



Universidad Nacional de La Plata

Facultad de Ciencias Exactas

Departamento de Ciencias Biológicas

Trabajo de Tesis Doctoral:

**Caracterización de baculovirus autóctonos para el control de la
oruga militar tardía (*Spodoptera frugiperda*)**

Tesista: Tomás Masson

Director: Dr. Prof. Víctor Romanowski

Directora: Dra. María Leticia Ferrelli

Año: 2021

El presente trabajo de tesis para optar al grado de Doctor de la Facultad de Ciencias Exactas (Area Ciencias Biológicas) de la Universidad Nacional de La Plata, ha sido realizado en el Instituto de Biotecnología y Biología Molecular (IBBM, UNLP-CONICET), Departamento de Ciencias Biológicas, Facultad de Ciencias Exactas, Universidad Nacional de La Plata bajo la dirección conjunta del Dr. Prof. Víctor Romanowski y la Dra. María Leticia Ferrelli.

Confirmando que el trabajo y los resultados presentados en esta tesis son propios. Toda información derivada de otras fuentes ha sido citada debidamente.

Resumen

El cogollero del maíz, también conocido como oruga militar tardía (*Spodoptera frugiperda*), es una plaga voraz que actualmente presenta una distribución global y representa una amenaza para los productores agrícolas. Dos miembros de la familia *Baculoviridae* son capaces de infectar a esta plaga, proporcionando una alternativa para su control biológico sin incurrir en el uso excesivo de pesticidas químicos.

En esta tesis se abordó la caracterización dos aislamientos argentinos de baculovirus capaces de infectar *S. frugiperda* (SfGV y SfMNPV) a través de distintas tecnologías ómicas. El principal objetivo de esta trabajo fue profundizar nuestro conocimiento sobre los factores genómicos y proteómicos involucrados en la adaptación de estos aislamientos nativos a su hospedador. Sumado a esto, se exploró la posibilidad de multiplicar el granulovirus SfGV en cultivos celulares, de forma similar a las metodologías empleadas con nucleopoliedrovirus. En última instancia, se espera que los conocimientos y datos generados en esta tesis contribuyan a mejorar la aplicabilidad de estos baculovirus como biopesticidas.

El estudio de la composición proteica de la partícula infectiva de granulovirus se desarrolla en el capítulo 2. En el capítulo 3 se discute la diversidad genética encontrada dentro de un aislamiento argentino de SfMNPV y sus posibles consecuencias en la adaptación del virus a su hospedador. En el capítulo 4 se describen los resultados preliminares de la caracterización del ciclo infectivo de SfGV en cultivos celulares de insecto. Por último, en el capítulo 5 se provee una discusión global de los resultados obtenidos y se proyectan las perspectivas de futuros estudios para caracterizar una proteína con capacidad agregante en SfGV y una colagenasa conservada en SfMNPV.

Abstract

The fall armyworm (*Spodoptera frugiperda*) is a voracious pest with a global distribution that poses a serious threat to agricultural producers. Members of the *Baculoviridae* family are capable of infecting this pest, providing an alternative for its biological control without the excessive use of chemical pesticides.

This thesis sought to characterize two Argentine baculovirus isolates infectious to *S. frugiperda* (SfGV and SfMNPV) through different omic technologies. The main goal of this job was to deepen our knowledge about the genomic and proteomic factors involved in the adaptation of these native isolates to its host. In addition to this, we explored the possibility of multiplying the granulovirus SfGV in cell cultures, in a similar fashion to the methodologies used with nucleopolyhedrovirus. Ultimately, we hope that this knowledge will contribute to improving the applicability of baculovirus as biopesticides.

The study of the protein composition of the granulovirus infective particle is developed in chapter 2. Chapter 3 discusses the genetic diversity found within an Argentine isolate of SfMNPV and its possible consequences on virus adaptation to its host. Chapter 4 describes the preliminary result of the characterization of the infective cycle of SfGV in insect cell cultures. Finally, chapter 5 provides a global discussion of the results obtained and points out the perspectives of future studies to characterize a protein with aggregating capacity in SfGV and a conserved collagenase in SfMNPV.

Agradecimientos

En primer lugar me gustaría hacer un reconocimiento al Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), al Instituto de Biotecnología y Biología Molecular (IBBM), y a la Universidad Nacional de La Plata (UNLP) por haber hecho posible esta trabajo a través de su financiamiento e infraestructura.

Al Dr. Santiago Haase, quien me acompañó en mis primeros pasos dentro la investigación científica y me ayudo a llevar a cabo mi trabajo de licenciatura. Por su paciencia y dedicación para ayudarme a incorporar los conceptos básicos del trabajo en el laboratorio.

Al Dr. Víctor Romanowski y la Dra. María Leticia Ferrelli, quienes me guiaron y aconsejaron durante la realización del presente trabajo. Por su predisposición a explorar las distintas ideas que surgieron en el camino y su rigurosidad científica en la discusión de los resultados.

A todo equipo de trabajo del laboratorio de Virología Molecular del IBBM. Por su generosidad y solidaridad para ayudar siempre que se presentó algún inconveniente.

A Romi, por su apoyo incondicional y el cariño a largo de este camino que estamos transitando juntos.

Abreviaturas

AcMNPV	<i>Autographa californica</i> Multiple Nucleopolyhedrovirus
AgMNPV	<i>Anticarsia gemmatalis</i> Multiple Nucleopolyhedrovirus
AHEV	<i>Adoxophyes honmai</i> Entomopoxvirus
AMEV	<i>Amsacta moorei</i> Entomopoxvirus
BLAST	Basic local alignment search tool
Bt	<i>Bacillus thuringiensis</i>
BV	Budded virus
CpGV	<i>Cydia pomonella</i> Granulovirus
DNA	Deoxyribonucleic acid
EMBRAPA	Empresa brasileira de pesquisa agropecuária
eGFP	Enhanced green fluorescent protein
EpapGV	<i>Epinotia aporema</i> Granulovirus
emPAI	Exponentially modified protein abundance index
IMV	Intranuclear microvesicles
ITRAQ	Isobaric tags for relative and absolute quantification
GV	Granulovirus
HearNPV	<i>Helicoverpa armigera</i> Nucleopolyhedrovirus
LC	Liquid chromatography
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MIP	Manejo integrado de plagas
MySEV	<i>Mythimna separata</i> Entomopoxvirus

MyunGV	<i>Mythimna unipuncta</i> Granulovirus
MSA	Multiple sequence alignment
NPV	Nucleopolyhedrovirus
NCBI	National Center for Biotechnology Information
OB	Occlusion body (cuerpo de oclusión)
ODV	Occlusion-derived virus
Pb	Pares de bases
PCR	Polymerase chain reaction
PIF	<i>Per os</i> infectivity factor
RT-PCR	Reverse transcription PCR
SDS-PAGE	Sodium dodecyl sulphate polyacrilamide gel electrophoresis
SfGV	<i>Spodoptera frugiperda</i> Granulovirus
SILAC	Stable isotope labeling by amino acids in cell culture
SeMNPV	<i>Spodoptera exigua</i> Multiple Nucleopolyhedrovirus
SfMNPV	<i>Spodoptera frugiperda</i> Multiple Nucleopolyhedrovirus
SpltNPV-II	<i>Spodoptera litura</i> Nucleopolyhedrovirus II
TCID	Tissue culture infectious dose

Índice general

Resumen	I
Abstract	III
Agradecimientos	v
Abreviaturas	vii
1. Introducción	1
1.1. El cogollero del maíz como plaga agronómica	2
1.2. Control biológico de plagas	3
1.3. Biología de los baculovirus	5
1.4. Composición proteica del virión	7
1.5. Ciclo infectivo de los baculovirus	9
1.6. Regulación génica en baculovirus	12
1.7. Baculovirus para el control de la oruga militar tardía	13
1.8. Diversidad genómica en poblaciones naturales de baculovirus	14
1.9. Fundamentos y alcance de este trabajo	16
2. Proteoma de los cuerpos de oclusión de granulovirus	17
2.1. Resumen	18
2.2. Introducción	19
2.3. Resultados	21
2.3.1. Purificación de los OB de EpapGV y SfGV	21

2.3.2.	Componentes estructurales detectados en los OB de EpapGV y SfGV	21
2.3.3.	Anotación proteogenómica de nuevos péptidos en el genoma de EpapGV y SfGV	36
2.3.4.	Conservación evolutiva de los componentes estructurales dentro del género <i>Betabaculovirus</i>	39
2.3.5.	Caracterización evolutiva y molecular de Spfr101	43
2.4.	Discusión	47
2.5.	Materiales y Métodos	51
2.5.1.	Larvas y virus	51
2.5.2.	Producción y purificación de los cuerpos de oclusión (OB)	51
2.5.3.	Microscopía electrónica de barrido y de transmisión de los OB	51
2.5.4.	Análisis por espectrometría de masas	52
2.5.5.	Análisis de los datos de espectrometría de masas	54
2.5.6.	Búsqueda de péptidos no anotados	54
2.5.7.	Agrupamiento de secuencias ortólogas	55
2.5.8.	Clonado de <i>spfr101</i> en el vector pIP	55
2.5.9.	Cultivo y transfección de células de insecto	56
2.5.10.	Predicción de la tendencia a la agregación	56
3.	Diversidad genómica en una población de SfMNPV	57
3.1.	Resumen	58
3.2.	Introducción	59
3.3.	Resultados	60
3.3.1.	Secuenciación profunda de una población de SfMNPV ARG-M	60
3.3.2.	Diversidad genómica presente en el aislamiento SfMNPV ARG-M	70
3.3.3.	Diversidad genética dentro de los aislamientos geográficos de SfMNPV	74
3.3.4.	Evolución molecular de <i>sf29</i> en virus grandes de insectos	76
3.4.	Discusión	79
3.5.	Materiales y Métodos	82
3.5.1.	Aislamiento del virus y secuenciación del genoma	82

3.5.2. Ensamblado del genoma y anotación	82
3.5.3. Asignación de variantes nucleotídicas puntuales y estructurales	83
3.5.4. Análisis filogenético	83
3.5.5. Estudio de la evolución molecular de <i>sf29</i>	84
3.5.6. Disponibilidad del código y los datos	84
4. Estudio de la replicación de SfGV en cultivo celular	85
4.1. Introducción	86
4.2. Resultados	88
4.2.1. Transfección de células Sf9 con DNA genómico de baculovirus	88
4.2.2. Efectos citopáticos y transcripcionales asociados a la entrada de DNA genómico de SfGV	91
4.2.3. Desarrollo de herramientas para facilitar el estudio SfGV en cultivo celular	93
4.3. Discusión	96
4.4. Materiales y Métodos	98
4.4.1. Células y virus	98
4.4.2. Producción y purificación de los cuerpos de oclusión (OB)	98
4.4.3. Purificación de DNA viral a partir de OB	99
4.4.4. Transfección de células Sf9	99
4.4.5. Extracción de DNA total partir de cultivo celular	100
4.4.6. PCR cuantitativa (qPCR)	100
4.4.7. Retrotranscripción seguida de PCR (RT-PCR)	100
4.4.8. Clonado molecular de los genes <i>ie-1</i> , <i>gp64</i> y <i>p35</i>	101
4.4.9. Oligonucleótidos	101
5. Conclusiones y perspectivas finales	103
5.1. Conclusiones	104
5.2. Perspectivas	106
6. Material suplementario	107
6.1. Grupos de genes ortólogos presentes en los proteomas de baculovirus	108

6.2. Variantes nucleotídicas puntuales detectadas dentro de la población de SfMNPV . . 117

Referencias **155**

Índice de figuras

1.1. Distintos estadios de desarrollo del cogollero del maíz	2
1.2. Estructural general del virión baculoviral	6
1.3. Árbol filogenético de la historia evolutiva de los virus entomopatógenos grandes con genoma a DNA específicos de invertebrado	8
1.4. Ciclo infectivo de los baculovirus	11
1.5. Programa transcripcional utilizado por los baculovirus durante su ciclo infectivo . . .	12
2.1. Morfología de los OB de SfGV	22
2.2. Morfología de los OB de EpapGV	23
2.3. Regiones codificantes correspondientes a las proteínas estructurales presentes en los OBs de EpapGV	33
2.4. Regiones codificantes correspondientes a las proteínas estructurales presentes en los OBs de SfGV	34
2.5. Retención de una region colineal enriquecida en genes codificantes para componentes estructurales del ODV	35
2.6. Homólogo de <i>ac110</i> presente en EpapGV	37
2.7. Proteína de fusión putativa entre Epap48 y Epap49	39
2.8. Conservación evolutiva del proteoma de los OB dentro del género Betabaculovirus .	41
2.9. Modelo semi-cuantitativo del OB de granulovirus	42
2.10. Historia evolutiva de la proteína Spfr101	43
2.11. Construcción y expresión de Spfr101::eGFP en células de insecto	45
2.12. Formación de agregados por Spfr101::eGFP en células de insecto	46

3.1. Diversidad genómica dentro de la población de SfMNPV ARG-M	71
3.2. Diversidad genética presente en los aislamientos geográficos de SfMNPV	75
3.3. Diversidad dentro del proteoma de SfMNPV	77
3.4. Evolución molecular de <i>sf29</i>	78
4.1. Preparación de DNA genómico baculoviral para los ensayos de transfección en células Sf9	89
4.2. Evaluación de distintos reactivos para la transfección de células Sf9 con DNA genómico de AcMNPV-DsRed	90
4.3. Transfección de células Sf9 con DNA de SfGV	91
4.4. Expresión de <i>ie-1</i> en células Sf9 transfectadas con DNA de SfGV	92
4.5. Cuantificación de los niveles de DNA viral por qPCR	93
4.6. Construcciones diseñadas para expresar genes auxiliares en células de insecto	95

Índice de cuadros

2.1. Componentes estructurales del cuerpo de oclusión de EpapGV	25
2.2. Componentes estructurales del cuerpo de oclusión de SfGV	28
2.3. Péptidos detectados por MS sin anotación previa	38
3.1. Secuencias codificantes anotadas en el genoma de SfMNPV ARG-M	62
3.2. Variantes estructurales detectadas con Delly (D) y Lumpy (L)	73
4.1. Cuantificación de los niveles de DNA virales para distintas muestras de células Sf9 transfectadas con DNA viral de AcMNPV-DsRed	94
4.2. Oligonucleótidos utilizados para amplificar las regiones codificantes de los genes ie-1, g64 y p35, y en las reacciones de qPCR/RT-PCR.	101
6.1. Grupos de ortología entre proteínas estructurales presentes en los OBs de la familia Baculoviridae. En cada celda se indica el número de ORF para cada genoma.	109
6.2. Variantes nucleotídicas puntuales intrahospedador (iSNV) detectadas con Lofreq	118

Capítulo 1

Introducción

1.1. El cogollero del maíz como plaga agronómica

El cogollero del maíz (*Spodoptera frugiperda* J. E. Smith, Lepidoptera: Noctuidae), también conocido como la oruga militar tardía, es una plaga polífaga altamente destructiva que ataca una gran variedad de cultivos, entre ellos algodón, maíz, soja, sorgo y caña de azúcar.¹ El mayor daño a los cultivos es causado por las larvas que consumen su follaje y brotes (Figura 1.1), limitando el crecimiento de los mismos durante las etapas vegetativa avanzada y reproductiva.² Esta especie es altamente invasiva y puede cubrir largas distancias gracias a sus fuertes capacidades de vuelo.³



Figura 1.1: Distintos estadios de desarrollo de *S. frugiperda*. Estadío larval causante del daño por defoliación de los cultivos (izquierda) y fase adulta, o polilla, donde ocurre la reproducción (derecha).

Si bien *S. frugiperda* es una especie endémica del continente americano, recientemente se ha reportado su expansión hacia Africa y Asia, despertando preocupación a nivel mundial sobre su posible impacto agronómico en estas regiones.^{2,4,5} Basados en estimaciones del año 2018 para 12 países africanos, cerca de 17.7 millones de toneladas de maíz podrían perderse anualmente debido a esta plaga, afectando principalmente a los pequeños agricultores.⁶ En Argentina, *S. frugiperda* se encuentra ampliamente distribuida en las regiones Noroeste, Noreste y Centro del país, por lo que representa una de las principales plagas del cultivo de maíz. Los cultivos tardíos (sembrados durante diciembre-enero) de la zona núcleo argentina (norte de la provincia de Buenos Aires, sur de la provincia de Santa Fe y sureste de la provincia de Córdoba) resultan los más afectados por esta plaga, ya que durante la época de verano es donde aparecen el mayor número de cogolleros.⁷

1.2. Control biológico de plagas

Históricamente, las estrategias empleadas para el control de *S. frugiperda*, al igual que para otros insectos plaga, se han centrado principalmente en el uso de pesticidas de síntesis químicos.^{8,9} Sin embargo, el uso intensivo de este tipo de productos ha contribuido a la aparición de poblaciones resistentes a estos tratamientos,^{10,11} motivando la búsqueda de opciones alternativas. A esto se suma también el impacto ambiental que acarrearán estos productos químicos, principalmente por la contaminación que generan en alimentos y napas de agua. En el caso de Europa, ya existen legislaciones que restringen los niveles residuales aceptables para la mayoría de los pesticidas sintéticos comerciales.¹²

Un enfoque alternativo al uso de productos químicos ha sido el desarrollo de variedades de cultivos que expresan la toxina insecticida Cry (derivada de la bacteria *Bacillus thuringiensis*), los cuales suelen referirse como cultivos Bt.¹³ Estos cultivos permitieron mejorar el manejo de la resistencia a insecticidas químicos y reducir su uso, con la consiguiente disminución del impacto ambiental.¹⁴ Sumado a esto, los agricultores han incorporado refugios con cultivos no Bt como forma de ralentizar la aparición de eventos de resistencia por parte de diferentes plagas.¹⁵ Sin embargo, estos esfuerzos no han sido suficientes y en años recientes se han reportado varios eventos de resistencia a esta toxina,¹⁶⁻¹⁸ lo que dificulta la aplicación de estrategias de control efectivas basadas exclusivamente en cultivos Bt.

Dentro de este panorama complejo, el manejo integrado de plagas (MIP) basado en bioinsumos es una estrategia atractiva y sustentable que podría complementar los métodos de control más tradicionales mencionados anteriormente.^{19,20} El MIP se centra en un enfoque holístico para el control de las plagas que afectan cultivos de relevancia económica, a la vez que se busca minimizar la utilización de productos químicos.²¹ Para alcanzar estos objetivos, el MIP busca atenuar el impacto de las plagas en el rendimiento de los cultivos por debajo del umbral de daño, sin que esto implique la eliminación completa de la especie plaga. Como principios activos, el MIP emplea organismos vivos, generalmente microorganismos, entre los cuales se destacan bacterias, hongos y virus entomopatógenos.²²⁻²⁵ Las principales ventajas de este enfoque se centran en la especificidad del tratamiento. Al verse afectadas un número reducido de especies blanco, se minimiza el impacto

en el agro-ecosistema. Además, al no ser patógenos humanos, estos productos presentan elevados niveles de seguridad para los trabajadores. A estas ventajas se suma su elevada degradabilidad en comparación con los productos de síntesis química. Actualmente, la mayor parte del mercado de global de bioinsecticidas corresponde a formulados en base a diferentes *Bt*, seguido por pesticidas basados en hongos y virus entomopatógenos.¹²

Los virus pertenecientes a la familia *Baculoviridae* son el principio activo más utilizado en el mercado para el desarrollo de biopesticidas virales.¹² Su rango de hospedadores estrecho y su inocuidad hacia los humanos los ha posicionado como una alternativa extremadamente segura y efectiva.^{26,27} Existen numerosos proyectos de control biológico alrededor del mundo centrados en la aplicación de baculovirus como agente pesticida. Uno de los factores fundamentales para asegurar el éxito y la sostenibilidad de estos programas es la participación conjunta de todos los actores involucrados en la cadena productiva (organismos estatales, entes reguladores, centros de investigación y productores agrícolas). El caso más paradigmático de control biológico basado en baculovirus es el de Brasil, donde la aplicación del baculovirus AgMNPV (*Anticarsia gemmatalis* Multiple Nucleopolyhedrovirus) para el control de la oruga de las leguminosas (*Anticarsia gemmatalis*, una plaga que afecta principalmente los cultivos de soja) llegó a cubrir 2 millones de hectáreas en el año 2003. Esto representa uno de los casos de uso más exitosos para un baculovirus y se debió en gran parte a la activa participación de la corporación brasileña de investigación agrícola (EMBRAPA) en tareas de producción y difusión de material educativo.^{28,29} En América y Europa, el granulovirus de la polilla del manzano (*Cydia pomonella* granulovirus, CpGV) también ha sido ampliamente adoptado para la protección de diversos frutos (durazno, manzana, nuez y pera).^{30,31} En la region del Alto Valle del Río Negro, la zona de mayor producción de manzanas y peras en Argentina, los productos basados en CpGV han demostrado altos niveles de protección y no se han evidenciado eventos de resistencia, a diferencia de las poblaciones europeas de *C. pomonella* donde han aparecido múltiples eventos de resistencia a este tipo de pesticidas comerciales.³²

Para el caso de *S. frugiperda*, su control se ha basado en la aplicación de insecticidas químicos y/o cultivos Bt, con el consiguiente impacto ambiental y el riesgo de aparición de eventos de resistencia, como se destacó previamente. En este escenario, el uso de los baculovirus para el control de esta plaga supone una alternativa atractiva, la cual ha sido explorada en diversos países de América

Latina.²⁰ Uno de los baculovirus más prometedores para el control de *S. frugiperda* es SfMNPV (*S. frugiperda* Multiple Nucleopolyhedrovirus). Una de las características distintivas de esta especie viral es su diversidad genética dentro de los diferentes aislamientos naturales, evidenciada por la presencia de genotipos con extensas deleciones.^{33,34} Mediante ensayos a campos se ha comprobado que las formulaciones basadas en SfMNPV presentan una capacidad de control similar a la de los insecticidas químicos. Sin embargo, los elevados costos de producción y la baja persistencia ambiental de estas formulaciones limitan la aplicación masiva de estos productos. Expandir nuestro conocimiento acerca de los factores biológicos y ambientales que determinan los niveles de virulencia de los baculovirus a campo resulta fundamental para facilitar su adopción por parte de los productores.^{35,36} Actualmente existen varias empresas que comercializan productos basados en baculovirus para el control de diversas plagas, lo que demuestra las posibilidades reales de desarrollo de este tipo de productos. Por ejemplo, la empresa suiza Andermatt Biocontrol ofrece una amplia variedad de productos amigables con el ambiente, entre los cuales se destacan más de 10 insecticidas basados en baculovirus. Por otra parte, la empresa australiana AgBiTech comercializa el pesticida Vivus Max, una formulación basada en el baculovirus *Helicoverpa armigera* Nucleopolyhedrovirus (HearNPV) para el control del gusano del algodón (*Helicoverpa armigera*).

1.3. Biología de los baculovirus

La familia *Baculoviridae* comprende un grupo diverso de virus entomopatógenicos que infectan mayoritariamente a insectos pertenecientes al orden *Lepidoptera*, y en menor medida a los órdenes *Himenoptera* y *Diptera*.^{37,38} El nombre baculovirus deriva de la palabra latina *baculum* (báculo, bastón), haciendo referencia a la forma alargada de sus viriones.³⁹

La información genética en los baculovirus es codificada por un genoma a DNA doble cadena circular superenrollado, cuyo tamaño se ubica en el rango de 80-180 kilo pb y codifica alrededor de 90-180 genes.^{40,41} De estos, un grupo de 38 genes involucrados en funciones esenciales para la infección viral (denominados *core genes*) han sido identificados en todos los genomas de baculovirus.^{42,43} Este genoma viral es empaquetado dentro de una nucleocápside cilíndrica de 200-400 nm de largo y 40-60 nm de diámetro. Los viriones pueden encontrarse en dos formas, como viriones ocluidos (*occlusion-derived virus*, ODV) y como viriones brotantes (*budded virus*, BV).⁴⁴⁻⁴⁶ A pesar que la estructura

de la nucleocápside resulta similar para ambos morfotipos, estos difieren en la composición de su envoltura lipídica y en su función dentro del ciclo infeccioso (Figura 1.2).

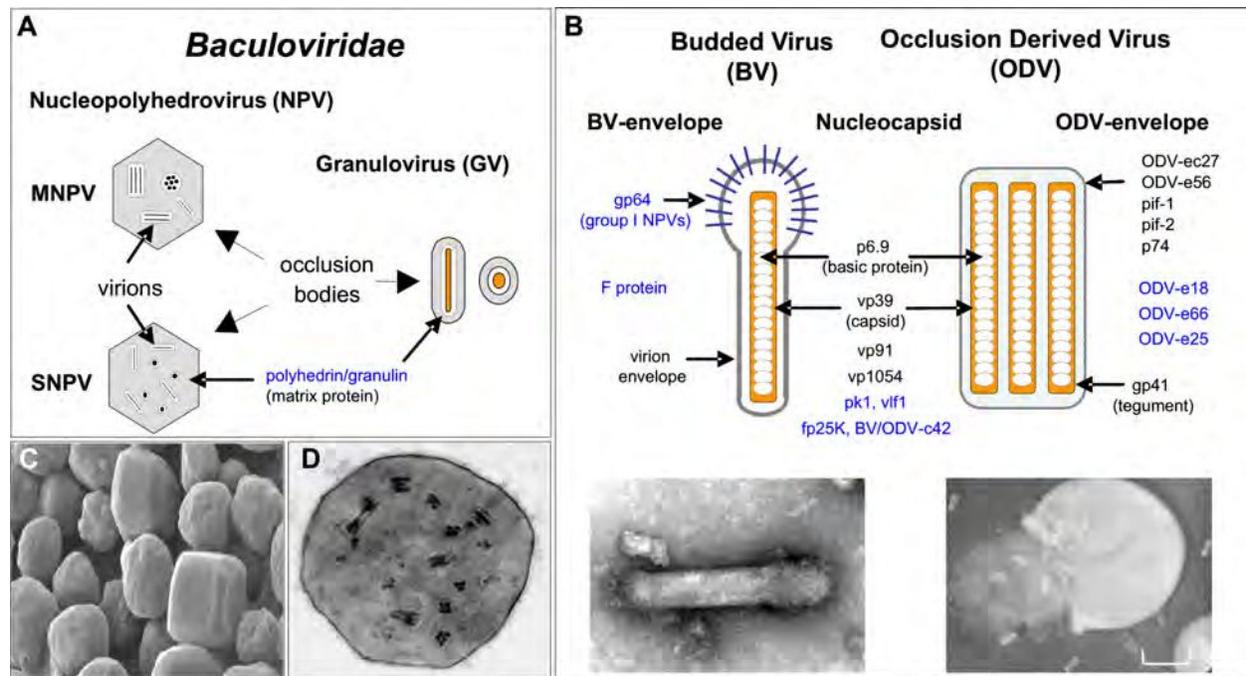


Figura 1.2: Estructura general del virión baculoviral. **A** Organización de los OB de nucleopoliedrovirus (izquierda) y granulovirus (derecha). En el caso de los NPV, los viriones pueden encontrarse como simples o múltiples. **B** Organización y proteínas mayoritarias de los dos morfotipos baculovirales, BV (izquierda) y ODV (derecha). En la parte inferior del panel se pueden observar dos micrografías electrónicas de las nucleocápsides de AcMNPV. **C** Micrografía electrónica de barrido correspondiente a los OB de AcMNPV. **D** Micrografía electrónica de transmisión del OB de AcMNPV. Adaptado de.⁴⁰

Los miembros de la familia *Baculoviridae* se caracterizan por formar estructuras paracristalinas de naturaleza proteica conocidas como cuerpos de oclusión (*occlusion bodies*, OB).⁴⁷⁻⁴⁹ Los OB son estructuras estables que pueden resistir condiciones ambientales como la desecación, lo que permite que los viriones mantengan su infectividad en el ambiente. A partir de la morfología del OB, los baculovirus fueron clasificados inicialmente en dos grupos, los nucleopoliedrovirus (NPV) de forma poliédrica y los granulovirus (GV) que adoptan una forma granular. Los poliedros presentan un tamaño en el orden de los 0.6-2 μM de diámetro, por lo que son visibles en el microscopio óptico convencional, mientras que los gránulos son más pequeños (0.2-0.4 μM de diámetro) y requieren de un microscopio de campo oscuro para poder visualizarse. A esta organización estructural de los OB se suma el hecho de que las nucleocápsides de los NPV pueden organizarse dentro de su envoltura

como agregados simples (SNPV) o múltiples (MNPV).⁵⁰

El origen de los baculovirus ha sido situado hace 310 millones de años en el período paleozoico, junto con la aparición de otras familias de virus entomopatógenos grandes con un genoma a DNA, como los nudivirus y los bracovirus.^{51,52} Mas aún, 20 *core genes* baculovirales son compartidos con estas familias virales, lo que sugiere un origen compartido a partir de un ancestro común (Figura 1.3). Recientemente, análisis filogenómicos basados en los 38 *core genes* conservados en todas las especies de baculovirus han permitido una clasificación a nivel molecular de esta familia en 4 grupos taxonómicos con elevado soporte estadístico: *Alphabaculovirus* (NPVs específicos de lepidópteros), *Betabaculovirus* (GVs), *Gammabaculovirus* (NPVs específicos de himenópteros) y *Deltabaculovirus* (NPVs específicos de dípteros).⁵³

1.4. Composición proteica del virión

La formación de las nucleocápsides baculovirales puede dividirse en dos etapas separadas que se llevan a cabo dentro del núcleo celular: el ensamblaje de cápsides vacías y el empaquetado del genoma viral dentro de estas.⁴⁴ La proteína de cápside mayoritaria, VP39, es la encargada de formar las estructuras helicoidales huecas que recubrirán el genoma viral.⁵⁴ Acompañando a VP39, aparecen otras proteínas de cápside minoritarias, como PP78/83 (involucrada en la translocación mediada por actina de los viriones hacia el núcleo celular),⁵⁵ P40/C42 (factor viral que inhibe la degradación de PP78/83),⁵⁶ 38K (necesaria para el ensamblaje de la nucleocápside),⁵⁷ VP80 (involucrada en la salida de las nucleocápsides del núcleo),⁵⁸ VP1054 (necesario para el transporte de VP39 hacia el estroma virogénico)⁵⁹ y P24, de función desconocida.⁶⁰ El genoma baculoviral empaquetado dentro de la nucleocápside se encuentra bajo la forma de un complejo nucleoproteico junto con la proteína básica P6.9, la cual es necesaria para neutralizar las cargas negativas presentes en los grupos fosfato del ADN.⁶¹

Además de los componentes estructurales mayoritarios, se ha detectado una plétora de proteínas accesorias acompañando a la nucleocápside, probablemente como una forma de aportar funciones auxiliares durante el proceso de infección. Si bien se creía que las nucleocápsides de los BV y ODV eran prácticamente idénticas, estudios proteómicos recientes han evidenciado componentes propios

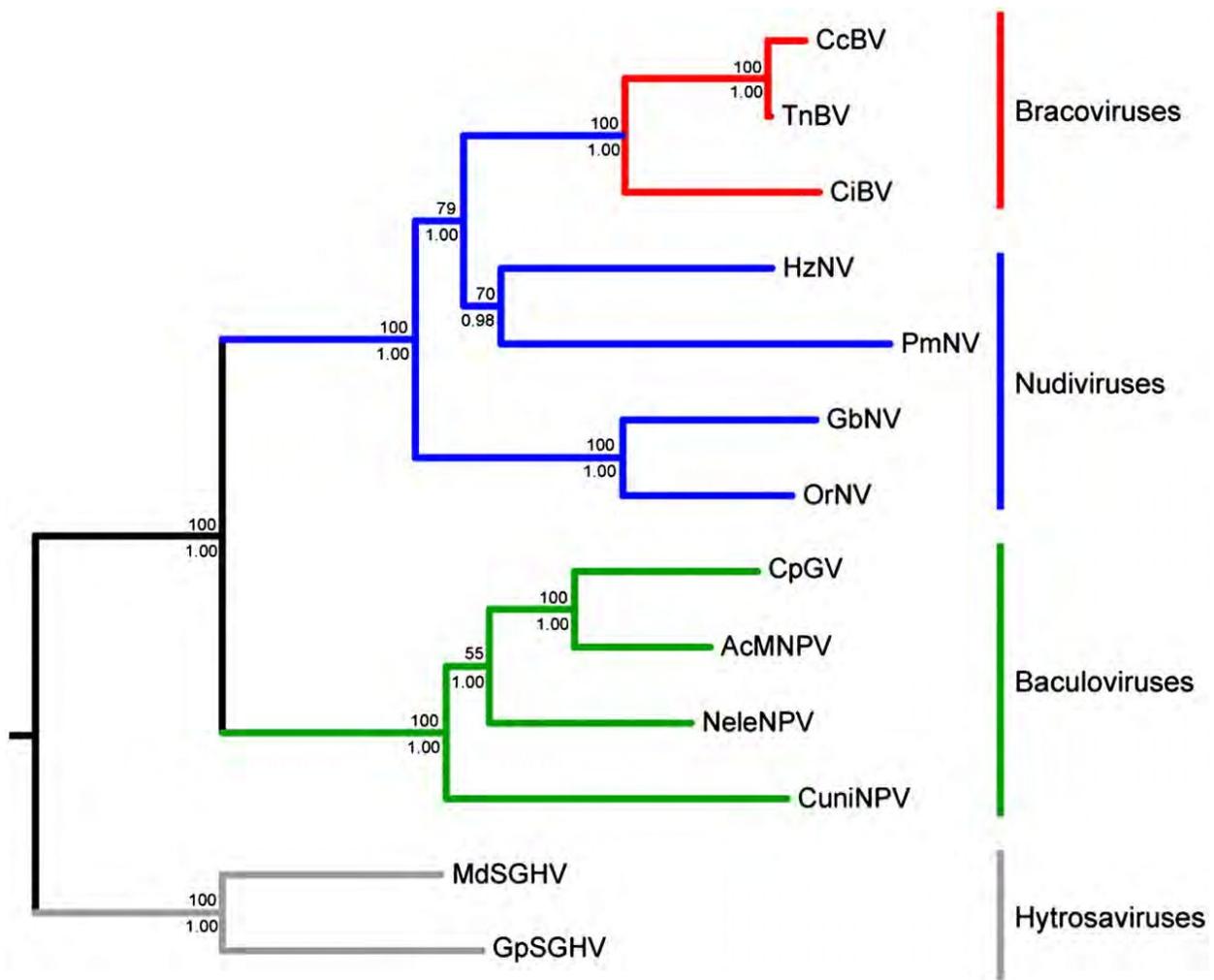


Figura 1.3: Árbol filogenético de la historia evolutiva de los virus entomopatógenos grandes con genoma a DNA específicos de invertebrados. La reconstrucción filogenética se llevó a cabo utilizando un grupo de 20 *core genes* compartido por todas las especies mediante criterios de máxima verosimilitud e inferencia bayesiana. Los valores de soporte para cada nodo ancestral dentro de la filogenia fueron calculados a partir de 1000 réplicas de bootstrap. Los hytrosavirus fueron empleados como grupo externo para enraizar el árbol. Adaptado de.⁵²

para cada fenotipo,⁶²⁻⁶⁵ lo que refleja en parte la especialización funcional de cada uno. A partir de diversos estudios proteómicos realizados en distintas especies de baculovirus⁶⁶⁻⁷² se ha hecho evidente la elevada variabilidad del proteoma de la partícula viral (tanto BV como ODV), lo que dificulta discernir entre componentes *bona fide* y contaminantes que no han sido removidos por los pasos de purificación. La aplicación en el futuro de técnicas cuantitativas que presenten mayor sensibilidad y robustez proveerá una visión más completa y confiable de estas estructuras.

El tegumento de la partícula viral, definido como el componente proteico que se localiza entre la nucleocápside y la envoltura lipídica, esta compuesto mayoritariamente por la proteína O-glicosilada GP41. Aunque su función permanece en mayor parte desconocida, los trímeros de GP41 han sido implicados en la correcta formación de los ODV.⁷³

La envoltura lipídica que rodea a la nucleocápside y el tegumento es uno de los puntos de mayor diferencias entre los BV y los ODV. Mientras que en los BV la envoltura se encuentra decorada casi exclusivamente por la proteína de fusión F, y en algunas especies de alphabaculovirus por GP64, los ODV presentan un complejo macromolecular de más de 500 kDa conformado por un grupo de proteínas, los *per os* infectivity factors (PIF), necesarios para la infección de las células columnares del intestino.^{74,75} Los componentes PIF1, PIF2 y PIF3 son los encargados de formar un complejo núcleo de ~230 kDa al cual luego se agregan PIF0, PIF4, PIF6, PIF7, PIF8 y PIF9 para formar el complejo completo, el cual se ensambla a nivel de la membrana nuclear interna o del retículo endoplasmático y luego se libera como microvesículas que se unen a las nucleocápsides.

1.5. Ciclo infectivo de los baculovirus

La transmisión horizontal en el ambiente de los baculovirus se da cuando los OB presentes en follaje contaminado son ingeridos (Figura 1.4).^{76,77} Una vez dentro del tracto gastrointestinal de la larva, los OB son capaces de atravesarlo hasta llegar al lumen del intestino medio. El elevado pH en esta región del intestino conduce a la desintegración de los OB y la presencia de proteinasas favorece la eliminación de la envoltura del poliedro, lo que desemboca en la liberación de los ODV.³⁹ La membrana peritrófica es una capa formada por quitina, mucopolisacáridos y proteínas que contribuye a la protección del epitelio del intestino medio. Varias especies de alpha y betabaculovirus incluyen

metaloproteasas (conocidas como Enhancins, factores virales potenciadores (VEF) o factores sinérgicos) dentro de los OB con el fin de acelerar la degradación de la membrana peritrófica^{78,79} y facilitar la entrada de los ODV a las células epiteliales del intestino. Este proceso denominado infección primaria, se da a través de la fusión de sus membranas mediada por el complejo supramolecular PIF.⁷⁵

Las nucleocápsides liberadas hacia el citosol celular son translocadas hacia el núcleo celular a través de un transporte mediado por actina.⁸¹ Una vez dentro del núcleo celular, la expresión temprana de varias proteínas replicativas (IE1, LEF3 y DNA-Helicase)⁸² y la presencia de modificaciones epigenéticas⁸³ induce la remodelación de la cromatina celular y la formación del estroma virogénico. Una vez completada la replicación del ADN viral comienza el ensamblaje de nuevas nucleocápsides dentro del estroma virogénico, las cuales luego serán destinadas a formar BV y ODV. Debido a que las células columnares presentan varios mecanismos antivirales (principalmente el desprendimiento del epitelio y la inducción de apoptosis celular), la infección baculoviral debe propagarse hacia los tejidos periféricos para poder sostenerse, lo que se conoce como infección secundaria o sistémica.⁸⁴ Los BV producidos en las células del epitelio intestinal son los encargados de llevar a cabo la infección sistémica hacia el resto de los tejidos del hospedador, un proceso que se ve facilitado por la presencia de una proteína de fusión en su envoltura lipídica.⁸⁵ Adicionalmente, se ha hipotetizado que una fracción de las nucleocápsides que entran a la célula epitelial podrían saltar la etapa replicativa dentro de esta y movilizarse a través de la membrana basal del epitelio hacia la hemolinfa.⁸⁶

Durante las fases tardías de la infección, una fracción de las nucleocápsides presentes en el núcleo celular son retenidas para formar nuevos ODV. Estas, adquieren su membrana lipídica a partir de la envoltura nuclear, la cual se encuentra decorada con una gran variedad de proteínas virales.⁸⁷ La elevada concentración de poliedrina durante las etapas tardías de la infección favorece la cristalización de esta sobre los ODV, lo que conduce a la formación de nuevos OB. Dado que los OB se producen dentro del núcleo de las células infectadas, el cadáver del hospedador debe ser degradado para que los OB puedan ser liberados al ambiente. Varias especies de baculovirus han desarrollado mecanismos de licuefacción basados en enzimas degradativas (como catepsinas y quitinasas), los cuales les permiten una liberación más rápida y eficiente de los OB al ambiente.⁸⁸

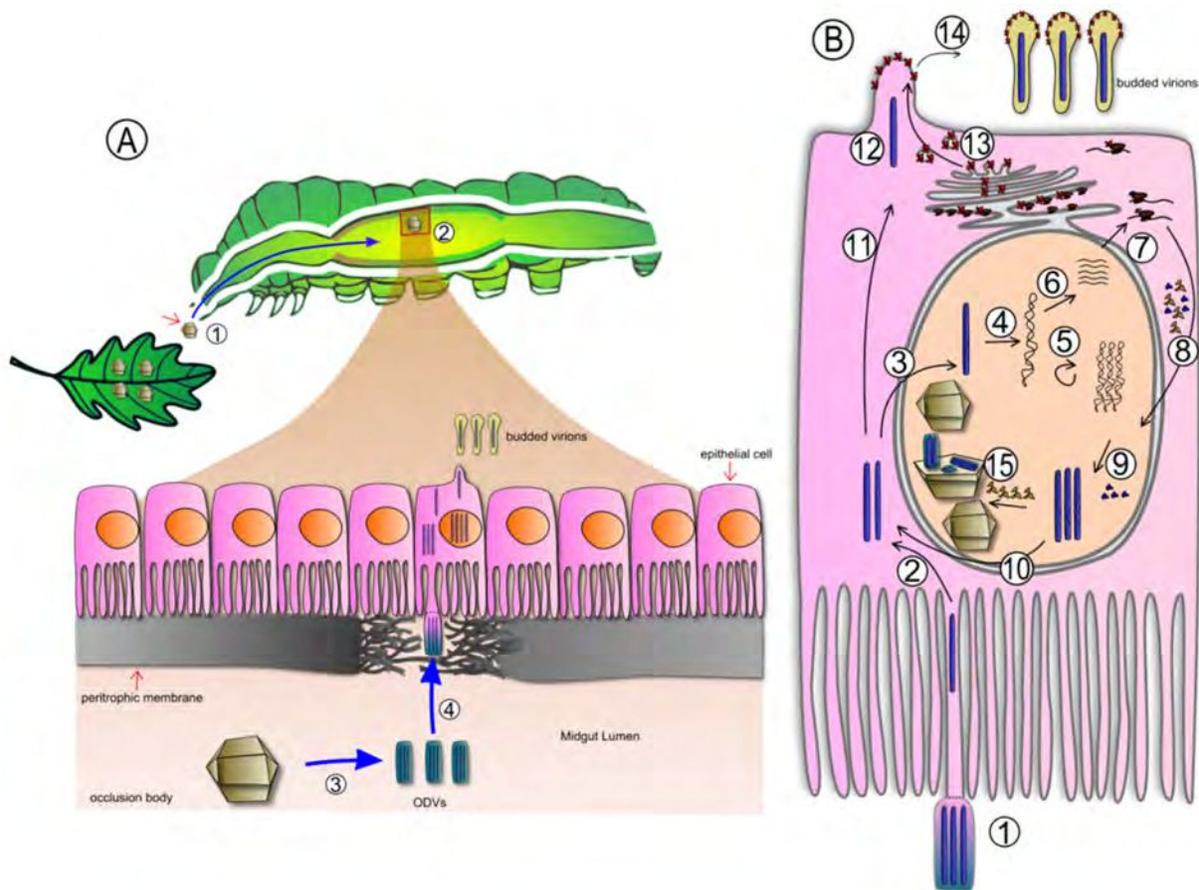


Figura 1.4: Ciclo infeccioso de los baculovirus. **A** Representación de una larva de lepidóptero. Un cuerpo de oclusión (OB) presente en el ambiente es ingerido por la larva comenzando un nuevo ciclo de infección (1). Cuando el OB alcanza el intestino medio (2) se disuelve en el medio alcalino, liberando ODV (3). La membrana peritrófica, que protege a las células epiteliales del contacto directo con el contenido intestinal, es degradada por la acción de proteínas codificadas por el virus y la célula hospedadora presentes en el OB (4), permitiendo el ingreso de los ODV en las células (5). **B** Representación del ciclo de vida de un baculovirus. Los ODV ingresan en la célula a través de la fusión con las microvellosidades (1), liberando las nucleocápsides (NC) en el citoplasma (2). Las NC pueden ingresar en el núcleo (3) donde se desensamblan liberando el genoma viral (4). De esta forma los genes tempranos del virus son transcritos (6) y traducidos en complejos ribosomiales en el citoplasma (7). Algunos productos de estos genes traslocan al núcleo para participar en la replicación viral (5), la transcripción de los genes tardíos (6) y el ensamblado de nuevas NC (9). En las primeras etapas de la infección, las NC son transportadas al citoplasma (10) y emergen de la célula por brotación de la membrana citoplasmática basolateral (11) (12) en forma de viriones brotantes (14) en las regiones donde se acumula la proteína de fusión viral (13), previamente expresada y direccionada por la ruta secretoria. En la etapa tardía de la infección, en cambio, las NC son envueltas en el núcleo y ocluidas en la matriz proteica de poliedrina (15). Adaptado de.⁸⁰

1.6. Regulación génica en baculovirus

El programa transcripcional de los baculovirus se encuentra segmentado en tres fases sucesivas: temprana, tardía y muy tardía (Figura 1.5). La transcripción de los genes virales durante la fase temprana es llevada a cabo por la RNA polimerasa II del hospedador y ocurre inmediatamente luego de la entrada a la célula. Estos genes pueden subdividirse en dos categorías: inmediatos y retrasado. La expresión de los genes tempranos inmediatos no requiere de ningún factor adicional y conduce a la aparición de reguladores transcripcionales que modulan a los genes tempranos retrasados involucrados en la replicación y transcripción viral. Los genes tardíos y muy tardíos están asociados con la aparición de BV y ODV en la célula pero necesitan de la replicación viral para poder ser expresados. En estas etapas una RNA polimerasa viral es la encargada de conducir la expresión de genes estructurales necesarios para el ensamblaje de los viriones y posteriormente los OB.⁸⁹

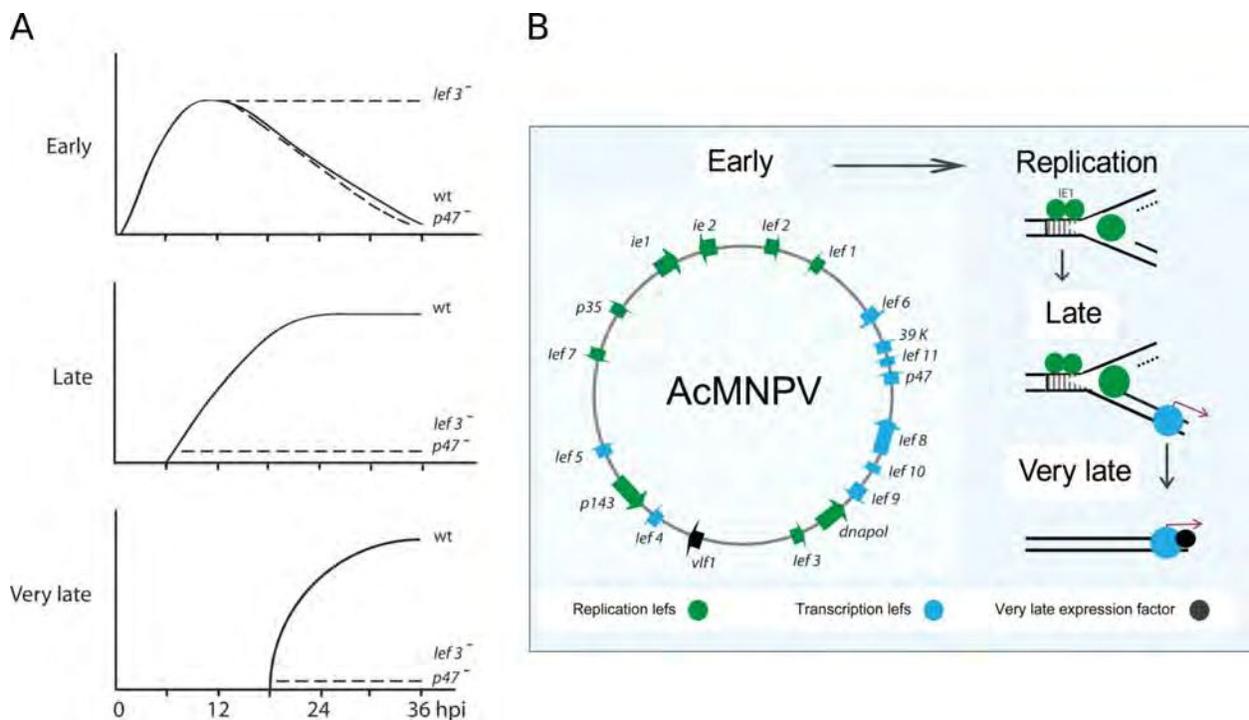


Figura 1.5: Programa transcripcional utilizado por los baculovirus durante su ciclo infeccioso. A Perfiles de expresión para los distintos grupos temporales de genes para AcMNPV en células Sf21. **B** Cascada de eventos transcripcionales durante el ciclo infeccioso. Adaptado de.⁸⁹

1.7. Baculovirus para el control de la oruga militar tardía

A la fecha se han descrito 2 especies de baculovirus capaces de infectar a la oruga militar tardía: *Spodoptera frugiperda* Multiple Nucleopolyhedrovirus (SfMNPV) y *Spodoptera frugiperda* Granulovirus (SfGV).^{90,91} Para SfMNPV se han logrado aislar y secuenciar completamente cinco aislamientos de distintas regiones geográficas (EEUU, Brasil, 2 de Nicaragua y Colombia), lo que ha permitido estudiar los patrones de conservación y variabilidad del genoma viral en relación con su adaptación a su hospedador. Sumado a esto, se ha descrito que las poblaciones naturales de SfMNPV presentan una gran variedad de genomas defectivos, los cuales presentan extensas deleciones en su genoma.⁹² Si bien no se conocen en profundidad las implicancias funcionales de este fenómeno, estudios recientes sugieren que estas deleciones podrían modular la relación entre velocidad de muerte del hospedador y la cantidad de progenie viral que se produce.³⁴ En este sentido, el gen viral ecdysteroid UDP-glucosyltransferase (*egt*), involucrado en la inhibición del proceso de muda y la prolongación del estadio larval del hospedador, juega un papel importante en los genotipos defectivos. El locus donde se ubica *egt* es un “hotspot” que es frecuentemente eliminado del genoma viral a través de deleciones, lo que conduce a un fenotipo de muerte rápida de la larva a costa de un menor rendimiento de OB por larva infectada.⁹³ En el aislamiento colombiano también se encontró evidencia de un evento de transferencia génica horizontal con una especie de baculovirus cercana (*Spodoptera litura* Nucleopolyhedro virus, SIMNPV), lo que enfatiza las características dinámica y altamente flexibles de las poblaciones naturales de baculovirus.⁹⁴

En el caso de SfGV, se han reportado 2 secuencias genómicas completas a la fecha, una correspondiente a un aislamiento colombiano⁹⁵ y otra a uno argentino.⁹¹ Si bien ambos aislamientos presentan una composición génica similar, SfGV-ARG presenta una copia trunca del gen *photolyase* y carece del gen *lef7*. Mas aún, a partir de búsquedas de similitud a nivel de secuencia (BLASTp) se ha propuesto que varios genes presentes en el genoma de SfGV han sido adquiridos a través de varios eventos de transferencia génicas horizontales. El proceso infectivo de SfGV se caracteriza por ser relativamente lento (aproximadamente 30 días para la muerte del hospedador) y por presentar lesiones severas en la región abdominal de la larva.⁹⁶

Estos baculovirus han sido empleados en distintos contextos para el control biológico de *S. frugiperda*.

