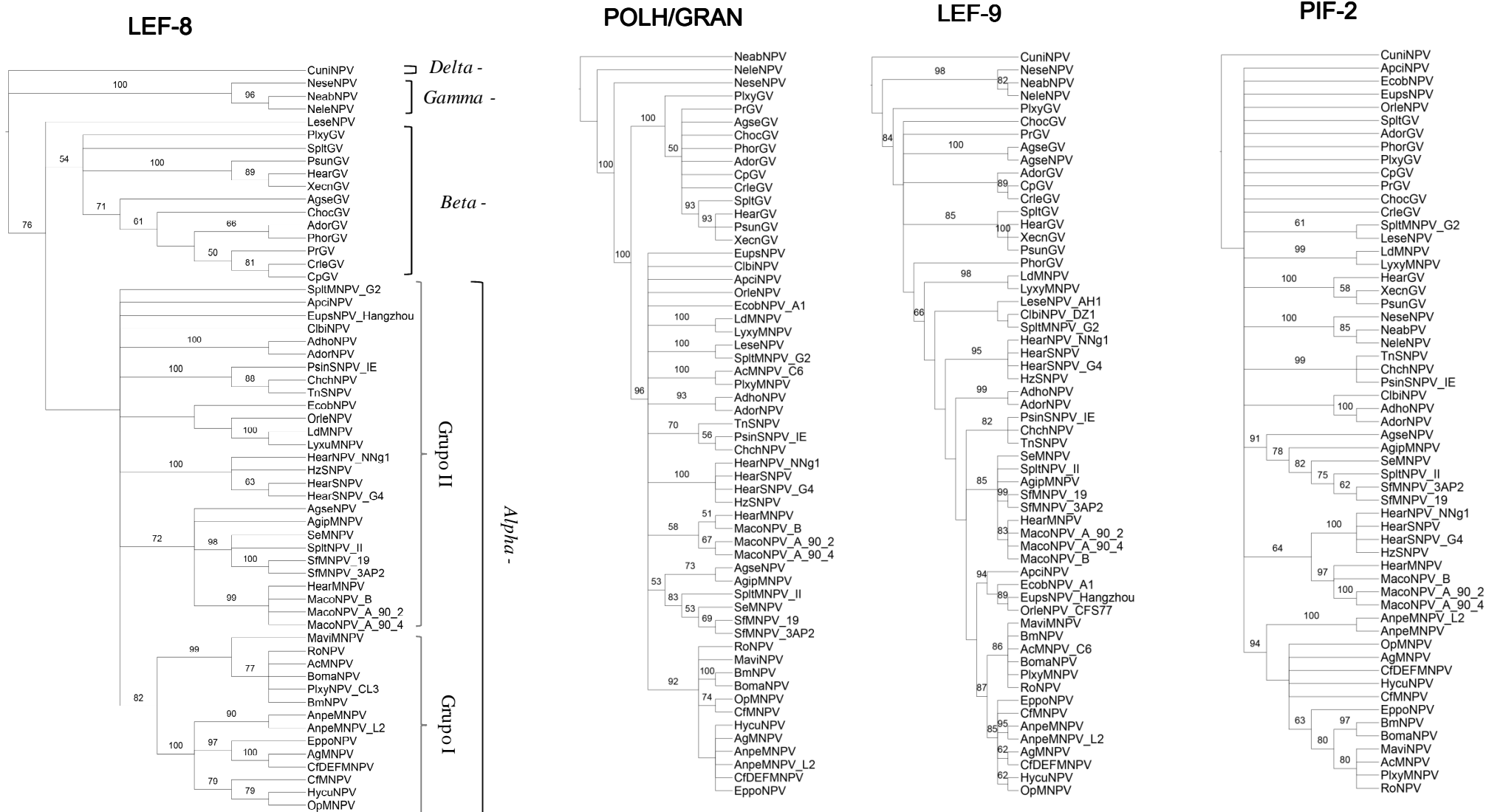


Figura 24. Árvores filogenéticas representando as relações filogenéticas da família *Baculoviridae* obtidas da análise MP não ponderada baseada nas sequências individuais LEF-8, LEF-9, PIF-2 e POLH/GRAN. (A) Cladograma do consenso de 458 árvores MP obtidas da análise da sequência peptídica do gene *lef-8*. (B) Cladograma do consenso de 1592 árvores MP obtidas da análise da sequência peptídica do gene *polh/gran*. (C) Cladograma do consenso de 500 árvores MP obtidas da análise da sequência peptídica do gene *lef-9*. (D) Cladograma do consenso de 281 árvores MP obtidas da análise da sequência peptídica do gene *pif-2*. Os números nos ramos representam os valores de *bootstrap* (> 50%) para 1.000 replicatas.