SfMNPV ha tenido una historia de aplicación más extensa debido a su mayor eficacia a campo. Las estrategias de mejoramiento de este virus se han centrado mayoritariamente en el descubrimiento y caracterización de aislamientos con propiedades de control deseables, como ser tiempo de muerte rápido, baja dosis letal y estabilidad a campo.^{90,92,97,98} El uso de abrillantadores ópticos, los cuales proveen protección a los OB contra los rayos UV, y otros medios de formulación han sido explorados como un medio para extender la potencia de las formulaciones basadas en baculovirus.^{99,100} Por otra parte, la idea de complementar formulaciones de nucleopoliedrovirus con granulovirus que actúan como potenciadores, debido a la presencia de proteínas de virulencia en su OB, ha sido postulada como un método para incrementar su capacidad pesticida. Recientemente, se ha demostrado que la formulación conjunta de SfMNPV como principio activo con dosis más bajas de SfGV como potenciador presentan una elevada eficacia para el control de *S. frugiperda*.¹⁰¹

1.8. Diversidad genómica en poblaciones naturales de baculovirus

La caracterización de nuevos aislamientos de baculovirus se ha desarrollado históricamente a través de la secuenciación de los genes *polh*, *lef8*, y *lef9*, seguido del cálculo de la distancia genética entre aislamientos empleando una matriz de sustitución Kimura 2-parámetros.⁵³ Más recientemente, un método basado en un modelo probabilístico de Poisson ha sido utilizado para clasificar más de 500 aislamientos de baculovirus y estudiar sus procesos de diversificación.³⁸ Una observación relevante que surge a partir de estos estudios es que la mayoría de las especies de baculovirus presentan un elevado nivel de adaptación a su nicho ecológico, lo que se traduce en un rango de hospedador muy estrecho. Por otra parte, los patrones de restricción obtenidos a través de la digestión de DNA genómico viral con una enzima de restricción (en general BamHI, EcoRI y HindIII) también han sido de gran utilidad para caracterizar a grandes rasgos la composición genética de los aislamientos de baculovirus y analizar posibles determinantes de patogenicidad.¹⁰²⁻¹⁰⁴

Gracias al rápido desarrollo de las plataformas de secuenciación masiva (High-Throughput Sequencing, HTS), la secuenciación de genomas completos a pasado a ser un componente común para varios campos dentro de la virología, lo que se ha traducido en un incremento acelerado de la infor-

mación genómica disponible. Dado que el nivel de detalle que alcanza un proyecto de secuenciación depende en gran parte de los objetivos de investigación, se han propuesto cinco categorías para clasificar el grado de completitud de un genoma viral: borrador (fragmentado), alta calidad (ORF incompletos), ORF completos, completo (secuencias codificantes y no codificantes) y terminado (variantes poblacionales).¹⁰⁵ En el caso de los baculovirus, gran parte del esfuerzo en investigación ha sido direccionado hacia el descubrimiento y la caracterización de los genomas completos de nuevas especies virales a partir de muestras de campo. En este sentido, resulta evidente el incremento en los últimos años de la cantidad de genomas completos depositados para miembros de la familia *Baculoviridae*, en su mayoría a través de técnicas de HTS basadas en la plataforma Illumina, aunque también hay reportes recientes donde se emplean las tecnologías Nanopore y PacBio.¹⁰⁶⁻¹⁰⁸

Las tecnologías HTS no solo han permitido la secuenciación de un mayor número de genomas virales, sino que también han posibilitado el estudio de los aislamientos virales como poblaciones compuestas por diversos genomas que portan polimorfismos y variantes estructurales.¹⁰⁹ Dentro de la familia *Baculoviridae*, la aplicación de estas metodologías sugiere que los aislamientos baculovirales presentan niveles apreciables de heterogeneidad genética, un fenómeno que podría contribuir directamente a su capacidad adaptativa al hospedador.^{31,110,111} Sumado a esto, el número creciente de secuencias genómicas correspondientes a distintos aislamientos geográficos y temporales para una misma especie de baculovirus han contribuido a nuestra comprensión de los factores génicos que contribuyen a las interacciones virus-hospedador.¹¹²⁻¹¹⁴

1.9. Fundamentos y alcance de este trabajo

Los productos biológicos basados en baculovirus representan una alternativa prometedora para el control de diversas plagas a la vez que se disminuye de manera paulatina el uso de productos químicos. Sin embargo, su elevada especificidad y los largos tiempos de acción complejizan su aplicación a campo. El estudio de las bases y mecanismos moleculares que contribuyen a la virulencia y el rango de hospedador de los diferentes aislamientos baculovirales es central para acelerar su aplicación como bioinsumos. El aislamiento de dos baculovirus autóctonos (SfMNPV y SfGV) con potencial para el control de *S. frugiperda* representa una gran oportunidad para el desarrollo de pesticidas con menor impacto ambiental.

En este contexto, el objetivo principal de esta tesis doctoral es caracterizar los factores involucrados en la adaptación de aislamientos autóctonos de SfMNPV y SfGV a la oruga militar tardía. Además, este trabajo busca generar herramientas para facilitar el estudio de SfGV a través de cultivos celulares. A partir de esto, se propusieron los siguientes objetivos:

- Caracterizar y comparar el proteoma de los cuerpos de oclusión de SfGV y EpapGV, dos granulovirus con características infectivas marcadamente disímiles.
- Realizar un análisis comparativo del perfil de proteínas estructurales descritas para cuerpos de oclusión en la familia *Baculoviridae*, resaltando los componentes conservados y virus-específicos.
- Investigar la diversidad genética presente en un aislamiento autóctono de SfMNPV y realizar un análisis comparativo con otros aislamientos de este virus.
- Evaluar la capacidad infectiva de SfGV en la línea celular Sf9 mediante transfecciones de cultivos celulares con DNA genómico viral.

Capítulo 2

Proteoma de los cuerpos de oclusión de granulovirus

Parte de los resultados presentados en este capítulo han sido adaptados para su publicación en el siguiente artículo:

Masson T, Fabre ML, Ferrelli ML, Pidre ML, Romanowski V. Protein composition of the occlusion bodies of *Epinotia aporema* granulovirus. PLoS One. 2019;14(2):e0207735. Published 2019 Feb 12. doi:10.1371/journal.pone.0207735

2.1. Resumen

Varios miembros del género *Betabaculovirus* son ampliamente utilizados como agentes de control biológico contra diversas plagas de lepidópteros. Sumado a esto, también han demostrado utilidad como agentes potenciadores de la virulencia de los *Alphabaculovirus* en bioensayos de laboratorio. Debido a que los cuerpos de oclusión (OB) juegan un papel clave en la transmisión horizontal de los baculovirus y el inicio de la infección oral en el epitelio intestinal de la larva, decidimos abordar el estudio comparativo de la composición proteica del OB en granulovirus con distintos grados de virulencia. *Epinotia aporema* Granulovirus (EpapGV), un granulovirus que infecta al barrenador del brote (*Epinotia aporema*), representa un biopesticida prometedor debido a su corto tiempo de acción (2 a 3 días). En contraposición, el granulovirus de *S. frugiperda* (SfGV) tiene un tiempo de acción relativamente lento (alrededor de 30 días) pero presenta resultados preliminares prometedores como agente sinergizante del nucleopoliedrovirus de *S. frugiperda* (SfMNPV). En primer lugar, se llevó a cabo la caracterización del proteoma de los OB correspondientes a EpapGV y SfGV a través de un enfoque proteómico basado en espectrometría de masas. Los datos espectrales presentados aquí proporcionan una validación experimental de varias secuencias codificantes predichas solamente a través de métodos computacionales. Mediante un análisis proteogenómico se detectaron en EpapGV un *core gene* ortólogo a *ac110*, el cual no había sido anotado previamente, y un producto de fusión traduccional putativo entre los ORF *epap48* y *epap49*. El estudio comparativo de los proteomas disponibles para el género *Betabaculovirus* enfatizó el rol de los *core genes* como componentes mayoritarios de los OB, aunque también se observó la conservación de dos proteínas específicas de granulovirus (Epap48/Spfr40 y Epap95/Spfr101). Además, la cuantificación basada en el índice de abundancia proteica¹¹⁵ mostró que Epap95/Spfr101 es uno de los componentes más abundantes en los OB de granulovirus. Los ensayos de expresión en células de insecto y estudios bioinformáticos sugieren que esta proteína podría formar agregados intracelulares.

2.2. Introducción

La familia *Baculoviridae* comprende un grupo diverso de virus grandes capaces de infectar las larvas de insectos pertenecientes a los órdenes *Lepidoptera*, *Hymenoptera* y *Diptera*.³⁷ El virión de los baculovirus se caracteriza por presentar una nucleocápside en forma de bastón envuelta en una membrana lipídica, con un genoma a DNA circular cerrado cuyo tamaño oscila entre los 80-180 kbp.⁴⁰ Estos viriones se encuentran en el ambiente incluidos en una matriz proteica conocida como cuerpos de oclusión (OB), una estructura que les confiere protección frente a condiciones ambientales adversas (por ej. exposición a luz ultravioleta o desecación).

La mayoría de los productos comerciales formulados a base de baculovirus emplean aislamientos virales que pertenecen a los géneros *Alphabaculovirus* (nucleopoliedrovirus) y *Betabaculovirus* (granulovirus).²⁰ Con el fin de comprender mejor los determinantes moleculares que contribuyen a la infectividad de los granulovirus, nuestro laboratorio ha caracterizado previamente la secuencia genómica de dos aislamientos autóctonos con potencial biopesticida: *Epinotia aporema* Granulovirus (EpapGV)^{116,117} y *Spodoptera frugiperda* Granulovirus (SfGV).^{91,96} EpapGV es un betabaculovirus de acción rápida capaz de conducir a la muerte del barrenador del brote¹¹⁸ en pocos días. En contraposición, SfGV se caracteriza por tener un tiempo de infección prolongado (~ 30 días). Sin embargo, este virus ha sido empleado como agente potenciador del nucleopoliedrovirus de *Spodoptera frugiperda* (SfMNPV).^{101,119} En este sentido, los factores genéticos y proteicos que determinan la letalidad y la capacidad potenciadora de los granulovirus permanecen todavía poco claros. Dentro del género *Betabaculovirus*, es posible identificar dos grupos diferentes a nivel filogenético, clado *a* y clado *b*.¹²⁰ El clado *a*, al cual pertenece SfGV, comprende un conjunto de granulovirus que solo infectan lepidópteros pertenecientes a la familia *Noctuidae*, además de compartir una elevada similitud en sus genomas y presentar un conjunto exclusivo de genes (¹²¹). En cambio, los miembros del clado *b*, el cual incluye a EpapGV, son mucho más disímiles entre sí a nivel genómico e infectan insectos pertenecientes a varias familias (*Tortricidae*, *Pieridae*, *Sphingidae*, *Crambidae*, *Gelichiidae* y *Pyralidae*).

La proteómica basada en espectrometría de masas (MS) representa una técnica poderosa para interrogar la composición proteica global de las partículas virales.¹²² Además de la identificación

directa de proteínas, los datos espectrales derivados de experimentos de MS pueden ser utilizados para corroborar o expandir la anotación de datos genómicos y transcriptómicos disponibles en bases de datos públicas. Esta metodología, conocida como proteogenómica, se basa en la comparación de una base de datos compuesta por todas las regiones codificantes putativas presentes en una secuencia genómica con los datos espectrales de MS, lo que proporciona un enfoque no sesgado que permite refinar modelos génicos y a su vez descubrir nuevas secuencias codificantes de proteínas.¹²³ Sumado a la descripción cualitativa del conjunto de proteínas que componen la partícula viral, el desarrollo y validación de diversos estimadores de la concentración de proteínas dentro de la muestra han permitido el análisis cuantitativo de los datos proteómicos.¹²⁴ Específicamente, la fórmula para estimar el índice de abundancia proteica (emPAI),¹¹⁵ permite obtener una estimación semicuantitativa de los niveles de abundancia de cada proteína presente dentro de una muestra, sin necesidad de añadir marcadores costosos, como aquellos empleados en las metodologías de marcaje por etiquetas isobáricas (*isobaric tags for relative and absolute quantification*, ITRAQ) o a través de aminoácidos marcados con isótopos (*stable isotope labeling by amino acids in cell culture*, SILAC).^{125,126}

El análisis proteómico de la partícula viral se ha aplicado para estudiar numerosas familias de virus grandes con genomas a DNA (*Ascoviridae*,¹²⁷ *Herpesviridae*,¹²⁸ *Iridoviridae*,¹²⁹ *Nudiviridae*¹³⁰ y *Poxviridae*).¹³¹ En el caso de la familia *Baculoviridae*, se han analizado a la fecha ocho proteomas distintos (AcMNPV,⁶² AgMNPV,⁶⁹ ChchNPV,⁶⁷ HearNPV,⁶⁴ MabNPV,⁷¹ ClanGV,⁷⁰ PiraGV⁶⁸ y CuniNPV).¹³² Estos conjuntos de datos apuntan a una composición compleja que comprende una gran variedad de proteínas involucradas en procesos como la morfogénesis del virión, la formación del OB y la posterior infección de las células intestinales del hospedador.

Este capítulo explora la composición proteica del OB de los granulovirus a través de un enfoque comparativo, donde se contrastan los proteomas estructurales de EpapGV y SfGV utilizando un enfoque basado en MS. El análisis de dos granulovirus marcadamente disímiles en sus propiedades biológicas (ciclo infectivo rápido vs lento) y su posición filogenética dentro del género *Betabaculovirus* (SfGV corresponde al clado *a* y EpapGV al clado *b*) fue el punto de partida para identificar proteínas diferenciales candidatas que podrían contribuir a las diferencias fenotípicas observadas para estos granulovirus. Este estudio comparativo fue complementado con la información proveniente

de los estudios proteómicos publicados previamente para otras especies de granulovirus, permitiéndonos obtener resultados generalizables para todo el género *Betabaculovirus*. Si bien nuestros datos derivan del análisis proteómico de la partícula de OB completa, en contraposición con otros estudios basados solo en ODV purificados, se asume que la composición proteica del OB y el ODV son altamente similares basados en reportes previos.⁷⁰ Los datos espectrales reportados en este capítulo proporcionan una fuente de validación experimental para varios productos génicos cuya anotación estaba basada solo en métodos computacionales, lo que muestra la utilidad de este enfoque para mejorar la anotación de genoma virales. La utilización del emPAI como estimador de la abundancia proteica facilitó el desarrollo de un modelo cualitativo del proteoma completo de la partícula madura del OB para ambos granulovirus y permitió identificar sus componentes mayoritarios. En ensayos subsecuentes, se demostró que uno de los componentes mayoritarios del OB de granulovirus, Epap95/Spfr101, es capaz de formar agregados al ser expresada *in vitro*.

2.3. Resultados

2.3.1. Purificación de los OB de EpapGV y SfGV

La obtención de OB puros se realizó mediante una purificación por ultracentrifugación en gradiente de sacarosa de un lisado proveniente de larvas moribundas infectadas con el granulovirus correspondiente. Con el fin de evaluar la pureza e integridad estructural de las preparaciones de OB correspondientes a SfGV, se llevaron a cabo observaciones mediante microscopía electrónica de transmisión y barrido de los mismos (Figura 2.1). A partir de estas imágenes se pudo apreciar que los gránulos purificados presentaron la morfología oval típica del género *Betabaculovirus* y no se evidenciaron contaminantes obvios dentro de la muestra. En el caso de EpapGV (Figura 2.2), se empleó una preparación de OB purificados previamente.¹¹⁷

2.3.2. Componentes estructurales detectados en los OB de EpapGV y SfGV

Se caracterizó la composición proteica de las partículas de OB perteneciente a dos especies de granulovirus, EpapGV y SfGV, a través de espectrometría de masas en tandem acoplada a cromatografía líquida (LC-MS/MS). Para el caso de EpapGV, procesamos dos réplicas biológicas que fueron sometidas a un protocolo de preparación de muestra líquida y luego analizadas por LC-MS/MS. En

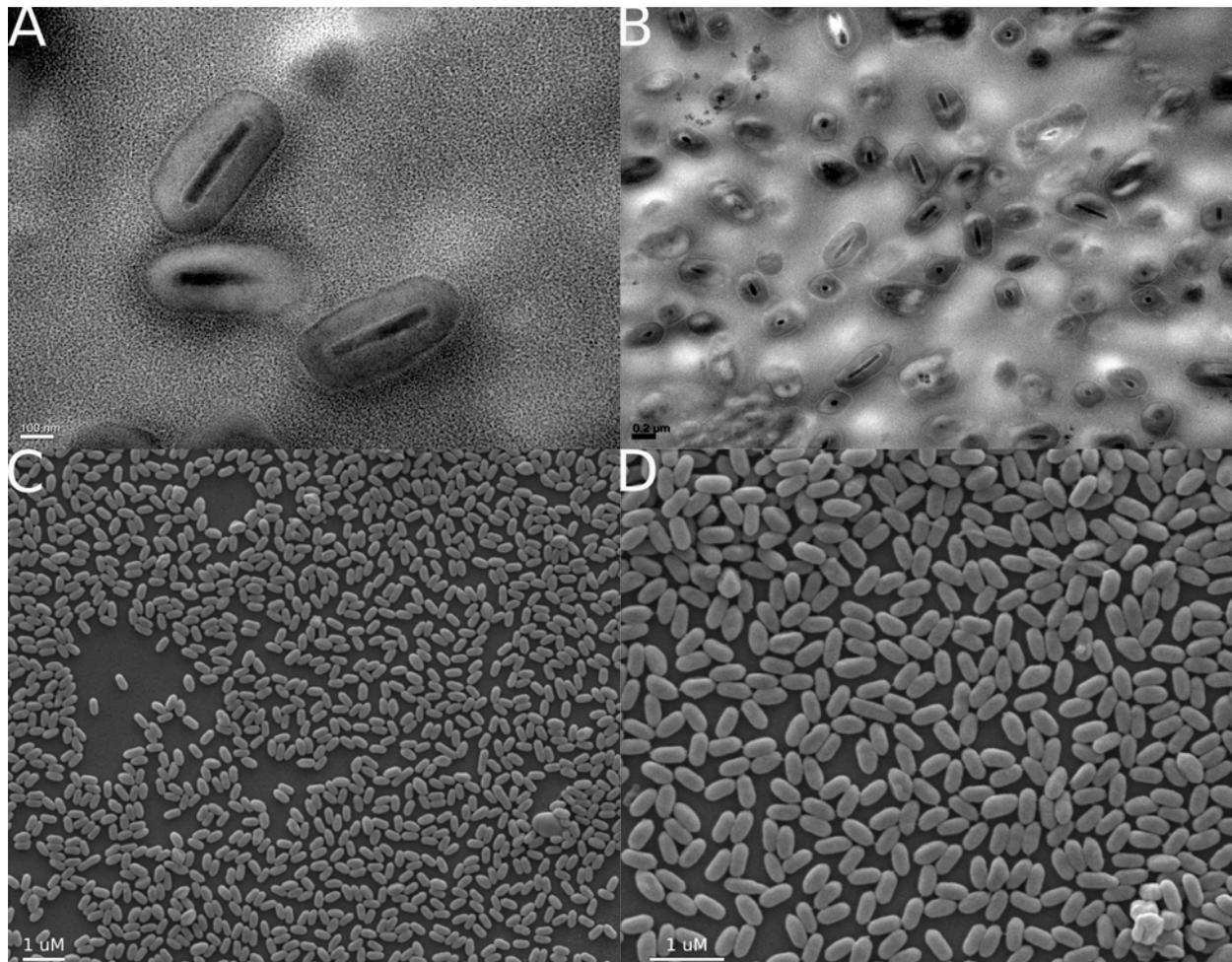


Figura 2.1: Morfología de los OB de SfGV. Micrografías electrónicas de **A-B** transmisión y **C-D** barrido de los OB purificados mediante ultracentrifugación. Las barras en cada micrografía muestran la escala correspondiente.

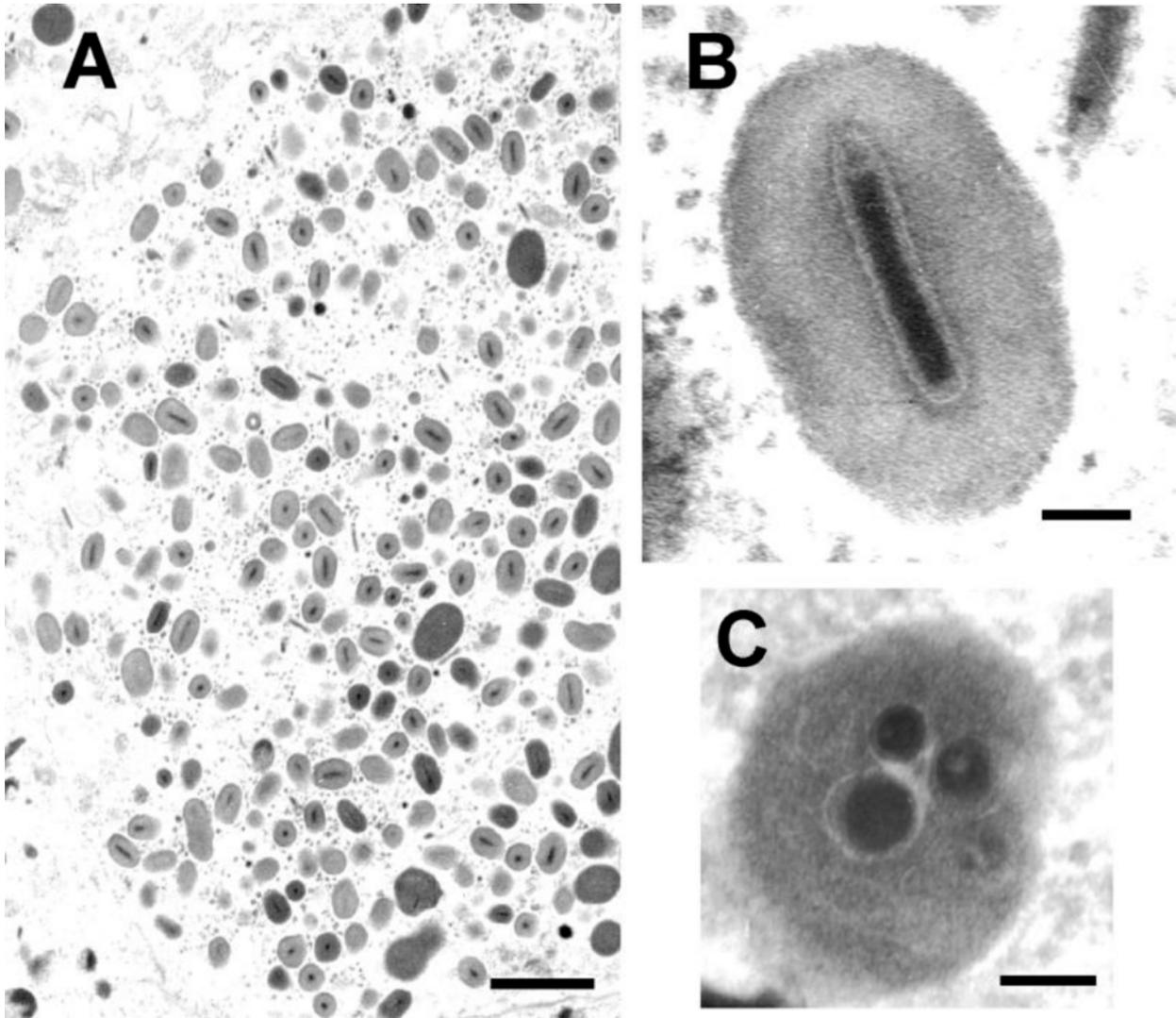


Figura 2.2: Morfología de los OB de EpapGV. Micrografías electrónicas ilustrando las propiedades morfológicas principales de los OB purificados de EpapGV. **A** Sección a través de una célula epidérmica que muestra viriones y gránulos (barra, 1 μ m). **B** Sección longitudinal a través de un OB que contiene un solo virión incluido (barra, 100 nm). **C** Gránulo con múltiple viriones incluidos (barra, 100 nm). Adaptado de Sciocco-Cap et al. 2001.

el caso de SfGV, se analizaron dos réplicas biológicas, una sometida al mismo procesamiento que las muestras de EpapGV y otra muestra separada por electroforesis en gel de poliacrilamida con dodecil sulfato de sodio (SDS-PAGE) antes de ser analizadas por LC-MS/MS. La introducción de un paso adicional de separación nos permitió transformar una muestra compleja de proteínas en fracciones más simples para su caracterización proteómica. Para EpapGV se detectaron un total 56 proteínas formando parte del OB, mientras que para SfGV se identificaron un total de 72 proteínas (Cuadros 2.1 y 2.2). Si comparamos las muestras analizadas para ambos granulovirus se puede observar que en general existe buen acuerdo entre réplicas para EpapGV (se detectaron 52 proteínas en la muestra 1 y 47 en la muestra 2, de las cuales 43 son compartidas entre ambas), mientras que para SfGV se evidenció una mayor diferencia entre los conjuntos de proteínas identificadas (se confirmaron 42 y 70 proteínas para cada muestra, respectivamente, con 40 proteínas comunes a ambas muestras). Para ambos granulovirus, el conjunto de proteínas que forman parte de los OB comprende una fracción cercana a la mitad de los ORF anotados en su genoma (43.93 % para EpapGV y 47.68 % para SfGV), lo que demuestra que una gran parte de las secuencias codificantes dentro del genoma viral cumplen una función estructural.

Cuadro 2.1: Componentes estructurales del cuerpo de oclusión de EpapGV

ORF	Proteína	NCBI ID	Tamaño (aa)	# Péptidos M1	# Péptidos M2	emPAI M1	emPAI M2	emPAI (%VP39)
1	Granulin	YP_006908509.1	248	25	25	1,995,261	794,327	1,539,317.82
5	P78/83	YP_006908513.1	137	1	2	0.33	0.78	0.62
6	PK1	YP_006908514.1	276	9	5	3.44	1.25	2.54
10	Epap10	YP_006908518.1	90	1	-	0.33	0.00	0.00
14	EFP	YP_006908522.1	541	1	-	0.08	0.00	0.00
21	PEP2	YP_006908529.1	142	5	3	4.62	1.37	3.08
22	PEP/P10	YP_006908530.1	308	7	8	9.00	5.81	8.84
25	PEP1	YP_006908533.1	178	6	6	5.31	2.98	4.86
27	PIF5	YP_006908535.1	354	7	4	5.58	2.51	4.58
28	Ac142	YP_006908536.1	457	18	14	4.84	2.98	4.65
29	ODV-E18	YP_006908537.1	88	4	6	55.23	176.83	120.84
31	v-Cath	YP_006908539.1	329	-	2	0.00	0.29	0.00
39	ODV-E66	YP_006908547.1	654	14	10	5.31	1.75	3.73
40	Epap40	YP_006908548.1	102	6	5	16.78	9.00	15.03
47	PIF2	YP_006908555.1	374	5	7	0.69	1.08	1.05
48	Epap48	YP_006908556.1	446	13	12	2.54	1.83	2.63
49	Epap49	YP_006908557.1	1,465	35	27	2.87	1.32	2.38
52	v-Ubi	YP_006908560.1	93	3	1	1.37	0.78	1.26

ORF	Proteína	NCBI ID	Tamaño (aa)	# Péptidos M1	# Péptidos M2	emPAI M1	emPAI M2	emPAI (%VP39)
53	ODV-EC43	YP_006908561.1	348	12	12	6.20	6.20	7.58
58	SOD	YP_006908566.1	183	9	6	24.12	5.31	13.84
59	PIF0	YP_006908567.1	653	5	1	0.47	0.08	0.24
62	Epap62	YP_006908570.1	106	1	1	0.39	0.39	0.48
66	P24	YP_006908574.1	165	2	2	0.87	0.52	0.82
69	PIF1	YP_006908577.1	538	3	2	0.52	0.23	0.43
71	Epap71	YP_006908579.1	104	1	-	0.47	0.00	0.00
74	LEF6	YP_006908582.1	82	-	1	0.00	0.47	0.00
78	P48/45	YP_006908586.1	380	1	-	0.11	0.00	0.00
79	P12	YP_006908587.1	115	3	1	4.62	0.78	2.32
80	P40	YP_006908588.1	373	9	5	1.68	0.73	1.36
81	P6.9	YP_006908589.1	56	1	1	2.16	2.16	2.64
83	38K	YP_006908591.1	295	5	3	1.45	0.47	1.01
84	PIF4	YP_006908592.1	162	2	3	0.67	1.15	1.07
85	Hel-1	YP_006908593.1	1,085	1	-	0.00	0.00	0.00
86	ODV-E25	YP_006908594.1	213	7	8	9.00	9.00	11.00
87	P18	YP_006908595.1	158	-	1	0.00	0.00	0.00
88	P33	YP_006908596.1	254	7	7	1.98	1.08	1.79
90	ChaB	YP_006908598.1	75	2	1	2.98	1.51	2.60
92	VP39	YP_006908600.1	293	21	16	232.57	28.76	100.01

ORF	Proteína	NCBI ID	Tamaño (aa)	# Péptidos M1	# Péptidos M2	emPAI M1	emPAI M2	emPAI (%VP39)
93	ODV-EC27	YP_006908601.1	284	8	6	3.39	2.16	3.31
94	BRO	YP_006908602.1	359	2	-	0.23	0.00	0.00
95	Epap95	YP_006908603.1	73	5	5	176.83	999.00	513.91
96	PIF8	YP_006908604.1	567	6	7	1.22	1.22	1.49
98	Ac81	YP_006908606.1	191	5	1	1.61	0.21	0.71
99	GP41	YP_006908607.1	286	23	16	87.59	19.69	50.78
100	Ac78	YP_006908608.1	88	2	2	16.78	4.62	10.77
101	VLF-1	YP_006908609.1	368	11	10	3.06	3.06	3.74
103	Ac75	YP_006908611.1	149	9	4	7.11	1.85	4.43
106	DNA Pol	YP_006908614.1	1,068	1	-	0.00	0.00	0.00
109	PIF6	YP_006908617.1	148	5	2	2.59	0.67	1.61
113	FP25K	YP_006908621.1	146	1	3	0.23	0.87	0.55
115	DNA Lig	YP_006908623.1	534	1	-	0.08	0.00	0.00
123	Epap123	YP_006908630.1	102	-	1	0.00	0.47	0.00
126	Epap126	YP_006908633.1	347	11	8	3.49	1.72	3.00
127	Epap127	YP_006908634.1	69	1	2	1.15	3.64	2.51
129	VP1054	YP_006908636.1	344	1	1	0.11	0.11	0.13
133	ME53	YP_006908640.1	373	1	-	0.22	0	0.00
-	Ac110	This study	47	-	-	-	-	-

Cuadro 2.2: Componentes estructurales del cuerpo de oclusión de SfGV

ORF	Proteína	NCBI ID	Tamaño (aa)	# Péptidos M1	# Péptidos M2	emPAI M1	emPAI M2	emPAI (%VP39)
1	Granulin	AXS01020.1	248	20	26	79431.82	99999.00	62582.97
2	PK1	AXS01021.1	287	1	13	0.16	12.34	0.97
3	P78/83	AXS01022.1	203	3	3	1.68	4.18	1.86
5	P10	AXS01024.1	73	4	12	24.12	463.16	74.22
7	Spfr007	AXS01026.1	189	-	5	0.00	2.16	0.00
9	Spfr009	AXS01028.1	190	-	7	0.00	1.74	0.00
11	ODV-E18	AXS01030.1	85	6	5	176.83	315.23	165.79
12	49k	AXS01031.1	454	21	33	5.72	99.00	16.71
14	PIF5	AXS01033.1	353	5	15	2.16	55.23	7.67
16	PEP1	AXS01035.1	196	8	11	14.85	78.43	23.96
17	PEP2	AXS01036.1	152	4	12	6.50	748.89	48.99
18	PEP/P10	AXS01037.1	383	5	14	14.85	3980.07	170.71
25	Spfr025	AXS01044.1	228	-	6	0.00	2.51	0.00
26	Spfr026	AXS01045.1	193	4	7	1.15	9.00	2.26
27	PIF3	AXS01046.1	192	-	6	0.00	2.16	0.00
28	Spfr028	AXS01047.1	100	1	-	9.00	0.00	0.00
29	Spfr029	AXS01048.1	111	6	4	6.50	2.16	2.63
36	p13	AXS01055.1	280	-	6	0.00	1.40	0.00

ORF	Proteína	NCBI ID	Tamaño (aa)	# Péptidos M1	# Péptidos M2	emPAI M1	emPAI M2	emPAI (%VP39)
38	PIF2	AXS01057.1	380	1	11	0.12	4.62	0.53
40	Spfr040	AXS01059.1	1143	20	62	1.13	14.85	2.88
43	v-Ubi	AXS01062.1	82	2	6	0.93	12.90	2.43
44	ODV-EC43	AXS01063.1	351	11	23	2.02	38.81	6.22
45	Spfr045	AXS01064.1	101	1	2	0.39	0.93	0.42
50	dUTPase	AXS01069.1	141	-	2	0.00	0.36	0.00
52	SOD	AXS01071.1	142	3	3	2.162	5.81	2.49
63	PIF0	AXS01083.1	746	2	18	0.14	2.93	0.45
64	Spfr064	AXS01084.1	90	-	4	0.00	4.18	0.00
68	P24	AXS01088.1	176	-	5	0.00	3.64	0.00
69	38.7 k	AXS01089.1	173	-	2	0.00	0.43	0.00
71	P10	AXS01091.1	173	6	12	5.00	214.44	22.98
72	PIF1	AXS01092.1	541	-	6	0.00	1.40	0.00
74	Chtb-2a	AXS01094.1	162	1	8	0.26	6.94	0.94
75	Chtb-2b	AXS01095.1	159	-	2	0.00	0.78	0.00
77	DBP	AXS01097.1	281	-	3	0.00	0.64	0.00
79	P48/45	AXS01099.1	372	-	8	0.00	1.61	0.00
80	Spfr080	AXS01100.1	223	-	3	0.00	0.54	0.00
81	P12	AXS01101.1	111	1	2	0.39	0.93	0.42
82	P40	AXS01102.1	372	5	22	0.69	22.10	2.74

ORF	Proteína	NCBI ID	Tamaño (aa)	# Péptidos M1	# Péptidos M2	emPAI M1	emPAI M2	emPAI (%VP39)
83	P6.9	AXS01103.1	58	1	-	2.16	0.00	0.00
84	38K	AXS01104.1	304	-	5	0.00	1.15	0.00
87	PIF4	AXS01107.1	157	2	4	0.59	1.51	0.66
88	ODV-E25	AXS01108.1	219	8	12	20.54	176.83	42.32
89	P33	AXS01109.1	303	4	9	0.69	3.81	1.14
90	P18	AXS01110.1	158	-	3	0.00	1.15	0.00
91	ChaB	AXS01111.1	79	-	2	0.00	9.00	0.00
96	VP39	AXS01116.1	324	18	30	16.30	1244.20	100.01
97	ODV-EC27	AXS01117.1	290	7	13	1.45	7.80	2.36
98	BRO-f	AXS01118.1	370	-	12	0.00	3.64	0.00
100	Spfr100	AXS01120.1	385	-	4	0.00	0.62	0.00
101	Spfr101	AXS01121.1	124	8	8	9.00	30.62	11.66
105	PIF8	AXS01125.1	718	7	24	0.55	9.00	1.56
106	Ac81	AXS01126.1	191	2	8	0.59	9.00	1.61
107	GP41	AXS01127.1	294	12	20	17.48	214.44	42.99
108	Ac78	AXS01128.1	103	4	3	13.68	20.54	11.77
109	VLf-1	AXS01129.1	368	-	15	0.00	4.48	0.00
112	Spfr112	AXS01132.1	145	4	11	1.31	17.74	3.38
114	Desmoplakin	AXS01134.1	637	-	5	0.00	0.28	0.00
115	PIF6	AXS01135.1	139	-	6	0.00	9.00	0.00

ORF	Proteína	NCBI ID	Tamaño (aa)	# Péptidos M1	# Péptidos M2	emPAI M1	emPAI M2	emPAI (%VP39)
120	FP-25K	AXS01140.1	170	5	9	1.42	6.02	2.06
125	Alk-exo	AXS01145.1	405	-	9	0.00	1.42	0.00
126	Hel-2	AXS01146.1	472	-	2	0.00	0.17	0.00
127	Spfr127	AXS01147.1	333	-	3	0.00	0.35	0.00
130	ODV-E66	AXS01150.1	674	8	15	1.99	14.51	3.78
131	Enhancin-1	AXS01151.1	867	14	27	1.24	7.91	2.20
137	Enhancin-2	AXS01157.1	857	18	37	1.83	18.31	4.07
140	Ac53	AXS01160.1	196	-	2	0.00	0.67	0.00
141	Spfr141	AXS01161.1	417	17	29	6.31	135.89	20.55
143	Spfr143	AXS01163.1	67	3	4	2.98	9.00	3.64
144	VP1054	AXS01164.1	324	-	9	0.00	2.34	0.00
145	Spfr145	AXS01165.1	77	-	2	0.00	2.16	0.00
150	ME53	AXS01170.1	298	-	8	0.00	1.82	0.00
151	Spfr151	AXS01171.1	99	-	2	0.00	0.78	0.00

A partir del mapeo genómico de los genes correspondientes a estas proteínas estructurales, se observó una elevada sintenia en este grupo de genes de EpapGV y SfGV (Figuras 2.3 y 2.4). Más aún, en ambos granulovirus se confirmó la presencia de una región colineal densamente poblada con *core genes* estructurales (región colineal baculoviral), que ha sido descrita previamente en otras especies de baculovirus^{133,134} y se postula como un elemento genómico conservado en los géneros *Alphabaculovirus* y *Betabaculovirus*. La comparación entre los proteomas de distintos miembros de la familia *Baculoviridae* evidenció que esta región, en comparación con el resto del genoma viral, se encuentra densamente ocupado por secuencias codificantes asociadas con funciones estructurales (Figuras 2.5).

Si bien nuestro enfoque apuntó a caracterizar todos los componentes presentes en el OB, en nuestras muestras no se detectaron ciertas proteínas (PIF3 y desmoplakin para EpapGV, Ac110/PIF7 en el caso de SfGV) las cuales han sido confirmadas previamente como componentes del virión de AcMNPV por western blot y MS.⁶² Otra ausencia llamativa en EpapGV es la de GP37, una proteína que comparte un ancestro común con las fusolinas de los entomopoxvirus, las cuales han sido caracterizadas como factores de virulencia oral.¹³⁵ Se ha demostrado que GP37 es capaz de unirse a quitina y causar daño en la membrana peritrófica larval, por lo que sería esperable que esta proteína se encontrara dentro de la estructura del OB.¹³⁶ Su ausencia en nuestras preparaciones podría tener varias explicaciones posibles: degradación proteolítica parcial, bajos niveles de expresión de estas proteínas en el OB, una ionización deficiente de esos componentes en nuestras muestras o incluso diferencias en la localización celular entre distintas especies.

Además de identificar los componentes de los OB, la abundancia relativa de cada proteína fue estimada a través del cálculo del valor de emPAI, como fue propuesto previamente.¹¹⁵ El valor de emPAI de la proteína mayoritaria de la cápside viral, VP39, se tomó como valor de referencia para la normalización de la concentración proteica. Tomando como punto de corte una abundancia de al menos 10 % emPAI(VP39), se encontraron de forma consistente ODV-E18, GP41, Granulin, SOD y Epap95/Spfr101, junto con VP39, como las proteínas más abundantes en los OB de ambos granulovirus. La proteína mayoritaria de cápside, VP39, la proteína que forma el tegumento localizado entre el virión y la envoltura viral (GP41) y el componente proteico mayoritario de la matriz del OB (Granulin) son tres proteínas que era esperable encontrar en grandes cantidades debido a

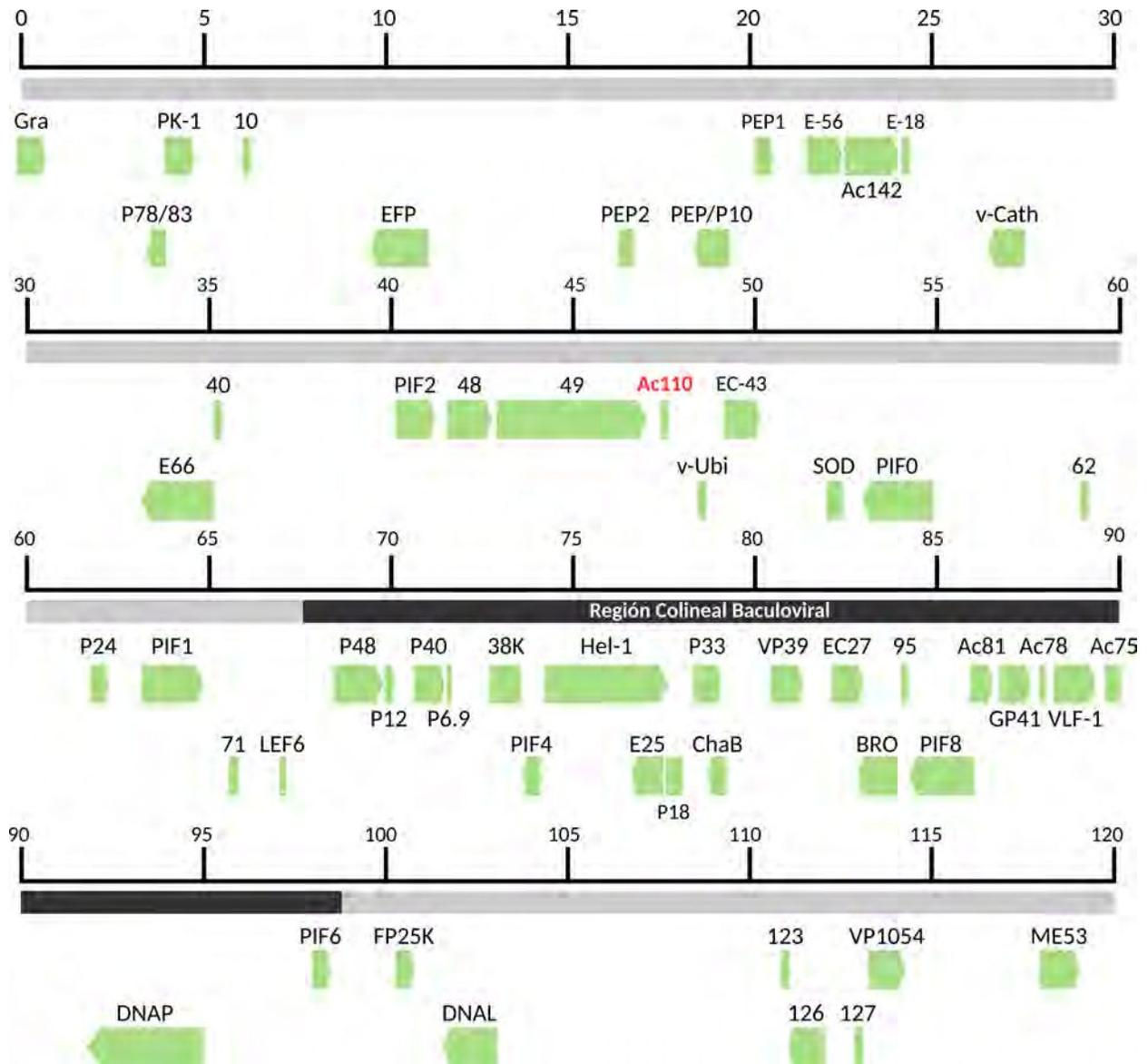


Figura 2.3: Regiones codificantes correspondientes a las proteínas estructurales presentes en los OBs de EpapGV. En rojo se muestra un gen que presenta homología a *ac110*, descubierto a partir de una búsqueda protegenómica.

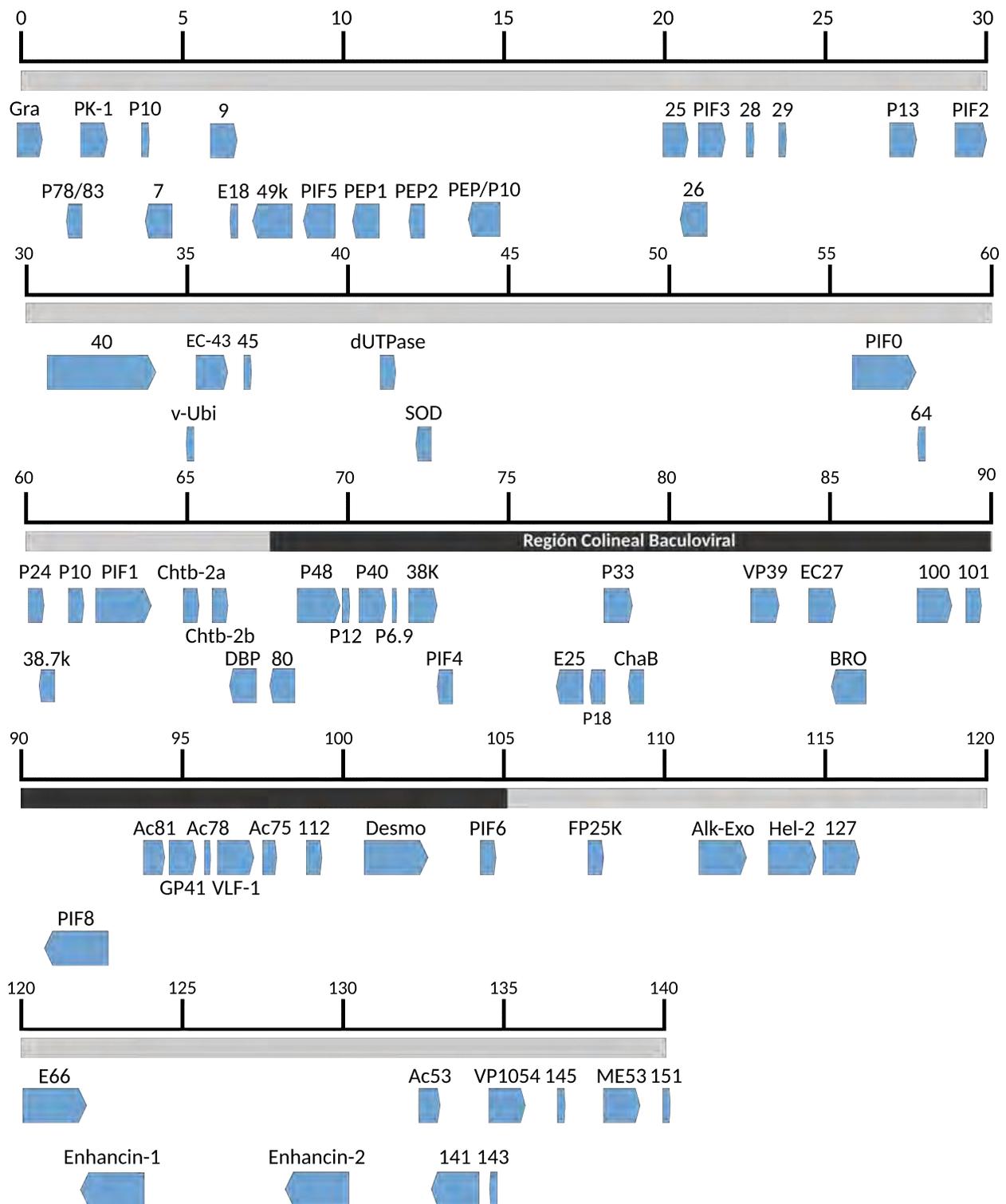


Figura 2.4: Regiones codificantes correspondientes a las proteínas estructurales presentes en los OBs de SfGV.

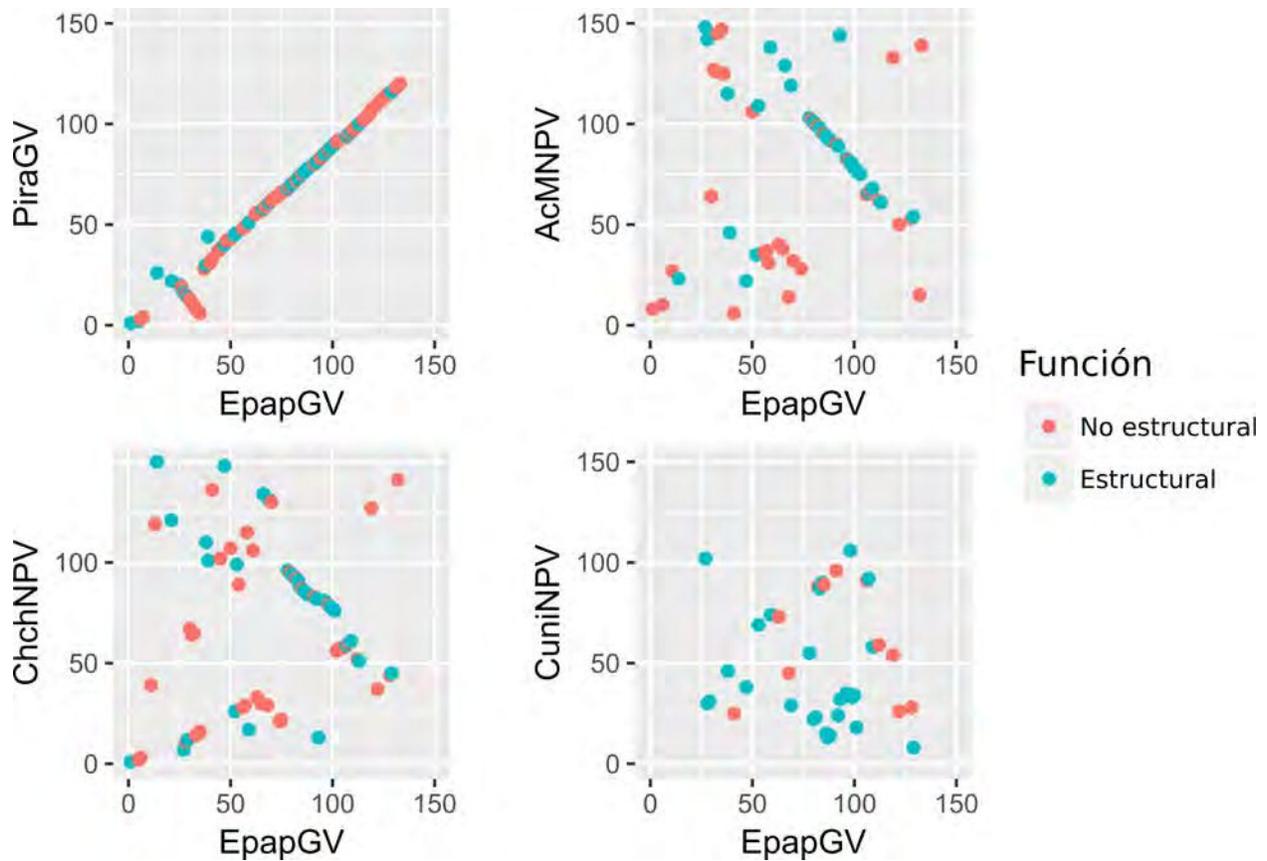


Figura 2.5: Retención de una region colineal enriquecida en genes codificantes para componentes estructurales del ODV. Gráfico de paridad génica para las secuencias codificantes de proteínas estructurales del ODV presentes en AcMNPV, ChchNPV, CuniNPV y PiraGV, comparadas contra EpapGV

su función estructural básica. Por otra parte, ODV-E18 es un producto de un *core gene* necesario para la producción de BV y el cual se asocia con la membranas intranucleares que dan origen a la membrana que envuelve a las nucleocápsides.^{137,138} Finalmente, *epap95* y *spfr101* pertenecen a una familia génica presente solo en el género *Betabaculovirus* y sus productos proteicos han sido detectados de forma consistente como un componente de los ODV de ClanGV y PiraGV.^{68,70} Sin embargo, no existe información disponible acerca de la posible función de estas proteínas.

En este sentido, el contenido génico del género *Betabaculovirus* permanece poco caracterizado, con un gran número de regiones codificantes que cuentan con evidencia derivada de predictores automáticos basados solamente en métodos computacionales. Nuestros datos proteómicos nos permitieron asignar evidencia experimental *bona fide* para más de 20 genes hipotéticos en ambos granulovirus y confirmar la presencia de su producto de traducción: para EpapGV se identificaron Epap10, Epap40, Epap48, Epap49, Epap62, Epap71, Epap95, Epap123, Epap126 y Epap127, mientras que para SfGV se detectaron Spfr7, Spfr9, Spfr25, Spfr26, Spfr28, Spfr29, Spfr40, Spfr45, Spfr64, Spfr80, Spfr100, Spfr101, Spfr112, Spfr127, Spfr138, Spfr141 and Spfr151.

2.3.3. Anotación proteogenómica de nuevos péptidos en el genoma de EpapGV y SfGV

Con el fin de identificar proteínas previamente no anotadas en el genoma viral, se realizó una búsqueda de péptidos a partir de los datos espectrales recolectados frente a una base de datos teórica compuesta por todos los productos de traducción derivados de los seis marcos de lectura del genoma viral (todos los ORFs putativos fueron incluidos, sin tomar en cuenta un tamaño mínimo como valor de corte). Empleando esta metodología se detectaron siete péptidos derivados del genoma de EpapGV los cuales no se encontraban anotados previamente,¹¹⁷ mientras que para SfGV no se hallaron péptidos no anotados. La secuencia peptídica y la localización genómica de estos péptidos novedosos en EpapGV se encuentran detallada en la Cuadro 2.3. Uno de esos péptidos, de 47 aminoácidos de longitud, es coincidente con un ORF que se solapa con *epap51* pero presenta una orientación inversa. Análisis posteriores confirmaron la presencia de este nuevo ORF en otros miembros de la familia *Baculoviridae* y mostraron que se trata de un *core gene* ortólogo de *ac110*.⁴³ Este gen ha sido descrito como un factor de infectividad oral (*per os* infectivity factor 7, *pif7*) y su

producto de traducción solo ha sido detectado en el proteoma del ODV de HearNPV⁶⁴ y EpapGV (reportado aquí). La localización genómica y la orientación de este gen ortólogo a *ac110* ha sido mantenida dentro del género *Betabaculovirus*, lo que provee evidencia adicional de su conservación evolutiva (Figura 2.6). En cuanto al gen *epap51*, que codifica para 69 amino ácidos, no encontramos evidencia de su traducción en nuestros datos espectrales. Por otra parte, una búsqueda de secuencias homólogas a través de BLASTp no logró recuperar ningún homólogo para este gen, lo que podría indicar que este gen ha surgido específicamente en EpapGV. Los seis péptidos restantes solapan ORF anotados (*chitinase*, *DNA-ligase* y *granulin*) o afectan regiones intergénicas. Dos péptidos se encontraron entre los ORFs *epap48* y *epap49*, y un péptido entre *epap61* y *epap62*. Búsquedas de similitud de secuencia con el software TBLASTN mostraron que solo los péptidos presentes en la secuencia de *chitinase* y *granulin* se encuentran conservados en *loci* homólogos de *Alphabaculovirus* y *Betabaculovirus*. Sin embargo estos péptidos se encuentra en ORF diferentes a los de *chitinase* y *granulin* (Cuadro 2.3).

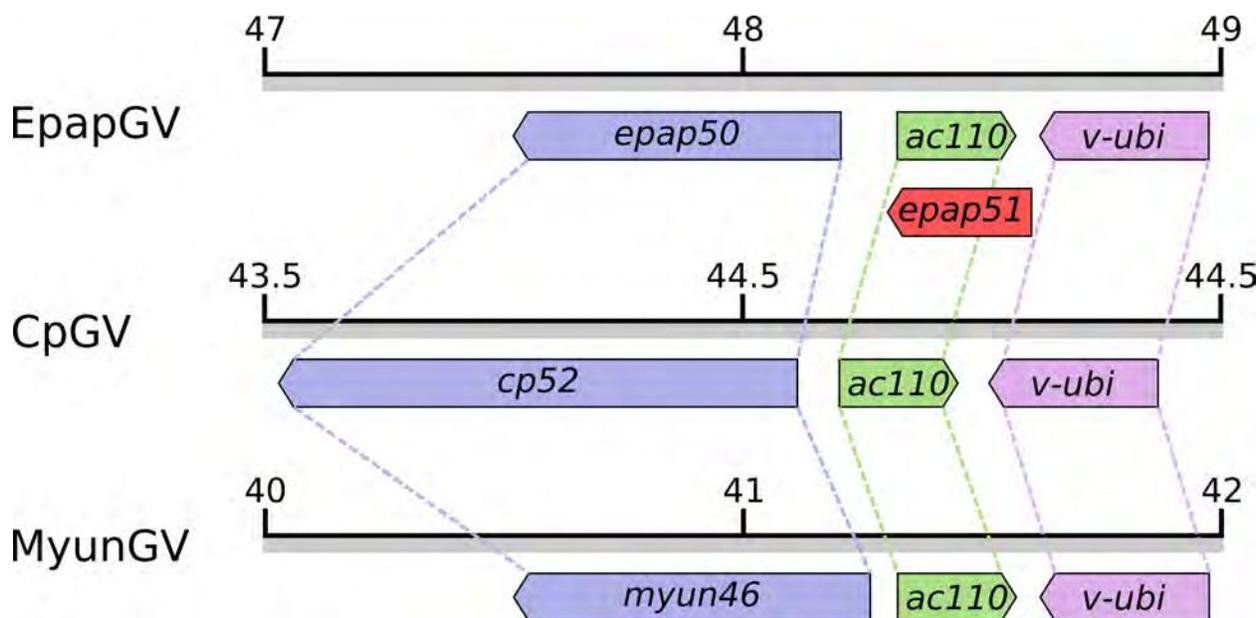


Figura 2.6: Homólogo de *ac110* presente en EpapGV. Conservación evolutiva en distintos betabaculovirus (*Cydia pomonella* Granulovirus, CpGV, y *Mythimna unipuncta* Granulovirus, MyunGV) de la orientación y sintenia del ORF homólogo a *ac110* detectado en EpapGV. La presencia del mismo color indica la existencia de homología entre los distintos genes.

Cuadro 2.3: Péptidos detectados por MS sin anotación previa

Secuencia peptídica	Tamaño (aa)	Gene cercanos/solapados	Coordenadas genómicas
SWTCGSSSSLDR	12	granulin (misma cadena)	350 > 391
AVWIGGDGVV	10	chitinase (cadena opuesta)	28005 < 28037
DTIDLLTEYVNKDDSGAER	19	Intergénico epap48-epap49	43128 > 43190
NLLLTNIGGSDTVK	14	Intergénico epap48-epap49	43200 > 43247
EQEIQYMLYEHR*	12	epap51 (cadena opuesta)	48549 > 48400
ITIKFIK	7	Intergénico epap61-epap62	58524 > 58652
SILILDGR	8	DNA-ligase (cadena opuesta)	101551 > 101577

Sorprendentemente, los péptidos comprendidos entre *epap48* y *epap49* cubren casi completamente esta región intergénica de 145 pb (Figura 2.7). *Epap48* codifica una proteína de 446 amino ácidos que se encuentra conservada en el género *Betabaculovirus*. El producto de traducción putativo de *epap49* es una proteína compuesta por 1465 residuos que no presenta ningún ortólogo detectable en otros baculovirus. Los péptidos detectados mediante este enfoque proteogenómico se localizan en el mismo marco de lectura que Epap49, pero no presentan un codon Metionina de inicio en marco (Figura 2.7). Una hipótesis que podría explicar la presencia de estos péptidos es que *Epap48* y *Epap49* podrían ser expresados como una proteína de fusión debido a un evento de cambio en el marco de lectura +1 (*frameshifting*) cerca del extremo C-terminal de *Epap48*. Para confirmar esta hipótesis será necesario detectar directamente el mRNA o la proteína producto de este evento.

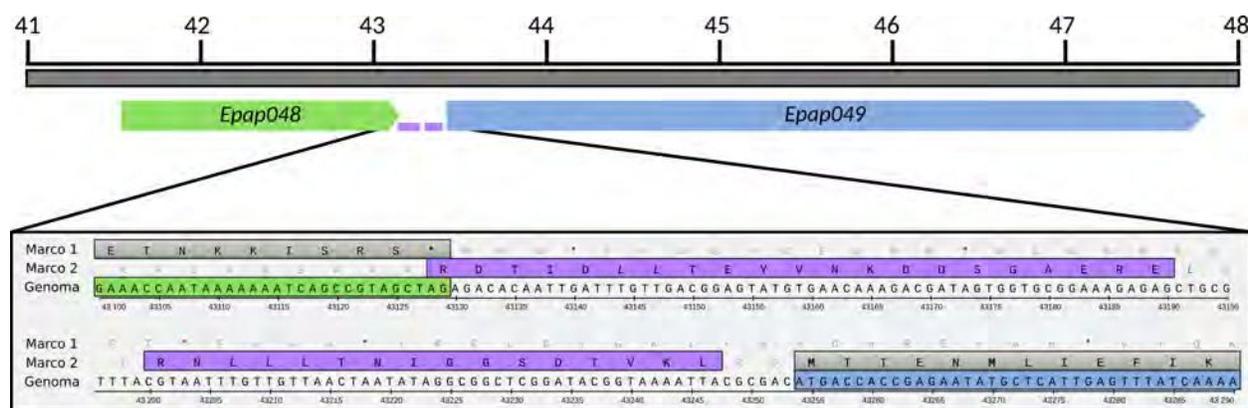


Figura 2.7: Proteína de fusión putativa entre Epap48 y Epap49. Locus genómico de los genes *epap48* (verde) y *epap49* (celeste). Los péptidos detectados a través de MS dentro de esta región intergénica están coloreados en violeta. Las coordenadas genómicas corresponden a la secuencia NC_018875.1

2.3.4. Conservación evolutiva de los componentes estructurales dentro del género *Betabaculovirus*.

Los reportes previos de proteomas correspondientes a OB y ODV de granulovirus fueron utilizados para evaluar la conservación evolutiva de los distintos componentes de la partícula viral dentro de este género. A la fecha, solo dos estudios proteómicos han sido reportados para este grupo.^{68,70} Este trabajo expande estos resultados previos con los proteomas de los OB de EpapGV y SfGV. Con el fin de establecer las posibles relaciones evolutivas entre las proteínas detectadas en los OB de baculovirus, decidimos llevar a cabo un análisis proteómico comparativo. Las distintas secuencias

de aminoácidos de las proteínas detectadas en los estudios proteómicos de baculovirus fueron agrupadas en conjuntos de proteínas ortólogas (ortogrupos) usando la estrategia *reciprocal best hit* entre proteomas (Cuadro suplementario 6.1). Basándonos en las relaciones de ortología establecidas por nuestro análisis previo, los componentes de los proteomas de betabaculovirus (ClanGV, EpapGV, PiraGV y SfGV) fueron comparados utilizando un diagrama de Venn con el fin de identificar componentes compartidos y únicos dentro de cada proteoma (Figura 2.8).

A partir del gráfico de Venn podemos observar que existe un núcleo de 31 componentes conservados que son detectados consistentemente en todos los granulovirus. Sin embargo, se observó que varias proteínas reportadas previamente como componentes del OB en el virus modelo AcMNPV⁶² no fueron detectadas en alguno de los proteomas de granulovirus (ej. ODV-E18, ODV-E25, PIF3, P24). Basados en esta observación y teniendo en cuenta la complejidad del contenido proteico del OB, se decidió tomar un criterio de conservación mas laxo e incluir como componentes conservados también a las proteínas presentes en al menos 3 proteomas de granulovirus, lo que eleva el conjunto total de proteínas conservadas a 44. Por otra parte, se observó un gran número de proteínas virus-específicas para ClanGV y SfGV (16 y 14, respectivamente), mientras que EpapGV y PiraGV presentaron un número reducido de estas proteínas (3 en cada caso). Cuando se analizó la función asignada a estas proteínas virus-específicas, se evidenció que en su mayoría son proteína con funciones auxiliares (Chitinase, Gp37, P35, CBP, Enhancin, dUTPasa) o con función desconocida.

Para el caso de SfGV, se evidenció la presencia en el OB de las dos Enhancin codificadas en su genoma, las cuales son metaloproteasas que han sido involucradas en la degradación de la membrana peritrófica durante las fases iniciales de la infección oral,⁸⁸ lo que podría una de las fuentes del efecto sinergizante observado para SfGV. A su vez, este es el primer trabajo donde se reporta la presencia de varios miembros de esta familia de proteínas dentro del OB baculoviral. Resulta interesante destacar que en ambas replicas se observó que la abundancia de Enhancin-2 corresponde aproximadamente al doble de la de Enhancin-1.

Tomando ventaja de las estimaciones semi-cuantitativas provistas por los valores normalizados de emPAI para EpapGV y SfGV, se construyó un modelo del OB de granulovirus (Figura 2.9) utilizando las 44 proteínas conservadas en proteomas de granulovirus que se describen en la Figura 2.8 (estas proteínas están presentes en al menos 3 de los 4 proteomas). En primer lugar, puede obser-

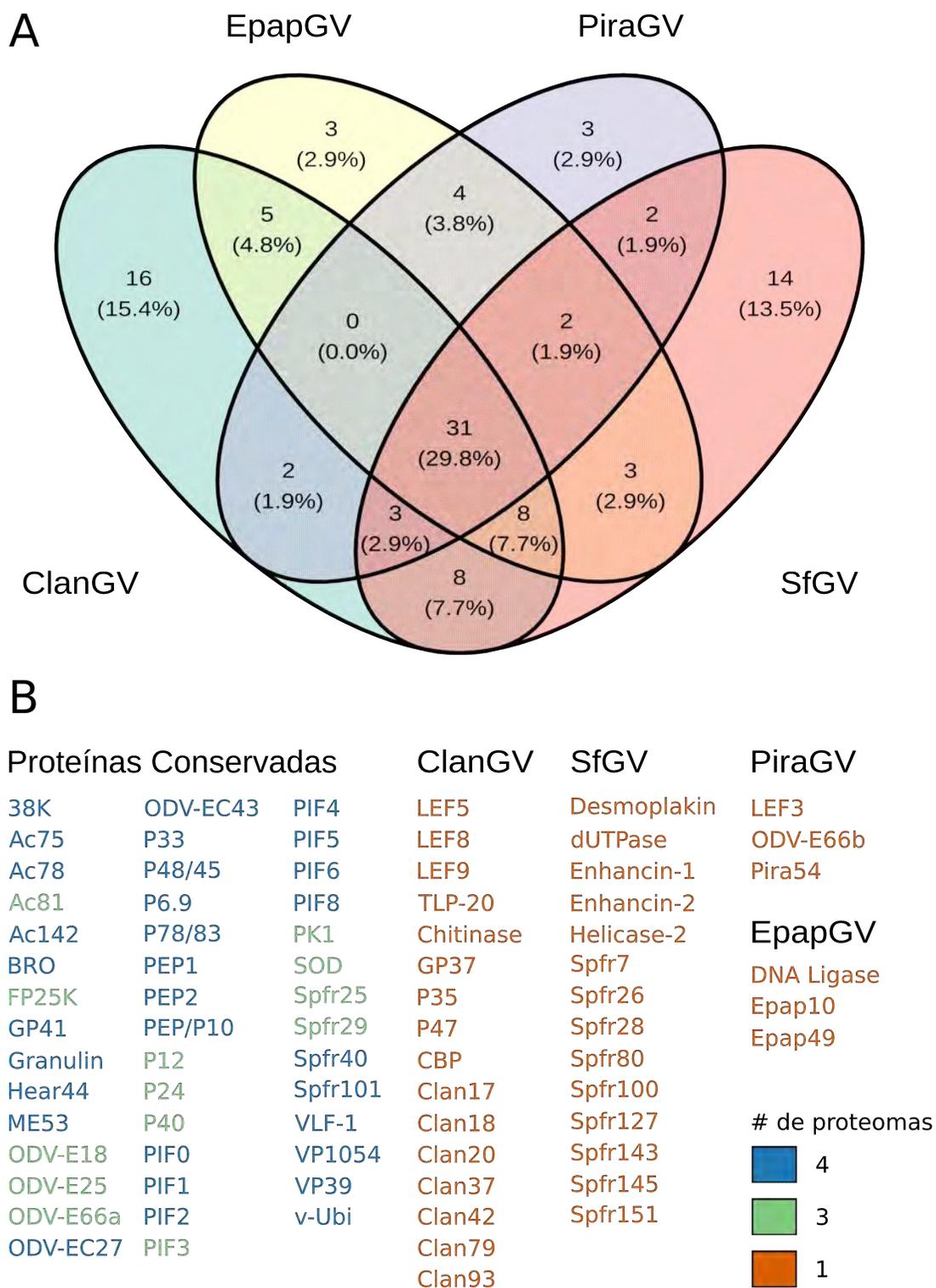


Figura 2.8: Conservación evolutiva del proteoma de los OB dentro del género *Beta-baculovirus*. **A** Las proteínas detectadas en los proteomas de granulovirus fueron agrupadas en ortogrupos y visualizadas empleando un diagrama de Venn. **B** Conjuntos de proteínas conservadas y virus-específicas identificadas en proteomas de granulovirus. Dentro de las proteínas conservadas, aquellas identificadas en los 4 proteomas se muestran en color azul, mientras que aquellas detectadas solo en 3 se muestran en verde y las proteínas virus-específicas en naranja.

vase que los componentes mayoritarios de las principales estructuras del OB se encuentra en gran abundancia (proteína mayoritaria de matriz, Granulina; de cápside, VP39; y de tegumento, GP41). Por otra parte, puede observarse que los componentes del complejo de proteínas *per os* infectivity factors (PIF), presentes en la membrana del virión, muestran diferencias en sus abundancias, con PIF5 y PIF8 como los componentes más abundantes. ODV-E66, una condritín sulfatasa involucrada en la degradación de la membrana peritrófica de la larva también apareció como un componente conservado del OB. En cuanto a las proteínas del OB con función desconocida, es interesante destacar que Spfr101 es uno de los componentes más abundantes junto con SOD y Hear44. El grupo de proteínas ortólogas al que pertenecen Epap95 y Spfr101 es exclusivo del género *Betabaculovirus*, mientras que Hear44 solo se encuentra presente en nucleopoliedrovirus de grupo II y en granulovirus (Epap126 y Spfr141).

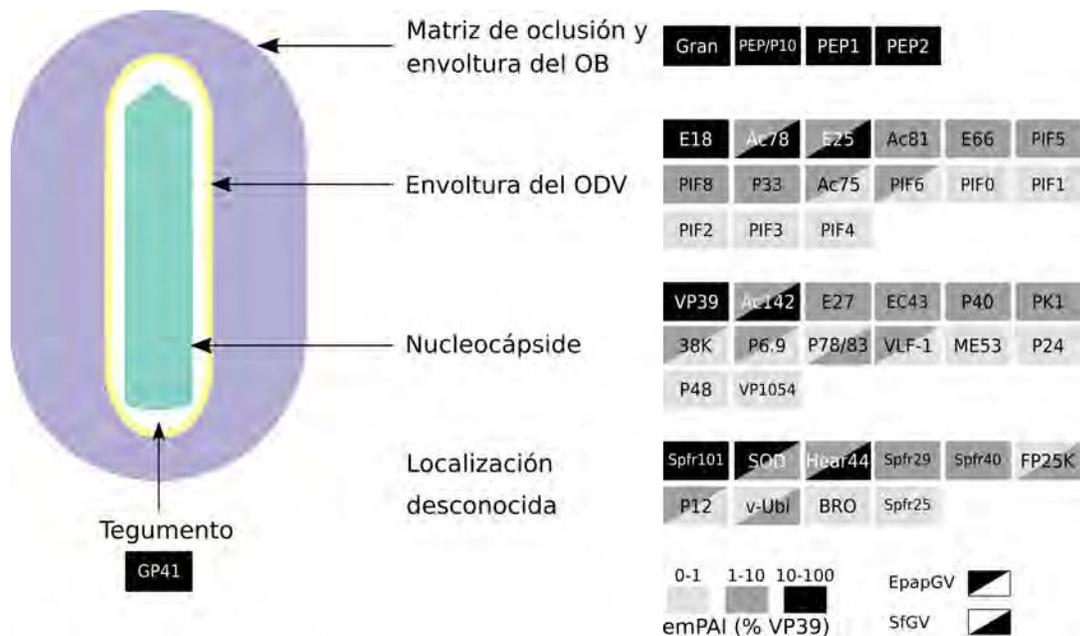


Figura 2.9: Modelo semi-cuantitativo del OB de granulovirus. Basados en datos proteómicos y las localizaciones descritas en reportes previos, propusimos este modelo de la partícula de OB. La cantidad aproximada de cada proteína fue estimada a partir de su valor de emPAI y se expresa como valor relativo con respecto a la proteína mayoritaria de cápside, VP39.

2.3.5. Caracterización evolutiva y molecular de Spfr101

Motivados por los elevados niveles de Spfr101 en el OB, decidimos profundizar el estudio de esta proteína a través de estudios bioinformáticos y celulares. En primer lugar, se realizó una búsqueda de similitud de secuencia con los programas BLASTp y HMMER^{139,140} con el fin de detectar secuencias ortólogas de Spfr101 presentes en la base de datos de proteínas de NCBI y luego reconstruir la historia evolutiva de esta familia de proteínas. Se encontraron un total de 20 secuencias, de las cuales todas han sido anotadas como proteínas hipotéticas, con la excepción de Hycu88 presente en *Hyphantia cunea* Granulovirus. Esta proteína presenta un dominio DASH_Dad3, el cual se encuentra en la proteína Dad3 del complejo DASH, un complejo de 10 subunidades capaz de unir microtúbulos.^{141,142} A partir de una reconstrucción filogenética se confirmó que Spfr101 es una proteína presente solo en granulovirus y en general la evolución de esta proteína no parece estar fuertemente relacionada con el clado de su granulovirus (Figura 2.10).

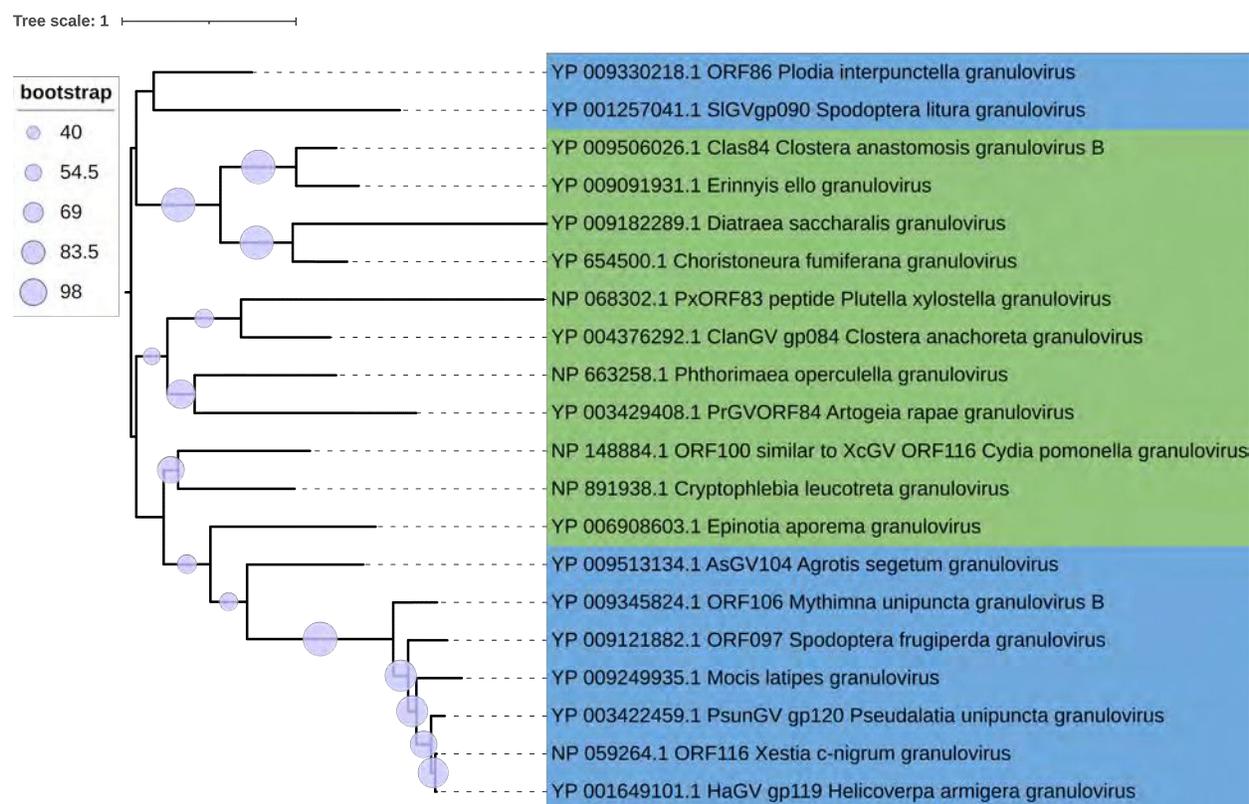


Figura 2.10: Historia evolutiva de la proteína Spfr101. Sombreado en azul se muestran los granulovirus correspondientes al clado *a*, mientras que en verde se muestran aquellos del clado *b*. Los círculos celestes representan los valores de soporte para cada rama de la filogenia obtenidos a partir de 1000 réplicas de UltraFast Bootstrap.

Como el objetivo de esclarecer la función de esta proteína, se decidió estudiar los posibles interactores celulares de Spfr101 en un modelo de cultivo celular de insecto. Para esto, se diseñó y construyó una proteína de fusión entre Spfr101 y la proteína fluorescente verde (eGFP) utilizando un vector que permite su expresión en células de insecto a partir del promotor *ie-2* (Figura 2.11 A). Como control negativo se incluyó el vector pIP-eGFP vacío, el cual fue digerido con *KpnI* y *SacI*, y luego religado para eliminar el codón de inicio del ORF eGFP, por lo que esta proteína no es traducida durante los ensayos de transfección. Por otra parte, como control positivo se utilizó un vector pIP-Z::eGFP, el cual expresa la proteína Z del virus Junín fusionada en su extremo C-terminal con eGFP. Mediante transfecciones de células *Sf9* con las construcciones mencionadas anteriormente se corroboró que la construcción correspondiente a Spfr101::eGFP es expresada exitosamente en células de insecto, mientras que el control negativo no mostró niveles de expresión significativos y el control positivo evidenció la presencia de la proteína Z::eGFP (Figura 2.11 B).

Llamativamente, se observó la formación de agregados fluorescentes en las células transfectadas con la construcción Spfr101::GFP (Figura 2.12 B), mientras que en el control positivo de expresión solo se observó una señal de fluorescencia homogénea (Figura 2.11 B). Utilizando el software TANGO, el cual permite predecir la tendencia a formar estructuras agregantes de tipo amiloide de una secuencia proteica, corroboramos que la región central de Spfr101 (residuos 57 a 61) contiene efectivamente un motivo proclive a formar estructuras amiloides (Figura 2.12 A).

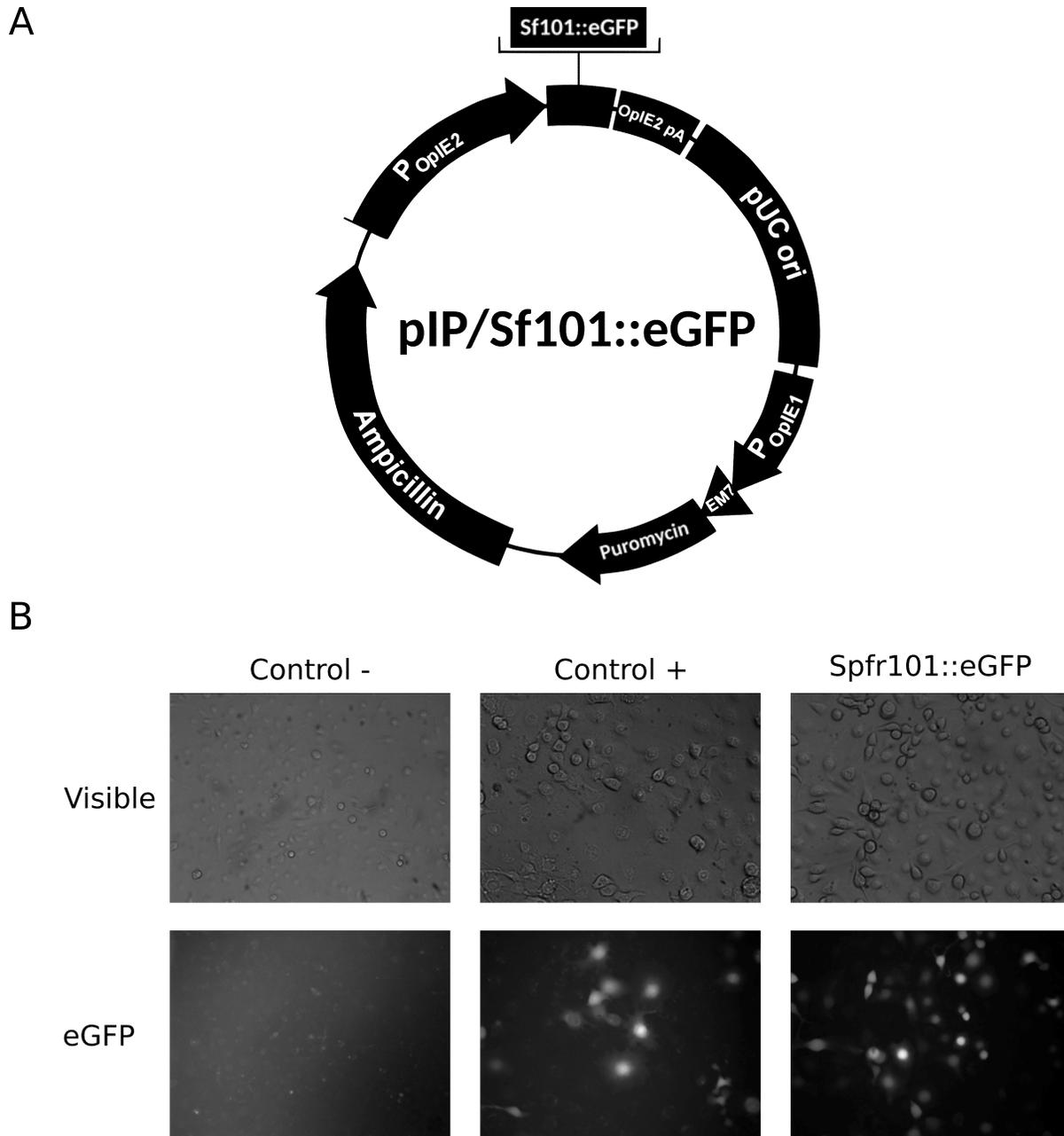


Figura 2.11: Construcción y expresión de Spfr101::eGFP en células de insecto. A Construcción empleada para expresar la proteína de fusión Spfr101::GFP. **B** Microscopía de fluorescencia de células Sf9 cinco días post-transfección.

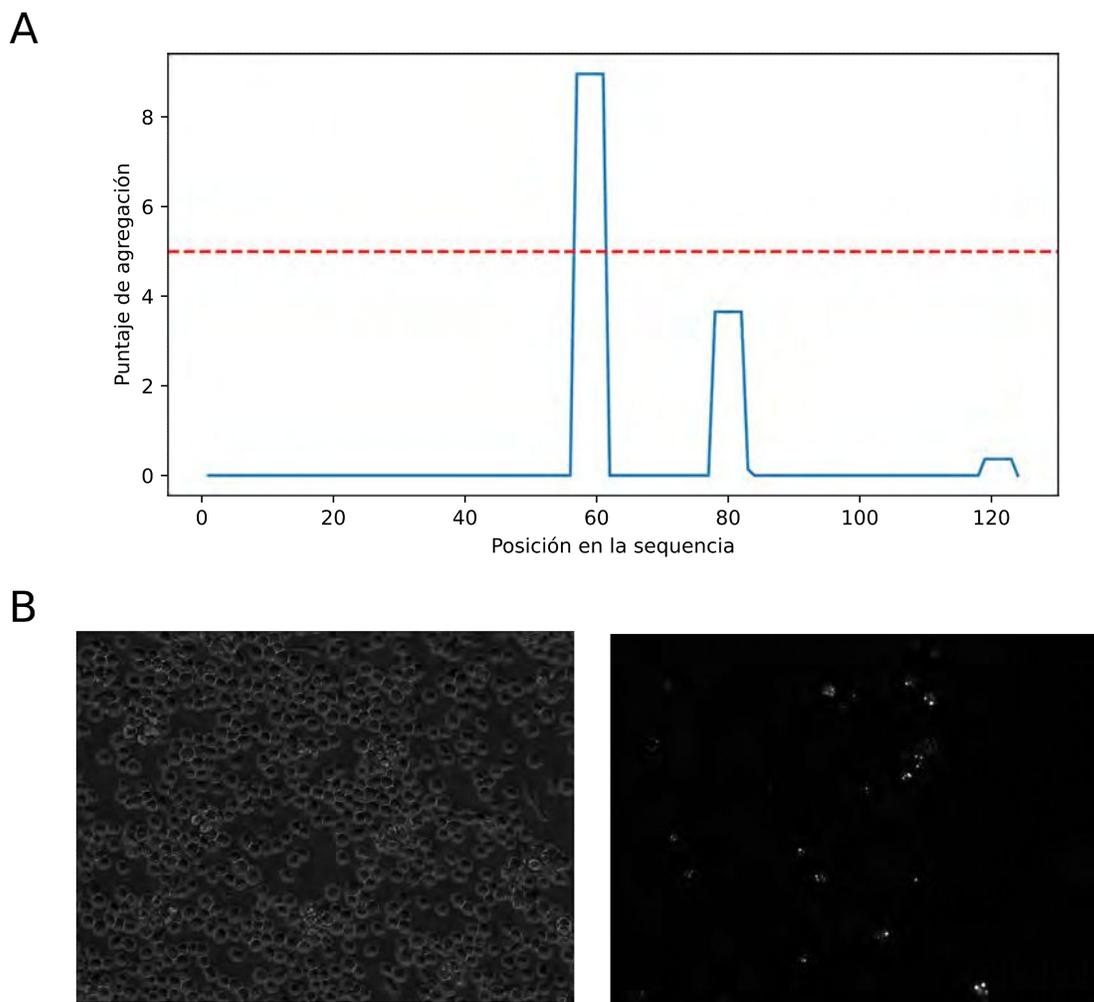


Figura 2.12: Formación de agregados por Spfr101::eGFP en células de insecto. A Perfil de agregación para la secuencia de Spfr101 obtenido con TANGO. La línea roja indica el puntaje mínimo a partir del cual se considera como amiloide una región dentro de la secuencia proteica. **B** Microscopía de fluorescencia de células Sf9 cinco días post-transfección donde se aprecian los agregados de Spfr101::GFP.