## Árvores MP com reponderação progressiva – Sequência peptídica

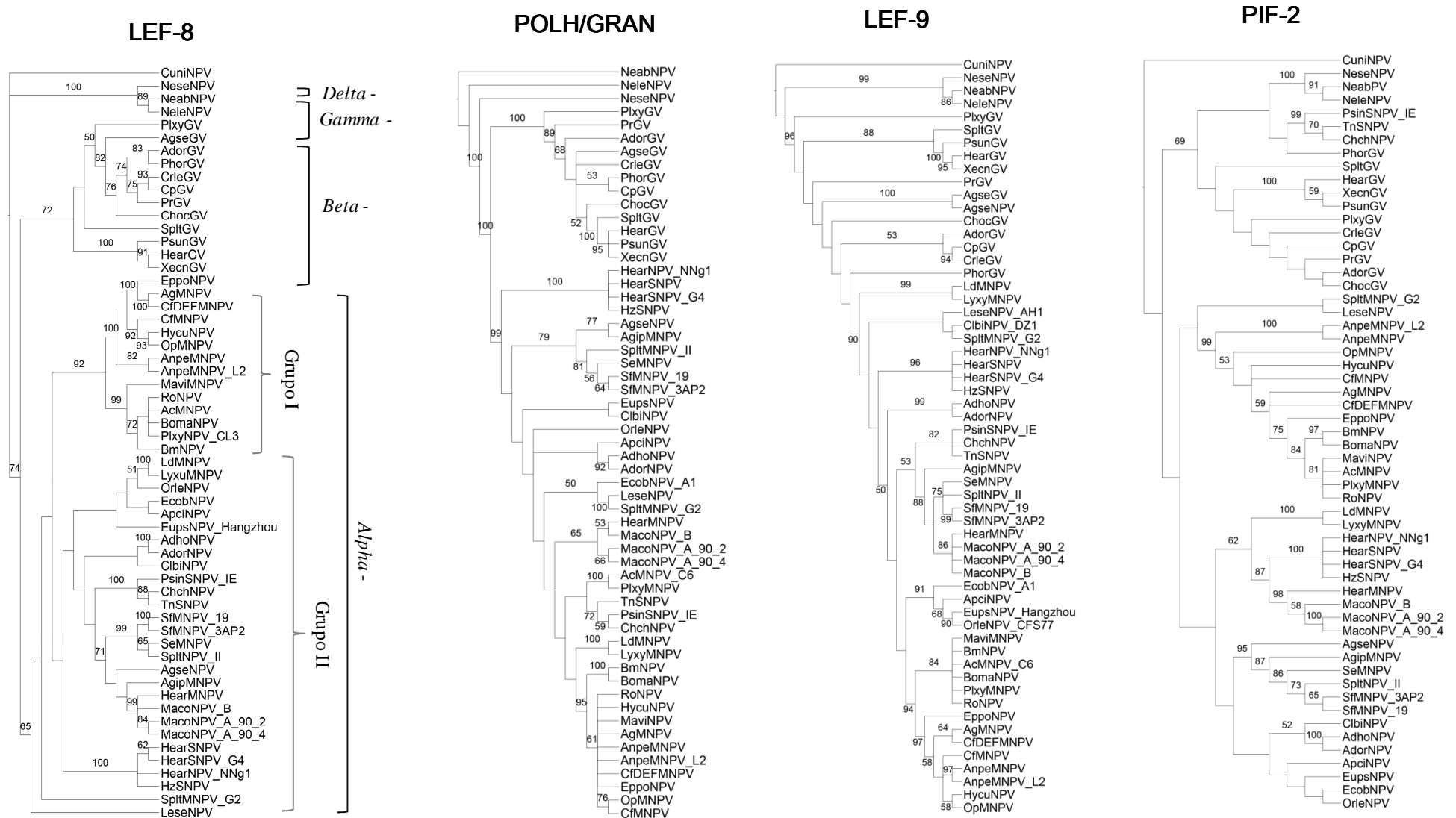
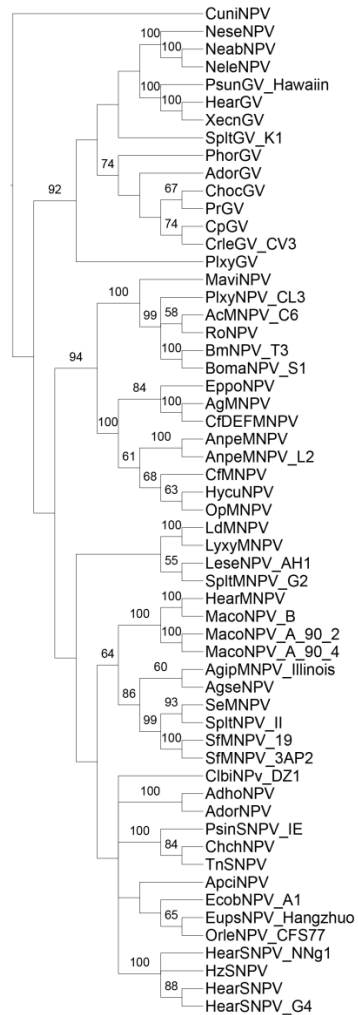


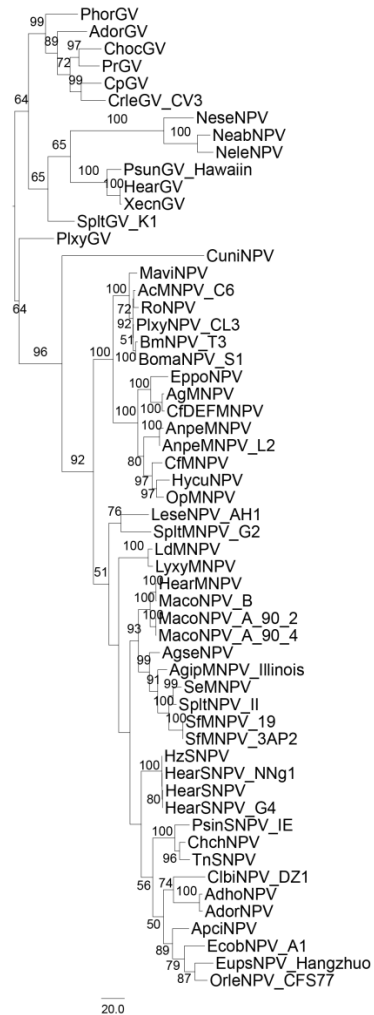
Figura 25. Árvores filogenéticas representando as relações filogenéticas da família *Baculoviridae* obtidas da análise MP com reponderação progressiva baseada nas sequências peptídicas LEF-8, LEF-9, PIF-2 e POLH/GRAN. (A) Cladograma do consenso de 2 árvores MP obtidas da análise da sequência peptídica do gene *lef-8*. (B) Cladograma do consenso de 120 árvores MP obtidas da análise da sequência peptídica do gene *polh/gran*. (C) Cladograma do consenso de 18 árvores MP obtidas da análise da sequência peptídica do gene *lef-9*. (D) Cladograma do consenso de 4 árvores MP obtidas da análise da sequência peptídica do gene *pif-2*. Os números nos ramos representam os valores de *bootstrap* (> 50%) para 1.000 replicatas.

## Análise filogenética da sequência nucleotídica concatenada dos genes: *lef-8*, *lef-9* e *pif-2*

### MP não ponderada



### MP reponderada



### Inferência Bayesiana

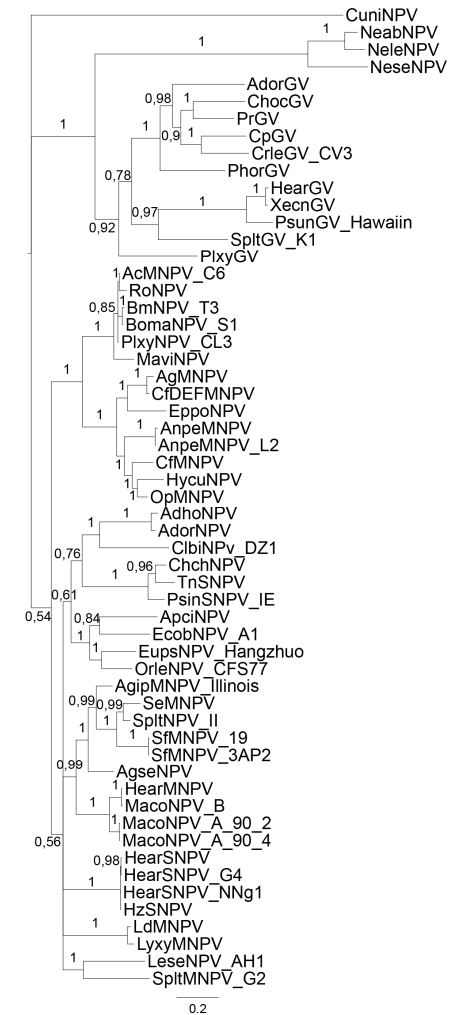
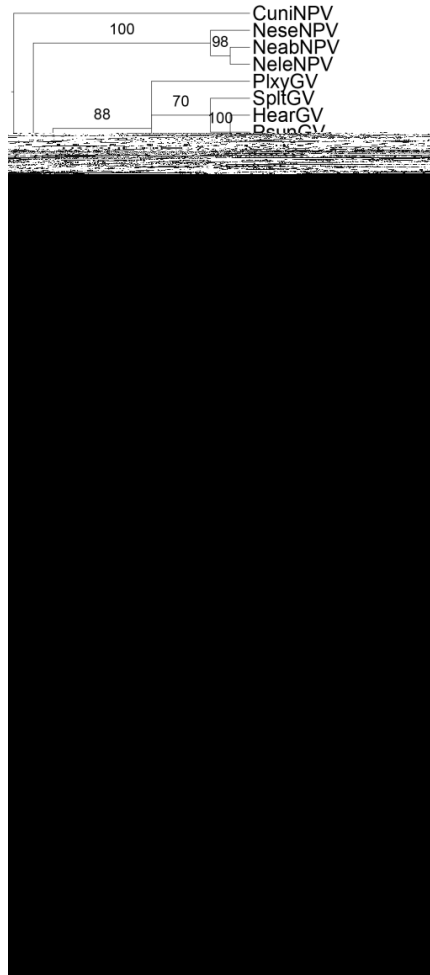