2.4. Discusión

El OB es la estructura baculoviral encargada de iniciar la infección oral primaria en el intestino medio de la larva. Para lograr esto, esta partícula viral requiere de un complejo conjunto de proteínas virales que participan en las diferentes etapas del proceso infectivo. A su vez, la conservación de ciertos componentes dentro del proteoma del OB aporta información acerca de proteínas poco caracterizadas previamente. Hasta el momento, solo los proteomas de los betabaculovirus ClanGV y PiraGV habían sido interrogados a través de técnicas basadas en MS,^{68,70} lo que nos dejaba un escenario incompleto y con datos limitados acerca de la composición proteómica dentro de este género. Los resultados presentados en este capítulo expanden este área a través de la incorporación de los proteomas de EpapGV y SfGV. Más aun, el estudio de SfGV representa el primer estudio de un granulovirus perteneciente al clado *a*, dado que ClanGV, EpapGV y PiraGV son miembros del clado *b*.

Es interesante destacar que aunque la familia *Baculoviridae* presentan una gran diversidad en cuanto a su contenido génico y su organización genómica, nuestros datos sustentan la noción de que a lo largo de su historia evolutiva una porción significativa de su genoma ha sido dedicada a codificar la mayoría de los componentes estructurales necesarios para el OB. Más aún, la mayoría de estos genes estructurales colineales son a su vez *core genes*, lo que remarca su importancia a nivel funcional. Estas proteínas estructurales forman parte de la envoltura, la nucleocápside y el tegumento del OB, como se observa en nuestro modelo semi-cuantitativo.

Ha sido reportado previamente que las proteínas virales relacionadas con el metabolismo del DNA pueden ser retenidas dentro del virión, lo que podría contribuir a acelerar el ciclo infectivo una vez que el virus ingresa a la célula hospedadora.⁶² Nuestros datos derivados del estudio comparativo entre los distintos proteomas de granulovirus apoyan la presencia de este tipo de enzimas en el OB. Sin embargo, no se identificaron de forma consistente los mismos conjuntos de enzimas entre los distintos proteomas (LEF-5, LEF-8 y LEF-9 para ClanGV, DNA-ligase, DNA-polimerase y Helicase-1 para EpapGV, LEF-3 para PiraGV, dUTPasa y Helicase-2 para SfGV), lo que podría explicarse si asumimos que estas proteínas, al no desempeñar un rol estructural *per se*, podrían incorporarse al OB de forma estocástica y sin un patrón determinado. Otra explicación plausible sería que

las diferencias bioquímicas presentes entre las proteínas de distintas especies podría modificar su afinidad para ser empaquetadas en el OB. Cuál es la explicación subyacente deberá ser investigada en mayor detalle en estudios futuros.

La envoltura que rodea al ODV está adaptada para el proceso de infección primaria dado que debe establecer interacciones moleculares con el epitelio intestinal del insecto e iniciar el proceso de ingreso a la célula. Las proteínas de envoltura pueden clasificarse en dos grandes grupos funcionales; aquellas requeridas para la envoltura del virión y aquellas relacionadas con la infectividad oral. La morfogénesis de la envoltura viral comienza con la formación de las microvesículas intranucleares (IMV) derivadas de la membrana nuclear interna y su asociación con la cápside viral. Las proteínas de membrana del ODV Ac75 y P18 son necesarias para la generación de las IMV.^{143,144} Subsecuente, el recubrimiento de las nucleocápsides ensambladas requiere de las proteínas Ac78, Ac81, Ac142, ODV-E25, ODV-EC43, P33 y P48 para formar el OB maduro.¹⁴⁵⁻¹⁵¹ Por otra parte, el otro grupo de proteínas de membrana del ODV pertenecen al complejo PIF (PIF0, PIF1, PIF2, PIF3, PIF4, PIF5, PIF6, PIF7/Ac110 y PIF8); este complejo molecular es el efector principal de la infección oral en el intestino del hospedador. Esas proteínas son codificadas por *core genes* conservados en todos los miembros de la familia *Baculoviridae*.⁷⁵ Nuestro modelo del OB de granulovirus da cuenta de estos componentes necesarios para la envoltura del ODV, con la excepción de P18. Un análisis detallado de la conservación de P18 muestra que esta proteína ha sido detectada tanto en EpapGV como en SfGV, por lo que su ausencia en los proteomas de ClanGV y PiraGV podría estar relacionado con un problema de ionización/detección más que con su ausencia propiamente dicha.

Los genomas de los baculovirus están densamente poblados con secuencias codificantes, las cuales en muchos casos se encuentran solapadas, además de poseer regiones intergénicas cortas.⁴⁰ Un estudio reciente ha descrito el perfil transcripcional durante la infección baculoviral en células de insecto, demostrando la existencia de varios transcriptos virales de naturaleza policistrónica y que a su vez se presentan solapados.¹⁰⁷ Junto con otras tecnologías ómicas, el mapeo proteogenómico es una herramienta valiosa para mejorar la anotación de regiones codificantes complejas de forma abarcativa. Este enfoque ha sido utilizado en el estudio de varias familias de virus,¹⁵² pero en el caso de los baculovirus solo ha sido empleado en el estudio de AgMNPV.¹¹² A partir de nuestros datos espectrales, pudimos identificar siete péptidos que no corresponden a ninguna región previa-

mente anotada en el genoma de EpapGV. Uno de esos péptidos es un ortólogo de *ac110*, un *core gene* involucrado en la infectividad oral. Mas aún, la presencia de péptidos derivados de marcos de lectura alternativos dentro de las regiones codificantes de *granulin* y *chitinase* plantean preguntas acerca de la complejidad subyacente de los procesos de transcripción y traducción en baculovirus. Sorprendentemente, se encontraron dos péptidos codificados por la región intergénica de *epap48* y *epap49*, lo que sugiere la presencia de un producto de fusión putativa entre ambas proteínas. Epap49 es la proteína más larga del genoma de EpapGV, compuesta por 1465 amino ácidos. Como se reportó previamente, fue difícil anotar esta como una proteína hipotética dado su atípico gran tamaño, la ausencia de homólogos detectable en la base de datos Genbank y la falta de motivos asociados con promotores.¹¹⁷ Además, se evidenció que otras proteínas de gran tamaño son codificadas en una posición similar en los genomas de ChocGV y HearGV.^{37,153} Nuestros datos en este capítulo proveen evidencia de que Epap49 es efectivamente traducida.

Dentro de las proteínas identificadas de forma específica en el OB de SfGV, se encontraron las dos Enhancin codificadas en su genoma (Enhancin-1 y Enhancin-2). Reportes previos han demostrado que las proteínas Enhancin se localizan dentro del OB para el caso de *Pseudatelia unipuncta* granulovirus (PsunGV) y *Lymantria dispar* Multiple Nucleopoliedrovirus (LdMNPV), lo que permitiría la interacción entre estas enzimas degradativas y la membrana peritrófica del intestino medio.^{79,154} Esto sugiere que la función potenciadora de SfGV cuando es aplicado en conjunto con SfMNPV podría ser explicado en parte gracias a la presencia de ambas Enhancin en su OB.^{101,119}

Nuestros datos proteómicos sugieren que las proteínas pertenecientes a la familia de Epap95/Spfr101 no solo son componentes específicos que los granulovirus han adaptado para su uso en el OB, sino que también muestra que estas proteínas se encuentran en elevados niveles de abundancia, sugiriendo un rol importante para estas proteínas. El hecho de que en un contexto *in vitro* esta proteína es capaz de formar agregados potencialmente amiloides (como lo demuestra su perfil de agregación *in silico*), podría indicar que esta proteína tiene un comportamiento similar al de LEF-10, la cual se ha demostrado recientemente que puede actuar como una proteína tipo prión.¹⁵⁵ En este sentido, es curioso que tanto Epap95/Spfr101 como LEF-10 no comparten un origen evolutivo común, lo que implicaría que este comportamiento amiloide se originó al menos dos veces de forma independiente dentro de los baculovirus. Aunque la función de este fenómeno aún no está claro, parece que estaría

involucrado en la regulación de la expresión de genes tardíos.¹⁵⁵

2.5. Materiales y Métodos

2.5.1. Larvas y virus

Larvas de *S. frugiperda* se obtuvieron a partir de los insectarios del Instituto Nacional de Tecnología Agropecuaria (INTA) y la Estación Experimental Agroindustrial Obispo Colombres (EEAOC-CONICET-ITANOVA), Tucumán, y fueron utilizadas para establecer una colonia dentro del Instituto de Biotecnología y Biología Molecular (IBBM). La colonia se mantuvo con una dieta artificial y un período de luz controlada (16 horas de luz). En el caso de EpapGV, se emplearon OB preparados anteriormente.¹¹⁷ Los aislamientos virales utilizados en este estudio son los reportados previamente por nuestro laboratorio:^{91,117} EpapGV (Refseq ID NC_018875) y SfGV (Refseq ID MH170055).

2.5.2. Producción y purificación de los cuerpos de oclusión (OB)

Las larvas de *E. aporema* en su 4° estadio y *S. frugiperda* en 1°-2° estadio fueron infectadas por vía oral usando dieta artificial contaminada con una suspensión de OB de EpapGV y SfGV, respectivamente. Las larvas moribundas con signos de infección compatibles con baculovirus fueron almacenadas y procesadas como se describió en.^{91,117} Brevemente, las larvas infectadas se colocaron en agua destilada y fueron homogeneizadas empleando un homogenizador Dounce de vidrio. La solución resultante fue filtrada a través de tres capas de gasa para eliminar restos de insecto insolubles. Este extracto se clarificó por 3 pasos de centrifugación a 10000 g por 10 minutos seguidos de un lavado con solución 0.05 % v/v de SDS. La solución clarificada se sometió a una ultracentrifugación en un gradiente continuo de sacarosa 30-60 % p/p (50000 g, 60 minutos a 4°C, rotor Beckman SW 41 Ti). La banda blancuzca/opalescente correspondiente a los OB fue recolectada, diluida 10 veces en agua destilada y precipitadas por centrifugación a 14000 g por 10 minutos. El precipitado final fue resuspendido en agua destilada y almacenado a -20°C. Se procesaron dos replicas biológicas de EpapGV y SfGV. La cantidad de proteína total en cada muestra fue cuantificada a través del ensayo de Bradford.¹⁵⁶

2.5.3. Microscopía electrónica de barrido y de transmisión de los OB

Para la fijación de los OB previo a ambas técnicas de microscopía se empleó un buffer de fijación compuesto por paraformaldehído 4 % en buffer fosfato salino. Este se preparó agregando 4 g de

paraformaldehído en 80 mL buffer fosfato salino a 55°C, luego se agregó 1 mL de NaOH 1M y se agitó hasta disolución; se ajustó el pH a 7.4, se llevó a 100 mL de volumen final y se esterilizó por filtración. Las concentraciones de iones empleados en la solución de buffer fosfato salino fueron las siguientes: con NaCl 137 mM, KCl 2.7 mM, Na₂HPO₄ 10 mM y KH₂PO₄ 1.8 mM, a pH 7.4. En el caso de la microscopía de barrido, una muestra de OB se fijó por con paraformaldehído 4% a 4°C por 24 horas. Luego de esta incubación, se centrifugaron los OB por 5 minutos a 8000 g para decantarlos y se resuspendieron en etanol 100% y se sometieron a un secado por punto crítico antes de realizar un metalizado con oro. La adquisición se realizó con un microscopio electrónico de barrido ambiental (ESEM), marca FEI, modelo Quanta 200 del servicio de microscopía electrónica de barrido de la Facultad de Ingeniería, UNLP.

Para la microscopía de transmisión, una alícuota de la suspensión de OB de SfGV se resuspendió en agua de manera de obtener una solución translúcida. Esta solución se incubó por 120 minutos en buffer de fijación diluido 4 veces. Luego de esta incubación, se centrifugaron los OB por 5 minutos a 8000 g para decantarlos y se descartó el sobrenadante. Este proceso de fijación se repitió otra vez pero empleando el buffer de fijación diluido 2 veces. Una fijación secundaria se realizó con tetróxido de osmio al 1% durante 1 hora a 4°C y posteriormente, las muestras se deshidrataron en una serie creciente de alcoholes y se incluyeron en resina epoxi. Los cortes ultrafinos (90 nm) se contrastaron con acetato de uranilo y citrato de plomo y se examinaron en un microscopio electrónico de transmisión JEM 1200 EX II (JEOL Ltd., Tokio, Japón) y se fotografiaron con una cámara Erlangshen ES1000W, Modelo 785 (Gatan Inc., Pleasanton, California, USA) del servicio central de microscopía electrónica de la Facultad de Ciencias Veterinarias, UNLP.

2.5.4. Análisis por espectrometría de masas

Para la electroforesis en gel desnaturizante de poliacrilamida (SDS-PAGE) se preparó un sistema con dos geles contiguos, uno separador 12% p/v de acrilamida en Tris-HCl 0.375 mM pH 8.8 y uno de siembra 4% p/v de acrilamida en Tris-HCl 0.125 mM pH 6.8. Ambos se prepararon con SDS a una concentración de 0.1% p/v. La muestra fue tratada con un buffer de siembra conteniendo SDS 0.1% p/v y β-mercapto etanol, se hirvió 5 minutos y se corrió a 120 voltios por 30 minutos. La calle con las proteínas se cortó en cuatro fracciones, las cuales se sometieron al procesamiento

de muestra para MS.

La digestión proteica y el análisis proteómico fue realizado en las instalaciones del Centro de Estudios Químicos y Biológicos por Espectrometría de Masas (CEQUIBIEM), perteneciente a la Universidad de Buenos Aires y el Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). Las muestras de proteínas fueron reducidas empleando en un buffer de bicarbonato de amonio 50 mM pH 8 y 10 mM de ditioneol (45 minutos a 56°C) y posteriormente carbamidometiladas con buffer bicarbonato de amonio 50 mM pH 8 con 20 mM de iodoacetamida (40 minutos a temperatura ambiente y en oscuridad). La solución de proteínas fue precipitada con 0.2 volúmenes de ácido tricloroacético 100% w/v a -20°C por 120 minutos y centrifugada a 12000 g a 4°C por 10 minutos. El precipitado resultante se lavó dos veces con acetona fría y se dejó secar a temperatura ambiente. El precipitado de proteínas se resuspendió en buffer bicarbonato de amonio 50 mM pH 8 y se digirió con tripsina (Promega V5111). Los péptidos resultantes fueron desalados con una columna ZipTip C18 (Millipore).

La mezcla de péptidos fue analizada por cromatografía líquida acoplada a LC-MS/MS en un espectrómetro de masas Thermo Scientific Q Exactive conectado a un cromatógrafo EASY-nLC 1000 (Thermo Scientific). Para el análisis por LC-MS/MS, se cargó aproximadamente 1 µg de péptidos en una columna de fase reversa Easy-Spray Column PepMap RSLC (P/N ES8001) (C18, 2 µm x 10 nm particle size, 50 µm x 150 mm) y se eluyó por 120 minutos para separar con un alto grado de resolución. El flujo empleado para la nano columna fue de 300 nL por minuto y el rango de solvente fue desde 7% B (5 minutos) hasta 35% B (120 minutos). Solvente A fue 0.1% ácido fórmico en agua mientras que el solvente B fue 0.1% ácido fórmico en acetonitrilo. El volumen de inyección fue de 2 µL. El equipamiento de MS tiene una celda de disociación por colisión elevada (high collision dissociation, HCD) para la fragmentación de los péptidos y un analizador Orbitrap (Thermo Scientific, Q-Exactive). Se utilizó un voltaje de 3.5 kV se utilizó para la ionización por electro-spray (Thermo Scientific, EASY-SPRAY).

El software XCalibur 3.0.63 (Thermo Scientific) fue utilizado para la adquisición de datos y para configurar del equipo de forma tal que permita la identificación peptídica en simultáneo con su separación cromatográfica. Se obtuvieron escaneos completos de los espectros de masa utilizando el analizador Orbitrap. El rango de masas escaneadas fue de 400-1800 m/z, con una resolución de

70000 a 400 m/z. Los 12 iones de mayor intensidad en cada ciclo fueron aislados secuencialmente, fragmentados por HCD y medidos en el analizador Orbitrap. Los péptidos con una carga de +1 o con un estado de carga no asignado fueron excluidos de la fragmentación por MS2.

2.5.5. Análisis de los datos de espectrometría de masas

Los datos crudos del Q Exactive fueron procesados con el software Proteome Discoverer™ (version 2.1.1.21, Thermo Scientific) y utilizados para interrogar una base de datos de proteínas de EpapGV descargada desde NCBI (números de acceso NC_018875 y MH170055, National Center for Biotechnology Information; www.ncbi.nlm.nih.gov) las cuales fueron digeridas *in silico* con un máximo de un corte fallido por péptido. Las búsquedas con Proteome Discoverer™ se realizaron con una tolerancia al precursor de masa de 10 ppm y una tolerancia al producto de ionización de 0.05 Da. Como modificaciones estáticas se seleccionaron carbamidometilación de los residuos Cys, y como modificaciones dinámicas se seleccionaron la oxidación de los residuos Met y la acetilación N-terminal. Los resultados positivos de la búsqueda se filtraron en base a la presencia de péptidos con alto nivel de confianza y un valor máximo de la tasa de falsos positivos del 1 %, calculada usando una estrategia de base de datos reversa. El emPAI se calculó automáticamente con el software Proteome Discoverer™ y se utilizó para estimar la abundancia relativa de las proteínas identificadas en cada muestra. Para obtener una estimación de los niveles de cada proteína que fueran comparable entre sí, los valores de emPAI de cada proteína fueron promediados entre réplicas a través de una media geométrica y luego fueron normalizados con respecto al valor de la proteína mayoritaria de cápside VP39 ($emPAI_n = \frac{\sqrt{emPAI_1 \times emPAI_2}}{emPAI_{VP39}}$).

2.5.6. Búsqueda de péptidos no anotados

Los genomas completos de EpapGV y SfGV fueron traducidos *in silico* para los seis marcos de lectura posibles con el software de búsqueda Mascot. Los datos espectrales fueron interrogados para descartar todos los hits de péptidos pertenecientes a ORFs anotados previamente. Los péptidos remanentes (candidatos a ser parte de un ORFs no anotado) fueron mapeados a su posición genómica correspondiente y se extendieron hasta encontrar un codón de terminación en su extremo C-terminal, y un codón de inicio o terminación en su extremo N-terminal. Se buscaron secuencias homólogas en otros baculovirus con la herramienta TBLASTN.

2.5.7. Agrupamiento de secuencias ortólogas

Se generó una base de datos conteniendo todas las proteínas detectadas en los OB de baculovirus fue generada en base a los datos reportados previamente.^{62,64,67-71,132} La herramienta BLASTP¹⁵⁷ y HMMER¹⁴⁰ se emplearon para identificar grupos de proteínas ortólogas entre los distintos proteomas mediante la estrategia *reciprocal best hit*. Brevemente, cada proteína de interés se utilizó como query en una búsqueda de similitud de secuencia y se anotó el hit con mayor puntaje para cada proteoma de la base de datos. Dos proteínas se anotaron como miembros del mismo ortogrupo si y solo si correspondían al par de proteínas con mayor puntaje entre ambos proteomas (por ej. la proteína A en el proteoma 1 y la proteína B en el proteoma 2 se identificaron como ortólogos si al usar como query a la proteína A, el hit de mayor puntaje en el proteoma 2 es la proteína B, y viceversa, si al usar como query a la proteína B, el hit de mayor puntaje en el proteoma 1 es la proteína A).

2.5.8. Clonado de *spfr101* en el vector pIP

Los OB purificados se disolvieron por lisis alcalina y la extracción del DNA viral se realizó como se describió previamente.⁹¹ El DNA genómico correspondiente a SfGV fue empleado para amplificar la secuencia de *spfr101* utilizando un par de oligonucleótidos específicos con sitios de reconocimiento para las enzimas *KpnI* y *SacI* (oligo *forward* CTGCAAGGTACC-CCAATTATTTGCGAGACCAA, oligo *reverse* CTGCAAGAGCTC-TTATTTAAACACATATTGAC). Este producto de amplificación fue purificado mediante un kit de purificación por columna de sílica y digerido con *KpnI* y *SacI* en simultáneo durante 60 minutos. El vector pIP, desarrollado previamente en nuestro laboratorio [Fabre_2020_tesis], también fue digerido con las mismas enzimas y adicionalmente fue defosforilado con fosfatasa alcalina termosensible (PromegaTM) para disminuir la frecuencia de eventos de religación. El inserto conteniendo la secuencia de *spfr101* y el vector pIP linealizado, ambos digeridos, se ligaron utilizando T4 DNA ligase (*New England Biolabs*TM) en una incubación de 60 minutos a temperatura ambiente.

Un microlitro (1 μ L) de la mezcla de ligación se utilizó para electroporar 50 μ L de bacterias competentes DH5 α (*F*⁻, Φ 80*dlacZKD* Δ *M15*, Δ (*lacZYA-argF*)*U169*, *deoR*, *recA1*, *endA1*, *hsdR17*(*r_k*, *m_k*+), *phoA*, *supE44*, λ , *thi-1*, *gyrA96*, *relA1*) bajo las siguientes condiciones: 25 μ F de capacitancia y 200 ohm de resistencia en un equipo Gene Pulser TM (BioRad, Hercules, CA, EEUU).

Las bacterias se recuperaron en 1 mL de medio LB por 60 minutos luego de ser electroporadas y finalmente se plaquearon 100 μ l de este cultivo en una placa de LB sólido con ampicilina (100 μ g/mL). Estas placas se incubaron 18 horas a 37°C y las colonias resultantes se analizaron por PCR para confirmar la presencia de la construcción de interés.

2.5.9. Cultivo y transfección de células de insecto

Las células Sf9 fueron cultivadas en medio de Grace (Thermo) suplementado al 10 % con suero fetal bovino (Internegocios, Argentina) y mantenidas en estufa a 27°C.

Para la transfección de células de insecto se empleó Cellfectin (Thermo) de acuerdo al protocolo sugerido por el fabricante. Brevemente, se plaquearon 8×10^5 células resuspendidas en medio Grace's con 1.5 % de suero fetal bovino por pocillo en una placa de 6 pocillos. Se diluyeron 8 μ L de reactivo Cellfectin (previamente homogeneizado) en 100 μ L de medio de Grace y 1 μ g de DNA plasmídico en 100 μ L de medio de Grace sin suplementar. Ambos tubos se mezclaron por separado mediante vortex, luego se combinó el DNA diluido con el Cellfectin, se volvieron a mezclar mediante vortex y se incubaron 15 minutos a temperatura ambiente. Los 200 μ L de mezcla de transfección se agregaron por goteo al pocillo con las células y se incubaron 3 a 5 días a 27°C.

2.5.10. Predicción de la tendencia a la agregación

El programa TANGO¹⁵⁸ se utilizó con su configuración por defecto para predecir las regiones propensas a formar estructuras β dentro de Spfr101. Este programa se obtuvo a través de una licencia académica y se descargó desde la siguiente página web (<http://tango.crg.es/about.jsp>).

Capítulo 3

Diversidad genómica en una población de SfMNPV

Parte de los resultados presentados en este capítulo han sido adaptados para su publicación en el siguiente artículo:

Tomás Masson María Laura Fabre, Matias Luis Pidre , José María Niz, Marcelo Facundo Berretta, Víctor Romanowski, María Leticia Ferrelli. Genomic Diversity in a Population of *Spodoptera frugiperda* Nucleopolyhedrovirus. *Infection, Genetics and Evolution*. Volume 90, June 2021, 104749. doi: 10.1016/j.meegid.2021.104749.

3.1. Resumen

El virus *Spodoptera frugiperda* Multiple Nucleopolyhedrovirus (SfMNPV) es un fuerte candidato para el desarrollo de pesticidas de bajo impacto ambiental contra la oruga militar tardía (*Spodoptera frugiperda*), una plaga ampliamente distribuida que representa un gran riesgo para diversos cultivos alrededor del mundo. Sin embargo, disponemos de poca información acerca de la diversidad genómica presente dentro de los aislamientos naturales de SfMNPV y su posible impacto en la interacción virus-hospedador. En este capítulo, abordamos por primera vez el estudio de la diversidad genómica dentro de un aislamiento autóctono de SfMNPV a través de secuenciación masiva de elevada profundidad. Se identificaron 704 variantes nucleotídicas puntuales intrahospedador, de las cuales 184 son mutaciones no sinónimas distribuidas entre 82 secuencias codificantes diferentes. A su vez detectamos variantes estructurales en distintas ubicaciones del genoma de SfMNPV, incluidas dos deleciones en la región *egt* las cuales han sido reportadas en otros aislamientos a partir de patrones de restricción. Un análisis comparativo entre los polimorfismos presentes entre diferentes aislamientos geográficos de SfMNPV y las variantes intrahospedador reportadas en este trabajo sugiere que las regiones codificantes con niveles de diversidad genómica más elevados están asociados con funciones de infectividad oral o desconocida. En este contexto, a través de estudios de evolución molecular mostramos evidencia de selección adaptativa actuando sobre el gen *sf29*, el cual codifica una colagenasa putativa que podría contribuir a la infectividad oral de SfMNPV. En suma, estos resultados contribuyen a profundizar nuestro conocimiento del proceso de adaptación de SfMNPV a su hospedador, un paso necesario para el diseño de estrategias direccionadas a mejorar su aplicabilidad como agente de control biológico.

3.2. Introducción

La oruga militar tardía, *Spodoptera frugiperda* (J.E. Smith, Lepidoptera: Noctuidae), es una plaga agrícola invasiva que ataca una gran variedad de cultivos y se encuentra distribuida naturalmente en el continente americano. Reportes recientes han descripto la presencia de *S. frugiperda* en los continentes de África y Asia.^{159,160} Los métodos de control tradicional para esta peste involucran el uso de productos químicos o de cultivos transgénicos capaces de expresar la toxina Bt.¹⁶¹ Sin embargo, los casos de resistencia a campo frente a la toxina Bt y el impacto ambiental de los pesticidas químicos han forzado el desarrollo de nuevas estrategias de control.¹⁶² Los baculovirus son virus específicos de insectos con un rango de hospedadores muy estrecho y una elevada virulencia, lo que los convierte en candidatos potencialmente seguros para el desarrollo de pesticidas.²⁰ Específicamente, un nucleopoliedrovirus capaz de infectar a *S. frugiperda* (SfMNPV) ha sido implementado para el control biológico de poblaciones de la oruga militar.¹⁶³

A la fecha, cinco genomas completos de SfMNPV han sido reportados,^{90,93,94,164,165} y recientemente tres nuevos genomas han sido depositados en la base de datos Genbank de NCBI, abriendo las puertas a análisis comparativos que exploren la relación entre la evolución del genoma viral y las distintas propiedades infectivas de cada aislamiento. Las deleciones y los genotipos mixtos son comunes en las poblaciones naturales de SfMNPV y han sido propuestas como moduladores de su velocidad de muerte y potencia.^{166,167} Mas aún, ensayos biológicos comparativos entre distintas variantes genotípicas presentes dentro de un mismo aislamiento de SfMNPV han demostrado que estos genotipos pueden diferir en sus propiedades biológicas.^{33,34,90,92,168}

La innovación genética es una de las mayores fuentes de adaptación durante el proceso de coevolución entre el virus y su hospedador.^{169,170} Para el caso del baculovirus modelo, *Autographa californica* Multiple Nucleopolyhedrovirus (AcMNPV), los polimorfismos presentes en la población viral se han demostrado como fuente relevante de diversidad funcional para las proteínas virales.¹¹⁰ Mas aún, se ha propuesto que el escape viral de la supresión del hospedador es dirigido por un grupo de genes con elevada diversidad génica y bajo un régimen de selección positiva recurrente.¹⁷¹ Sin embargo, ningún estudio ha explorado la ocurrencia y frecuencia de polimorfismos genómicos dentro de las poblaciones de SfMNPV a través de secuenciación de alto rendimiento (*high-throughput sequencing*,

HTS). Esta información nos ayudaría a expandir nuestro conocimiento sobre el proceso de adaptación de SfMNPV a la oruga militar tardía y complementarían los datos de diversidad genética disponibles de otros estudios.

En este trabajo, abordamos este problema a través de la caracterización de la diversidad genómica presente dentro de un aislamiento argentino de SfMNPV (ARG-M)¹⁷² mediante HTS. A través del análisis de las variantes no sinónimas, identificamos genes relacionados con la infectividad oral o con función desconocida como los dos grupos de genes más diversos dentro del genoma de SfMNPV. Mas aún, un estudio evolutivo del gen *sf29* sugiere que este gen codifica una colagenasa putativa con un motivo catalítico conservado y presenta evidencia de selección positiva pervasiva. Además, pudimos detectar varias variantes estructurales (en su mayoría deleciones y duplicaciones) que contribuyen a la diversidad genómica de SfMNPV. El análisis genómico comparativo basado en un alineamiento de genoma completo posicionó a ARG-M como un aislamiento cercano a los de Nicaragua y sugiere que las variantes nucleotídicas puntuales (*single nucleotide variants*, SNP) son una herramienta útil para discriminar la composición genómica de diferentes aislamientos, como se demostró previamente para CpGV.³¹ Creemos que estos datos ayudarán a comprender mejor el proceso de adaptación de SfMNPV a su hospedador natural y podrían ser útiles para mejorar su aplicabilidad como biopesticida.

3.3. Resultados

3.3.1. Secuenciación profunda de una población de SfMNPV ARG-M

Un aislamiento argentino autóctono de SfMNPV, nombrado ARG-M, fue caracterizado a través de secuenciación masiva profunda utilizando tecnología Illumina por primera vez, lo que nos permitió abordar el estudio de su diversidad genómica. Nuestros datos de secuenciación profunda alcanzaron una cobertura promedio de 52313X (desviación estándar de 4052) para cada posición del genoma. La secuencia genómica consenso de SfMNPV ARG-M está compuesta por 132696 pares de bases y comprende 144 regiones codificantes (Accession number MW162628, Cuadro 3.1), de las cuales 143 están presentes en el aislamiento de referencia 3AP2⁹⁰ y el gen restante corresponde al ORF *sf110a*, presente solo en los aislamientos provenientes de Nicaragua (NicB and NicG).¹⁶⁵ Cuando

se examinaron las diferencias estructurales entre los genomas de SfMNPV ARG-M y el aislamiento 3AP2, solo encontramos un pequeño número de inserciones/deleciones de una longitud menor a 6 pb en regiones no codificantes, con la excepción de una deleción de 31 pb en la región comprendida por el gen *me53* y la secuencia repetitiva *hr1*. Las discrepancias restantes entre ambas secuencias genómicas se debieron a SNP.

Cuadro 3.1: Secuencias codificantes anotadas en el genoma de SfMNPV ARG-M

Número_ORF	Inicio	Terminación	Polaridad	Gen	Proteína	SfMNPV-3AP2	AcMNPV
ORF1	1	741	+	polh	Polyhedrin	sf1	ac8
ORF2	806	2194	-	orf1629	PP78/82	sf2	ac9
ORF3	2193	3029	+	pk-1	PK-1	sf3	ac10
ORF4	3086	4501	-	hoar	HOAR	sf4	-
ORF5	4557	4727	-	sf5	Sf5	sf5	-
ORF6	4727	4900	-	sf6	Sf6	sf6	-
ORF7	4837	5124	+	sf7	Sf7	sf7	-
ORF8	5135	6589	+	sf8	Sf8	sf8	-
ORF9	6773	7888	+	pif-5	PIF-5	sf9	ac148
ORF10	8082	9173	+	me53	ME53	sf10	ac139
ORF11	11019	11177	-	sf11	Sf11	sf11	-
ORF12	11176	13284	+	efp	F-protein	sf12	ac23
ORF13	13403	13693	-	gp16	GP16		sf13
ORF14	13710	14423	-	p24	P24	sf14	ac129
ORF15	14471	14806	+	sf15	Sf15	sf15	-
ORF16	14763	15404	+	lef-2	LEF-2	sf16	ac6
ORF17	15450	16601	-	38.7k	38.7K	sf17	ac13
ORF18	16605	17324	-	lef-1	LEF-1	sf18	ac14

Número_ORF	Inicio	Terminación	Polaridad	Gen	Proteína	SfMNPV-3AP2	AcMNPV
ORF19	17315	17734	+	sf19	Sf19	sf19	-
ORF20	17715	18734	-	v-cath	Cathepsin	sf20	ac127
ORF21	18852	19907	-	lef-7	LEF-7	sf21	ac125
ORF22	19937	21631	+	chiA	Chitinase	sf22	ac126
ORF23	21888	22553	+	sf23	Sf23	sf23	-
ORF24	22664	23446	+	gp37	GP37	sf24	ac64
ORF25	23443	23946	-	ptp-2	PTP-2	sf25	-
ORF26	24042	25619	+	egt	EGT	sf26	ac15
ORF27	25806	26330	+	sf27	Sf27	sf27	-
ORF28	26337	26984	+	sf28	Sf28	sf28	ac17
ORF29	27004	29643	-	sf29	Sf29	sf29	-
ORF30	29754	30308	+	sf30	Sf30	sf30	-
ORF31	30387	30914	+	pkip-1	PKIP-1	sf31	ac24
ORF32	30946	31464	-	sf32	Sf32	sf32	-
ORF33	31427	32323	-	arif-1	ARIF-1	sf33	ac20/21
ORF34	32220	33416	+	pif-2	PIF-2	sf34	ac22
ORF35	33434	35023	+	pif-1	PIF-1	sf35	ac119
ORF36	35031	35288	+	sf36	Sf36	sf36	ac120
ORF37	35285	36427	-	fgf	FGF	sf37	ac32
ORF38	36670	37386	+	sf38	Sf38	sf37	-

Número_ORF	Inicio	Terminación	Polaridad	Gen	Proteína	SfMNPV-3AP2	AcMNPV
ORF39	37428	38645	-	alk-exo	Alkaline-exonuclease	sf39	ac133
ORF40	39182	39520	-	sf40	Sf40	sf40	ac19
ORF41	39519	40673	+	sf41	Sf41	sf41	ac18
ORF42	40688	41059	-	sf42	Sf42	sf42	-
ORF43	41123	42310	-	sf43	Sf43	sf43	-
ORF44	42518	43309	+	sf44	Sf44	sf44	-
ORF45	43324	44322	-	pep	PEP	sf45	ac131
ORF46	44410	44718	-	sf46	Sf46	sf46	ac117
ORF47	44648	44956	-	sf47	Sf47	sf47	-
ORF48	45017	45478	-	sod	SOD	sf48	ac31
ORF49	45547	46176	+	pif-3	PIF-3	sf49	ac115
ORF50	46178	46609	+	sf50	Sf50	sf50	-
ORF51	46591	48144	+	sf51	Sf51	sf51	-
ORF52	48167	48820	+	sf52	Sf52	sf52	ac106/107
ORF53	48857	49963	-	sf53	Sf53	sf53	ac33
ORF54	50507	50941	+	dutpase	dUTPase	sf54	-
ORF55	51007	51627	+	sf55	Sf55	sf55	-
ORF56	51694	52527	+	p13	P13	sf56	-
ORF57	52587	54677	+	odv-e66	ODV-E66	sf57	ac46
ORF58	54670	55014	-	sf58	Sf58	sf58	ac108

Número_ORF	Inicio	Terminación	Polaridad	Gen	Proteína	SfMNPV-3AP2	AcMNPV
ORF59	55043	56113	-	odv-ec43	ODV-EC43	sf59	ac109
ORF60	56097	56276	-	pif-7	PIF-7	sf60	ac110
ORF61	56273	57988	-	vp80	VP80	sf61	ac104
ORF62	58030	59160	+	p45	P45	sf62	ac103
ORF63	59135	59440	+	p12	P12	sf63	ac102
ORF64	59463	60614	+	p40	P40	sf64	ac101
ORF65	60685	60906	+	p6.9	P6.9	sf65	ac100
ORF66	60903	61733	-	lef-5	LEF-5	sf66	ac99
ORF67	61632	62534	+	38k	38K	sf67	ac98
ORF68	62546	63001	+	sf68	ChtB2	sf68	ac150
ORF69	62998	64002	-	bro	BRO	sf69	-
ORF70	64062	64490	-	sf70	Sf70	sf70	-
ORF71	64554	65060	-	pif-4	PIF-4	sf71	ac96
ORF72	65032	68718	+	dnahel	DNA-helicase	sf72	ac95
ORF73	68796	69446	-	odv-e25	ODV-E25	sf73	ac94
ORF74	69443	69922	-	p18	P18	sf74	ac93
ORF75	69924	70682	+	p33	P33	sf75	ac92
ORF76	70726	72105	-	lef-4	LEF-4	sf76	ac90
ORF77	72104	73090	+	vp39	VP39	sf77	ac89
ORF78	73303	74679	+	cg30	CG30	sf78	ac88

Número_ORF	Inicio	Terminación	Polaridad	Gen	Proteína	SfMNPV-3AP2	AcMNPV
ORF79	74763	77207	-	pif-8	PIF-8	sf79	ac83
ORF80	77176	77766	+	tlp20	TLP20	sf80	ac82
ORF81	77573	78307	+	ac81-like	Ac81-like	sf81	ac81
ORF82	78273	79271	+	gp41	GP41	sf82	ac80
ORF83	79296	79643	+	ac78-like	Ac78-like	sf83	ac78
ORF84	79645	80748	+	vlf-1	VLF-1	sf84	ac77
ORF85	81189	81356	-	sf85	Sf85	sf85	-
ORF86	81470	82228	-	p26a	P26a	sf86	-
ORF87	82276	83139	-	iap-0	IAP-2	sf87	ac71
ORF88	83042	83842	-	mtase-1	MTase	sf88	ac69
ORF89	83835	84209	-	pif-6	PIF-6	sf89	ac68
ORF90	84208	85374	+	lef-3	LEF-3	sf90	ac67
ORF91	85436	87574	-	desmoplakin	Desmoplakin	sf91	ac66
ORF92	87576	90695	+	dnapol	DNA-polymerase	sf92	ac65
ORF93	90731	91117	-	sf93	Sf93	sf93	ac75
ORF94	91144	91401	-	sf94	Sf94	sf94	ac76
ORF95	91494	91865	+	sf95	Sf95	sf95	-
ORF96	91840	92340	+	sf96	Sf96	sf96	-
ORF97	92393	93889	-	lef-9	LEF-9	sf97	ac62
ORF98	93948	94535	+	fp25k	FP25K	sf98	ac61

Número_ORF	Inicio	Terminación	Polaridad	Gen	Proteína	SfMNPV-3AP2	AcMNPV
ORF99	94726	96849	+	94k	94K	sf99	ac134
ORF100	96933	97205	+	sf100	Sf100	sf100	ac60
ORF101	97222	97800	+	sf101	Sf101	sf101	ac58/59
ORF102	97793	98278	-	sf102	Sf102	sf102	ac57
ORF103	98513	98890	-	sf103	Sf103	sf103	-
ORF104	98781	99125	-	sf104	Sf104	sf104	ac55
ORF105	99088	100098	-	vp1054	VP1054	sf105	ac54
ORF106	99959	100186	-	lef-10	LEF-10	sf106	ac53a
ORF107	100170	100373	+	sf107	Sf107	sf107	-
ORF108	100391	101290	+	sf108	Sf108	sf108	-
ORF109	101295	101708	-	ac53-like	Ac53-like	sf109	ac53
ORF110	101764	102303	+	sf110	Sf110	sf110	ac52
ORF111	102451	102804	-	sf110a	Sf110a	-	-
ORF112	103206	104066	+	iap-3	IAP-3	sf111	ac27
ORF113	104114	105307	-	bjdp	BJDP	sf112	ac51
ORF114	105328	107973	+	lef-8	LEF-8	sf113	ac50
ORF115	108057	108236	-	sf114	Sf114	sf114	ac43
ORF116	108239	110368	-	odv-e66b	ODV-E66b	sf115	-
ORF117	110398	111597	+	p47	P47	sf116	ac40
ORF118	111680	112306	+	sf117	Sf117	sf117	-

Número_ORF	Inicio	Terminación	Polaridad	Gen	Proteína	SfMNPV-3AP2	AcMNPV
ORF119	112354	112905	+	sf118	Sf118	sf118	-
ORF120	112974	113708	+	adprase	ADPRase	sf119	ac38
ORF121	113603	114004	+	lef-11	LEF-11	sf120	ac37
ORF122	113976	114905	+	39k	39K	sf122	ac36
ORF123	114934	115272	+	sf122	Sf122	sf122	-
ORF124	115330	115533	-	sf123	Sf123	sf123	-
ORF125	115527	115769	-	v-ubi	Ubiquitin	sf124	ac35
ORF126	115812	116348	+	sf125	Sf125	sf125	ac34
ORF127	116760	117161	-	sf126	Sf126	sf126	ac26
ORF128	117255	118247	+	dbp	DBP	sf128	ac25
ORF129	118264	118731	+	lef-6	LEF-6	sf128	ac28
ORF130	118764	119048	-	sf129	Sf129	sf129	-
ORF131	118875	119051	+	sf130	Sf130	sf130	ac29
ORF132	119122	119961	+	p26b	P26b	sf131	ac136
ORF133	119948	120265	+	p10	P10	sf132	ac137
ORF134	120342	121232	-	sf133	Sf133	sf133	-
ORF135	121251	123191	-	pif-0	PIF-0	sf134	ac138
ORF136	123244	123564	+	sf135	Sf135	sf135	-
ORF137	123614	125662	-	ie-1	IE-1	sf136	ac147
ORF138	125692	126336	+	sf137	Sf137	sf137	ac146

Número_ORF	Inicio	Terminación	Polaridad	Gen	Proteína	SfMNPV-3AP2	AcMNPV
ORF139	126373	126651	-	sf138	ChtB1	sf138	ac145
ORF140	126655	127506	-	odv-ec27	ODV-EC27	sf139	ac144
ORF141	127531	127776	-	odv-e18	ODV-E18	sf140	ac143
ORF142	127782	129164	-	p49	P49	sf141	ac142
ORF143	129179	129889	-	ie-0	IE-0	sf142	ac141
ORF144	129976	132492	-	rr-1	RR-1	sf143	-

3.3.2. Diversidad genómica presente en el aislamiento SfMNPV ARG-M

A partir de nuestros datos de secuenciación se detectaron 704 variantes nucleotídicas puntuales intrahospedador (iSNV) y su frecuencia alélica asociada mediante el software LoFreq (Material suplementario 6.2). La distribución global de las iSNV resultó ser homogénea dentro del genoma viral y abarcó un amplio rango de frecuencias alélicas (Figura 3.1 A). La diversidad genómica global de la población de SfMNPV ARG-M fue de 5.3×10^{-3} iSNV por pb. Las iSNV fueron clasificadas en tres grupos de acuerdo a su impacto predicho en secuencias codificantes (no sinónimas, sinónimas e intergénicas). Aproximadamente, la mitad de las iSNVs corresponden a variantes sinónimas (384), mientras que el resto se distribuye entre no sinónimas (184) e intergénicas (131) (Figura 3.1 B). Adicionalmente, se encontraron cuatro variantes que introducen codones de terminación prematuros (Glu357* *sf41*, Glu392* *sf43*, Tyr73* *sf47* y Cys6* *sf110a*) y una variante que modifica un codon de inicio (*sf45*). Como es esperable, las variantes sinónimas presentaron una frecuencia alélica media mayor frente a los otros grupos, probablemente debido a su menor impacto estructural en comparación con las variantes no sinónimas que afectan regiones codificantes o las variantes intergénicas dentro de regiones reguladoras (Figura 3.1 C). Sorprendentemente, 31 % (54 de 184) de las variantes no sinónimas mostraron una frecuencia alélica mayor a 0.25, lo que sugiere podrían coexistir proteoformas alternativas del mismo gen durante la infección por SfMNPV.

Basados en el número de iSNV no sinónimas, los genes de SfMNPV ARG-M fueron ordenados de acuerdo a su diversidad genética. Identificamos *sf29*, *odv-e66a*, *odv-e66b*, *sf23*, *p40*, *pif-0*, *sf110a*, *lef-6*, *lef-7*, *sf68*, *vp80*, *rr1*, *94k* y *p47* como el 10 % de genes más variables. Otra observación que emerge de nuestros datos es que aproximadamente el 57 % (105 de 184) de las iSNV no sinónimas introducen un cambio en la polaridad del aminoácido que afectan, lo que podría llevar a modificaciones en las propiedades biofísicas de las proteínas virales, como ha sido descrito anteriormente ⁽¹¹⁰⁾. Dentro de este grupo de variantes que modifican la polaridad de ciertos aminoácidos, solo 27 afectaron a *core genes* ($n=38$), mientras que las restantes 78 se localizaron dentro de genes *non-core* ($n=106$).