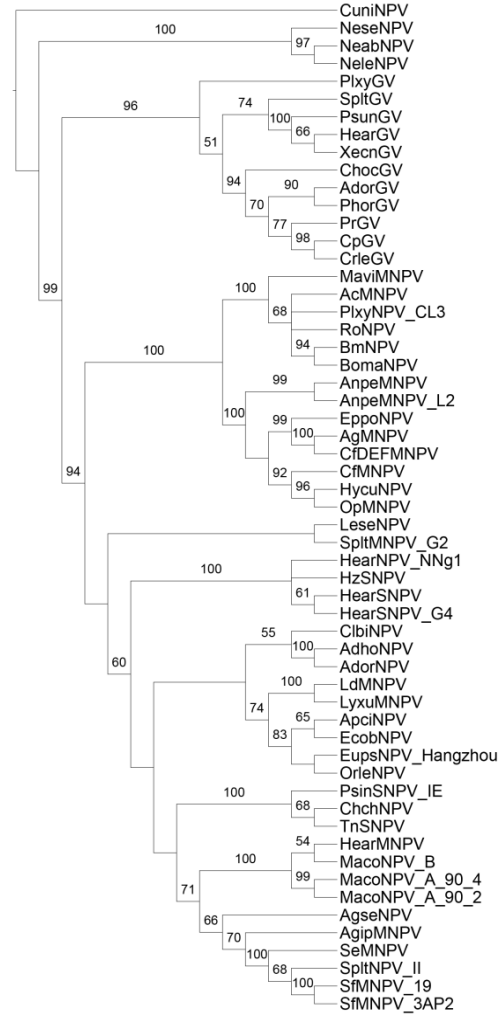
Figura 26. Árvores filogenéticas representando as relações filogenéticas da família *Baculoviridae* obtidas da análise MP e IB baseada na concatenação das sequências nucleotídicas dos genes *lef-8*, *lef-9* e *pif-2*. (A) Cladograma do consenso de 9 árvores MP não ponderadas (B) Filograma obtido da análise MP com reponderação progressiva. (C) Filograma obtido da análise IB utilizando modelo de evolução GTR+G+I. Os números nos ramos representam: análise MP - valores de *bootstrap* (> 50%) para 1.000 replicatas e análise IB - probabilidade posterior. A barra de escala representa o número de substituições nucleotídicas por base.

## Análise filogenética da concatenação das sequências peptídicas LEF-8, LEF-9 e PIF-2

### MP não ponderada



### MP reponderada



### Inferência Bayesiana

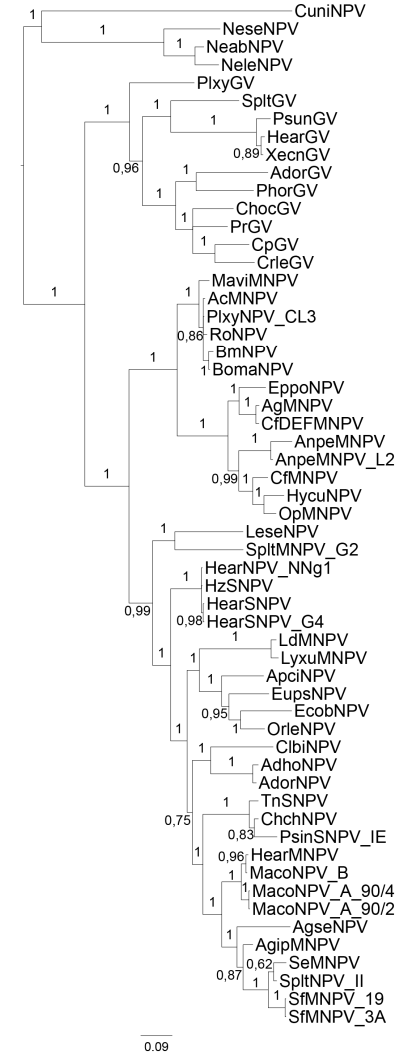
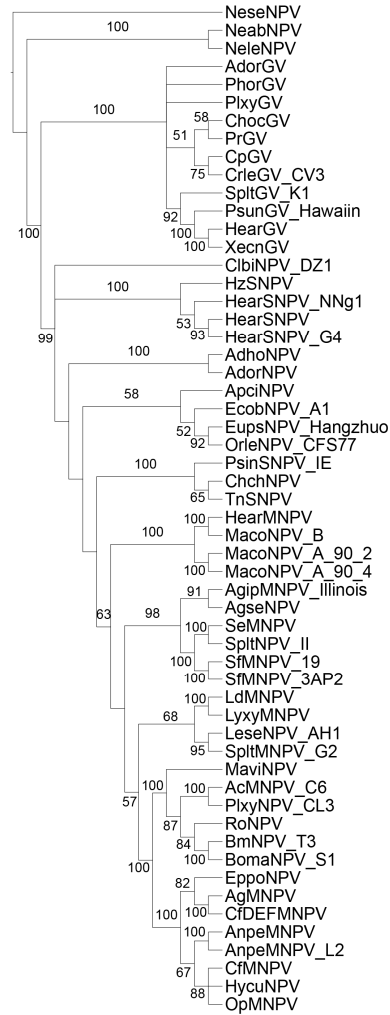


Figura 27. Árvores filogenéticas representando as relações filogenéticas da família *Baculoviridae* obtidas da análise MP e IB baseada na concatenação das sequências peptídicas LEF-8, LEF-9 e PIF-2. (A) Cladograma do consenso de 120 árvores MP não ponderadas (B) Cladograma do consenso de 2 árvores MP com reponderação progressiva. (C) Filograma obtido da análise IB utilizando modelo de evolução WAG. Os números nos ramos representam: análise MP - valores de *bootstrap* (> 50%) para 1.000 replicatas e análise IB - probabilidade posterior. A barra de escala representa o número de substituições nucleotídicas por base.