Variantes estructurales genómicas (SV), las cuales comprenden deleciones, inserciones, inversiones, duplicaciones y translocaciones, representan otra fuente de diversidad en poblaciones de baculovirus y otros virus grandes con genoma a DNA.¹⁷³ Para SfMNPV, se han reportado previamente deleciones que afectan el locus *egt* como las principales SV.^{92,174} Decidimos investigar la presencia de otras SV

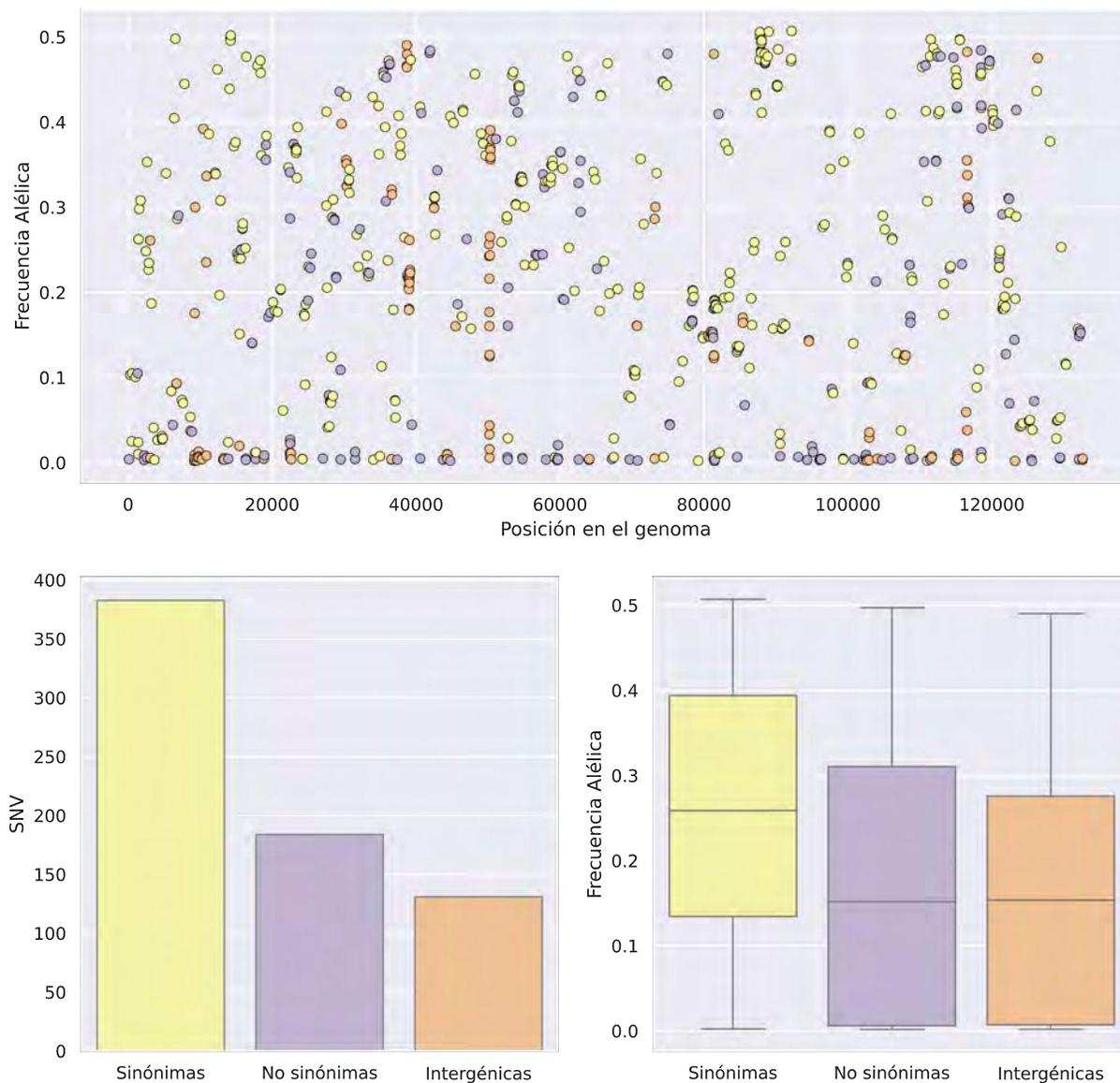


Figura 3.1: Diversidad genómica dentro de la población de SfmNPV ARG-M. **A** Distribución de iSNV sinónimas (amarillo), no sinónimas (violeta) e intergénicas (naranja) a lo largo del genoma. **B** Número de iSNV detectadas para cada categoría. **C** Gráfico de caja mostrando la frecuencia alélicas para cada categoría de iSNV.

en nuestro conjunto de datos con el fin de complementar el cuerpo de conocimiento existente. Los programas Delly y Lumpy lograron detectar 13 SV con precisión de 1 pb (Cuadro 3.2). Recuperamos dos deleciones largas que comprenden la region 22000-26000 y llevan a la eliminación completa del gene *egt*, en acuerdo con los reportes previos. Además, SV no descriptas previamente comprenden dos deleciones largas que eliminan porciones significativas del genoma (36.5 y 54.2 Kpb), cinco deleciones pequeñas (afectando a *f-protein*, *pif-1*, *odv-e66a* y *cg30*), una inversión de 52.8 Kpb y tres duplicaciones (4.3, 5.0 y 9.8 Kpb). Utilizando el soporte de lecturas para cada SV y la cobertura media de nuestros datos, se realizó una estimación de la frecuencia por genoma de cada SV, de manera similar al enfoque propuesto previamente.¹⁷⁵ Las duplicaciones y deleciones presentaron una frecuencia en el rango de 1.9×10^{-5} - 1.7×10^{-3} , pero las deleciones afectando los genes *egt* y *odv-e66a*, junto con el evento de inversión, presentaron frecuencias superiores (2.3×10^{-3} - 1.2×10^{-2}).

Cuadro 3.2: Variantes estructurales detectadas con Delly (D) y Lumpy (L)

Clase_SV	Pos_D	Pos_L	Tamaño_D	Tamaño_L	Freq_D	Freq_L	Lecturas_D	Lecturas_L
Delección	6505	6506	36498	36496	0.0003	0.00044	16	23
Delección	9548	9562	324	309	0.00069	0.00019	36	10
Delección	13603	13603	54189	54188	0.0013	0.00051	69	30
Delección	12600	12609	316	306	0.00022	0.00011	12	6
Delección	21336	21336	4756	4755	0.0089	0.0035	465	181
Delección	24401	24402	1691	1689	0.012	0.0056	639	293
Delección	33629	33639	466	455	0.0017	0.0008	89	42
Inversión	50168	50191	52832	52808	0.16	0.0089	8610	467
Delección	53651	53666	529	513	0.0066	0.0023	345	121
Delección	73264	73275	403	391	0.0021	0.00076	110	40
Duplicación	80710	80689	9826	9853	0.000095	0.00059	5	31
Duplicación	104708	104707	4330	4331	0.00019	0.0004	10	21
Duplicación	121445	121444	5001	5007	0.000095	0.00038	5	20

3.3.3. Diversidad genética dentro de los aislamientos geográficos de SfmNPV

Cinco secuencias genómicas derivadas de distintos aislamientos geográficos de SfmNPV se encuentran disponibles a la fecha. Los genomas B y G del aislamiento de Nicaragua presentan características distintivas con respecto a su contenido génico en comparación con el genoma de los aislamientos 3AP2 y 19.¹⁶⁵ Basados en estos datos preliminares, decidimos comparar las características genómicas de ARG-M con el resto de los aislamientos. Existen tres ORFs reportados solo en NicB y NicG, que se encuentran ausentes en 3AP2 y 19, estos son *sf39a*, *sf57a* y *sf110a*. De estos, encontramos solo a *sf110a* presente en ARG-M. *Sf39a* codifica para una proteína de 63 amino ácidos de longitud en NicB y NicG. Búsquedas de similitud de secuencias por Tblastn revelaron que ARG-M tiene un gen homólogo más corto de 33 amino ácidos, por lo que no fue anotado. *Sf57a* se localiza dentro de una inserción reportada para NicB, la cual está ausente en 3AP2 y 19,¹⁶⁵ así como en ARG-M. Por otra parte, ARG-M codifica *sf129*, un gen que es exclusivo de 3AP2 y está ausente en los genomas de Nicaragua y el 19. Entonces, ARG-M presenta una composición de genes diferencial en comparación con otros genomas de SfmNPV.

Se construyó un alineamiento múltiple de genoma completo (MSA) para explorar la diversidad genómica subyacente entre los diferentes aislamientos de SfmNPV. La reconstrucción filogenética resultante de este MSA demostró que existe un elevado grado de similitud entre los aislamientos de SfmNPV, con distancias evolutivas en el orden de 1×10^{-3} sustituciones por pb. En este contexto, el aislamiento ARG-M resultó ser más similar a los aislamiento de Nicaragua, mientras que los aislamientos 3AP2 y 19 se agruparon juntos, lo que deja al aislamiento colombiano como el genoma más divergente (Figura 3.2 A). Basándonos en una metodología reportada previamente,¹¹³ donde los polimorfismos fueron utilizados para agrupar distintos aislamientos de *Cydia pomonella* Granulovirus (CpGV), extrajimos 1502 SNP presentes en el MSA y analizamos su distribución a lo largo del genoma de SfmNPV. Se evidenció una distribución homogénea con aproximadamente 100 SNP por cada intervalo de 10.34 Kpb, exceptuando la región 20000-30000 donde se apreció un pico (Figura 3.2 B). A partir de investigaciones subsecuentes, este número elevado de SNP podría explicarse por la presencia de un segmento genómico altamente divergente de *Spodoptera litura* II Nucleopolyhedrovirus (SplNPV-II), el cual se introdujo a través de un evento de transferencia génica horizontal en el genoma del aislamiento colombiano.⁹⁴ También se observó un número

apreciable de SNP específicos de aislamiento: 577 para el colombiano, 210 para el 19, 162 para el 3AP2, 87 para ARG-M y 15 para NicG (Figura 3.2 C). Aunque encontramos 93 SNP compartidas por los aislamientos NicB y NicG, no pudimos identificar SNP específicos para NicB. En general, se observó que los SNP compartidos entre distintos aislamientos reflejan la información obtenida mediante la reconstrucción filogenética: los aislamientos ARG-M, NicB y NicG comparten 92 SNP, mientras que los aislamientos 19+COL y 19+3AP2 tienen 59 y 56 SNP en común, respectivamente.

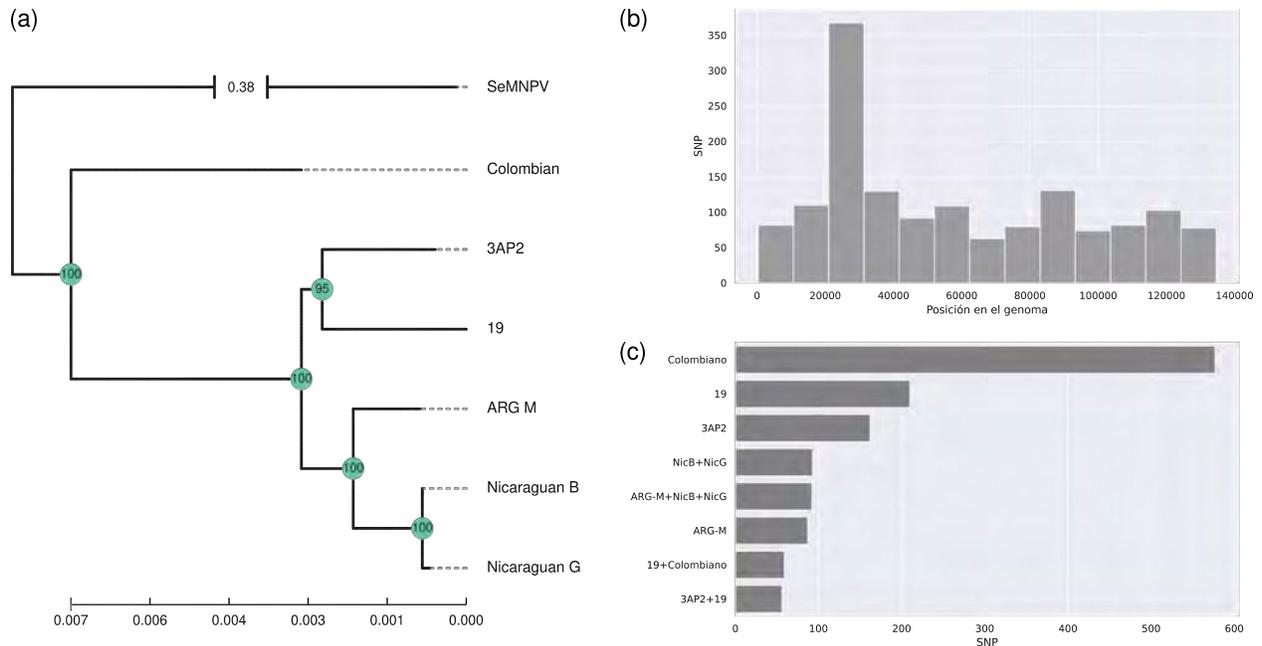


Figura 3.2: Diversidad genética presente en los aislamientos geográficos de SfMNPV. A Filogenia de máxima verosimilitud construida a partir de un alineamiento de genoma completo. Los valores de soporte se calcularon utilizando 1000 réplicas de ultrafast bootstrap. *Spodoptera exigua* Multiple Nucleopolyhedrovirus (SeMNPV) se utilizó como grupo externo. **B** Histograma mostrando el número de SNP por ventana detectados dentro del alineamiento. El alineamiento se dividió en 13 ventanas, lo que corresponde aproximadamente a 10.38 Kpb por ventana. **C** Gráfico de barras mostrando el número de SNP detectados para cada aislamiento y las combinaciones detectadas.

Dado que estábamos interesados en estudiar qué secuencias codificantes poseen la mayor variabilidad, se analizaron los niveles de mutaciones no sinónimas presentes en cada región codificante a nivel intra- e inter-aislamiento (Figura 3.3). En general, notamos que las iSNV y los SNP tienden a afectar las mismas posiciones dentro de las regiones codificantes. La mayor parte de los genes de SfMNPV mostraron niveles bajos de variación, en promedio menos de 3×10^{-3} mutaciones no sinónimas por pb. Sin embargo, 15 genes presentaron niveles de diversidad por encima de 3×10^{-3} mutaciones no sinónimas por pb a nivel intra- e inter-aislamiento (*lef-7*, *sf23*, *sf29*, *odv-e66a*,

p40, *sf68*, *gp41*, *sf85*, *sf110a*, *sf118*, *sf122*, *ubiquitin*, *sf126*, *lef-6* y *sf135*). Varios de esos genes son coincidentes con nuestro ranking de diversidad basado en el número iSNV no sinónimas.

3.3.4. Evolución molecular de *sf29* en virus grandes de insectos

Basados en los resultados previos donde *sf29* exhibió niveles de diversidad elevados, decidimos evaluar la posibilidad de que gen se encuentre bajo un régimen de selección adaptativa. Esta proteína ha sido descrita como un factor viral involucrado en el número de viriones presentes dentro de la partícula del OB y se distribuye en los alphabaculovirus de grupo II y entomopoxvirus.¹⁷⁶ Una filogenia de máxima verosimilitud construida para *sf29* sugiere un posible origen bacteriano de esta proteína y un evento de transferencia posterior hacia un virus de insecto ancestral antes de pasar hacia los baculovirus y entomopoxvirus (Figura 3.4 A). Mas aún, empleando el programa aBSREL se encontró evidencia de selección diversificadora actuando en varias ramas de entomopoxvirus (MySEV, AHEV y dos especies de AMEV, todas marcadas en naranja en la filogenia), lo que sugiere un régimen sostenido de selección positiva dentro de este clado. Las ramas de los baculovirus PespNPV y MyunNPV también mostraron evidencia de selección diversificadora.

El análisis de la secuencia peptídica de Sf29 reveló la presencia de un dominio Peptidase M9 (collagenase, residuos 299-437) y una región desordenada (760-800) (Figura 3.4 B). Adicionalmente se detectaron un péptido señal y una región transmembrana C-terminal, lo que sugiere que esta podría ser una proteína de membrana. El test de selección positiva realizado con FEL mostró evidencia de selección positiva en cinco residuos (D225, N696, Y575, I818 y I827). Estos residuos bajo selección positiva y las iSNV detectadas en la población de SfMNPV ARG-M se localizan principalmente por fuera del dominio peptidasa, lo que sugiere que el proceso de evolución adaptativa de *sf29* está dirigido por sus regiones no catalíticas. Mas aún, varios residuos bajo selección positiva episódica, evidenciado a través del método MEME, fueron coincidentes con las iSNV presentes en SfMNPV ARG-M. El análisis de los residuos presentes en el sitio activo putativo demostró que el motivo HEXXH, involucrado en la unión de Zinc y necesario para el funcionamiento de las metaloproteasas,¹⁷⁷ se encuentra conservado. Esto indicaría que *sf29* podría codificar una peptidasa funcional (Figura 3.4 C).

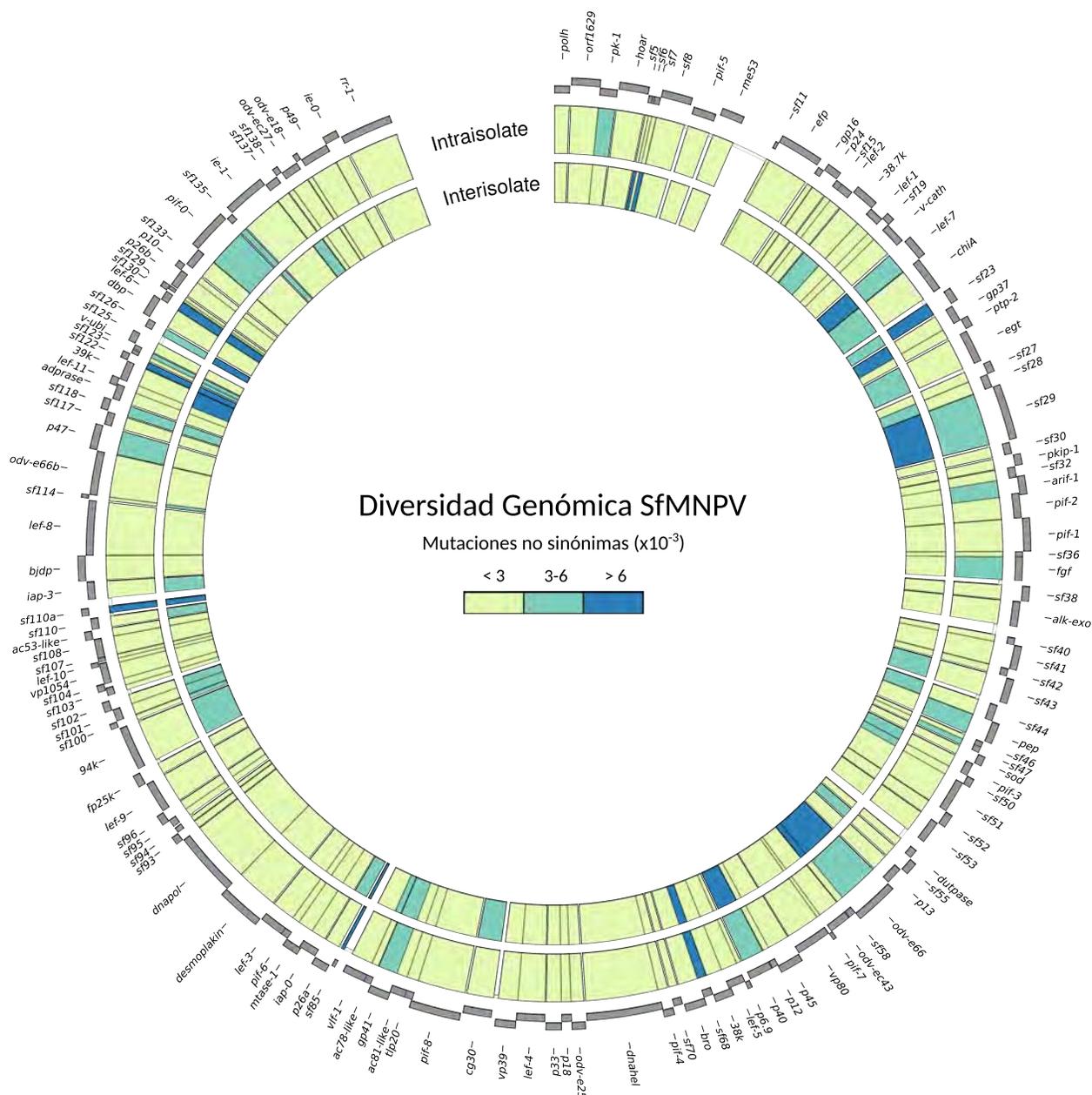


Figura 3.3: Diversidad dentro del proteoma de SfmNPV. Para cada región codificantes en el genoma de SfmNPV, el número de iSNV y SNP fue calculado. El círculo interno representa los SNP derivados del alineamiento de genoma completo de todos los aislamientos de SfmNPV, mientras que el círculo externo corresponde a las iSNV detectadas dentro de la población de SfmNPV ARG-M.

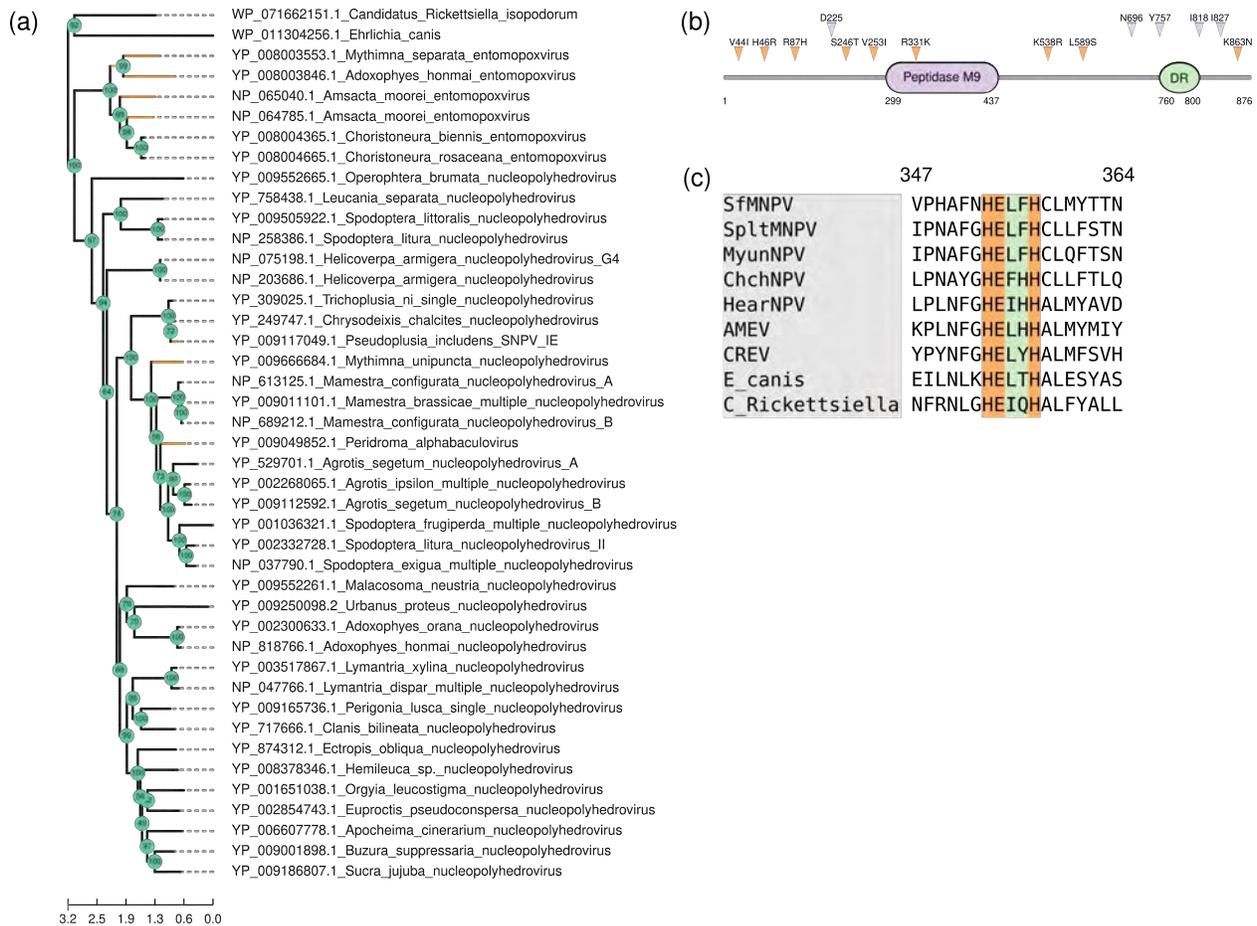


Figura 3.4: Evolución molecular de *sf29*. **A** Filogenia de máxima verosimilitud de *sf29*. Los valores de soporte se calcularon a partir de 1000 replicas de ultrafast bootstrap. Las secuencias ortólogas derivadas de las bacterias *E. canis* y *C. rickettsiella* fueron incluidas como grupos externos. Las ramas de la filogenia que presentaron evidencia de selección positiva de acuerdo al método aBSREL se encuentran marcadas en naranja. **B** Análisis de secuencia de *sf29*. Los sitios bajo selección positiva se encuentran marcados en azul, mientras que las iSNV se muestran en naranja. El dominio peptidasa M9 se muestra en violeta y la región desordenada en verde. **C** Conservación de secuencia del motivo peptidasa del sitio activo (HEXXH), extraído del alineamiento de secuencias de *sf29*.

3.4. Discusión

Cada vez es más reconocida la relevancia de la diversidad genómica como un determinante clave de la dinámica evolutiva de los virus y su virulencia,¹⁷⁸ especialmente para el caso de los virus grandes con genomas a DNA.¹⁰⁹ Este trabajo presenta una foto de la diversidad genómica subyacente dentro de un aislamiento natural de SfMNPV, expandiendo el conocimiento provisto por estudios previos donde se reportaron las secuencias consenso de diversos aislamientos geográficos.^{90,94,164,165}

La diversidad genética encontrada dentro de SfMNPV ARG-M (5.3×10^{-3} iSNV por pb) es similar a los reportes previos de otros baculovirus.^{31,111,112,114} Dado que las variantes no sinónimas tienen el potencial para introducir cambios drásticos en la estructura y las propiedades biofísicas de las proteínas, en este trabajo se focalizó mucho en el análisis de este tipo de mutaciones. Ambas, iSNV y SNP, señalan a un conjunto común de genes (Figura 3.3) que concentran la mayor cantidad de los sitios variables. La mayoría de esos genes (*sf23*, *sf85*, *sf110a*, *sf118*, *sf122*, *sf126* y *sf135*) codifican proteínas con funciones desconocidas, lo que constituye una gran fuente de adaptabilidad en SfMNPV que permanece inexplorada. El examen de esos genes con la herramienta InterProScan solo permitió encontrar una región desordenada dentro de *sf23* (residuos 180-209) y un dominio de función desconocida dentro de *sf126* (DUF424, e-value $9,5 \times 10^{-20}$). Además, dos genes asociados con la infectividad oral (*odv-e66a* y probablemente *sf29*) y *sf68*, el cual afecta la velocidad de muerte y la producción de OB,¹⁷⁹ también presentaron una elevada diversidad genética. *Odv-e66a* codifica una condroitín liasa que degrada la membrana peritrófica, mientras que la proteína derivada de *sf68* contiene un dominio de unión a quitina (Superfamily SSF57625, e-value 5.49×10^{-5}) y *sf29* posee un posible dominio colagenasa. En este escenario, es posible que la elevada diversidad genética de estos genes virales contribuya a la adaptación de SfMNPV al ambiente hostil impuesto por el intestino de su hospedador.^{180,181}

Las deleciones localizadas en el locus *egt* son comunes en las poblaciones naturales de SfMNPV y han sido relacionadas con la velocidad de muerte y potencia de los OBs.¹⁶⁷ Sin embargo, hay poca información acerca de la ocurrencia de SV a lo largo del genoma. Nuestros datos avalan la presencia de varias SV, abarcando deleciones, duplicaciones y una inversión. Una de las SV detectadas de forma más frecuente es la deleción del gen *egt*, una observación que podría estar relacionada con

la detección sostenida de esta delección en poblaciones naturales. Otro punto que surge de las SV descubiertas para SfMNPV es que las delecciones se localizan aproximadamente dentro de la primera porción de la secuencia genómica de SfMNPV (posición 6505 a 73275), mientras que los eventos de duplicación afectaron principalmente la segunda mitad del genoma de SfMNPV (posición 80710 a 121444). A través de la inspección del contenido génico de cada una de esas dos regiones se evidenció que los eventos de delección conducen a la pérdida de genes asociados con la infectividad oral, funciones auxiliares y funciones desconocidas, mientras que las duplicaciones llevan a un incremento en genes relacionados con funciones replicativas/transcripcional (*ie-1*, *dnapol*, *lef-3* y *lef-8*). La localización diferencial de estas delecciones y duplicaciones soporta la hipótesis de que el proceso de multiplicación viral podría verse acelerado gracias a la presencia de genomas de menor tamaño que a su vez se encuentran enriquecidos en genes replicativos, como se ha propuesto previamente.^{182,183}

Motivados por la presencia de un dominio Peptidase M9 dentro de *sf29* y por su elevada diversidad genética, se llevó a cabo un estudio evolutivo detallado de este gen para tratar de comprender mejor su posible función. La historia evolutiva de *sf29* ejemplifica el flujo génico recurrente entre baculovirus y otros organismos,^{184,185} en este caso particular, con entomopoxvirus y probablemente procariontes. Además de su dominio Peptidase M9, se demostró la conservación del motivo HEXXH característico de las metaloproteasas,¹⁷⁷ lo que refuerza la hipótesis de que *sf29* codifica una colagenasa funcional. Estudios bioquímicos complementarios son requeridos para confirmar esta hipótesis. A partir de diferentes métodos se evidenció la presencia de selección positiva actuando sobre la secuencia de *sf29*. Una posible explicación de este régimen de evolución adaptativa podría relacionarse con el cambio en el entorno celular de esta proteína luego de su captura por un virus ancestral de insectos. La incorporación de un número elevado de mutaciones en esta proteína podría haber facilitado la aparición de interacciones con las biomoléculas presentes en su nuevo contexto celular.¹⁸⁰

Los polimorfismos exhibidos por los distintos aislamientos de SfMNPV representan una herramienta útil para catalogarlos y asegurar su eficacia a campo. El genotipado de los aislamientos permite manejar la aparición de eventos de resistencia que son aislamiento-específicos, como se demostró para CpGV.³¹ A pesar de que el número de aislamientos de SfMNPV todavía es bajo, algunas generalidades pueden extraerse a partir de su patrón de SNP. Encontramos que la mayoría de los

SNP presentes en SfMNPV son bialélicos y específicos para solo un aislamiento, lo que sugiere que estos SNP podrían utilizarse como herramientas de genotipado. Recientemente, se ha empleado la variación genómica entre distintos aislamientos de nudivirus con títulos virales diferentes para inferir posiciones en el genoma que están directamente asociadas con la virulencia,¹⁸⁶ un resultado que resalta la utilidad de la información genómica para predecir propiedades infectivas. Con el mayor acceso a tecnologías de secuenciación de alto rendimiento y la disponibilidad de herramientas computacionales para analizar los datos generados, creemos que la secuenciación de nuevos aislamientos de SfMNPV a través de diferentes escalas temporales y geográficas nos permitirán comprender de manera más precisa los determinantes moleculares de la patogenicidad de este virus. Este trabajo provee una primera aproximación a la diversidad genómica subyacente en las poblaciones de SfMNPV y sus posibles consecuencias en la adaptación viral a la oruga militar tardía.

3.5. Materiales y Métodos

3.5.1. Aislamiento del virus y secuenciación del genoma

El aislamiento argentino de *Spodoptera frugiperda* Multiple Nucleopolyhedrovirus (SfMNPV ARG-M) se obtuvo a partir de una larva de *S. frugiperda* infectada que fue recolectada en un cultivo de maíz en Oliveros (provincia de Santa Fe) en 1985 y posteriormente fue purificada en las instalaciones del Instituto de Microbiología y Zoología Agrícola (IMyZA), Instituto Nacional de Tecnología Agrípecuaria (INTA), Castelar, Argentina.¹⁷² Para la amplificación de este aislamiento viral, una colonia de laboratorio de *S. frugiperda* se estableció en el Instituto de Biotecnología y Biología Molecular (IBBM), La Plata, Argentina. El stock viral fue amplificado en larvas de 3 días alimentadas con dieta artificial contaminada con virus y mantenidas individualmente a 26°C con un fotoperíodo de 16:8 horas (luz:oscuridad) hasta su muerte o pupación. Las larvas muertas fueron homogeneizadas en agua doblemente destilada (ddH₂O), filtradas a través de gasa y resuspendidas en una solución de SDS 0.1 % v/v. Los restos celulares se eliminaron por centrifugación a 1000 g por 2 minutos y el sobrenadante conteniendo los OB fue cargado en un gradiente discontinuo de sacarosa 35-60 % p/p para luego ser centrifugado a 20000 rpm por 1 hora a 4°C en un rotor Beckman SW41. La banda opalescente conteniendo los OB fue extraída, lavada dos veces con ddH₂O y el pellet viral final fue resuspendido en ddH₂O. Los OB purificados se disolvieron por lisis alcalina y la extracción del DNA viral se realizó como se describió previamente.⁹¹ El DNA genómico correspondiente a SfMNPV-ARG M fue utilizado para preparar una biblioteca de secuenciación (TruSeq DNA PCR-Free kit) y posteriormente secuenciado en una plataforma Illumina HiSeq 4000 a cargo de Macrogen Corporation, South Korea. El conjunto de datos crudos corresponde a 70239480 lecturas, las cuales fueron depositadas en la base de datos *Sequence Read Archive* (SRA), perteneciente al NCBI, bajo el número de acceso SRX8829998.

3.5.2. Ensamblado del genoma y anotación

Debido a que el procesamiento de las lecturas de secuenciación requiere el uso de varias piezas de software diferentes y la conexión de diversos comandos ejecutados en una terminal (por ej. Bash), optamos por desarrollar un pipeline para el ensamblado del genoma basado en la herramienta Snakemake con el fin de asegurar la consistencia y reproducibilidad de nuestro análisis.¹⁸⁷ El chequeo

de la calidad de las lecturas crudas y la eliminación de las secuencias adaptadores presentes en las mismas se llevo a cabo con fastp utilizando los parámetros por defecto.¹⁸⁸ Para acelerar el proceso de ensamblaje, tomamos como muestra representativa 2 millones de lecturas al azar de nuestro conjunto de datos originales con la herramienta seqtk. El ensamblado del genoma se realizó con Megahit¹⁸⁹ y se seleccionó el *scaffold* más largo para su refinamiento posterior. Una serie de scripts fue utilizada para circularizar esta molécula y fijar el codon ATG de inicio del gen *polyhedrin* como la posición inicial de la secuencia genómica. Los marcos de lectura abiertos (ORFs) putativos fueron predichos con la herramienta ORFinder, con los siguientes parámetros: -c t -s 0 -ml 150.¹⁹⁰ La anotación de las secuencias codificantes se llevo a cabo utilizando la herramienta InterProScan¹⁹¹ y a través de búsquedas de homología con BLASTp¹³⁹ contra todas las secuencias genómica disponibles para SfMNPV, seguidas de la recuperación de la información de anotación con el software Entrez Direct utilities.¹⁹⁰ Los datos correspondientes a la secuencia genómica han sido depositado en la base de datos Genbank bajo el número de acceso MW162628.

3.5.3. Asignación de variantes nucleotídicas puntuales y estructurales

Todas las lecturas derivadas del conjunto de datos completo fueron procesadas con fastp.¹⁸⁸ Las lecturas que pasaron el filtro de calidad mínima fueron alineadas contra el genoma de SfMNPV-ARG M utilizando el algoritmo BWA-MEM (¹⁹²) y posteriormente ordenadas de acuerdo a sus coordenadas con SAMtools (¹⁹³). Lofreq2 fue utilizado para asignar las variantes nucleotídicas puntuales (iSNV)¹⁹⁴ y el efecto de cada variante se evaluó mediante SnpEff,¹⁹⁵ para lo cual utilizamos nuestra anotación genómica previamente descripta. Delly¹⁹⁶ y Lumpy¹⁹⁷ fueron invocados con los parámetros por defecto para detectar evidencia de variantes estructurales (SV) presentes dentro de nuestro conjunto de lecturas mapeadas.

3.5.4. Análisis filogenético

Las secuencias genómicas y proteómicas disponibles para los distintos aislamientos de SfMNPV fueron descargadas desde las bases de datos de nucleótidos y proteínas de NCBI, respectivamente. Se construyó un alineamiento de genoma completo con MAFFT¹⁹⁸ y los polimorfismos puntuales de nucleótidos fueron extraídos desde este alineamiento con SNP-sites.¹⁹⁹ A partir del MSA de genomas completos se infirió una filogenia de máxima verosimilitud utilizando IQ-TREE con parámetros

por defecto. Para la selección del modelo evolutivo se empleó ModelFinder y para la estimación del valor de soporte de cada nodo en la filogenia se calcularon 1000 replicas utilizando ultrafast bootstrap.²⁰⁰⁻²⁰² La manipulación, coloreado y visualización del árbol filogenético se llevo a cabo con ToyTree²⁰³ e Inkscape. Las secuencias proteicas se agruparon utilizando BLASTp,¹³⁹ se alinearon con MAFFT¹⁹⁸ y luego se utilizaron para extraer las mutaciones no sinónimas dentro de las secuencias codificantes utilizadas en el gráfico Circos.

3.5.5. Estudio de la evolución molecular de *sf29*

Se emplearon búsquedas de homología por BLASTp¹³⁹ para identificar ortólogos de *sf29* presentes en baculovirus y entomopoxvirus. Se incluyeron secuencias ortólogas de origen bacteriano como grupos externos para el análisis filogenético. Se recolectaron las secuencias proteicas correspondientes a cada ortólogo y se emplearon para construir un alineamiento múltiple y un árbol filogenético, como se describió en la sección anterior. Se llevó a cabo el análisis de evolución molecular con alineamientos de codones empleando el paquete HyPhy.²⁰⁴ Chequeamos la ausencia de eventos de recombinación en nuestro alineamiento con GARD.²⁰⁵ Para la detección de selección diversificadora pervasiva o episódica en cada sitio del alineamiento de codones utilizamos FEL²⁰⁶ y MEME,²⁰⁷ respectivamente. Estudiamos la evidencia de selección diversificadora en las distintas ramas de nuestra filogenia con aBSREL.²⁰⁸ Debido a que algunos de los análisis evolutivos mencionados anteriormente son costosos computacionalmente, en algunos casos utilizamos el servidor Datamonkey.²⁰⁹ La predicción de secuencias desordenadas se realizó con IUPRED2²¹⁰ y para visualizar los alineamientos de secuencias empleamos Aliview.²¹¹

3.5.6. Disponibilidad del código y los datos

El conjunto de las lecturas crudas y la secuencia genómica ensamblada han sido depositados en las bases de datos SRA y Genbank de NCBI, respectivamente. Todo el código utilizado para procesar los datos y generar las figuras se encuentra disponible de forma libre en Github tomasMasson/sfmnpv_genomics. Los archivos generados por HyPhy pueden ser visualizados de forma interactiva utilizando la herramienta web HyPhy Vision.

Capítulo 4

Estudio de la replicación de SfGV en cultivo celular

4.1. Introducción

El desarrollo de sistemas de cultivo *in vitro* ha sido fundamental para el estudio y la comprensión del proceso infectivo de los baculovirus a nivel celular. En primer lugar, la posibilidad de reproducir el ciclo replicativo viral en condiciones controladas de laboratorio permite diseñar y probar distintas perturbaciones para conocer su funcionamiento a nivel molecular.^{212,213} Por otra parte, las herramientas de cultivo celular facilitan enormemente la aplicación de técnicas clásicas de ingeniería genética y virología para el estudio de baculovirus mutantes.^{214,215} En particular, el estudio de mutantes que portan deleciones y los correspondientes virus reparados con una nueva copia del gen, la cual ha sido insertada en un locus distinto al original, representan el enfoque genético más empleado a la fecha para la caracterización de la función molecular de genes baculovirales no anotados.^{74,216–218} A nivel biotecnológico, los cultivos celulares también representa una alternativa atractiva debido a su compatibilidad natural con procesos de escalado y estandarización industriales.^{219–222}

Para el caso de varios granulovirus (AgseGV, CpGV, PhopGV, PrGV y TnGV), se ha experimentado en el pasado con diversas líneas celulares para su producción *in vitro*. Para el caso de TnGV, se exploraron líneas celulares embrionarias que inicialmente permitieron la replicación de TnGV, pero luego de 20-25 pasajes perdieron su capacidad de soportar el ciclo infectivo de este granulovirus.^{223,224} También se han empleado líneas celulares derivadas de *Cydia pomonella* y de *Phthorimaea operculella* para la amplificación de CpGV y PhopGV, respectivamente. Sin embargo, estas líneas corresponden a cultivos primarios no inmortalizados que han sido repicados solo por 30 pasajes, lo que no garantiza que su susceptibilidad sea sostenible al aumentar el número de pasajes.²²⁵ Existen casos menos prometedores también, como el de las líneas celulares embrionarias derivadas de los huevos de *Pieris rapae*, las cuales no son capaces de soportar la replicación de PrGV ni siquiera durante los pasajes iniciales del cultivo,²²⁶ o la línea celular derivada de *Agrotis ipsilon*, la cual es permisiva para varios nucleopoliedrovirus, pero no soporta la replicación de AgseGV.³¹ Estos resultados demuestran que el desarrollo de líneas celulares susceptibles a granulovirus es un proceso dificultoso y representa un área vacante. Además, a esto se suma el hecho de que los títulos virales obtenidos para granulovirus son comparativamente más bajos (1×10^5 TCID₅₀/ml) con respecto a los alcanzados por varios nucleopoliedrovirus ($1 \times 10^{7-8}$ TCID₅₀/ml)²²⁷ y requiere cultivar las células a temperaturas más bajas (20°C) para asegurar que mantengan su susceptibilidad.^{224,227}

En un reporte reciente, se confirmó la presencia de actividad transcripcional tardía y producción de BV para el granulovirus PlxyGV en tres líneas celulares, entre ellas Sf9 y High Five, dos de las líneas comerciales de insecto más ampliamente utilizadas.²²⁸ Para esto, los investigadores construyeron un bÁcmido de PlxyGV capaz de expresar los genes *ie-1*, *gp64* y *p35* derivados de AcMNPV. El razonamiento para la selecci3n de estos genes se bas3 en su rol central en distintos procesos moleculares esenciales del ciclo infectivo de los baculovirus. Dentro del grupo de los nucleopoliedrovirus, el producto del gen inmediatamente temprano 1 (IE-1) desempeña un papel central en la regulaci3n de la expresi3n viral temprana y en el proceso de replicaci3n gen3mica,⁸⁹ mientras que P35 actúa como supresor del proceso apopt3tico disparado por la infecci3n viral²²⁹ y GP64 es la prote3na de fusi3n involucrada en la entrada de los BVs a la c3lula, aunque en los granulovirus esta prote3na es reemplazada por la prote3na de fusi3n de la envoltura, EFP,²³⁰ con la excepci3n de DisaGV, el cual codifica una copia del gen *gp64*.²³¹ Es posible que las prote3nas IE-1, GP64 y P35 codificados por AcMNPV sean mÁs eficientes al momento de suprimir la respuesta antiviral por parte del hospedador y disparar la subsecuente cascada de expresi3n viral, en comparaci3n con sus hom3logos de granulovirus.

Nuestro laboratorio tiene experiencia previa en el estudio de la expresi3n ge3nica de granulovirus en cultivos celulares y en el desarrollo de herramientas de ingenier3a gen3tica para la generaci3n de baculovirus recombinantes. Previamente, se analiz3 la actividad transcripcional de distintos promotores virales derivados de EpapGV en l3neas celulares comerciales y en un contexto de infecci3n por AgMNPV.²³² Por otra parte, tambi3n se profundiz3 la caracterizaci3n de la actividad transcripcional de EpapGV y AgMNPV, y se desarroll3 un sistema de recombinaci3n novedoso para facilitar la modificaci3n gen3tica de AgMNPV.⁸⁰ Teniendo en cuenta estos reportes previos y la experiencia de nuestro grupo en el estudio de baculovirus en sistemas *in vitro*, decidimos iniciar la caracterizaci3n del proceso replicativo de SfGV en cultivo celular. Para esto, tomamos provecho de la l3nea celular Sf9, la cual ha sido derivada de tejido ovÁrico de *S. frugiperda*, el hospedador natural de SfGV. En primer lugar, se llevaron a cabo transfecciones con DNA gen3mico de SfGV para evaluar su efecto citopÁtico en cultivo celular. Luego, se analiz3 la expresi3n de genes de SfGV pertenecientes a diferentes etapas del ciclo infectivo en c3lulas transfectadas. Por 3ltimo, buscamos desarrollar nuevas herramientas que faciliten el trabajo con SfGV en sistemas de cultivo celular.

Para esto, se inició la puesta a punto de un protocolo de PCR cuantitativa que permita obtener medidas cuantitativas de los niveles de DNA genómico en distintas etapas del ciclo viral. Además, se construyeron plásmidos de expresión en insecto con los genes *ie-1*, *gp64* y *p35*, las cuales han sido descritas previamente como provirales, y serán utilizados a futuro en ensayos celulares para sinergizar la replicación de SfGV.

4.2. Resultados

4.2.1. Transfección de células Sf9 con DNA genómico de baculovirus

Con el fin de evaluar la capacidad infectiva de SfGV en la línea celular Sf9, la cual ha sido derivada de su hospedador natural *S. frugiperda*, decidimos llevar a cabo una serie de transfecciones con DNA genómico viral. Como control positivo de la metodología de transfección empleamos DNA genómico de un virus AcMNPV marcado con una proteína roja fluorescente (DsRed), el cual es altamente infectivo para las células Sf9 y puede visualizarse fácilmente mediante microscopía de fluorescencia. La extracción del DNA genómico se llevó a cabo a partir de OB; para el caso de SfGV se obtuvieron a partir de larvas infectadas, mientras que para AcMNPV-DsRed fueron recuperados a partir de cultivos celulares infectados en etapas tardías. Estos OB fueron disgregados mediante una lisis alcalina (Figura 4.1 A), seguida de una digestión de los viriones liberados con proteinasa K, una extracción del DNA viral con fenol-cloroformo y precipitación con alcohol.²³³ Una vez resuspendido el DNA viral en ddH₂O, se confirmó su pureza y se cuantificó su concentración a través del perfil de absorbancia a 260 nm en Nanodrop. Como control de calidad adicional, se evaluó su integridad física mediante electroforesis en gel de agarosa, donde se observó la presencia de bandas de elevado peso molecular tanto para la muestra de AcMNPV-DsRed como para SfGV (Figura 4.1 B). Una vez obtenido el DNA genómico viral, procedimos a ensayar dos condiciones de transfección diferentes para nuestro control AcMNPV-DsRed, una empleando el reactivo comercial Cellfectin y otra con el polímero catiónico polietilenimina (PEI). En ambos casos observamos la presencia de fluorescencia roja a partir del cuarto día post-transfección, sin embargo, la eficiencia de transfección comparada al séptimo día post-transfección resultó mayor para el tratamiento con Cellfectin, basados en el número de células fluorescentes observadas (Figura 4.2 A). Apoyados en estos datos, decidimos utilizar Cellfectin en los ensayos subsecuentes. Es interesante destacar que en los tratamientos con

DNA de AcMNPV-DsRed se observó la formación de OB entre los 3-5 días post-transfección, lo que demuestra que el ciclo infeccioso se desarrolla completamente (Figura 4.2 B).

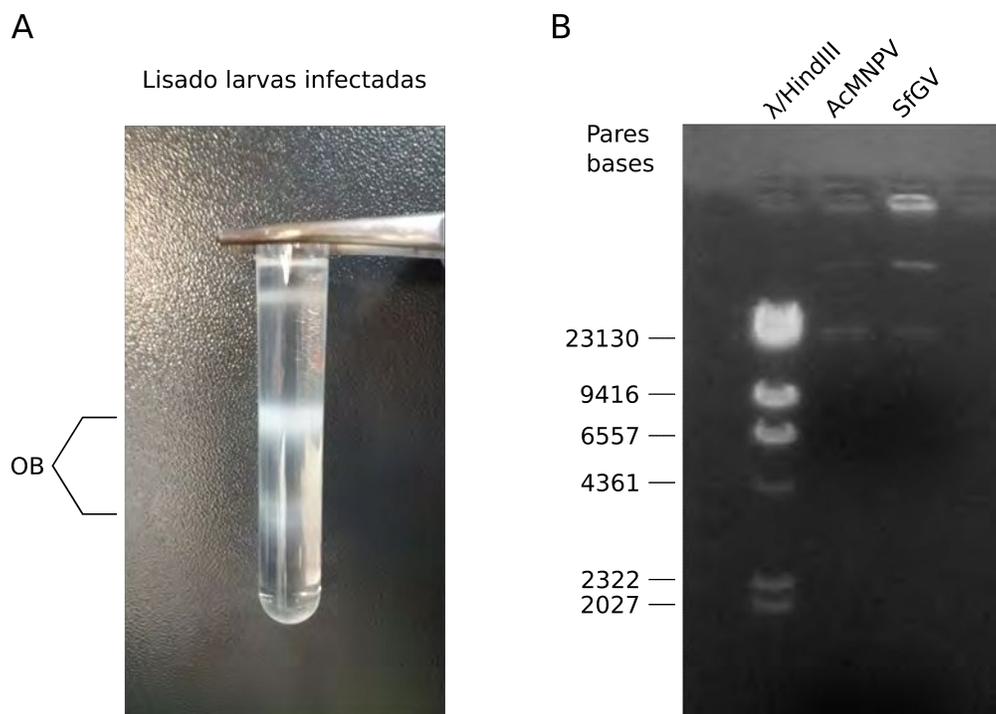


Figura 4.1: Preparación de DNA genómico baculoviral para los ensayos de transfección en células Sf9. **A** Bandas opalescentes correspondiente a la fracción de OB purificados de SfGV, obtenidos mediante ultracentrifugación. **B** Electroforesis en gel de agarosa (0.6% p/v) del DNA genómico viral purificado de AcMNPV-DsRed y SfGV. El marcador de peso molecular corresponde a DNA Lambda digerido con la enzima *HindIII*.

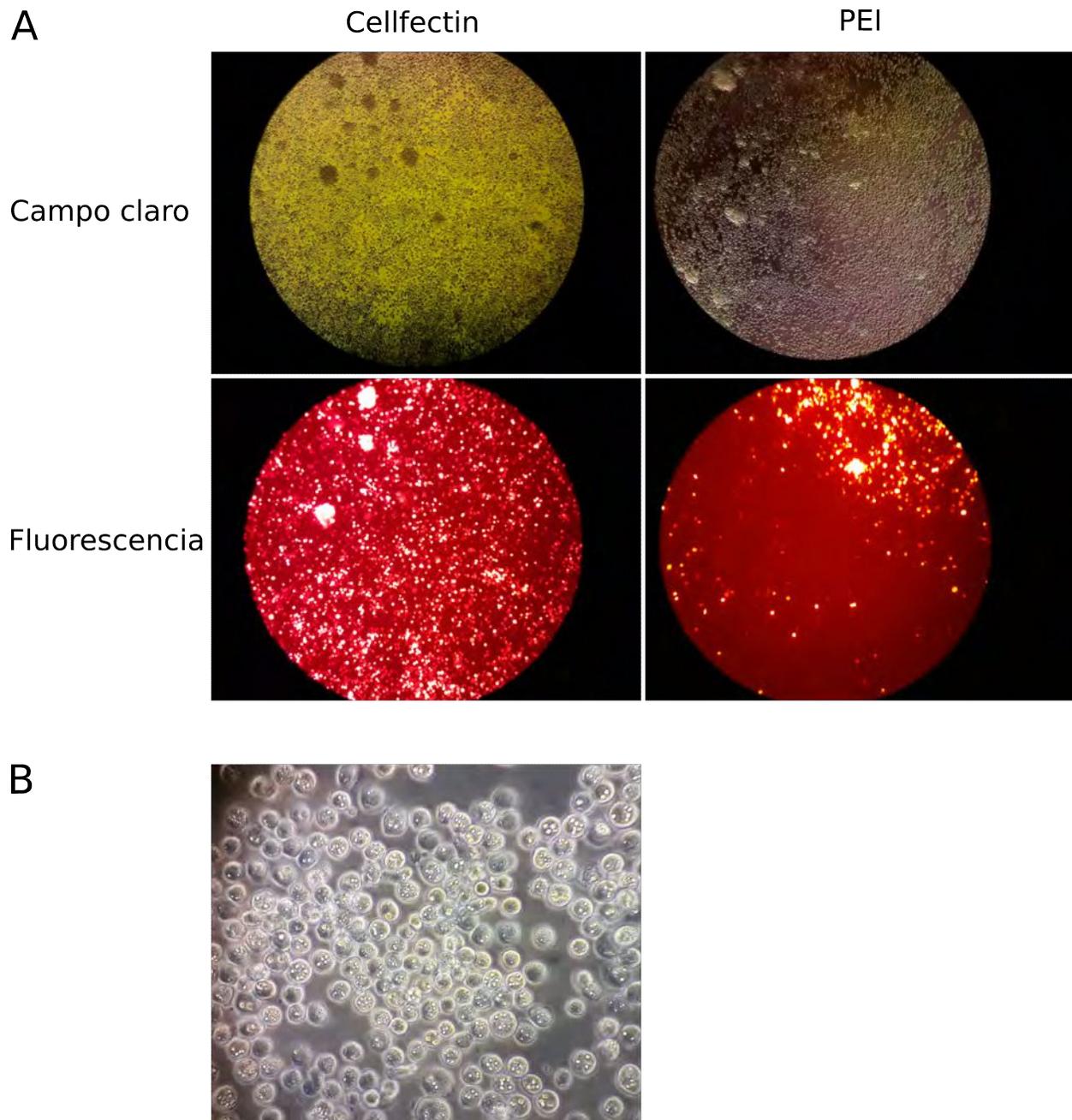


Figura 4.2: Evaluación de distintos reactivos para la transfección de células Sf9 con DNA genómico de AcMNPV-DsRed. A Células siete días post-transfección con Cellfectin (izquierda) o PEI (derecha). **B** OB presentes en células Sf9 a los siete días post-transfección.

4.2.2. Efectos citopáticos y transcripcionales asociados a la entrada de DNA genómico de SfGV

Una vez confirmada la capacidad del DNA genómico del virus control AcMNPV-DsRed para iniciar la infección en células Sf9, decidimos estudiar los efectos citopáticos inducidos por el ingreso de DNA genómico de SfGV, para lo cual empleamos el mismo procedimiento de transfección que se mencionó anteriormente. Sin embargo, las micrografías de células Sf9 a los siete días post-transfección muestran la ausencia de efectos citopáticos en las células tratadas con respecto al control (Figura 4.3). Aunque las células fueron incubadas por un período prolongado (hasta 14 días), en ningún caso observamos efectos diferenciales entre las células control y el tratamiento.

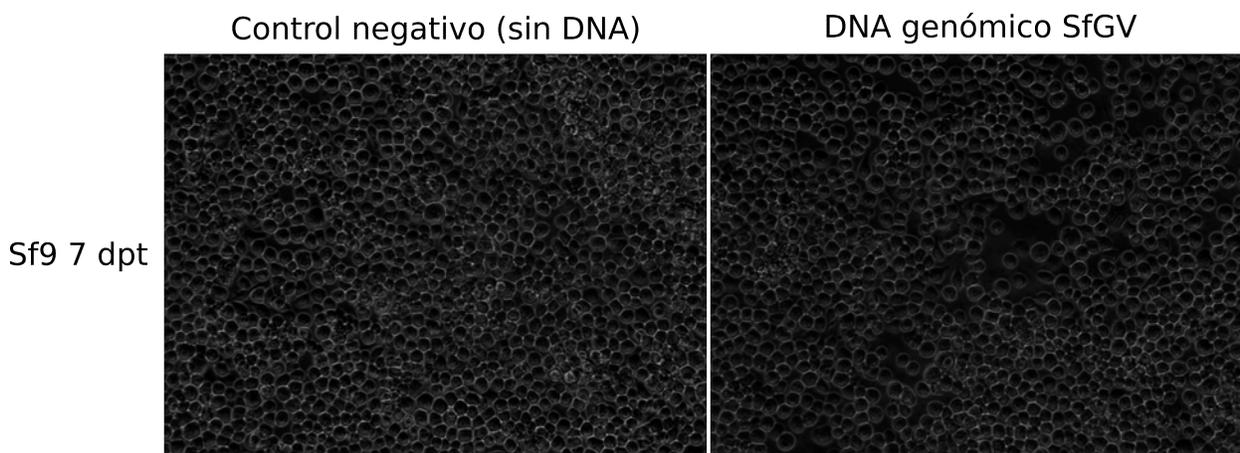


Figura 4.3: Transfección de células Sf9 con DNA de SfGV. Control negativo de transfección (izquierda) y DNA genómico de SfGV (derecha).

Debido a que no evidenciamos ningún cambio morfológico producto de la transfección con DNA genómico, decidimos explorar la presencia de mRNA virales por RT-PCR como una estrategia alternativa para estudiar el ciclo infeccioso de SfGV. En primer lugar, se diseñaron dos pares de oligonucleótidos para detectar los mRNA de los genes *ie-1* y *helicase* de SfGV. La especificidad de estos oligos fue confirmada visualmente mediante el análisis de los productos de amplificación a través electroforesis en gel de agarosa (Figura 4.4 A). Cuando evaluamos la presencia de transcritos virales 24 horas post-transfección con estos oligos, pudimos confirmar la expresión de *ie-1* en las células transfectadas con DNA de SfGV. Sin embargo, para el gen inmediato *helicase* no se observó un nivel apreciable de expresión génica (Figura 4.4 B). Esto indicaría que la cascada transcripcional de SfGV se ve interrumpida durante su etapa inmediata.

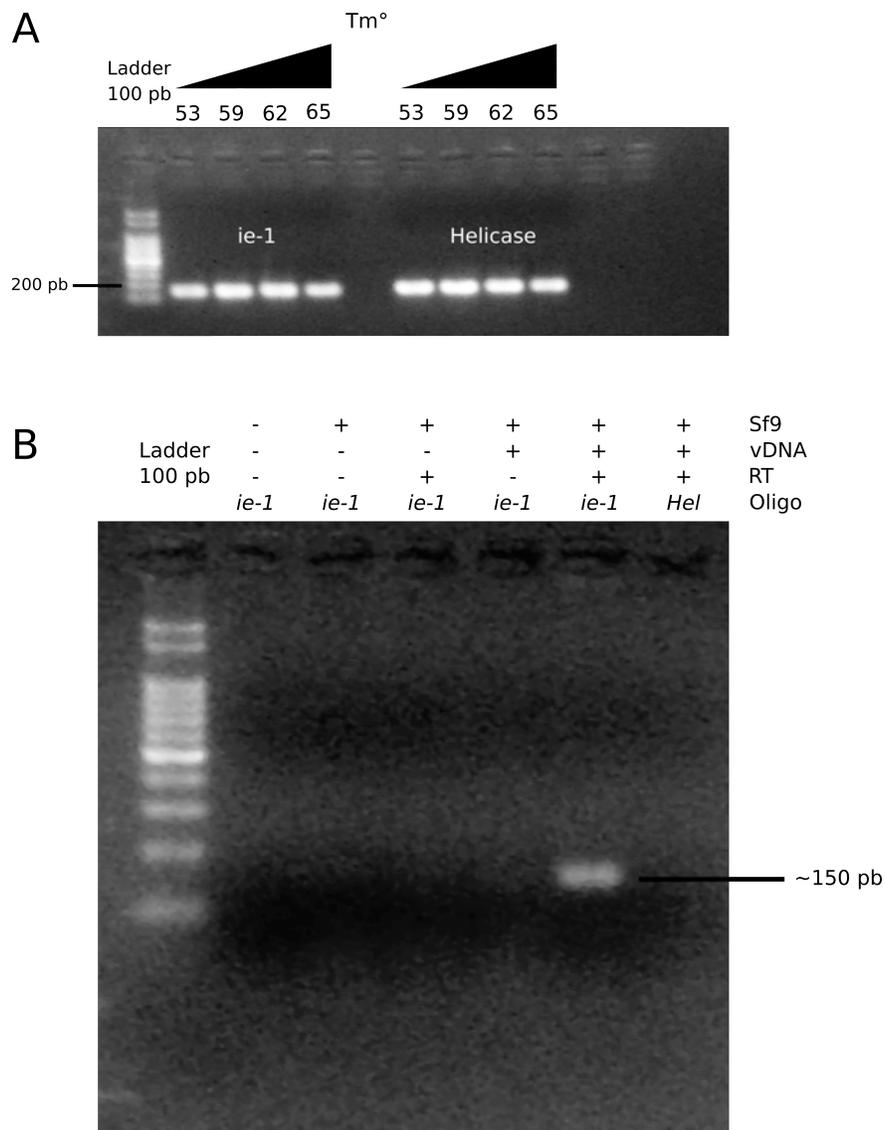


Figura 4.4: Expresión de *ie-1* en células Sf9 transfectadas con DNA de SfGV. A Electroforesis en gel de agarosa (1% p/v) de los productos de amplificación específicos de los oligos diseñados contra los mRNA de los genes *ie-1* y *hel* de SfGV. **B** Expresión del gen *ie-1* en células Sf9 transfectadas con DNA de SfGV (24 horas post-transfección).

4.2.3. Desarrollo de herramientas para facilitar el estudio SfGV en cultivo celular

Debido a que no pudo evidenciarse un efecto citopático apreciable en los ensayos previos, se pensó en desarrollar dos estrategias para, por un lado mejorar la detección de SfGV y por otro facilitar su replicación en cultivo celular. En primer lugar, se decidió poner a punto una metodología de PCR cuantitativa para evaluar los niveles de DNA genómico de SfGV a distintos tiempos post-transfección. Para esto, se empleó AcMNPV-DsRed para poner a punto las condiciones de reacción y se diseñó un par de oligonucleótidos para amplificar un fragmento específico del gen *vp39* de este virus. Se evaluó la reproducibilidad de esta técnica para muestras de DNA total extraído a distintos tiempos a partir de células transfectadas (Tabla 4.1). Como puede observarse, los resultados por triplicado mostraron un elevado nivel de variabilidad entre las distintas réplicas biológicas y no reflejaron un incremento en los niveles de DNA viral a mayores tiempos post-transfección. En contraposición, la curva de calibración mostró un nivel de correlación aceptable (coeficiente de determinación $R^2 = 0.9635$), lo que sugiere que los cebadores diseñados amplifican de forma eficiente la secuencia blanco (Figura 4.5).

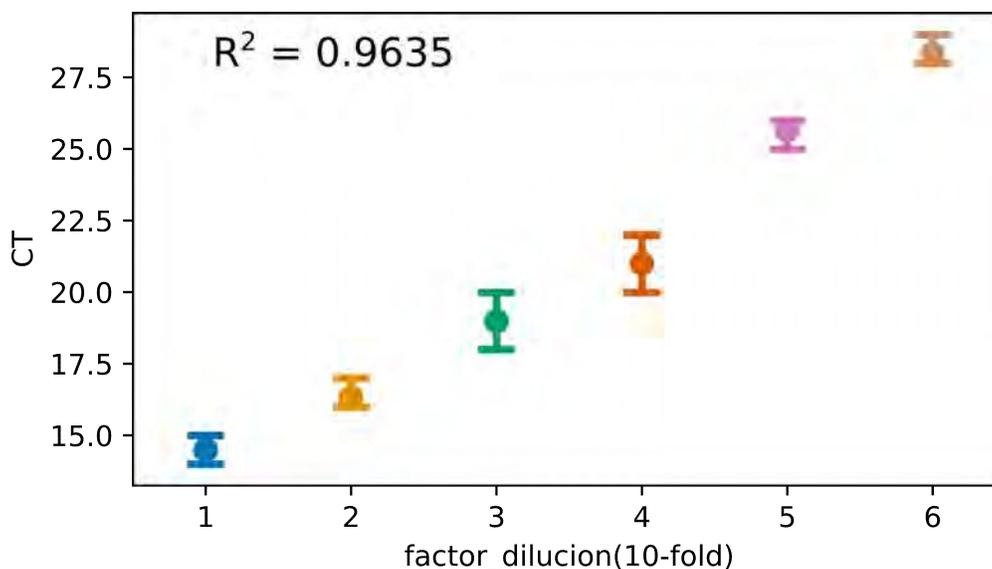


Figura 4.5: Cuantificación de los niveles de DNA viral por qPCR. Curva de calibración construida para cuantificar los niveles de DNA presentes en muestras de células Sf9 transfectadas con AcMNPV-DsRed.

Cuadro 4.1: Cuantificación de los niveles de DNA virales para distintas muestras de células Sf9 transfectadas con DNA viral de AcMNPV-DsRed

DNA AcMNPV	Tiempo postranfección (horas)	Ct
- (mock)	72	No detectado
- (mock)	72	32
- (mock)	72	33
+	24	33
+	24	30
+	24	29
+	48	27
+	48	28
+	48	28
+	72	25
+	72	34
+	72	28

Por otra parte, se construyeron plásmidos de expresión en células de insecto para las proteínas de AcMNPV descritas previamente como provirales, *ie-1*, *gp64* y *p35*.²²⁸ A partir de DNA genómico de AcMNPV *wt* se amplificaron los ORF correspondientes a estos genes utilizando oligonucleótidos específicos con los cuales se incorporó un sitio de restricción *XbaI*. Se lograron obtener productos de amplificación específicos, los cuales fueron clonados en el vector comercial pIB/V5-His (Figura 4.6). La secuencia de cada gen clonado fue corroborada a través de secuenciación Sanger.

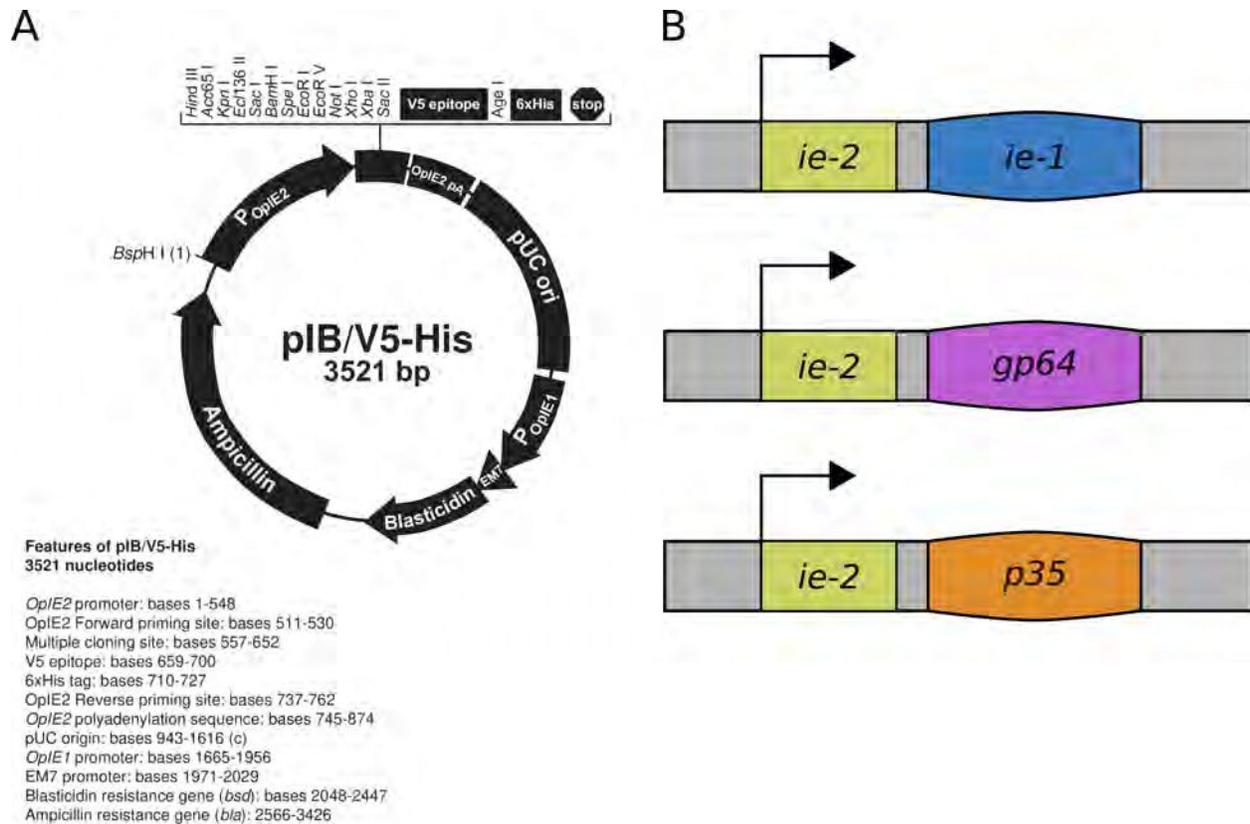


Figura 4.6: Construcciones diseñadas para expresar genes auxiliares en células de insecto. A Vector pIB/V5-His utilizado como esqueleto para los plásmidos de expresión. Adaptado del manual de Thermo ScientificTM. **B** Esquema de la disposición de los elementos génicos en los vectores construidos.

4.3. Discusión

En este capítulo abordamos el estudio del ciclo infectivo de SfGV en cultivo celular. Esfuerzos previos han demostrado que en general los granulovirus no son capaces de replicar de forma eficiente en cultivos celulares. Esto contrasta con la mayor permisividad que se ha observado para el caso de los nucleopoliedrovirus en general.³¹ En el caso de SfGV, demostramos que la introducción del DNA viral por sí sola no sería suficiente para desencadenar un proceso infectivo apreciable a nivel morfológico. Esto indicaría que esta restricción del ciclo infectivo es un fenómeno común dentro del género *Betabaculovirus*.

Por otra parte, pudimos confirmar la presencia de actividad transcripcional inmediatamente temprana en células transfectadas con DNA de SfGV. Esto resulta lógico ya que los genes perteneciente a la clase inmediatamente temprana son expresados por la RNA polimerasa II del hospedador, por lo que su expresión podría iniciarse una vez que el genoma viral alcanza el núcleo celular.⁸⁹ En contraposición, la ausencia de expresión del gen inmediato retrasado *hel* sugiere que ciertos factores transcripcionales se encuentran ausentes. Teniendo en cuenta que la línea Sf9 deriva del hospedador natural de SfGV, es factible suponer que los factores transcripcionales derivados del hospedador podrían estar presentes. Esto nos lleva a sugerir que ciertos factores de transcripción virales necesarios para la cascada transcripcional de SfGV no son expresados eficientemente en Sf9, evitando el avance del ciclo infectivo. Resultaría interesante cuantificar los niveles de expresión de todos los genes de SfGV durante las primeras horas post-transfección mediante RNA-seq para confirmar esta hipótesis. Por otra parte, estos resultados confirman que nuestro protocolo de transfección funciona correctamente en el caso de SfGV.

Debido a la ausencia de efectos citopáticos elicados por SfGV, decidimos poner a punto una metodología para cuantificar de forma sensible los niveles de DNA viral como una alternativa para estudiar el proceso infectivo de este granulovirus. Si bien logramos cuantificar de forma precisa y específica muestras control (derivadas de DNA purificado AcMNPV), la extracción y análisis de muestras de DNA derivadas de células resultó complejo y poco reproducible. En este sentido, la optimización en la preparación de muestras de DNA total resultará fundamental para asegurar la reproducibilidad en la medición de los niveles de DNA viral y su variación post-transfección.

Finalmente, buscamos complementar nuestro ensayos con una serie de plásmidos de expresión que potencialmente pudieran ayudar a la replicación de SfGV. Estas construcciones ya se encuentran disponibles y su secuencia ha sido confirmada, por lo que creemos serán de gran utilidad para evaluar si durante cotransfecciones del DNA de SfGV se observan incrementos en los niveles de replicación viral o cambios morfológicos.

4.4. Materiales y Métodos

4.4.1. Células y virus

Las células Sf9, (derivadas de *S. frugiperda*) fueron mantenidas a 27°C en medio de Grace (Invitrogen) suplementado con 10 % de suero fetal bovino (Internegocios). El stock viral de AcMNPV-DsRed fue propagado mediante infección en monocapa de células Sf9, mientras que SfGV fue amplificado y purificado a partir de larvas infectadas. Brevemente, larvas de *S. frugiperda* se obtuvieron a partir de los insectarios del Instituto Nacional de Tecnología Agropecuaria (INTA) y la Estación Experimental Agroindustrial Obispo Colombres (EEAOC-CONICET-ITANOVA), Tucumán, y fueron utilizadas para establecer una colonia dentro del Instituto de Biotecnología y Biología Molecular (IBBM). La colonia se mantuvo con una dieta artificial y un período de luz controlada (16 horas de luz).

4.4.2. Producción y purificación de los cuerpos de oclusión (OB)

Las larvas de *S. frugiperda* en 1°-2° estadio fueron infectadas por vía oral usando dieta artificial contaminada con una suspensión de OB de SfGV. Las larvas moribundas con signos de infección compatibles con baculovirus fueron almacenadas y procesadas como se describe en.^{91,117} Brevemente, las larvas infectadas se colocaron en agua destilada y fueron homogeneizadas empleando un homogenizador Dounce de vidrio. La solución resultante fue filtrada a través de tres capas de gasa para eliminar restos de insecto insolubles. Este extracto se clarificó por tres pasos de centrifugación a 10000 g por 10 minutos seguidos de un lavado con solución 0.05 % v/v de SDS. La solución clarificada se sometió a una ultracentrifugación en un gradiente continuo de sacarosa 30-60 % w/w (50000 g por 60 minutos a 4°C, rotor Beckman SW 41 Ti). La banda blancuzca/opalescente correspondiente a los OB fue recolectada, diluida 10 veces en agua destilada y precipitada por centrifugación a 14000 g por 10 minutos. El precipitado final fue resuspendido en agua destilada y almacenado a -20°C. Para el caso de AcMNPV-DsRed, los OB se purificaron a partir de cultivos en monocapa. Brevemente, un cultivo de células Sf9 al 70 % de confluencia fue incubado por 1 hora con una alícuota de viriones brotantes de AcMNPV-DsRed. Completada esta incubación se descartó el inóculo viral y se agregó medio de Grace complementado con SFB 10 % a las células. Luego de 7-10 días post-infección se recolectaron las células conteniendo los OB.

4.4.3. Purificación de DNA viral a partir de OB

En este caso, se siguió una metodología descrita previamente.²³³ Una alícuota concentrada de OB (opalescente) se centrifugó a 1000 g por 5 minutos y se resuspendieron los OB en 450 μL de ddH_2O . A esta suspensión se le agregaron 50 μL de buffer 1 M Na_2CO_3 y se incubó por 30 minutos a temperatura ambiente hasta que la solución se clarificó, signo indicativo de que se han disgregado los OB. Luego, esta mezcla se neutralizó con 70 μL de Tris-HCl pH 7.6 1 M y se centrifugó a 5000 g por 2 minutos para remover contaminantes. Para la digestión de las nucleocápsides presentes en solución, se agregaron 14 μL de EDTA 500 mM, 17.5 μL de SDS 10% p/v y 17.5 μL de proteinasa K 20 mg/mL (Productos Bio-lógicos), se llevó a un volumen final de 700 μL y se incubó por 3 horas a 37°C con agitación suave. Esta digestión se utilizó para llevar a cabo tres extracciones orgánicas, una con fenol equilibrado, otra con fenol/cloroformo/alcohol isoamílico y una extracción final con cloroformo/alcohol isoamílico. A la fase acuosa final se le agregó NaCl a una concentración final de 0.2 M y se precipitó el DNA con un volumen de isopropanol. El ovillo blancuzco formado por el DNA viral fue recogido, lavado con etanol 70%, secado por evaporación y resuspendido en 100 μL de buffer TE.

4.4.4. Transfección de células Sf9

Las transfecciones con Cellfectin II, se realizaron de acuerdo al protocolo sugerido por el fabricante. Se homogeneizó el reactivo antes de usarlo y luego se mezclaron 8 μL en 100 μL de medio de Grace sin complementar. En un tubo diferente, se mezcló 1 μg de DNA viral con 100 μL de medio Grace sin complementar. Luego se combinaron las soluciones conteniendo el medio con Cellfectin II y el medio con DNA viral, se homogeneizó con ayuda de un vortex y se incubó a temperatura ambiente por 15 minutos. La mezcla lípidos/DNA se agregó de a gotas a las células. Este momento se tomó como tiempo inicial de la transfección.

Las transfecciones con PEI se llevaron a cabo como se ha descrito previamente.²³⁴ Primero se preparó una mezcla de DNA viral (1 μg) en 100 μL de solución NaCl 100 mM y se vortexeó para homogenizarla. Esta mezcla se esterilizó por calentamiento a 72°C durante 10 minutos y se dejó enfriar. Luego se le agregaron 3,3 μL de PEI 10 μM , se mezcló con ayuda de vortex y se incubó por 10 minutos. Los 100 μL conteniendo los complejos DNA/PEI resultantes se agregaron lentamente

y de manera uniforme a las células. Este momento se tomó como tiempo inicial de la transfección a partir del cual las células fueron monitoreadas diariamente.

4.4.5. Extracción de DNA total partir de cultivo celular

Con el fin de eliminar los restos de DNA extracelular remanentes en el medio de cultivo post-transfección, las células se lavaron con medio Grace's tres veces antes de procesarlas. Las células fueron recolectadas en buffer Tris 10 mM EDTA 1 mM pH 8 y digeridas por 3 horas mediante el agregado de SDS 0.2% p/v y proteinasa K 0.1 mg/mL. Luego se procedió con una extracción de DNA total mediante solventes orgánicos, como se detalló en la sección "Purificación de DNA viral a partir de OB".

4.4.6. PCR cuantitativa (qPCR)

Para las reacciones de PCR cuantitativa se empleó una pre-mezcla de reacción 2X con SYBR Green (Roche) a la cual se le agregaron los oligonucleótidos y el molde previo al inicio de la reacción. Para la cuantificación de los niveles de DNA en las muestras se construyeron curvas de calibración a partir de diluciones seriadas de un stock de DNA genómico derivado de AcMNPV *wt* (concentración de 200 ng/ μ L, obtenida a partir de medidas de Nanodrop). Para la amplificación se utilizó el siguiente perfil de ciclado: desnaturalización inicial a 95°C por 10 minutos, seguido de 50 ciclos de desnaturalización a 95°C por 15 segundos, hibridación-extensión a 60°C por 1 minuto y un paso de extensión final a 60°C por 10 minutos.

4.4.7. Retrotranscripción seguida de PCR (RT-PCR)

El aislamiento del RNA total presente en las células transfectadas se llevó a cabo mediante extracción en fase orgánica con el reactivo Transzol (Trans) de acuerdo a las indicaciones del fabricante. La calidad y concentración del RNA purificado fue chequeado por absorbancia medida en Nanodrop. La reacción de retrotranscripción se llevó a cabo con una transcriptasa reversa derivada de M-MLV (Promega) de acuerdo a las indicaciones del fabricante. El cDNA resultante se empleó como molde para evaluar la presencia de transcriptos virales mediante PCR convencional con el siguiente perfil de ciclado: desnaturalización inicial a 95°C por 5 minutos, seguido de 35 ciclos de desnaturalización a 95°C por 10 segundos, hibridación a 55°C por 10 segundos, extensión a 72°C por 30 segundos y

un paso de extensión final a 72°C por 5 minutos. Para estas reacciones se emplearon reactivos de Productos Bio-Lógicos.

4.4.8. Clonado molecular de los genes *ie-1*, *gp64* y *p35*

Los genes mencionados se obtuvieron a partir de una PCR utilizando como molde DNA genómico viral derivado de AcMNPV *wt*. El producto de amplificación fue purificado mediante columnas de sílica, digerido con *XbaI* (New England Labs), y ligado con el vector pIB (Thermo) previamente linealizado con *XbaI* y defosforilado con fosfatasa alcalina termosensible (Promega). Se chequeó la orientación del gen mediante PCR y se corroboró su secuencia mediante secuenciación Sanger.

4.4.9. Oligonucleótidos

Cuadro 4.2: Oligonucleótidos utilizados para amplificar las regiones codificantes de los genes *ie-1*, *gp64* y *p35*, y en las reacciones de qPCR/RT-PCR.

Oligo	Virus	Sequence
<i>ie-1_XbaI_fw</i>	AcMNPV	ACTGTCTAGAGCGACAAGATCGTGAACAACC
<i>ie-1_XbaI_rv</i>	AcMNPV	ACTGTCTAGAGCGTCAGCGGCAAGTTAATG
<i>gp64_XbaI_fw</i>	AcMNPV	ACTGTCTAGCACACAAGCAAGATGGTAAG
<i>gp64_XbaI_rv</i>	AcMNPV	ACTGTCTAGATCGTATCGATCTTGCTAAGG
<i>p35_XbaI_fw</i>	AcMNPV	ACTGTCTAGAGTGAGCATTTGAGCTTTACC
<i>p35_XbaI_rv</i>	AcMNPV	ACTGTCTAGATCGTAAAGCGAGTTTATGAG
<i>ie-1_fw</i>	SfGV	TATGATGTGATGCGAGCCCC
<i>ie-1_rv</i>	SfGV	TCACCGGCAAGGAACACTAC
<i>dna_helicase_fw</i>	SfGV	CGCTGGACCAAAGTTTCACG
<i>dna_helicase_rv</i>	SfGV	GCAACGCCATCACATCCAAA
<i>vp39_fw</i>	AcMNPV	GCGTGTTGCTTGTGAAACCT
<i>vp39_rv</i>	AcMNPV	TGATGCAAGCCGAACAGCTA

Capítulo 5

Conclusiones y perspectivas finales

5.1. Conclusiones

En esta tesis buscamos generar nuevos conocimientos y herramientas con el fin de mejorar la aplicabilidad de baculovirus autóctonos para el control de la oruga militar tardía. Para lograr esto, decidimos adoptar un enfoque integrativo que combinara aproximaciones ómicas, herramientas computacionales y técnicas de virología clásica. Creemos que la intersección de estos campos nos ha permitido expandir el conocimiento disponible para esta familia de virus. Sumado a esto, esperamos que los conjuntos de datos generados en este trabajo (proteomas de EpapGV y SfGV, y genoma de SfMNPV) resulten útiles para analizar y guiar estudios futuros.

El capítulo 2 describe la caracterización proteómica del granulovirus de SfGV y proporciona un análisis comparativo frente a un betabaculovirus de acción rápida como EpapGV. Nuestros resultados mostraron que gran parte de los componentes estructurales presentes en los OB de granulovirus son compartidos entre distintas especies, aunque también aparecen proteínas virus-específicas que podrían estar involucradas en el tropismo de cada granulovirus. A partir de estimaciones semi-cuantitativas basadas en el valor de emPAI, esbozamos un modelo del OB de granulovirus, reflejando sus distintos componentes y su abundancia. Creemos que estos datos serán de gran utilidad para comprender mejor como la estructura de la partícula baculoviral contribuye a su infectividad oral. Por otra parte, de nuestros datos de abundancia proteica surgió una proteína (Epap95/Spfr101) con elevados niveles de expresión, la cual se encuentra conservada en todas las especies de granulovirus. Análisis bioinformáticos y la sobre-expresión de esta proteína en cultivos celulares mostró que podría formar agregados. Con el fin de caracterizar su función putativa, preparamos una construcción que expresa esta proteína fusionada a eGFP. A futuro se buscará completar el estudio proteómico de los posibles interactores de esta proteína mediante purificación por afinidad seguida de MS.

En el capítulo 3, caracterizamos la diversidad genética presente dentro de un aislamiento argentino de SfMNPV, con el fin de investigar cómo esta fuente de variabilidad contribuye a su adaptabilidad hacia su hospedador. Se descubrieron más de 700 variantes puntuales, de las cuales 184 introducen un cambio directo en las propiedades biofísicas de las proteínas virales. Mas aún, nuestro conjunto de datos nos permitió detectar variantes estructurales (deleciones, duplicaciones, inversiones) novedosas y recuperar una deleción común en poblaciones naturales de SfMNPV, lo que resalta

la utilidad de las técnicas de HTS para investigar la presencia de haplotipos disímiles dentro de las poblaciones baculovirales. Basados en los valores elevados de diversidad genética del gen *sf29*, llevamos a cabo un estudio filogenético de esta proteína, el cual arrojó evidencia de que podría tratarse de una peptidasa que todavía retiene su funcionalidad.

Por último, se abordó la caracterización inicial del ciclo replicativo del granulovirus SfGV en la línea celular Sf9, con el fin de explorar la posibilidad de producir este virus en cultivos celulares. Luego de poner a punto un protocolo de transfección eficiente a partir de DNA genómico viral, pudimos evaluar la respuesta celular frente a la entrada del DNA de SfGV. A nivel morfológico, no evidenciamos cambios apreciables en el aspecto celular ni tampoco observamos estructuras reminiscentes de OB. En cuanto a la expresión de genes virales, solamente se observó actividad transcripcional inmediatamente temprana. Esto está en línea con reportes previos para otros betabaculovirus, confirmando que a pesar de emplear un sistema celular derivado del mismo organismo hospedador los granulovirus presentan mayores requerimientos para el desarrollo de su ciclo infectivo. Debido a las dificultades inherentes al trabajo con granulovirus en cultivos celulares, se inició la puesta a punto de un método de qPCR para cuantificar los niveles de genoma viral. Además, se construyeron tres plásmidos de expresión con genes provirales los cuales serán probados a futuro como agentes sinergizantes de la infección de SfGV en cultivo celular.

En conclusión, los resultados presentados aquí muestran que a través de distintos enfoques experimentales es posible obtener información detallada con respecto a los diferentes aspectos de la biología de los baculovirus. Creemos que los datos presentados aquí ayudarán a comprender mejor el proceso infectivo de los baculovirus y, por extensión, repercutirán en su aplicación como biopesticidas.

5.2. Perspectivas

- Validar experimentalmente la presencia del producto de fusión Epap48-49 mediante el uso de anticuerpos específicos contra Epap48 o Epap49.
- Realizar los ensayos de purificación por afinidad seguida de análisis por espectrometría de masas para caracterizar los posibles interactores celulares de Epap95/Spfr101.
- Confirmar mediante expresión, purificación y ensayo bioquímico si el gen *sf29* de SfmNPV codifica una proteína con actividad peptidasa.
- Secuenciar y caracterizar los polimorfismos presentes dentro de distintos aislamientos geográficos y temporales de SfmNPV en Argentina.
- Mejorar la reproducibilidad en la extracción de DNA total desde cultivo celular previo a la reacción de qPCR.
- Evaluar mediante ensayos de co-transfección si los plásmidos de expresión de genes sinérgicos son capaces de incrementar los niveles de replicación viral en cultivo celular.

Capítulo 6

Material suplementario

6.1. Grupos de genes ortólogos presentes en los proteomas de baculovirus

Cuadro 6.1: Grupos de ortología entre proteínas estructurales presentes en los OBs de la familia Baculoviridae. En cada celda se indica el número de ORF para cada genoma.