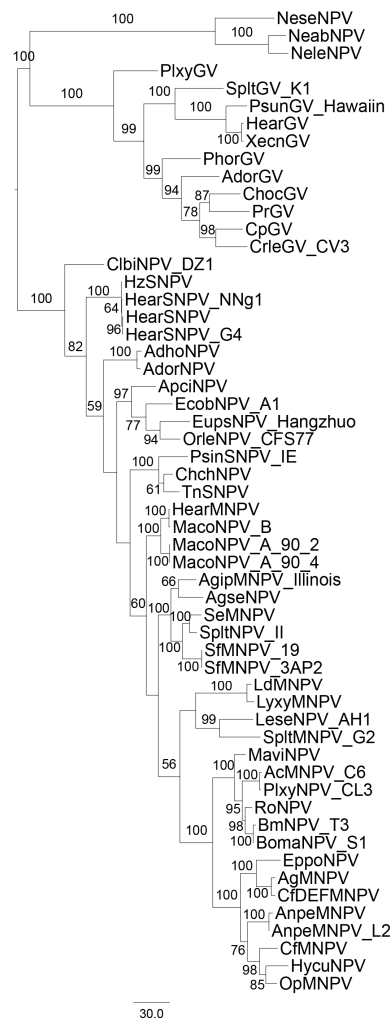


### Análise filogenética da sequência nucleotídica concatenada dos genes: *lef-8*, *polh/gran*, *lef-9* e *pif-2*

#### MP não ponderada



#### MP reponderada



#### Inferência Bayesiana

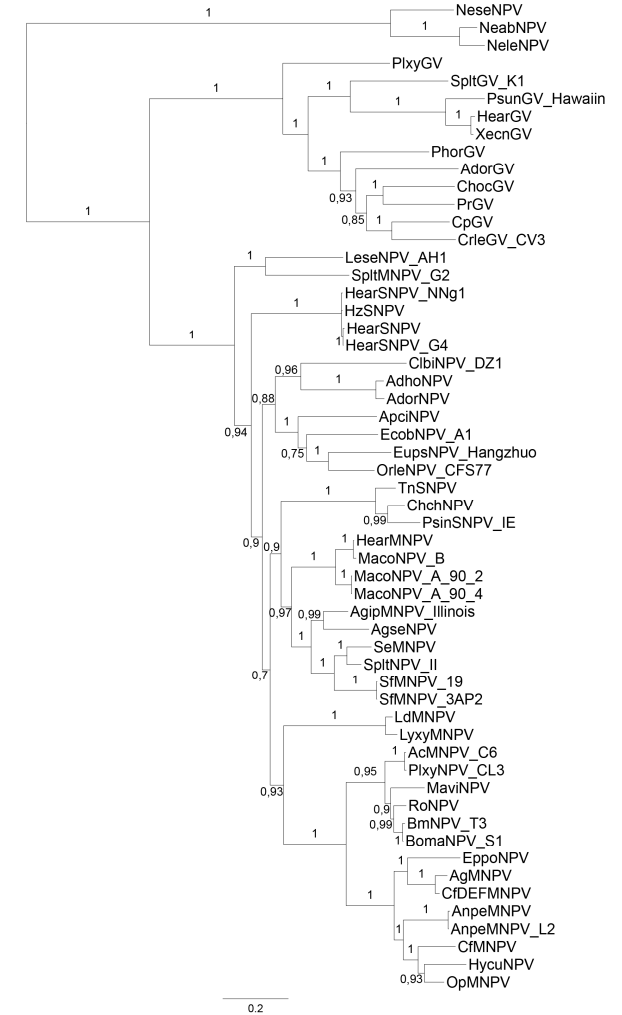


Figura 28. Árvores filogenéticas representando as relações filogenéticas da família *Baculoviridae* obtidas da análise MP e IB baseada na concatenação das sequências nucleotídicas dos genes *lef-8*, *lef-9*, *pif-2* e *polh/gran*. (A) Cladograma do consenso de 6 árvores MP não ponderadas (B) Filograma obtido da análise MP com reponderação progressiva. (C) Filograma obtido da análise IB utilizando modelo de evolução GTR+G+I. Os números nos ramos representam: análise MP - valores de *bootstrap* (> 50%) para 1.000 replicatas e análise IB - probabilidade posterior. A barra de escala representa o número de substituições nucleotídicas por base.

### Análise filogenética da concatenação das sequências peptídicas LEF-8, POLH/GRAN, LEF-9 e PIF-2

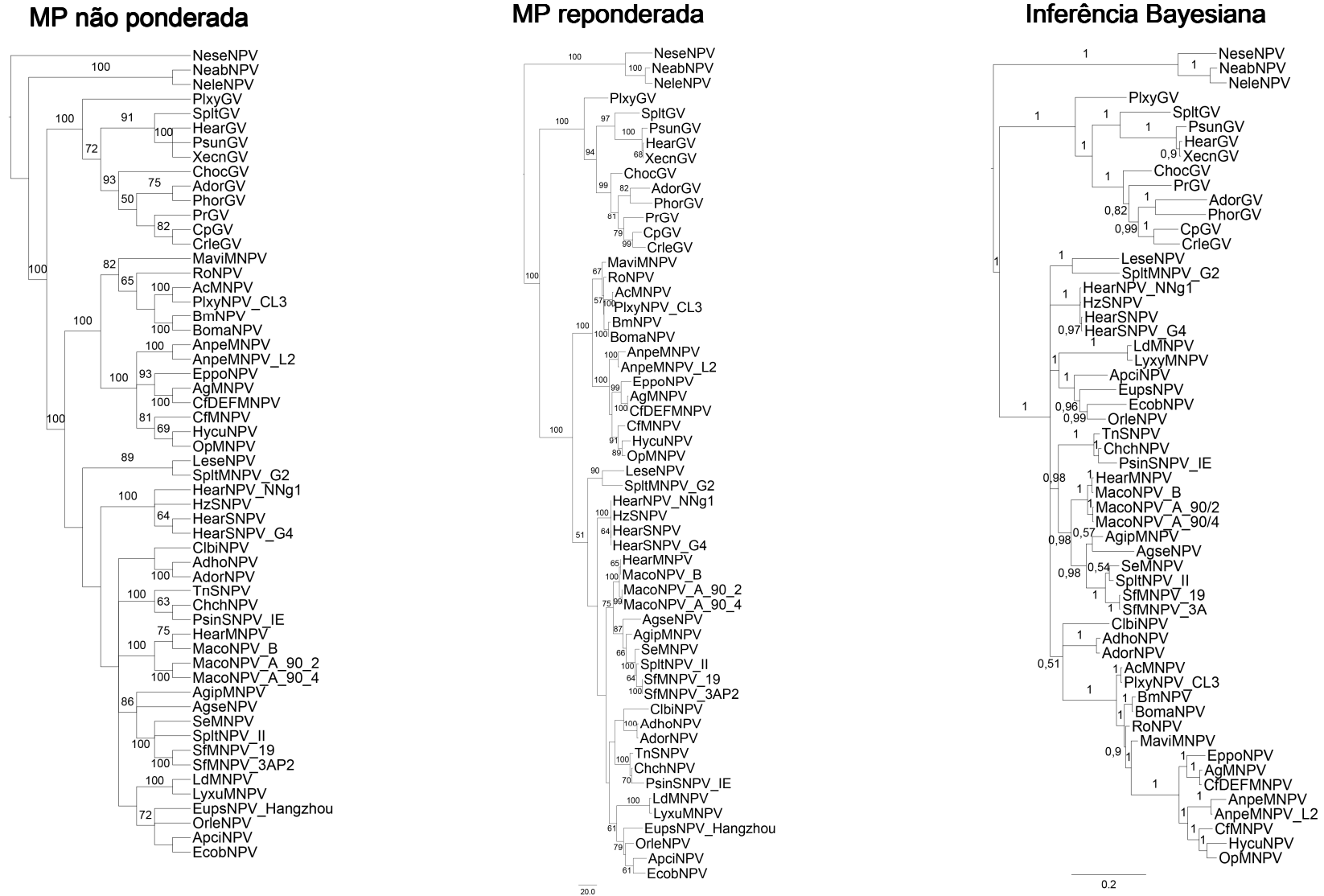


Figura 29. Árvores filogenéticas representando as relações filogenéticas da família *Baculoviridae* obtidas da análise MP e IB baseada na concatenação das sequências peptídicas LEF-8, POLH/GRAN, LEF-9 e PIF-2. (A) Cladograma do consenso de 22 árvores MP não ponderadas (B) Filograma obtido da análise MP com reponderação progressiva. (C) Filograma obtido da análise IB utilizando modelo de evolução WAG. Os números nos ramos representam: análise MP - valores de *bootstrap* (> 50%) para 1.000 replicatas e análise IB - probabilidade posterior. A barra de escala representa o número de substituições nucleotídicas por base.

## GLOSSÁRIO

**Algoritmo Monte Carlo via Cadeia de Markov (MCMC):** método de simulação baseados em Cadeias de Markov ergódicas onde a distribuição estacionária do processo estocástico é a distribuição *a posteriori* de interesse.

**Alinhamento de sequências:** hipótese de homologia posicional entre bases nucleotídicas ou aminoácidos que serve para comparar sequências supostamente ortólogas (sequência de uma mesma família gênica herdada diretamente de um ancestral comum) obtidas de diferentes indivíduos ou espécies.

**Árvore de consenso estrito:** contém apenas os agrupamentos encontrados em todas as topologias analisadas. É resultante da sumarização do grau de concordância de duas ou mais árvores obtidas em uma análise filogenética (Schneider, 2007).

**Bootstrap:** medida que testa o grau de confiabilidade de árvores filogenéticas. O teste *bootstrap* consiste na reamostragem de mesmo tamanho e com reposição dos dados da amostra original.

**Burn-in:** número de cadeias que serão descartadas que correspondem as amostras que não atingiram a fase estacionária na distribuição das probabilidades posteriores.

**Busca heurística:** busca da árvore mais parcimoniosa escolhendo caminhos que contêm a árvore com menor número de passos. Essa busca é utilizada quando é avaliado um elevado número de táxons. Outros métodos de busca de árvores são: busca exaustiva e *Branch and Bound*.

**Cladograma:** Representação gráfica utilizada em cladística que mostra as relações ancestrais entre os organismos estabelecidas em função dos caracteres homólogos.

**Coeficiente de similaridade:** estatística utilizada para comparação da similaridade entre duas amostras.

**Comprimento da árvore (nº de passos):** número total de mudanças necessárias para apoiar o relacionamento proposto para os táxons na árvore filogenética.

**Concatenação de sequências:** processo de acrescentar uma sequência de caracteres ao final de outra sequência de caracteres.

**Correlação cofenética:** é uma medida de validação utilizada, principalmente, nos métodos de agrupamento hierárquicos. A idéia básica é realizar uma comparação entre as distâncias efetivamente observadas entre os objetos e distâncias previstas a partir do processo de agrupamento (Barroso e Artes, 2003).

**Critério da informação Bayesiana:** método paramétrico (critério estatístico) que utiliza o argumento Bayesiano para seleção do modelo evolução.

**Critério de informação Akaike:** é uma estatística freqüentemente utilizada para a escolha da especificação ótima de uma equação de regressão quando não existem variáveis independentes comuns aos dois modelos em análise.

**Dendrograma:** é uma representação matemática e ilustrativa do procedimento de agrupamento através de uma estrutura de árvore (Everitt *et al.* 2001).

**Eletroferograma:** representação gerada por um seqüenciador automático que mostra o registro gráfico do sinal dos componentes da amostra em função do tempo permitindo a identificação da estrutura primária da molécula de DNA.