Ortogrupo	AcMNPV	AgMNPV	ChchNPV	HearNPV	MabrNPV	EpapGV	ClanGV	PiraGV	SfGV	CuniNPV
Polyhedrin	8	1	1	1	1	1	1	1	1	-
PEP	131	31	121	120	53	21, 25	19, 36	20, 22	16, 17	-
P10	137	24	18	21	152	-	16	17	5, 71	-
PEP/P10	-	-	-	-	-	22	35	21	18	-
ODV-E18	143	18	12	10	159	29	13	14	11	31
ODV-E25	94	71	86	82	87	86	76	76	88	15
ODV-E66	46	114	101	96	69, 136	39	-	39, 44	130	-
P33	92	73	84	80	88	88	78	78	89	14
PIF0	138	23	17	20	153	59	49	51	63	74
PIF1	119	43	131	111	43	69	61	61	72	29
PIF2	22	155	148	132	42	47	23	40	38	38
PIF3	115	46	110	98	59	38	33	30	27	46
PIF4	96	69	88	85	84	84	74	74	87	90
PIF5	148	13	7	15	6	27	15	16	11	102
PIF6	68	94	61	64	104	109	98	96	115	58
PIF8	83	80	81	76	93	96	86	85	105	35
AC75	75	88	57	69	108	103	94	92	112	-
Ac78	78	85	77	72	97	100	90	89	108	34

Ortogrupo	AcMNPV	AgMNPV	ChchNPV	HearNPV	MabrNPV	EpapGV	ClanGV	PiraGV	SfGV	CuniNPV
Ac81	81	82	79	74	95	98	88	87	106	106
Ac110	110	51	98	93	73	this study	20	44	42	70
GP41	80	83	78	73	96	99	89	88	107	33
F-protein	23	143	150	133	8	14	30	26	23	104
P78/83	9	2	2	2	2	5	2	2	3	-
VP39	89	76	82	78	91	92	81	81	96	24
p6.9	100	66	93	88	78	81	71	71	83	23
P40	101	65	94	89	77	80	70	70	82	22
P48/45	103	63	96	89	75	78	68	68	79	55
Ac142	142	19	11	9	160	28	14	15	12	30
ODV-EC43	109	53	99	94	72	53	44	46	44	69
P24	129	33	134	118	11	66	57	58	68	-
ODV-EC27	144	17	13	11	158	93	82	82	97	32
VLf-1	77	86	76	71	98	101	91	90	109	18
38K	98	68	91	86	80	83	73	73	84	87
VP1054	54	107	45	47	125	129	119	116	144	8
P18	93	72	85	81	86	87	77	77	90	13
ME53	139	21	8	16	7	133	123	120	150	-
PK1	10	3	3	3	3	6	3	3	2	-
VP80	105	62	97	92	74	-	-	-	-	-

Ortogrupo	AcMNPV	AgMNPV	ChchNPV	HearNPV	MabrNPV	EpapGV	ClanGV	PiraGV	SfGV	CumiNPV
FP25K	61	101	51	50	116	113	102	100	120	-
P12	102	64	95	90	76	79	69	69	81	-
Desmop	66	96	59	66	106	107	96	94	114	92
v-Ubi	35	138	26	28	145	52	43	45	43	-
SOD	31	131	115	106	57	58	48	50	52	-
LEF1	14	148	138	124	29	68	60	60	70	45
LEF3	67	95	60	65	105	108	97	95	116	-
LEF5	99	67	92	87	77	82	72	72	85	-
LEF6	28	124	21	24	149	74	65	65	76	-
LEF8	50	111	37	38	131	122	111	110	124	26
LEF9	62	100	52	52	114	112	101	99	119	59
DNA Pol	65	97	58	67	107	106	95	93	113	91
Hel-1	95	70	87	84	85	85	75	75	86	89
IE1	147	14	16	14	157	35	6	6	8	-
Alk-Exo	133	29	127	114	47	119-120	109	107	125	54
gp37	64	-	67	58	31	30	12	13	-	-
Hear44	-	-	42	44	128	126	116	114	141	-
Hear45	-	-	43	45	127	127	117	115	142	-
HCF-1	70	-	-	-	-	-	-	-	-	-
P43	39	-	-	-	-	-	-	-	-	-

Ortogrupo	AcMNPV	AgMNPV	ChchNPV	HearNPV	MabrNPV	EpapGV	ClanGV	PiraGV	SfGV	CuniNPV
PCNA	49	-	66	-	-	-	-	-	-	-
PNK/PNL	86	57	-	-	-	-	-	-	-	-
Ac30	30	126	-	-	-	-	-	-	-	-
Ac5	5	165	-	-	-	-	-	-	-	-
Ac74	74	89	-	-	-	-	-	-	-	-
Ac79	79	84	-	-	15	-	-	-	-	-
ChaB1	58/59	103	49	51	121	-	-	-	91	-
ODV-E26	16	150	-	-	-	-	-	-	-	-
CG30	88	77	-	77	92	-	-	-	-	-
Ac114	114	47	-	-	-	-	-	-	-	-
Ac132	132	30	-	-	-	-	-	-	-	-
PTP1/2	1	160	140	-	32	-	-	-	-	-
Ac18	18	152	125	-	49	-	-	-	-	-
EGT	15	149	141	126	33	132	122	119	-	-
38.7 kDa	13	147	137	123	30	67	59	59	69	-
PP31	36	140	28	31	143	-	-	-	-	-
IAP	several	123	39	several	131	several	100	several	118	-
CBP	145	16	14	12	157	33	8	8	10	-
Ac146	146	15	15	13	156	34	7	7	9	-
Ac34	34	137	25	27	146	-	-	-	-	-

Ortogrupo	AcMNPV	AgMNPV	ChchNPV	HearNPV	MabrNPV	EpapGV	ClanGV	PiraGV	SfGV	CumiNPV
BJDP	51	110	38	39	132	-	-	-	-	-
Ac48	48	113	-	-	-	123	113	111	-	-
Chch46	-	-	46	48	124	-	-	-	-	-
Ac56	56	105	47	49	123	-	-	-	-	-
ChaB2	60	102	50	52	120	90	-	-	-	-
Ac108	108	54	100	95	71	-	-	-	-	-
Chch105	-	-	105	-	-	-	-	-	-	-
PARG	-	-	108	100	61	-	-	-	-	-
Chch123	-	-	123	-	50	-	-	-	-	-
GP16	130	32	133	119	10	-	-	-	-	-
Chch135	-	-	135	-	-	-	-	-	-	-
HOAR	-	-	4	4	4	-	-	-	-	-
LEF12	41	118	-	36	-	-	-	-	-	-
Ac53	53	109	41	43	129	125	115	113	140	-
Ac76	76	87	56	70	109	102	92	91	111	-
Hear83	-	-	-	83	37	-	-	-	-	-
Ac63	63	-	31	121	-	-	-	-	-	-
Mabr52	151	-	-	-	17, 52, 110	-	-	-	-	-
Mabr21	-	-	-	-	21	-	-	-	-	-
Mabr23	-	-	-	-	23	-	-	-	-	-

Ortogrupo	AcMNPV	AgMNPV	ChchNPV	HearNPV	MabrNPV	EpapGV	ClanGV	PiraGV	SfGV	CuniNPV
Helicase-2	-	-	-	-	25	119-120	-	108	126	-
v-Cath	127	-	64	56	27	31	11	11	-	-
Ac4	4	164	-	-	38, 64	-	-	-	-	-
RR	-	-	122, 151	-	51, 162	-	-	-	-	-
Mabr56	-	-	-	-	56	-	-	-	-	-
Mabr58	-	-	111	-	58	-	-	-	-	-
NRK1	33	136	106	-	63	61	-	-	-	-
VEF	-	-	-	-	81	-	-	-	-	-
Mabr83	-	-	-	-	83	-	-	-	-	-
TLP-20	82	81	80	75	94	97	87	86	104	-
Mabr100	-	-	-	34	100	-	-	-	-	-
Mabr112	-	-	-	-	112	-	-	-	-	-
Ac43	43	116	35	37	135	-	-	-	-	-
Ac26	26	122	23	26	147	-	-	-	-	-
DBP	25	121	22	25	148	75	66	66	77	-
P26	136	25	63	22	151	-	-	-	-	-
Epap10	-	-	-	-	-	10	-	-	-	-
Epap40	-	-	-	-	-	40	39	31	29	-
Epap48	-	-	-	-	-	48	22	42	40	-
Epap49	-	-	-	-	-	49	-	-	-	-

Ortogrupo	AcMNPV	AgMNPV	ChchNPV	HearNPV	MabrNPV	EpapGV	ClanGV	PiraGV	SfGV	CuniNPV
Epap62	-	-	-	-	-	62	54	55	64	-
Epap71	-	-	-	-	-	71	63	63	-	-
Epap95	-	-	-	-	-	95	84	84	101	-
Spfr7	-	-	-	-	-	7	4	4	7	-
Spfr25	-	-	-	-	-	37	32	28	25	-
Spfr26	-	-	-	-	-	-	-	29	26	-
Spfr28	-	-	-	-	-	-	-	-	28	-
Spfr45	-	-	-	-	-	-	45	47	45	-
dUTPase	-	-	-	-	-	13	-	-	50	-
p13	-	-	-	-	-	40	24	38	36	-
Chtb-2a	-	-	-	-	-	73	64	64	74	-
Chtb-2b	-	-	-	-	-	73	64	64	75	-
Spfr80	-	-	-	-	-	-	-	-	80	-
BRO-n	-	-	-	-	-	-	-	83	-	-
BRO-f	-	-	-	-	-	-	108	106	98	-
Spfr100	-	-	-	-	-	-	-	-	100	-
Spfr101	-	-	-	-	-	95	84	84	100	-
Spfr112	-	-	-	-	-	103	94	92	112	-
Spfr127	112-113	-	-	-	-	-	-	-	127	-
Enhancin-1	-	-	-	-	-	-	-	-	131	-

Ortogrupo	AcMNPV	AgMNPV	ChchNPV	HearNPV	MabrNPV	EpapGV	ClanGV	PiraGV	SfGV	CumiNPV
Enhancin-2	-	-	-	-	-	-	-	-	137	-
Spfr143	-	-	-	-	-	-	-	-	143	-
Spfr145	-	-	-	-	-	-	-	-	145	-
Spfr151	-	-	-	-	-	-	-	-	151	-
Pira54	-	-	-	-	-	-	-	54	-	-
Clan17	-	-	-	-	-	-	17	18	-	-
Clan18	-	-	-	-	-	26	18	19	15	-
Clan20	-	-	-	-	-	-	20	-	-	-
Clan27	-	-	-	-	-	-	27	35	-	-
Clan37	-	-	-	-	-	-	37	-	139	-
Clan42	-	-	-	-	-	-	42	-	-	-
Clan79	-	-	-	-	-	-	79	-	-	-
Clan93	-	-	-	-	-	-	93	-	-	-

**6.2. Variantes nucleotídicas puntuales detectadas dentro de la
población de SfmNPV**

Cuadro 6.2: Variantes nucleotídicas puntuales intrahospedador (iSNV) detectadas con Lofreq

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
98	A	C	0.003845	sinsentido	polh	Non-core gene	p.Lys33Thr	+
273	C	T	0.103167	sinónima	polh	Non-core gene	p.Ser91Ser	-
474	G	A	0.024921	sinónima	polh	Non-core gene	p.Glu158Glu	-
543	T	C	0.104833	sinónima	polh	Non-core gene	p.Val181Val	-
1025	C	T	0.100709	sinónima	orf1629	Non-core gene	p.Leu390Leu	-
1309	T	C	0.104958	sinsentido	orf1629	Non-core gene	p.Ile296Val	-
1391	G	C	0.010092	sinónima	orf1629	Non-core gene	p.Pro268Pro	-
1391	G	T	0.023878	sinónima	orf1629	Non-core gene	p.Pro268Pro	-
1418	G	A	0.262583	sinónima	orf1629	Non-core gene	p.Ala259Ala	-
1430	G	A	0.262602	sinónima	orf1629	Non-core gene	p.Ser255Ser	-
1528	G	A	0.298123	sinónima	orf1629	Non-core gene	p.Leu223Leu	-
1724	C	T	0.308133	sinónima	orf1629	Non-core gene	p.Thr157Thr	-
2212	C	A	0.008445	sinsentido	pk-1	Non-core gene	p.Thr7Asn	-
2227	T	A	0.002946	sinsentido	pk-1	Non-core gene	p.Met12Lys	+
2459	G	A	0.248541	sinónima	pk-1	Non-core gene	p.Thr89Thr	-
2582	A	C	0.005703	sinsentido	pk-1	Non-core gene	p.Glu130Asp	-
2624	C	T	0.353297	sinónima	pk-1	Non-core gene	p.Asp144Asp	-
2837	C	T	0.226688	sinónima	pk-1	Non-core gene	p.Phe215Phe	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
2879	A	G	0.234629	sinónima	pk-1	Non-core gene	p.Leu229Leu	-
3068	C	T	0.261153	intergénica	-	-	-	-
3074	C	A	0.005035	intergénica	-	-	-	-
3076	T	A	0.00562	intergénica	-	-	-	-
3236	C	T	0.187142	sinónima	hoar	Non-core gene	p.Glu422Glu	-
3530	A	G	0.040964	sinónima	hoar	Non-core gene	p.Ile324Ile	-
3656	T	G	0.003187	sinónima	hoar	Non-core gene	p.Val282Val	-
4079	G	A	0.026458	sinónima	hoar	Non-core gene	p.Ile141Ile	-
4689	A	G	0.031087	sinónima	sf5	Non-core gene	p.Val13Val	-
4719	G	A	0.027038	sinónima	sf5	Non-core gene	p.Phe3Phe	-
4733	T	C	0.028542	sinónima	sf6	Non-core gene	p.Pro56Pro	-
4796	T	C	0.028295	sinónima	sf6	Non-core gene	p.Lys35Lys	-
5260	C	T	0.34018	sinónima	sf8	Non-core gene	p.Ile42Ile	-
5995	A	G	0.084011	sinónima	sf8	Non-core gene	p.Leu287Leu	-
6203	A	T	0.044011	sinsentido	sf8	Non-core gene	p.Thr357Ser	-
6401	C	T	0.404831	sinónima	sf8	Non-core gene	p.Leu423Leu	-
6550	A	G	0.498274	sinónima	sf8	Non-core gene	p.Lys472Lys	-
6702	A	T	0.092823	intergénica	-	-	-	-
6715	C	T	0.093137	intergénica	-	-	-	-
6826	A	C	0.286488	sinónima	pif-5	Core gene	p.Ser18Ser	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
6973	A	T	0.290255	sinsentido	pif-5	Core gene	p.Glu67Asp	-
7453	C	T	0.07357	sinónima	pif-5	Core gene	p.Asp227Asp	-
7630	T	C	0.06993	sinónima	pif-5	Core gene	p.Phe286Phe	-
7819	G	A	0.445187	sinónima	pif-5	Core gene	p.Gln349Gln	-
8642	G	A	0.053581	sinónima	me53	Non-core gene	p.Lys187Lys	-
8669	G	A	0.037305	sinónima	me53	Non-core gene	p.Lys196Lys	-
8851	C	A	0.036323	sinsentido	me53	Non-core gene	p.Thr257Asn	-
9183	A	G	0.001826	intergénica	-	-	-	-
9188	C	G	0.00299	intergénica	-	-	-	-
9245	T	C	0.175558	intergénica	-	-	-	-
9248	A	T	0.009056	intergénica	-	-	-	-
9283	T	C	0.004836	intergénica	-	-	-	-
9303	G	A	0.300225	intergénica	-	-	-	-
9547	T	C	0.002504	intergénica	-	-	-	-
9599	C	A	0.005956	intergénica	-	-	-	-
9803	A	C	0.012888	intergénica	-	-	-	-
9812	T	C	0.007088	intergénica	-	-	-	-
9827	A	T	0.005995	intergénica	-	-	-	-
10224	T	C	0.004801	intergénica	-	-	-	-
10244	A	G	0.003999	intergénica	-	-	-	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
10428	G	A	0.392324	intergénica	-	-	-	-
10830	G	A	0.235504	intergénica	-	-	-	-
10867	C	A	0.008106	intergénica	-	-	-	-
10879	T	A	0.336714	intergénica	-	-	-	-
11211	A	G	0.386287	sinónima	efp	Non-core gene	p.Thr12Thr	-
12084	C	T	0.340726	sinónima	efp	Non-core gene	p.Thr303Thr	-
12150	C	T	0.339107	sinónima	efp	Non-core gene	p.Val325Val	-
12387	T	C	0.46212	sinónima	efp	Non-core gene	p.Cys404Cys	-
12696	A	G	0.196697	sinónima	efp	Non-core gene	p.Leu507Leu	-
12823	T	C	0.307922	sinónima	efp	Non-core gene	p.Leu550Leu	-
12825	A	G	0.308165	sinónima	efp	Non-core gene	p.Leu550Leu	-
13204	C	A	0.003525	sinsentido	efp	Non-core gene	p.Pro677Thr	+
13329	G	T	0.005146	intergénica	-	-	-	-
13341	G	T	0.003413	intergénica	-	-	-	-
13900	C	T	0.004451	sinsentido	p24	Non-core gene	p.Ser175Asn	-
13911	A	G	0.023922	sinónima	p24	Non-core gene	p.Asn171Asn	-
14082	G	A	0.43926	sinónima	p24	Non-core gene	p.Arg114Arg	-
14202	G	A	0.498627	sinónima	p24	Non-core gene	p.Val74Val	-
14208	A	G	0.495934	sinónima	p24	Non-core gene	p.Pro72Pro	-
14244	A	G	0.502168	sinónima	p24	Non-core gene	p.Leu60Leu	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
14671	T	C	0.371901	sinónima	sf15	Non-core gene	p.Arg67Arg	-
14909	A	G	0.376492	sinónima	lef-2	Core gene	p.Lys49Lys	-
15143	C	T	0.245899	sinónima	lef-2	Core gene	p.Gly127Gly	-
15363	T	A	0.244515	sinsentido	lef-2	Core gene	p.Leu201Met	-
15371	T	C	0.239909	sinónima	lef-2	Core gene	p.Asn203Asn	-
15436	A	T	0.019533	intergénica	-	-	-	-
15480	G	A	0.151153	sinónima	38.7k	Non-core gene	p.Thr374Thr	-
15651	C	T	0.23988	sinónima	38.7k	Non-core gene	p.Arg317Arg	-
15678	T	C	0.250397	sinónima	38.7k	Non-core gene	p.Glu308Glu	-
15686	T	C	0.249964	sinsentido	38.7k	Non-core gene	p.Ser306Gly	+
15870	G	A	0.274794	sinónima	38.7k	Non-core gene	p.Ser244Ser	-
15882	C	T	0.275954	sinónima	38.7k	Non-core gene	p.Lys240Lys	-
15921	A	G	0.280873	sinónima	38.7k	Non-core gene	p.Asp227Asp	-
16293	A	G	0.252099	sinónima	38.7k	Non-core gene	p.Asp103Asp	-
16344	G	T	0.0045	sinónima	38.7k	Non-core gene	p.Ala86Ala	-
16349	A	C	0.003017	sinsentido	38.7k	Non-core gene	p.Phe85Val	-
16395	T	C	0.476988	sinónima	38.7k	Non-core gene	p.Thr69Thr	-
17211	G	A	0.140261	sinónima	lef-1	Core gene	p.Ser38Ser	-
17212	G	A	0.140637	sinsentido	lef-1	Core gene	p.Ser38Phe	+
17632	C	T	0.011395	sinónima	sf19	Non-core gene	p.Pro106Pro	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
17719	A	C	0.004156	sinsentido	sf19	Non-core gene	p.Lys135Asn	+
17739	C	T	0.012036	sinónima	v-cath	Non-core gene	p.Leu332Leu	-
17970	G	A	0.466761	sinónima	v-cath	Non-core gene	p.Val255Val	-
18345	T	G	0.47276	sinónima	v-cath	Non-core gene	p.Ala130Ala	-
18405	T	C	0.458082	sinónima	v-cath	Non-core gene	p.Leu110Leu	-
18507	C	T	0.361075	sinónima	v-cath	Non-core gene	p.Lys76Lys	-
18836	T	A	0.004641	intergénica	-	-	-	-
18839	G	T	0.004445	intergénica	-	-	-	-
18891	C	A	0.008375	sinsentido	lef-7	Non-core gene	p.Met339Ile	-
19116	C	T	0.356062	sinónima	lef-7	Non-core gene	p.Gly264Gly	-
19117	C	T	0.355982	sinsentido	lef-7	Non-core gene	p.Gly264Glu	+
19140	C	T	0.373421	sinsentido	lef-7	Non-core gene	p.Met256Ile	-
19200	C	T	0.384066	sinónima	lef-7	Non-core gene	p.Val236Val	-
19481	T	C	0.171192	sinsentido	lef-7	Non-core gene	p.Thr143Ala	+
19773	C	G	0.176121	sinsentido	lef-7	Non-core gene	p.Glu45Asp	-
20098	T	C	0.188535	sinónima	chiA	Non-core gene	p.Ser54Ser	-
20380	A	G	0.178233	sinónima	chiA	Non-core gene	p.Lys148Lys	-
20743	G	A	0.177161	sinónima	chiA	Non-core gene	p.Gly269Gly	-
21173	T	C	0.202901	sinsentido	chiA	Non-core gene	p.Ser413Pro	+
21200	C	T	0.203869	sinónima	chiA	Non-core gene	p.Leu422Leu	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
21515	T	C	0.061235	sinónima	chiA	Non-core gene	p.Leu527Leu	-
22253	T	C	0.346836	sinónima	sf23	Non-core gene	p.Tyr122Tyr	-
22470	G	A	0.010548	sinsentido	sf23	Non-core gene	p.Glu195Lys	+
22479	A	G	0.341355	sinsentido	sf23	Non-core gene	p.Lys198Glu	+
22482	A	G	0.286872	sinsentido	sf23	Non-core gene	p.Lys199Glu	+
22494	G	A	0.026888	sinsentido	sf23	Non-core gene	p.Glu203Lys	+
22499	T	A	0.011601	sinsentido	sf23	Non-core gene	p.Asp204Glu	-
22499	T	G	0.022402	sinsentido	sf23	Non-core gene	p.Asp204Glu	-
22538	C	A	0.005475	sinónima	sf23	Non-core gene	p.Ile217Ile	-
22604	A	G	0.01103	intergénica	-	-	-	-
22646	G	T	0.003882	intergénica	-	-	-	-
22973	T	C	0.374311	sinsentido	gp37	Non-core gene	p.Ser104Pro	+
23285	G	T	0.369534	sinsentido	gp37	Non-core gene	p.Ala208Ser	+
23314	G	A	0.364128	sinónima	gp37	Non-core gene	p.Arg217Arg	-
23371	T	C	0.368919	sinónima	gp37	Non-core gene	p.Val236Val	-
23377	A	C	0.367085	sinónima	gp37	Non-core gene	p.Gly238Gly	-
23434	T	C	0.334594	sinónima	gp37	Non-core gene	p.His257His	-
23446	G	A	0.358656	stop_retained	gp37	Non-core gene	p.Ter261Ter	-
23548	G	A	0.394233	sinónima	ptp-2	Non-core gene	p.Ile133Ile	-
24210	G	A	0.004293	sinsentido	egt	Non-core gene	p.Val57Ile	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
24212	C	T	0.004023	sinónima	egt	Non-core gene	p.Val57Val	-
24431	C	T	0.174474	sinónima	egt	Non-core gene	p.Ser130Ser	-
24515	T	C	0.172463	sinónima	egt	Non-core gene	p.Ile158Ile	-
24590	G	T	0.09148	sinónima	egt	Non-core gene	p.Ala183Ala	-
24677	G	C	0.186078	sinónima	egt	Non-core gene	p.Val212Val	-
24987	G	T	0.190517	sinsentido	egt	Non-core gene	p.Gly316Cys	+
25040	T	C	0.230385	sinónima	egt	Non-core gene	p.Leu333Leu	-
25283	G	A	0.228609	sinsentido	egt	Non-core gene	p.Met414Ile	-
25424	G	T	0.24549	sinsentido	egt	Non-core gene	p.Met461Ile	-
27046	C	A	0.00478	sinsentido	sf29	Non-core gene	p.Lys866Asn	+
27541	A	G	0.302098	sinónima	sf29	Non-core gene	p.Phe701Phe	-
27550	G	A	0.41244	sinónima	sf29	Non-core gene	p.Phe698Phe	-
27631	A	G	0.205475	sinónima	sf29	Non-core gene	p.Ala671Ala	-
27715	A	G	0.041252	sinónima	sf29	Non-core gene	p.Phe643Phe	-
27878	A	G	0.079134	sinsentido	sf29	Non-core gene	p.Leu589Ser	+
27937	G	A	0.042642	sinónima	sf29	Non-core gene	p.Ile569Ile	-
28031	T	C	0.072049	sinsentido	sf29	Non-core gene	p.Lys538Arg	-
28132	A	G	0.071355	sinónima	sf29	Non-core gene	p.Asn504Asn	-
28156	G	A	0.071912	sinónima	sf29	Non-core gene	p.Asn496Asn	-
28162	A	T	0.070563	sinónima	sf29	Non-core gene	p.Ile494Ile	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
28180	G	A	0.070196	sinónima	sf29	Non-core gene	p.Asn488Asn	-
28213	T	C	0.123917	sinónima	sf29	Non-core gene	p.Lys477Lys	-
28360	T	G	0.288335	sinónima	sf29	Non-core gene	p.Val428Val	-
28363	G	A	0.288101	sinónima	sf29	Non-core gene	p.Leu427Leu	-
28444	G	A	0.078264	sinónima	sf29	Non-core gene	p.Arg400Arg	-
28486	G	A	0.309071	sinónima	sf29	Non-core gene	p.Phe386Phe	-
28648	A	G	0.286029	sinónima	sf29	Non-core gene	p.Asn332Asn	-
28652	C	T	0.285023	sinsentido	sf29	Non-core gene	p.Arg331Lys	-
28887	C	T	0.218363	sinsentido	sf29	Non-core gene	p.Val253Ile	-
28907	C	G	0.217009	sinsentido	sf29	Non-core gene	p.Ser246Thr	-
29384	C	T	0.435962	sinsentido	sf29	Non-core gene	p.Arg87His	-
29507	T	C	0.108959	sinsentido	sf29	Non-core gene	p.His46Arg	-
29514	C	T	0.003512	sinsentido	sf29	Non-core gene	p.Val44Ile	-
29651	A	G	0.398228	intergénica	-	-	-	-
30287	G	A	0.430408	sinónima	sf30	Non-core gene	p.Leu178Leu	-
30310	T	C	0.354487	intergénica	-	-	-	-
30315	C	T	0.355499	intergénica	-	-	-	-
30320	A	G	0.351176	intergénica	-	-	-	-
30364	G	A	0.324765	intergénica	-	-	-	-
30372	T	C	0.324633	intergénica	-	-	-	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
30641	G	A	0.331372	sinónima	pkip-1	Non-core gene	p.Leu85Leu	-
30677	T	C	0.332477	sinónima	pkip-1	Non-core gene	p.Tyr97Tyr	-
30725	C	G	0.316994	sinónima	pkip-1	Non-core gene	p.Thr113Thr	-
30740	T	G	0.344544	sinónima	pkip-1	Non-core gene	p.Ala118Ala	-
31534	C	T	0.004094	sinsentido	arif-1	Non-core gene	p.Val264Ile	-
31537	C	T	0.012725	sinsentido	arif-1	Non-core gene	p.Ala263Thr	+
31769	C	T	0.267918	sinónima	arif-1	Non-core gene	p.Thr185Thr	-
31964	C	T	0.23013	sinónima	arif-1	Non-core gene	p.Thr120Thr	-
32173	T	C	0.274191	sinsentido	arif-1	Non-core gene	p.Thr51Ala	+
33167	G	A	0.243141	sinónima	pif-2	Core gene	p.Glu316Glu	-
33311	G	A	0.223268	sinónima	pif-2	Core gene	p.Ser364Ser	-
33386	C	T	0.21898	sinónima	pif-2	Core gene	p.Phe389Phe	-
33482	G	A	0.222625	sinsentido	pif-1	Core gene	p.Val17Ile	-
33985	C	T	0.429856	sinónima	pif-1	Core gene	p.Cys184Cys	-
34069	A	C	0.003378	sinónima	pif-1	Core gene	p.Ile212Ile	-
34783	C	T	0.419288	sinónima	pif-1	Core gene	p.Gly450Gly	-
34852	T	C	0.362472	sinónima	pif-1	Core gene	p.Asn473Asn	-
34975	T	C	0.007458	sinónima	pif-1	Core gene	p.Asn514Asn	-
35205	A	C	0.113127	sinónima	sf36	Non-core gene	p.Arg59Arg	-
35467	C	T	0.458846	sinsentido	fgf	Non-core gene	p.Val321Ile	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
35510	A	G	0.454448	sinónima	fgf	Non-core gene	p.Val306Val	-
35699	C	T	0.394421	sinónima	fgf	Non-core gene	p.Lys243Lys	-
35808	G	A	0.307102	sinsentido	fgf	Non-core gene	p.Ser207Leu	+
35870	G	T	0.452648	sinsentido	fgf	Non-core gene	p.Asp186Glu	-
36080	T	G	0.237535	sinónima	fgf	Non-core gene	p.Ala116Ala	-
36191	T	C	0.472331	sinónima	fgf	Non-core gene	p.Leu79Leu	-
36287	C	T	0.467874	sinónima	fgf	Non-core gene	p.Thr47Thr	-
36348	T	C	0.468627	sinsentido	fgf	Non-core gene	p.Asn27Ser	-
36565	T	A	0.003579	intergénica	-	-	-	-
36567	T	A	0.003926	intergénica	-	-	-	-
36625	A	G	0.320887	intergénica	-	-	-	-
36655	T	C	0.315082	intergénica	-	-	-	-
36849	G	A	0.179683	sinónima	sf38	Non-core gene	p.Ser60Ser	-
37083	T	C	0.073239	sinónima	sf38	Non-core gene	p.Thr138Thr	-
37110	C	T	0.07241	sinónima	sf38	Non-core gene	p.His147His	-
37146	A	C	0.052882	sinónima	sf38	Non-core gene	p.Ile159Ile	-
37383	G	T	0.003434	sinsentido	sf38	Non-core gene	p.Leu238Phe	-
37569	G	A	0.407699	sinónima	alk-exo	Core gene	p.Ile359Ile	-
37788	A	G	0.372487	sinónima	alk-exo	Core gene	p.Ser286Ser	-
37836	T	C	0.361678	sinónima	alk-exo	Core gene	p.Lys270Lys	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
37950	T	C	0.386699	sinónima	alk-exo	Core gene	p.Ser232Ser	-
38393	G	A	0.264549	sinónima	alk-exo	Core gene	p.Leu85Leu	-
38709	A	G	0.490534	intergénica	-	-	-	-
38718	C	T	0.466389	intergénica	-	-	-	-
38719	A	G	0.464598	intergénica	-	-	-	-
38751	A	T	0.479885	intergénica	-	-	-	-
38823	A	G	0.218346	intergénica	-	-	-	-
38827	T	G	0.217791	intergénica	-	-	-	-
38836	G	A	0.221082	intergénica	-	-	-	-
38991	A	C	0.206177	intergénica	-	-	-	-
39004	C	G	0.203746	intergénica	-	-	-	-
39013	T	C	0.180698	intergénica	-	-	-	-
39014	G	A	0.179394	intergénica	-	-	-	-
39033	C	T	0.214792	intergénica	-	-	-	-
39046	G	A	0.217527	intergénica	-	-	-	-
39058	C	T	0.217771	intergénica	-	-	-	-
39067	T	G	0.261583	intergénica	-	-	-	-
39075	G	T	0.211424	intergénica	-	-	-	-
39090	A	G	0.22165	intergénica	-	-	-	-
39101	C	T	0.221079	intergénica	-	-	-	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
39159	C	T	0.226253	intergénica	-	-	-	-
39172	G	C	0.223161	intergénica	-	-	-	-
39245	A	G	0.473333	sinónima	sf40	Non-core gene	p.Asp92Asp	-
39411	G	A	0.044465	sinsentido	sf40	Non-core gene	p.Ala37Val	-
40553	T	C	0.418162	sinónima	sf41	Non-core gene	p.Ile345Ile	-
40586	G	T	0.004846	sinsentido	sf41	Non-core gene	p.Trp356Cys	+
40587	G	T	0.005871	stop_gained	sf41	Non-core gene	p.Glu357*	-
40717	C	T	0.410701	sinsentido	sf42	Non-core gene	p.Asp115Asn	+
41137	C	A	0.005157	stop_gained	sf43	Non-core gene	p.Glu392*	-
41878	A	T	0.481501	sinsentido	sf43	Non-core gene	p.Leu145Ile	-
41953	G	T	0.484333	sinsentido	sf43	Non-core gene	p.Leu120Ile	-
42505	A	G	0.300087	intergénica	-	-	-	-
42507	A	C	0.299228	intergénica	-	-	-	-
42550	T	C	0.310803	sinónima	sf44	Non-core gene	p.Arg11Arg	-
42575	G	A	0.31287	sinsentido	sf44	Non-core gene	p.Gly20Ser	+
42583	G	A	0.312124	sinónima	sf44	Non-core gene	p.Lys22Lys	-
42649	C	T	0.268114	sinónima	sf44	Non-core gene	p.Ile44Ile	-
42956	G	A	0.343451	sinsentido	sf44	Non-core gene	p.Val147Ile	-
43743	T	G	0.003393	sinsentido	pep	Non-core gene	p.Asn194His	+
44297	T	A	0.009023	sinsentido	pep	Non-core gene	p.Gln9Leu	+

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
44314	C	A	0.006612	sinsentido	pep	Non-core gene	p.Leu3Phe	-
44318	G	A	0.00993	sinsentido	pep	Non-core gene	p.Ser2Leu	+
44320	C	A	0.007364	start_lost	pep	Non-core gene	p.Met1?	-
44327	T	A	0.005807	intergénica	-	-	-	-
44333	C	A	0.009545	intergénica	-	-	-	-
44358	C	A	0.004082	intergénica	-	-	-	-
44737	A	T	0.00209	sinsentido	sf47	Non-core gene	p.Cys74Ser	+
44738	A	C	0.00282	stop_gained	sf47	Non-core gene	p.Tyr73*	-
44750	A	G	0.40713	sinónima	sf47	Non-core gene	p.Pro69Pro	-
45206	T	C	0.399318	sinónima	sod	Non-core gene	p.Ser91Ser	-
45483	G	A	0.160319	intergénica	-	-	-	-
45769	A	C	0.186036	sinsentido	pif-3	Core gene	p.Lys75Gln	+
46375	A	G	0.171647	sinónima	sf50	Non-core gene	p.Val66Val	-
46495	T	C	0.414571	sinónima	sf50	Non-core gene	p.Ser106Ser	-
46525	T	C	0.412539	sinónima	sf50	Non-core gene	p.Tyr116Tyr	-
47026	A	G	0.262701	sinsentido	sf51	Non-core gene	p.Ile146Val	-
47625	C	T	0.157418	sinónima	sf51	Non-core gene	p.Phe345Phe	-
48172	G	A	0.456614	sinónima	sf52	Non-core gene	p.Glu2Glu	-
49034	T	C	0.386837	sinónima	sf53	Non-core gene	p.Lys310Lys	-
49316	T	C	0.375177	sinónima	sf53	Non-core gene	p.Lys216Lys	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
49640	G	A	0.361095	sinónima	sf53	Non-core gene	p.Phe108Phe	-
50155	T	C	0.2568	intergénica	-	-	-	-
50160	A	G	0.242752	intergénica	-	-	-	-
50169	A	C	0.216112	intergénica	-	-	-	-
50180	A	G	0.176967	intergénica	-	-	-	-
50185	T	C	0.160378	intergénica	-	-	-	-
50192	C	T	0.033181	intergénica	-	-	-	-
50197	A	G	0.005599	intergénica	-	-	-	-
50203	T	A	0.015817	intergénica	-	-	-	-
50208	G	C	0.043661	intergénica	-	-	-	-
50220	C	T	0.124647	intergénica	-	-	-	-
50221	A	G	0.126616	intergénica	-	-	-	-
50256	G	A	0.243595	intergénica	-	-	-	-
50264	A	G	0.265511	intergénica	-	-	-	-
50299	A	G	0.390639	intergénica	-	-	-	-
50326	A	G	0.370048	intergénica	-	-	-	-
50328	C	A	0.367239	intergénica	-	-	-	-
50334	T	G	0.358221	intergénica	-	-	-	-
50335	C	T	0.356471	intergénica	-	-	-	-
50336	A	G	0.359005	intergénica	-	-	-	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
50503	C	T	0.379786	intergénica	-	-	-	-
51093	A	T	0.380239	sinsentido	sf55	Non-core gene	p.Gln29His	+
51852	T	C	0.259138	sinónima	p13	Non-core gene	p.Ile53Ile	-
52622	T	C	0.28565	sinónima	odv-e66	Non-core gene	p.Phe12Phe	-
52631	C	T	0.289224	sinónima	odv-e66	Non-core gene	p.Phe15Phe	-
52728	A	T	0.006263	sinsentido	odv-e66	Non-core gene	p.Asn48Tyr	-
52737	A	T	0.003547	sinsentido	odv-e66	Non-core gene	p.Asn51Tyr	-
52757	T	C	0.028535	sinónima	odv-e66	Non-core gene	p.Asn57Asn	-
52760	A	C	0.160782	sinsentido	odv-e66	Non-core gene	p.Glu58Asp	-
52788	G	A	0.205476	sinsentido	odv-e66	Non-core gene	p.Asp68Asn	+
53222	T	C	0.377929	sinónima	odv-e66	Non-core gene	p.Tyr212Tyr	-
53363	G	A	0.455392	sinónima	odv-e66	Non-core gene	p.Lys259Lys	-
53507	T	C	0.459202	sinónima	odv-e66	Non-core gene	p.Leu307Leu	-
53642	C	T	0.424989	sinónima	odv-e66	Non-core gene	p.Ile352Ile	-
53656	A	G	0.425286	sinsentido	odv-e66	Non-core gene	p.Asp357Gly	+
53849	C	T	0.303646	sinónima	odv-e66	Non-core gene	p.Gly421Gly	-
53870	G	C	0.302763	sinónima	odv-e66	Non-core gene	p.Leu428Leu	-
54120	C	A	0.411667	sinsentido	odv-e66	Non-core gene	p.Leu512Ile	-
54239	C	T	0.443469	sinónima	odv-e66	Non-core gene	p.Asp551Asp	-
54275	T	C	0.440102	sinónima	odv-e66	Non-core gene	p.Ala563Ala	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
54281	G	A	0.441072	sinónima	odv-e66	Non-core gene	p.Gln565Gln	-
54340	T	C	0.436461	sinsentido	odv-e66	Non-core gene	p.Val585Ala	-
54368	T	G	0.44172	sinónima	odv-e66	Non-core gene	p.Thr594Thr	-
54458	G	A	0.330885	sinónima	odv-e66	Non-core gene	p.Val624Val	-
54461	T	C	0.329788	sinónima	odv-e66	Non-core gene	p.Leu625Leu	-
54532	G	A	0.331238	sinsentido	odv-e66	Non-core gene	p.Ser649Asn	-
54601	C	T	0.335931	sinsentido	odv-e66	Non-core gene	p.Ser672Leu	+
54638	T	C	0.336108	sinónima	odv-e66	Non-core gene	p.Phe684Phe	-
54682	T	C	0.334834	sinónima	sf58	Non-core gene	p.Pro111Pro	-
54751	T	G	0.005698	sinónima	sf58	Non-core gene	p.Ala88Ala	-
54784	G	A	0.330889	sinónima	sf58	Non-core gene	p.Phe77Phe	-
54789	T	G	0.004107	sinsentido	sf58	Non-core gene	p.Thr76Pro	+
55067	T	C	0.30062	sinónima	odv-ec43	Core gene	p.Ala349Ala	-
55094	T	C	0.23226	sinónima	odv-ec43	Core gene	p.Thr340Thr	-
56321	G	A	0.232109	sinónima	vp80	Non-core gene	p.Asp556Asp	-
56617	A	T	0.244473	sinsentido	vp80	Non-core gene	p.Phe458Ile	-
56821	T	C	0.243673	sinsentido	vp80	Non-core gene	p.Ile390Val	-
57588	C	A	0.244501	sinsentido	vp80	Non-core gene	p.Gly134Val	+
57643	C	T	0.33871	sinsentido	vp80	Non-core gene	p.Asp116Asn	+
57847	C	T	0.323614	sinsentido	vp80	Non-core gene	p.Val48Ile	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
57941	G	A	0.330743	sinónima	vp80	Non-core gene	p.Phe16Phe	-
58059	C	A	0.003507	sinsentido	p45	Core gene	p.Phe10Leu	-
58284	C	T	0.33102	sinónima	p45	Core gene	p.Phe85Phe	-
58692	T	C	0.329317	sinónima	p45	Core gene	p.Cys221Cys	-
58722	C	T	0.335355	sinónima	p45	Core gene	p.His231His	-
58800	T	C	0.3496	sinónima	p45	Core gene	p.Phe257Phe	-
58866	G	A	0.351791	sinónima	p45	Core gene	p.Val279Val	-
58959	C	T	0.35474	sinónima	p45	Core gene	p.Phe310Phe	-
58995	T	C	0.348658	sinónima	p45	Core gene	p.Tyr322Tyr	-
59281	C	T	0.005954	sinónima	p12	Non-core gene	p.Ala49Ala	-
59605	A	C	0.00269	sinsentido	p40	Core gene	p.Gln48Pro	+
59667	T	G	0.00375	sinsentido	p40	Core gene	p.Leu69Val	-
59681	T	C	0.00522	sinónima	p40	Core gene	p.Ala73Ala	-
59709	G	C	0.003516	sinsentido	p40	Core gene	p.Glu83Gln	+
59716	C	A	0.020429	sinsentido	p40	Core gene	p.Pro85Gln	+
60047	T	C	0.365078	sinónima	p40	Core gene	p.Asp195Asp	-
60053	C	T	0.364702	sinónima	p40	Core gene	p.Ile197Ile	-
60079	C	A	0.365208	sinsentido	p40	Core gene	p.Ser206Tyr	-
60357	T	C	0.345618	sinónima	p40	Core gene	p.Leu299Leu	-
60380	A	C	0.192186	sinónima	p40	Core gene	p.Ala306Ala	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
60564	A	G	0.191402	sinsentido	p40	Core gene	p.Ser368Gly	+
61017	C	T	0.477576	sinónima	lef-5	Core gene	p.Leu239Leu	-
61236	C	T	0.25257	sinónima	lef-5	Core gene	p.Pro166Pro	-
61822	C	T	0.429832	sinsentido	38k	Core gene	p.Ala64Val	-
62030	G	A	0.20195	sinónima	38k	Core gene	p.Thr133Thr	-
62402	T	C	0.460192	sinónima	38k	Core gene	p.Leu257Leu	-
62582	A	G	0.32858	sinsentido	sf68	Non-core gene	p.Ile13Val	-
62592	C	A	0.003489	sinsentido	sf68	Non-core gene	p.Thr16Lys	+
62791	G	A	0.449136	sinónima	sf68	Non-core gene	p.Pro82Pro	-
62799	A	T	0.449252	sinsentido	sf68	Non-core gene	p.Lys85Ile	+
62840	G	A	0.354685	sinsentido	sf68	Non-core gene	p.Val99Ile	-
62859	C	T	0.294688	sinsentido	sf68	Non-core gene	p.Ser105Leu	+
64029	G	T	0.00316	intergénica	-	-	-	-
64055	A	T	0.004147	intergénica	-	-	-	-
64659	G	A	0.341868	sinónima	pif-4	Core gene	p.Asn134Asn	-
64902	C	T	0.332821	sinónima	pif-4	Core gene	p.Ser53Ser	-
65138	T	C	0.22774	sinsentido	dnahel	Core gene	p.Val36Ala	-
65496	C	T	0.178261	sinónima	dnahel	Core gene	p.Tyr155Tyr	-
65655	C	T	0.431847	sinónima	dnahel	Core gene	p.Gly208Gly	-
65658	C	A	0.43098	sinónima	dnahel	Core gene	p.Arg209Arg	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
66096	T	C	0.236695	sinónima	dnahel	Core gene	p.Ile355Ile	-
66603	C	T	0.469387	sinónima	dnahel	Core gene	p.Thr524Thr	-
66861	G	A	0.198759	sinónima	dnahel	Core gene	p.Ser610Ser	-
67219	C	T	0.004185	sinsentido	dnahel	Core gene	p.His730Tyr	+
67473	C	T	0.006323	sinónima	dnahel	Core gene	p.Asn814Asn	-
67476	C	T	0.028531	sinónima	dnahel	Core gene	p.Asp815Asp	-
68007	T	C	0.203988	sinónima	dnahel	Core gene	p.Asp992Asp	-
69521	A	G	0.078452	sinónima	p18	Core gene	p.Ile134Ile	-
69673	C	A	0.003539	sinsentido	p18	Core gene	p.Asp84Tyr	+
69815	G	A	0.0763	sinónima	p18	Core gene	p.Ile36Ile	-
69821	G	A	0.076298	sinónima	p18	Core gene	p.Gly34Gly	-
70268	C	T	0.108604	sinónima	p33	Core gene	p.Phe115Phe	-
70330	T	C	0.103005	sinsentido	p33	Core gene	p.Ile136Thr	+
70331	T	C	0.103157	sinónima	p33	Core gene	p.Ile136Ile	-
70439	G	A	0.102449	sinónima	p33	Core gene	p.Ser172Ser	-
70455	C	T	0.107719	sinónima	p33	Core gene	p.Leu178Leu	-
70674	C	T	0.160938	sinónima	p33	Core gene	p.Leu251Leu	-
70701	T	A	0.160646	intergénica	-	-	-	-
70787	A	C	0.002343	sinsentido	lef-4	Core gene	p.Ile440Ser	+
70864	T	C	0.197234	sinónima	lef-4	Core gene	p.Glu414Glu	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
71008	C	T	0.205857	sinónima	lef-4	Core gene	p.Pro366Pro	-
71194	A	G	0.356794	sinónima	lef-4	Core gene	p.His304His	-
71647	A	G	0.280348	sinónima	lef-4	Core gene	p.Ile153Ile	-
73143	C	T	0.286475	intergénica	-	-	-	-
73148	A	T	0.004625	intergénica	-	-	-	-
73244	A	T	0.300489	intergénica	-	-	-	-
73425	C	T	0.340257	sinónima	cg30	Non-core gene	p.Cys41Cys	-
74262	A	T	0.448196	sinsentido	cg30	Non-core gene	p.Glu320Asp	-
74373	C	T	0.446745	sinónima	cg30	Non-core gene	p.Asp357Asp	-
74394	C	A	0.006434	sinónima	cg30	Non-core gene	p.Ser364Ser	-
74811	G	A	0.443551	sinónima	pif-8	Core gene	p.Cys799Cys	-
74945	C	T	0.480537	sinsentido	pif-8	Core gene	p.Val755Met	-
75237	A	C	0.044948	sinsentido	pif-8	Core gene	p.Asp657Glu	-
75257	T	C	0.044129	sinsentido	pif-8	Core gene	p.Ser651Gly	+
76515	G	T	0.095182	sinónima	pif-8	Core gene	p.Leu231Leu	-
77034	C	T	0.119219	sinónima	pif-8	Core gene	p.Gln58Gln	-
77920	C	T	0.160928	sinónima	ac81-like	Core gene	p.Phe116Phe	-
78345	G	T	0.189776	sinsentido	gp41	Core gene	p.Ala25Ser	+
78351	A	G	0.16719	sinsentido	gp41	Core gene	p.Thr27Ala	+
78353	A	G	0.163959	sinónima	gp41	Core gene	p.Thr27Thr	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
78354	A	G	0.165931	sinsentido	gp41	Core gene	p.Thr28Ala	+
78416	C	T	0.202666	sinónima	gp41	Core gene	p.Tyr48Tyr	-
78479	C	T	0.201705	sinónima	gp41	Core gene	p.Asn69Asn	-
78590	A	G	0.194661	sinónima	gp41	Core gene	p.Pro106Pro	-
79190	C	T	0.002216	sinónima	gp41	Core gene	p.Thr306Thr	-
79846	C	T	0.148243	sinónima	vlf-1	Core gene	p.Leu68Leu	-
79872	T	C	0.146148	sinónima	vlf-1	Core gene	p.Ile76Ile	-
80574	C	T	0.148839	sinónima	vlf-1	Core gene	p.Tyr310Tyr	-
80839	T	C	0.147832	intergénica	-	-	-	-
80949	C	T	0.153803	intergénica	-	-	-	-
81104	C	T	0.150691	intergénica	-	-	-	-
81113	G	A	0.150871	intergénica	-	-	-	-
81123	T	C	0.151379	intergénica	-	-	-	-
81206	G	A	0.153501	sinónima	sf85	Non-core gene	p.Leu51Leu	-
81229	G	A	0.151441	sinsentido	sf85	Non-core gene	p.Ser43Leu	+
81272	T	C	0.146783	sinsentido	sf85	Non-core gene	p.Asn29Asp	+
81335	A	C	0.180825	sinsentido	sf85	Non-core gene	p.Tyr8Asp	+
81359	A	C	0.004488	intergénica	-	-	-	-
81366	T	A	0.480329	intergénica	-	-	-	-
81373	T	C	0.126049	intergénica	-	-	-	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
81375	C	A	0.123302	intergénica	-	-	-	-
81378	T	G	0.122671	intergénica	-	-	-	-
81380	G	A	0.125711	intergénica	-	-	-	-
81411	A	T	0.008479	intergénica	-	-	-	-
81447	C	T	0.001979	intergénica	-	-	-	-
81454	C	T	0.191125	intergénica	-	-	-	-
81465	C	T	0.189883	intergénica	-	-	-	-
81516	T	G	0.003262	sinsentido	p26a	Non-core gene	p.Lys238Thr	+
81545	A	G	0.181923	sinónima	p26a	Non-core gene	p.Asn228Asn	-
81548	G	A	0.18261	sinónima	p26a	Non-core gene	p.Ile227Ile	-
81611	A	G	0.185822	sinónima	p26a	Non-core gene	p.His206His	-
81656	G	A	0.185335	sinónima	p26a	Non-core gene	p.Asp191Asp	-
81665	T	C	0.184925	sinónima	p26a	Non-core gene	p.Val188Val	-
81911	T	C	0.181336	sinónima	p26a	Non-core gene	p.Arg106Arg	-
82019	C	T	0.01066	sinónima	p26a	Non-core gene	p.Lys70Lys	-
82027	C	T	0.409756	sinsentido	p26a	Non-core gene	p.Asp68Asn	+
82157	C	T	0.011673	sinónima	p26a	Non-core gene	p.Thr24Thr	-
82810	C	T	0.19371	sinónima	iap-0	Non-core gene	p.Glu110Glu	-
82969	C	T	0.374843	sinónima	iap-0	Non-core gene	p.Arg57Arg	-
83390	C	T	0.367149	sinónima	mtase-1	Non-core gene	p.Ser151Ser	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
83486	T	G	0.194633	sinónima	mtase-1	Non-core gene	p.Val119Val	-
83552	G	A	0.211241	sinónima	mtase-1	Non-core gene	p.Phe97Phe	-
83606	G	A	0.222998	sinónima	mtase-1	Non-core gene	p.Asn79Asn	-
84573	G	A	0.12988	sinónima	lef-3	Non-core gene	p.Glu122Glu	-
84591	G	A	0.136717	sinónima	lef-3	Non-core gene	p.Lys128Lys	-
84617	A	C	0.006381	sinsentido	lef-3	Non-core gene	p.Glu137Ala	+
84723	C	T	0.132367	sinónima	lef-3	Non-core gene	p.Asn172Asn	-
84865	A	G	0.134896	sinsentido	lef-3	Non-core gene	p.Ile220Val	-
84885	C	T	0.136754	sinónima	lef-3	Non-core gene	p.Ile226Ile	-
85395	G	A	0.170382	intergénica	-	-	-	-
85415	G	A	0.16427	intergénica	-	-	-	-
85663	C	T	0.067384	sinsentido	desmoplakin	Core gene	p.Glu638Lys	+
86324	A	G	0.111203	sinónima	desmoplakin	Core gene	p.Ala417Ala	-
86561	T	C	0.192928	sinónima	desmoplakin	Core gene	p.Glu338Glu	-
86663	A	G	0.161624	sinónima	desmoplakin	Core gene	p.Ala304Ala	-
86906	A	G	0.249482	sinónima	desmoplakin	Core gene	p.Ser223Ser	-
86993	T	C	0.258877	sinónima	desmoplakin	Core gene	p.Thr194Thr	-
87176	C	T	0.434187	sinónima	desmoplakin	Core gene	p.Pro133Pro	-
87218	C	T	0.43189	sinónima	desmoplakin	Core gene	p.Gln119Gln	-
87725	C	T	0.473852	sinónima	dnapol	Core gene	p.Phe50Phe	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
87770	G	A	0.501638	sinónima	dnapol	Core gene	p.Glu65Glu	-
87779	T	C	0.505997	sinónima	dnapol	Core gene	p.Cys68Cys	-
87929	C	T	0.484145	sinónima	dnapol	Core gene	p.Ser118Ser	-
87932	C	T	0.483608	sinónima	dnapol	Core gene	p.Asp119Asp	-
87941	C	T	0.482098	sinónima	dnapol	Core gene	p.Asn122Asn	-
87944	C	T	0.481776	sinónima	dnapol	Core gene	p.Asn123Asn	-
87968	T	C	0.493228	sinónima	dnapol	Core gene	p.Leu131Leu	-
87980	C	T	0.497352	sinónima	dnapol	Core gene	p.Asn135Asn	-
88004	A	G	0.495971	sinónima	dnapol	Core gene	p.Leu143Leu	-
88034	A	G	0.411305	sinónima	dnapol	Core gene	p.Ser153Ser	-
88439	C	T	0.480711	sinónima	dnapol	Core gene	p.Gly288Gly	-
88484	C	T	0.469078	sinónima	dnapol	Core gene	p.Val303Val	-
88521	G	A	0.47033	sinsentido	dnapol	Core gene	p.Ala316Thr	+
88547	T	C	0.472372	sinónima	dnapol	Core gene	p.Ile324Ile	-
88853	T	C	0.477393	sinónima	dnapol	Core gene	p.Leu426Leu	-
89000	G	A	0.506622	sinónima	dnapol	Core gene	p.Lys475Lys	-
89597	C	T	0.15744	sinónima	dnapol	Core gene	p.Asn674Asn	-
90149	T	C	0.443978	sinónima	dnapol	Core gene	p.Arg858Arg	-
90182	T	C	0.441727	sinónima	dnapol	Core gene	p.Arg869Arg	-
90200	A	G	0.442802	sinónima	dnapol	Core gene	p.Ser875Ser	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
90308	G	A	0.485741	sinónima	dnapol	Core gene	p.Gln911Gln	-
90539	C	T	0.033841	sinónima	dnapol	Core gene	p.Asp988Asp	-
90572	G	A	0.022077	sinónima	dnapol	Core gene	p.Glu999Glu	-
90581	T	A	0.007464	sinsentido	dnapol	Core gene	p.Asp1002Glu	-
90593	A	G	0.242817	sinónima	dnapol	Core gene	p.Val1006Val	-
90812	A	G	0.157345	sinónima	sf93	Non-core gene	p.Tyr102Tyr	-
90815	G	A	0.158104	sinónima	sf93	Non-core gene	p.Asp101Asp	-
91007	C	T	0.163277	sinónima	sf93	Non-core gene	p.Arg37Arg	-
91281	G	A	0.161316	sinónima	sf94	Non-core gene	p.Leu41Leu	-
91314	G	A	0.25933	sinónima	sf94	Non-core gene	p.Leu30Leu	-
92106	T	C	0.47182	sinónima	sf96	Non-core gene	p.Tyr89Tyr	-
92127	A	G	0.474592	sinónima	sf96	Non-core gene	p.Lys96Lys	-
92166	A	G	0.507336	sinónima	sf96	Non-core gene	p.Lys109Lys	-
93093	T	G	0.006598	sinsentido	lef-9	Core gene	p.Asp266Ala	+
94355	A	C	0.003034	sinsentido	fp25k	Non-core gene	p.Glu136Asp	-
94571	T	A	0.144046	intergénica	-	-	-	-
94572	T	A	0.142547	intergénica	-	-	-	-
95005	G	A	0.019024	sinsentido	94k	Non-core gene	p.Asp94Asn	+
95161	A	C	0.012475	sinsentido	94k	Non-core gene	p.Thr146Pro	+
96045	G	T	0.002922	sinsentido	94k	Non-core gene	p.Met440Ile	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
96169	C	T	0.003867	sinónima	94k	Non-core gene	p.Leu482Leu	-
96189	T	C	0.004153	sinónima	94k	Non-core gene	p.Asp488Asp	-
96228	C	T	0.004472	sinónima	94k	Non-core gene	p.Ser501Ser	-
96253	G	A	0.003887	sinsentido	94k	Non-core gene	p.Val510Ile	-
96489	A	G	0.276797	sinónima	94k	Non-core gene	p.Lys588Lys	-
96750	A	G	0.27983	sinónima	94k	Non-core gene	p.Arg675Arg	-
97452	A	G	0.390274	sinónima	sf101	Non-core gene	p.Leu77Leu	-
97482	A	G	0.388382	sinónima	sf101	Non-core gene	p.Pro87Pro	-
97551	C	T	0.34499	sinónima	sf101	Non-core gene	p.Asn110Asn	-
97759	G	A	0.086599	sinsentido	sf101	Non-core gene	p.Ala180Thr	+
97928	T	G	0.081788	sinsentido	sf102	Non-core gene	p.Glu117Asp	-
97943	T	C	0.081666	sinónima	sf102	Non-core gene	p.Lys112Lys	-
99337	T	G	0.004739	sinsentido	vp1054	Core gene	p.Gln254His	+
99358	T	G	0.003573	sinónima	vp1054	Core gene	p.Arg247Arg	-
99412	C	T	0.353756	sinónima	vp1054	Core gene	p.Lys229Lys	-
99781	G	C	0.217991	sinónima	vp1054	Core gene	p.Arg106Arg	-
99946	G	T	0.234547	sinsentido	vp1054	Core gene	p.Asp51Glu	-
99955	A	G	0.231541	sinónima	vp1054	Core gene	p.Pro48Pro	-
100744	C	T	0.140005	sinónima	sf108	Non-core gene	p.Ile118Ile	-
100864	C	A	0.003806	sinsentido	sf108	Non-core gene	p.Asn158Lys	+