**Score:** comprimento ou número de passos da árvore.

**Filograma:** representação gráfica da relação de ancestralidade entre os organismos em que o tamanho dos ramos representa o número de mudanças ocorridas entre os nós.

**Gaps:** caracteres com uma deleção para pelo menos um táxon presentes no alinhamento das sequências de diferentes organismos.

**Grupo externo:** táxons próximos ao grupo interno, mas que não pertencem a ele e que servirão de referência na polarização dos caracteres e no enraizamento da árvore.

**Grupo monofilético:** conjunto de organismos que inclui uma espécie ancestral e todas as suas descendentes. **Clado** é sinônimo de grupo monofilético.

**Grupo parafilético:** grupo que não inclui todos os descendentes de um ancestral comum.

**Homoplasia:** característica que evoluiu independentemente por reversão ou paralelismo.

**ILD:** teste que mede a significância da incongruência entre os conjuntos de dados. Este teste foi descrito por Farris e colaboradores (1995a;1995b) e trata-se da diferença do número de passos da árvore obtida com os dados combinados e a soma dos comprimentos das árvores obtidos com os dados de cada partição. A diferença pode ser comparada a partições ao acaso para gerar uma probabilidade associada e assim avaliar a hipótese nula de que os dados são congruentes. Este teste é também chamado de teste de homogeneidade de partição (THP).

**Índice de consistência (IC):** quantidade relativa de homoplasias em um único caráter ou em uma árvore como um todo.

**Índice de consistência rescalonado (RC):** é o produto do índice de consistência pelo índice de retenção para a minimização do efeito das autopomorfias sobre o IC.

**Índice de homoplasia (IH):** é igual a 1 menos o índice de consistência.

**Índice de retenção (IR):** cálculo requerido para a estimativa do índice de consistência rescalonado utilizado para contornar o problema de superestimativa do valor de IC causado pela presença de autopomorfias.

**Matriz binária:** matriz em qual cada entrada é zero (0) ou um (1).

**Matriz de similaridade:** matriz de pontuação, contagem dos dados, que expressa a similaridade entre dois pontos dados.

**Modelos de substituição nucleotídica:** define uma probabilidade para cada substituição de um nucleotídeo para outro.

**Número de caracteres informativos para parcimônia:** sítios que apresentam no mínimo duas bases diferentes duplicadas.

**Oligonucleotídeos degenerados:** utilizados quando as seqüências de DNA não estão disponíveis ou são difíceis de serem alinhadas. Neste caso, é determinada as possíveis seqüências de nucleotídicas de um grupo de proteínas que serão utilizadas como molde para desenhar os oligonucleotídeos (iniciadores ou *primers*).

**PABA:** teste da adição de partição alterando o *bootstrap*, desenvolvido por Stuck et al. (2006), que se baseia na investigação da alteração do suporte dos ramos (*bootstrap*) com a adição das diferentes partições dos dados. Esse teste avalia se as partições interagem entre si, assim, se a adição for congruente, causará aumento no valor de *bootstrap* e, se for incongruente, causará diminuição desse valor de suporte do ramo.

**Politomia:** quando um ancestral comum dá origem a mais de dois descendentes o que significa que qualquer relação de parentesco é possível entre os táxons envolvidos. As politomias podem ser “macias” (*soft polytomy*), ou seja, resultado de dados insuficientes ou conflitantes e portanto um artefato de técnica ou “duras” (*hard polytomy*) que indicam um evento real de especiação praticamente simultânea de mais de dois grupos, como no caso de uma radiação adaptativa.

**Probabilidade posterior:** probabilidade de que a árvore (ou o clado) seja a correta (Schneider, 2007). Também chamado de credibilidade do clado ou probabilidade *a posteriori*.

**Reweighted:** método que aplica pesos diferentes para cada caráter com o intuito de diminuir a influência das homoplasias na análise filogenética.

**Stepwise Addition:** processo de adição aleatória e seqüencial dos táxons utilizado como estratégia da busca heurística para obtenção da árvore mais parcimoniosa que não seja em um ponto local.

**Teste de Mantel:** teste estatístico da correlação entre duas matrizes.

**Topologia:** padrão de ramificação da árvore filogenética.

**Transição:** substituição nucleotídica de uma purina → purina ou pirimidina → pirimidina. Purinas = citosina e timina; Pirimidinas = adenina e guanina.

**Transversões:** substituição nucleotídica de uma purina → pirimidina.

**Tree bisection and reconnection:** processo de secção da árvore em dois pedaços e ligação em outras partes (Schneider, 2007). Utilizado como estratégia da busca heurística.

**Variante genotípicas:** mistura de genótipos que existe entre os indivíduos de uma população.