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
101561	A	G	0.387304	sinónima	ac53-like	Core gene	p.Leu50Leu	-
102152	T	G	0.003054	sinsentido	sf110	Non-core gene	p.Ile130Arg	+
102487	A	G	0.004426	sinónima	sf110a	Non-core gene	p.Ser106Ser	-
102500	A	T	0.007727	sinsentido	sf110a	Non-core gene	p.Met102Lys	+
102503	A	G	0.008893	sinsentido	sf110a	Non-core gene	p.Leu101Ser	+
102516	A	C	0.008977	sinsentido	sf110a	Non-core gene	p.Cys97Gly	+
102672	T	A	0.003462	sinsentido	sf110a	Non-core gene	p.Asn45Tyr	-
102678	C	G	0.003909	sinsentido	sf110a	Non-core gene	p.Asp43His	+
102700	T	C	0.002266	sinónima	sf110a	Non-core gene	p.Thr35Thr	-
102734	T	C	0.09367	sinsentido	sf110a	Non-core gene	p.Tyr24Cys	+
102787	G	T	0.094978	stop_gained	sf110a	Non-core gene	p.Cys6*	-
102949	T	C	0.003011	intergénica	-	-	-	-
102955	T	C	0.029278	intergénica	-	-	-	-
102978	G	A	0.035517	intergénica	-	-	-	-
103066	G	A	0.09443	intergénica	-	-	-	-
103289	T	C	0.092239	sinónima	iap-3	Non-core gene	p.Ala28Ala	-
103894	A	G	0.212861	sinsentido	iap-3	Non-core gene	p.Lys230Arg	-
104070	G	T	0.004508	intergénica	-	-	-	-
104748	T	G	0.002881	sinsentido	bjdp	Non-core gene	p.Asn187Thr	-
104854	A	G	0.29014	sinónima	bjdp	Non-core gene	p.Leu152Leu	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
105125	T	C	0.274057	sinónima	bjdp	Non-core gene	p.Thr61Thr	-
105933	G	T	0.005338	sinsentido	lef-8	Core gene	p.Leu202Phe	-
105960	C	T	0.410035	sinónima	lef-8	Core gene	p.Ile211Ile	-
106141	T	C	0.263404	sinónima	lef-8	Core gene	p.Leu272Leu	-
106155	G	A	0.26391	sinónima	lef-8	Core gene	p.Glu276Glu	-
106170	T	C	0.262087	sinónima	lef-8	Core gene	p.Ile281Ile	-
106782	C	T	0.128402	sinónima	lef-8	Core gene	p.Tyr485Tyr	-
107379	C	T	0.037552	sinónima	lef-8	Core gene	p.Tyr684Tyr	-
107754	G	A	0.122189	sinónima	lef-8	Core gene	p.Pro809Pro	-
107766	T	C	0.121025	sinónima	lef-8	Core gene	p.Phe813Phe	-
107859	T	C	0.12613	sinónima	lef-8	Core gene	p.Asn844Asn	-
108020	G	A	0.125891	intergénica	-	-	-	-
108556	T	C	0.232266	sinsentido	odv-e66b	Non-core gene	p.Thr605Ala	+
108570	G	A	0.232145	sinsentido	odv-e66b	Non-core gene	p.Ser600Leu	+
108664	T	C	0.003707	sinsentido	odv-e66b	Non-core gene	p.Asn569Asp	+
108682	A	C	0.171742	sinsentido	odv-e66b	Non-core gene	p.Tyr563Asp	+
108685	T	A	0.164589	sinsentido	odv-e66b	Non-core gene	p.Asn562Tyr	-
108695	A	G	0.015003	sinónima	odv-e66b	Non-core gene	p.Asn558Asn	-
108842	G	A	0.218221	sinónima	odv-e66b	Non-core gene	p.Val509Val	-
108920	A	G	0.213753	sinónima	odv-e66b	Non-core gene	p.Phe483Phe	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
109069	T	G	0.005129	sinsentido	odv-e66b	Non-core gene	p.Asn434His	+
110305	A	G	0.465129	sinónima	odv-e66b	Non-core gene	p.Leu22Leu	-
110563	A	G	0.353634	sinsentido	p47	Core gene	p.Ser56Gly	+
110763	G	A	0.467074	sinónima	p47	Core gene	p.Pro122Pro	-
110764	G	A	0.466832	sinsentido	p47	Core gene	p.Asp123Asn	+
110835	G	A	0.413186	sinónima	p47	Core gene	p.Leu146Leu	-
111003	C	T	0.002329	sinónima	p47	Core gene	p.Asp202Asp	-
111033	G	A	0.307072	sinónima	p47	Core gene	p.Lys212Lys	-
111415	T	C	0.477776	sinónima	p47	Core gene	p.Leu340Leu	-
111422	A	C	0.004488	sinsentido	p47	Core gene	p.Lys342Thr	+
111423	A	C	0.003493	sinsentido	p47	Core gene	p.Lys342Asn	+
111513	C	T	0.497497	sinónima	p47	Core gene	p.Asp372Asp	-
111659	A	T	0.007392	intergénica	-	-	-	-
111662	A	T	0.005583	intergénica	-	-	-	-
111883	A	G	0.487529	sinónima	sf117	Non-core gene	p.Lys68Lys	-
112246	A	G	0.354867	sinónima	sf117	Non-core gene	p.Ala189Ala	-
112247	A	G	0.354077	sinsentido	sf117	Non-core gene	p.Thr190Ala	+
112371	A	G	0.480963	sinónima	sf118	Non-core gene	p.Leu6Leu	-
112444	G	A	0.477317	sinsentido	sf118	Non-core gene	p.Val31Ile	-
112569	C	T	0.409946	sinónima	sf118	Non-core gene	p.Tyr72Tyr	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
112725	T	C	0.413098	sinónima	sf118	Non-core gene	p.Asn124Asn	-
112870	A	G	0.477472	sinsentido	sf118	Non-core gene	p.Ile173Val	-
113282	C	T	0.174068	sinónima	adprase	Non-core gene	p.Gly103Gly	-
113321	A	G	0.210277	sinónima	adprase	Non-core gene	p.Pro116Pro	-
114144	C	T	0.228095	sinsentido	39k	Non-core gene	p.Leu57Phe	-
114278	T	C	0.230703	sinónima	39k	Non-core gene	p.Asn101Asn	-
114670	T	C	0.476104	sinsentido	39k	Non-core gene	p.Ile232Thr	+
114940	G	T	0.002089	sinsentido	sf122	Non-core gene	p.Val3Phe	-
114990	C	T	0.461114	sinónima	sf122	Non-core gene	p.Pro19Pro	-
115083	C	T	0.452217	sinónima	sf122	Non-core gene	p.Gly50Gly	-
115119	C	T	0.444624	sinónima	sf122	Non-core gene	p.Asp62Asp	-
115152	T	C	0.416646	sinónima	sf122	Non-core gene	p.Phe73Phe	-
115153	C	A	0.418172	sinsentido	sf122	Non-core gene	p.Leu74Ile	-
115167	A	G	0.445316	sinónima	sf122	Non-core gene	p.Glu78Glu	-
115218	C	T	0.005133	sinónima	sf122	Non-core gene	p.Tyr95Tyr	-
115222	T	A	0.010412	sinsentido	sf122	Non-core gene	p.Leu97Met	-
115278	A	T	0.007789	intergénica	-	-	-	-
115287	A	T	0.004557	intergénica	-	-	-	-
115540	G	A	0.497475	sinsentido	v-ubi	Non-core gene	p.Thr77Ile	+
115545	G	A	0.496689	sinónima	v-ubi	Non-core gene	p.Gly75Gly	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
115835	C	G	0.233209	sinsentido	sf125	Non-core gene	p.Asn8Lys	+
116432	C	T	0.059342	intergénica	-	-	-	-
116528	T	C	0.355287	intergénica	-	-	-	-
116576	A	C	0.482846	intergénica	-	-	-	-
116585	T	A	0.037924	intergénica	-	-	-	-
116591	C	A	0.337771	intergénica	-	-	-	-
116604	A	T	0.31138	intergénica	-	-	-	-
116700	A	G	0.301947	intergénica	-	-	-	-
116797	G	A	0.299534	sinsentido	sf126	Non-core gene	p.Pro122Leu	+
116800	A	G	0.29909	sinsentido	sf126	Non-core gene	p.Leu121Pro	+
117899	G	A	0.08841	sinónima	dbp	Non-core gene	p.Ala215Ala	-
118139	A	G	0.109541	sinónima	dbp	Non-core gene	p.Gln295Gln	-
118428	C	T	0.454993	sinónima	lef-6	Non-core gene	p.Ser55Ser	-
118458	C	T	0.457585	sinónima	lef-6	Non-core gene	p.Gly65Gly	-
118490	A	G	0.464437	sinsentido	lef-6	Non-core gene	p.Lys76Arg	-
118501	C	T	0.484682	sinsentido	lef-6	Non-core gene	p.His80Tyr	+
118506	T	C	0.457517	sinónima	lef-6	Non-core gene	p.His81His	-
118518	C	T	0.41684	sinónima	lef-6	Non-core gene	p.His85His	-
118519	C	T	0.4195	sinsentido	lef-6	Non-core gene	p.His86Tyr	+
118522	C	T	0.012227	sinsentido	lef-6	Non-core gene	p.His87Tyr	+

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
118531	A	C	0.00543	sinónima	lef-6	Non-core gene	p.Arg90Arg	-
118586	C	T	0.393074	sinsentido	lef-6	Non-core gene	p.Ser108Leu	+
119616	G	A	0.467575	sinónima	p26b	Non-core gene	p.Glu165Glu	-
119679	C	T	0.47263	sinónima	p26b	Non-core gene	p.Ile186Ile	-
119684	A	T	0.471695	sinsentido	p26b	Non-core gene	p.Tyr188Phe	+
120094	C	T	0.414834	sinónima	p10	Non-core gene	p.Asn49Asn	-
120127	A	G	0.410182	sinónima	p10	Non-core gene	p.Leu60Leu	-
120175	T	C	0.401561	sinónima	p10	Non-core gene	p.Asn76Asn	-
120190	A	G	0.39588	sinónima	p10	Non-core gene	p.Leu81Leu	-
120369	G	A	0.401701	sinónima	sf133	Non-core gene	p.Thr288Thr	-
120842	C	T	0.398583	sinsentido	sf133	Non-core gene	p.Glu131Lys	+
121009	T	A	0.238451	sinsentido	sf133	Non-core gene	p.Tyr75Phe	+
121023	A	G	0.229363	sinónima	sf133	Non-core gene	p.Asn70Asn	-
121026	G	A	0.229826	sinónima	sf133	Non-core gene	p.Asp69Asp	-
121077	A	G	0.245803	sinónima	sf133	Non-core gene	p.Ile52Ile	-
121122	G	A	0.249173	sinónima	sf133	Non-core gene	p.His37His	-
121466	C	T	0.291815	sinsentido	pif-0	Core gene	p.Ala576Thr	+
121476	C	T	0.180743	sinónima	pif-0	Core gene	p.Leu572Leu	-
121512	C	T	0.182381	sinónima	pif-0	Core gene	p.Leu560Leu	-
121581	C	T	0.185906	sinónima	pif-0	Core gene	p.Val537Val	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
121676	G	A	0.180646	sinónima	pif-0	Core gene	p.Leu506Leu	-
121782	A	G	0.194921	sinónima	pif-0	Core gene	p.Ala470Ala	-
121887	A	G	0.183347	sinónima	pif-0	Core gene	p.Ile435Ile	-
121949	T	C	0.127512	sinsentido	pif-0	Core gene	p.Ile415Val	-
122220	C	T	0.211252	sinónima	pif-0	Core gene	p.Ala324Ala	-
122325	A	G	0.310602	sinónima	pif-0	Core gene	p.Leu289Leu	-
122369	C	G	0.310041	sinsentido	pif-0	Core gene	p.Asp275His	+
122375	C	T	0.309932	sinsentido	pif-0	Core gene	p.Asp273Asn	+
122432	T	G	0.069615	sinsentido	pif-0	Core gene	p.Thr254Pro	+
122457	C	T	0.29377	sinónima	pif-0	Core gene	p.Pro245Pro	-
123134	A	G	0.144336	sinsentido	pif-0	Core gene	p.Tyr20His	+
123217	A	T	0.001818	intergénica	-	-	-	-
123249	C	T	0.19248	sinónima	sf135	Non-core gene	p.Phe2Phe	-
123363	G	A	0.28926	sinónima	sf135	Non-core gene	p.Ser40Ser	-
123372	G	C	0.414429	sinsentido	sf135	Non-core gene	p.Glu43Asp	-
123889	G	A	0.042645	sinónima	ie-1	Non-core gene	p.Leu592Leu	-
123914	A	G	0.042813	sinónima	ie-1	Non-core gene	p.Thr583Thr	-
124154	T	C	0.043318	sinónima	ie-1	Non-core gene	p.Lys503Lys	-
124268	G	A	0.04617	sinónima	ie-1	Non-core gene	p.Phe465Phe	-
125148	T	C	0.048989	sinsentido	ie-1	Non-core gene	p.Asn172Ser	-

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
125279	T	C	0.049788	sinónima	ie-1	Non-core gene	p.Ser128Ser	-
125285	T	C	0.05012	sinónima	ie-1	Non-core gene	p.Ser126Ser	-
125505	T	G	0.003589	sinsentido	ie-1	Non-core gene	p.His53Pro	+
125550	T	G	0.001727	sinsentido	ie-1	Non-core gene	p.Lys38Thr	+
125621	C	T	0.038587	sinónima	ie-1	Non-core gene	p.Leu14Leu	-
125881	A	G	0.072167	sinsentido	sf137	Non-core gene	p.Arg64Gly	+
126222	T	G	0.43673	sinónima	sf137	Non-core gene	p.Ser177Ser	-
126371	C	T	0.47539	intergénica	-	-	-	-
128085	G	A	0.377424	sinónima	p49	Core gene	p.Phe360Phe	-
128963	G	A	0.028338	sinónima	p49	Core gene	p.Leu68Leu	-
129075	A	G	0.049325	sinónima	p49	Core gene	p.Asp30Asp	-
129200	C	T	0.050061	sinónima	ie-0	Non-core gene	p.Gln230Gln	-
129448	T	G	0.004963	sinsentido	ie-0	Non-core gene	p.Lys148Gln	+
129449	T	G	0.005914	sinsentido	ie-0	Non-core gene	p.Glu147Asp	-
129497	G	A	0.052867	sinónima	ie-0	Non-core gene	p.Ile131Ile	-
129647	T	C	0.25323	sinónima	ie-0	Non-core gene	p.Leu81Leu	-
130234	C	T	0.11643	sinónima	rr-1	Non-core gene	p.Thr753Thr	-
130318	A	G	0.114966	sinónima	rr-1	Non-core gene	p.Asn725Asn	-
131959	C	T	0.157303	sinónima	rr-1	Non-core gene	p.Glu178Glu	-
132102	G	T	0.003654	sinsentido	rr-1	Non-core gene	p.Pro131Thr	+

Posición	Referencia	Variante	Frecuencia	Clase	Gen	Clase_gen	Cambio_AA	Cambio_polaridad
132117	C	T	0.148816	sinsentido	rr-1	Non-core gene	p.Glu126Lys	+
132223	C	T	0.153576	sinónima	rr-1	Non-core gene	p.Gln90Gln	-
132242	C	T	0.155009	sinsentido	rr-1	Non-core gene	p.Ser84Asn	-
132322	G	C	0.152427	sinsentido	rr-1	Non-core gene	p.Asn57Lys	+
132599	T	G	0.003296	intergénica	-	-	-	-
132607	T	G	0.0048	intergénica	-	-	-	-

Referencias

1. Montezano, D. G. *et al.* Host plants of *spodoptera frugiperda* (lepidoptera: Noctuidae) in the americas. *African Entomology* **26**, 286–300 (2018).
2. Sparks, A. N. A review of the biology of the fall armyworm. *The Florida Entomologist* **62**, 82 (1979).
3. Johnson, S. J. Migration and the life history strategy of the fall armyworm, *spodoptera frugiperda* in the western hemisphere. *International Journal of Tropical Insect Science* **8**, 543–549 (1987).
4. Goergen, G., Kumar, P. L., Sankung, S. B., Togola, A. & Tamò, M. First report of outbreaks of the fall armyworm *spodoptera frugiperda* (j e smith) (lepidoptera, noctuidae), a new alien invasive pest in west and central africa. *PLOS ONE* **11**, e0165632 (2016).
5. Jing, D.-P. *et al.* Initial detections and spread of invasive *spodoptera frugiperda* in china and comparisons with other noctuid larvae in cornfields using molecular techniques. *Insect Science* **27**, 780–790 (2019).
6. Day, R. *et al.* Fall armyworm: Impacts and implications for africa. *Outlooks on Pest Management* **28**, 196–201 (2017).
7. Murúa, M. G., Juárez, M. L., Prieto, S., Gastaminza, G. & Willink, E. Distribución temporal y espacial de poblaciones larvarias de *spodoptera frugiperda* (smith) (lep.: Noctuidae) en diferentes hospederos en provincias del norte de la argentina. *Revista industrial y agricola de Tucuman* **86**, e46146 (2009).

-
8. Al-Sarar, A., Hall, F. R. & Downer, R. A. Impact of spray application methodology on the development of resistance to cypermethrin and spinosad by fall armyworm *spodoptera frugiperda* (JE smith). *Pest Management Science* **62**, 1023–1031 (2006).
 9. Carvalho, R. A., Omoto, C., Field, L. M., Williamson, M. S. & Bass, C. Investigating the molecular mechanisms of organophosphate and pyrethroid resistance in the fall armyworm *spodoptera frugiperda*. *PLoS ONE* **8**, e62268 (2013).
 10. Bolzan, A. *et al.* Selection and characterization of the inheritance of resistance of *spodoptera frugiperda* (lepidoptera: Noctuidae) to chlorantraniliprole and cross-resistance to other diamide insecticides. *Pest Management Science* **75**, 2682–2689 (2019).
 11. Lira, E. C. *et al.* Resistance of *spodoptera frugiperda* (lepidoptera: Noctuidae) to spinetoram: Inheritance and cross-resistance to spinosad. *Pest Management Science* **76**, 2674–2680 (2020).
 12. Lacey, L. A. *et al.* Insect pathogens as biological control agents: Back to the future. *Journal of Invertebrate Pathology* **132**, 1–41 (2015).
 13. Wu, K.-M., Lu, Y.-H., Feng, H.-Q., Jiang, Y.-Y. & Zhao, J.-Z. Suppression of cotton bollworm in multiple crops in china in areas with bt toxin-containing cotton. *Science* **321**, 1676–1678 (2008).
 14. Hutchison, W. D. *et al.* Areawide suppression of european corn borer with bt maize reaps savings to non-bt maize growers. *Science* **330**, 222–225 (2010).
 15. Wan, P. *et al.* Hybridizing transgenic bt cotton with non-bt cotton counters resistance in pink bollworm. *Proceedings of the National Academy of Sciences* **114**, 5413–5418 (2017).
 16. Yang, F., Kerns, D. L., Head, G. P., Price, P. & Huang, F. Cross-resistance to purified bt proteins, bt corn and bt cotton in a Cry2Ab2-corn resistant strain of *spodoptera frugiperda*. *Pest Management Science* **73**, 2495–2503 (2017).

-
17. Huang, F. Resistance of the fall armyworm, *spodoptera frugiperda* , to transgenic bacillus thuringiensis Cry1F corn in the americas: Lessons and implications for bt corn IRM in china. *Insect Science* (2020) doi:10.1111/1744-7917.12826.
 18. Burtet, L. M. *et al.* Managing fall armyworm, *spodoptera frugiperda* (lepidoptera: Noctuidae), with bt maize and insecticides in southern brazil. *Pest Management Science* **73**, 2569–2577 (2017).
 19. Brewer, M. J. & Goodell, P. B. Approaches and incentives to implement integrated pest management that addresses regional and environmental issues. *Annual Review of Entomology* **57**, 41–59 (2012).
 20. Haase, S., Sciocco-Cap, A. & Romanowski, V. Baculovirus insecticides in latin america: Historical overview, current status and future perspectives. *Viruses* **7**, 2230–2267 (2015).
 21. Stenberg, J. A. A conceptual framework for integrated pest management. *Trends in Plant Science* **22**, 759–769 (2017).
 22. Fira, D., Dimkić, I., Berić, T., Lozo, J. & Stanković, S. Biological control of plant pathogens by bacillus species. *Journal of Biotechnology* **285**, 44–55 (2018).
 23. Bamisile, B. S., Dash, C. K., Akutse, K. S., Keppanan, R. & Wang, L. Fungal endophytes: Beyond herbivore management. *Frontiers in Microbiology* **9**, (2018).
 24. Lu, D., Baiocchi, T. & Dillman, A. R. Genomics of entomopathogenic nematodes and implications for pest control. *Trends in Parasitology* **32**, 588–598 (2016).
 25. Szewczyk, B., Hoyos-Carvajal, L., Paluszek, M., Skrzecz, I. & Souza, M. L. de. Baculoviruses — re-emerging biopesticides. *Biotechnology Advances* **24**, 143–160 (2006).
 26. Moscardi, F. ASSESSMENT OF THE APPLICATION OF BACULOVIRUSES FOR CONTROL OF LEPIDOPTERA. *Annual Review of Entomology* **44**, 257–289 (1999).

-
27. Popham, H. J. R., Nusawardani, T. & Bonning, B. C. Introduction to the use of baculoviruses as biological insecticides. in *Methods in molecular biology* 383–392 (Springer New York, 2016). doi:10.1007/978-1-4939-3043-2_19.
 28. Kogan, M., Turnipseed, S. G., Shepard, M., Oliveira, E. B. D. & Borgo, A. Pilot insect pest management program for soybean in southern Brazil. *Journal of Economic Entomology* **70**, 659–663 (1977).
 29. Moscardi, F. Use of viruses for pest control in brazil: The case of the nuclear polyhedrosis virus of the soybean caterpillar, *anticarsia gemmatalis*. *Memórias do Instituto Oswaldo Cruz* **84**, 51–56 (1989).
 30. Lacey, L. A., Vail, P. V. & Hoffmann, D. F. Comparative activity of baculoviruses against the codling moth *cydia pomonella* and three other tortricid pests of tree fruit. *Journal of Invertebrate Pathology* **80**, 64–68 (2002).
 31. Alletti, G. *et al.* Using next generation sequencing to identify and quantify the genetic composition of resistance-breaking commercial isolates of *cydia pomonella* granulovirus. *Viruses* **9**, 250 (2017).
 32. Asser-Kaiser, S. *et al.* Rapid emergence of baculovirus resistance in codling moth due to dominant, sex-linked inheritance. *Science* **317**, 1916–1918 (2007).
 33. Simón, O., Williams, T., Lopez-Ferber, M. & Caballero, P. Genetic structure of a *spodoptera frugiperda* nucleopolyhedrovirus population: High prevalence of deletion genotypes. *Applied and Environmental Microbiology* **70**, 5579–5588 (2004).
 34. Barrera, G., Williams, T., Villamizar, L., Caballero, P. & Simón, O. Deletion genotypes reduce occlusion body potency but increase occlusion body production in a colombian *spodoptera frugiperda* nucleopolyhedrovirus population. *PLoS ONE* **8**, e77271 (2013).
 35. Gomez, J., Guevara, J., Cuartas, P., Espinel, C. & Villamizar, L. Microencapsulated *spodoptera frugiperda* nucleopolyhedrovirus: Insecticidal activity and effect on arthropod populations in maize. *Biocontrol Science and Technology* **23**, 829–846 (2013).
-

-
36. Zamora-Avilés, N. *et al.* Effects of a nucleopolyhedrovirus in mixtures with azadirachtin on *Spodoptera frugiperda* (J. E. Smith) (Lepidoptera: Noctuidae) larvae and viral occlusion body production. *Biocontrol Science and Technology* **23**, 521–534 (2013).
 37. Harrison, R. L. *et al.* ICTV virus taxonomy profile: baculoviridae. *Journal of General Virology* **99**, 1185–1186 (2018).
 38. Thézé, J., Lopez-Vaamonde, C., Cory, J. & Herniou, E. Biodiversity, evolution and ecological specialization of baculoviruses: A treasure trove for future applied research. *Viruses* **10**, 366 (2018).
 39. Rohrmann, G. F. *Baculovirus molecular biology 4th edition*. (National Center for Biotechnology Information (US), 2019).
 40. Oers, M. van & Vlak, J. Baculovirus genomics. *Current Drug Targets* **8**, 1051–1068 (2007).
 41. Ferrelli, M. L. *et al.* The baculoviral genome. in *Viral genomes - molecular structure, diversity, gene expression mechanisms and host-virus interactions* (InTech, 2012). doi:10.5772/32209.
 42. Garavaglia, M. J., Miele, S. A. B., Iserte, J. A., Belaich, M. N. & Ghiringhelli, P. D. The ac53, ac78, ac101, and ac103 genes are newly discovered core genes in the family baculoviridae. *Journal of Virology* **86**, 12069–12079 (2012).
 43. Javed, M. A. *et al.* *Autographa californica* multiple nucleopolyhedrovirus AC83 is a per os infectivity factor (PIF) protein required for occlusion-derived virus (ODV) and budded virus nucleocapsid assembly as well as assembly of the PIF complex in ODV envelopes. *Journal of Virology* **91**, (2016).
 44. Zhao, S., He, G., Yang, Y. & Liang, C. Nucleocapsid assembly of baculoviruses. *Viruses* **11**, 595 (2019).
 45. Wang, Q. *et al.* Budded baculovirus particle structure revisited. *Journal of Invertebrate Pathology* **134**, 15–22 (2016).
-

-
46. Slack, J. & Arif, B. M. The baculoviruses occlusion-derived virus: Virion structure and function. in *Advances in virus research* 99–165 (Elsevier, 2006). doi:10.1016/s0065-3527(06)69003-9.
 47. Coulibaly, F. *et al.* The atomic structure of baculovirus polyhedra reveals the independent emergence of infectious crystals in DNA and RNA viruses. *Proceedings of the National Academy of Sciences* **106**, 22205–22210 (2009).
 48. Ji, X. *et al.* How baculovirus polyhedra fit square pegs into round holes to robustly package viruses. *The EMBO Journal* **29**, 505–514 (2009).
 49. Gati, C. *et al.* Atomic structure of granulin determined from native nanocrystalline granulovirus using an x-ray free-electron laser. *Proceedings of the National Academy of Sciences* **114**, 2247–2252 (2017).
 50. Rohrmann, G. F. Baculovirus nucleocapsid aggregation (MNPV vs SNPV): An evolutionary strategy, or a product of replication conditions? *Virus Genes* **49**, 351–357 (2014).
 51. Wang, Y., Kleespies, R. G., Huger, A. M. & Jehle, J. A. The genome of gryllus bimaculatus nudivirus indicates an ancient diversification of baculovirus-related nonoccluded nudiviruses of insects. *Journal of Virology* **81**, 5395–5406 (2007).
 52. Theze, J., Bezier, A., Periquet, G., Drezen, J.-M. & Herniou, E. A. Paleozoic origin of insect large dsDNA viruses. *Proceedings of the National Academy of Sciences* **108**, 15931–15935 (2011).
 53. Jehle, J. A. *et al.* Molecular identification and phylogenetic analysis of baculoviruses from lepidoptera. *Virology* **346**, 180–193 (2006).
 54. Thiem, S. M. & Miller, L. K. Identification, sequence, and transcriptional mapping of the major capsid protein gene of the baculovirus autographa californica nuclear polyhedrosis virus. *Journal of Virology* **63**, 2008–2018 (1989).

-
55. Goley, E. D. *et al.* Dynamic nuclear actin assembly by Arp2/3 complex and a baculovirus WASP-like protein. *Science* **314**, 464–467 (2006).
 56. Wang, Y. *et al.* Identification of a novel regulatory sequence of actin nucleation promoting factor encoded by autographa californica multiple nucleopolyhedrovirus. *Journal of Biological Chemistry* **290**, 9533–9541 (2015).
 57. Lai, Q. *et al.* The 38K-mediated specific dephosphorylation of the viral core protein P6.9 plays an important role in the nucleocapsid assembly of autographa californica multiple nucleopolyhedrovirus. *Journal of Virology* **92**, (2018).
 58. Marek, M., Merten, O.-W., Galibert, L., Vlak, J. M. & Oers, M. M. van. Baculovirus VP80 protein and the f-actin cytoskeleton interact and connect the viral replication factory with the nuclear periphery. *Journal of Virology* **85**, 5350–5362 (2011).
 59. Guan, Z. *et al.* The autographa californica multiple Nucleopolyhedrovirus ac54 Gene is crucial for localization of the major capsid protein VP39 at the site of nucleocapsid assembly. *Journal of Virology* **90**, 4115–4126 (2016).
 60. Wolgamot, G. M., Gross, C. H., Russell, R. L. Q. & Rohrmann, G. F. Immunocytochemical characterization of p24, a baculovirus capsid-associated protein. *Journal of General Virology* **74**, 103–107 (1993).
 61. Tweeten, K. A., Bulla, L. A. & Consigli, R. A. Characterization of an extremely basic protein derived from granulosis virus nucleocapsids†. *Journal of Virology* **33**, 866–876 (1980).
 62. Braunagel, S. C., Russell, W. K., Rosas-Acosta, G., Russell, D. H. & Summers, M. D. Determination of the protein composition of the occlusion-derived virus of autographa californica nucleopolyhedrovirus. *Proceedings of the National Academy of Sciences* **100**, 9797–9802 (2003).
 63. Wang, R. *et al.* Proteomics of the autographa californica nucleopolyhedrovirus budded virions. *Journal of Virology* **84**, 7233–7242 (2010).
-

-
64. Hou, D. *et al.* Comparative proteomics reveal fundamental structural and functional differences between the two progeny phenotypes of a baculovirus. *Journal of Virology* **87**, 829–839 (2012).
65. Blissard, G. W. & Theilmann, D. A. Baculovirus entry and egress from insect cells. *Annual Review of Virology* **5**, 113–139 (2018).
66. Deng, F. *et al.* Proteomics analysis of *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus identified two new occlusion-derived virus-associated proteins, HA44 and HA100. *Journal of Virology* **81**, 9377–9385 (2007).
67. Xu, F., Ince, İ. A., Boeren, S., Vlak, J. M. & Oers, M. M. van. Protein composition of the occlusion derived virus of *Chrysodeixis chalcites* nucleopolyhedrovirus. *Virus Research* **158**, 1–7 (2011).
68. Wang, X.-F. *et al.* ODV-associated proteins of the *Pieris rapae* Granulovirus. *Journal of Proteome Research* **10**, 2817–2827 (2011).
69. Braconi, C. T. *et al.* Proteomic analyses of *Baculovirus anticarsia gemmatalis* multiple nucleopolyhedrovirus budded and occluded virus. *Journal of General Virology* **95**, 980–989 (2014).
70. Zhang, X., Liang, Z., Yin, X. & Shao, X. Proteomic analysis of the occlusion-derived virus of *Clostera anachoreta* granulovirus. *Journal of General Virology* **96**, 2394–2404 (2015).
71. Hou, D., Chen, X. & Zhang, L.-K. Proteomic analysis of *Mamestra brassicae* nucleopolyhedrovirus progeny virions from two different hosts. *PLOS ONE* **11**, e0153365 (2016).
72. Masson, T., Fabre, M. L., Ferrelli, M. L., Pidre, M. L. & Romanowski, V. Protein composition of the occlusion bodies of *Epinotia aporema* granulovirus. *PLOS ONE* **14**, e0207735 (2019).
73. Li, Y. *et al.* The functional oligomeric state of tegument protein GP41 is essential for baculovirus budded virion and occlusion-derived virion assembly. *Journal of Virology* **92**, (2018).
-

-
74. Wang, X. *et al.* Baculovirus per os infectivity factor complex: Components and assembly. *Journal of Virology* **93**, (2019).
75. Boogaard, B., Oers, M. V. & Lent, J. V. An advanced view on baculovirus per os infectivity factors. *Insects* **9**, 84 (2018).
76. Myers, J. H. & Cory, J. S. Ecology and evolution of pathogens in natural populations of lepidoptera. *Evolutionary Applications* **9**, 231–247 (2015).
77. Williams, T., Virto, C., Murillo, R. & Caballero, P. Covert infection of insects by baculoviruses. *Frontiers in Microbiology* **8**, (2017).
78. Wang, P. & Granados, R. R. An intestinal mucin is the target substrate for a baculovirus enhancin. *Proceedings of the National Academy of Sciences* **94**, 6977–6982 (1997).
79. Slavicek, J. M. & Popham, H. J. R. The lymantria dispar nucleopolyhedrovirus enhancins are components of occlusion-derived virus. *Journal of Virology* **79**, 10578–10588 (2005).
80. Haase, S. Desarrollo de herramientas moleculares para incrementar la aplicabilidad del baculovirus de anticarsia gemmatalis en el control biológico de plagas. (Universidad Nacional de La Plata, 2015). doi:10.35537/10915/48298.
81. Ohkawa, T., Volkman, L. E. & Welch, M. D. Actin-based motility drives baculovirus transit to the nucleus and cell surface. *Journal of Cell Biology* **190**, 187–195 (2010).
82. Nagamine, T., Kawasaki, Y., Abe, A. & Matsumoto, S. Nuclear marginalization of host cell chromatin associated with expansion of two discrete virus-induced subnuclear compartments during baculovirus infection. *Journal of Virology* **82**, 6409–6418 (2008).
83. Kong, X. *et al.* Dynamic chromatin accessibility profiling reveals changes in host genome organization in response to baculovirus infection. *PLOS Pathogens* **16**, e1008633 (2020).
84. Engelhard, E. K., Kam-Morgan, L. N., Washburn, J. O. & Volkman, L. E. The insect tracheal system: A conduit for the systemic spread of autographa californica nuclear polyhedrosis virus. *Proceedings of the National Academy of Sciences* **91**, 3224–3227 (1994).
-

-
85. Wang, M. *et al.* Unraveling the entry mechanism of baculoviruses and its evolutionary implications. *Journal of Virology* **88**, 2301–2311 (2013).
 86. Granados, R. R. & Lawler, K. A. In vivo pathway of autographa californica baculovirus invasion and infection. *Virology* **108**, 297–308 (1981).
 87. Braunagel, S. & Summers, M. Molecular biology of the baculovirus occlusion-derived virus envelope. *Current Drug Targets* **8**, 1084–1095 (2007).
 88. Ishimwe, E., Hodgson, J. J., Clem, R. J. & Passarelli, A. L. Reaching the melting point: Degradative enzymes and protease inhibitors involved in baculovirus infection and dissemination. *Virology* **479-480**, 637–649 (2015).
 89. Berretta, F. M., Ferrelli, M. L., Salvador, R., Sciocco-Cap, A. & Romanowski, V. Baculovirus gene expression. in *Current issues in molecular virology - viral genetics and biotechnological applications* (InTech, 2013). doi:10.5772/56955.
 90. Harrison, R. L., Puttler, B. & Popham, H. J. R. Genomic sequence analysis of a fast-killing isolate of spodoptera frugiperda multiple nucleopolyhedrovirus. *Journal of General Virology* **89**, 775–790 (2008).
 91. Ferrelli, M. L. *et al.* Genomic analysis of an argentinean isolate of spodoptera frugiperda granulovirus reveals that various baculoviruses code for lef-7 proteins with three f-box domains. *PLOS ONE* **13**, e0202598 (2018).
 92. Niz, J. M. *et al.* Genetic variants in argentinean isolates of spodoptera frugiperda multiple nucleopolyhedrovirus. *Virus Genes* **56**, 401–405 (2020).
 93. Simón, O., Palma, L., Williams, T., López-Ferber, M. & Caballero, P. 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 (2012).

-
94. Barrera, G. P., Belaich, M. N., Patarroyo, M. A., Villamizar, L. F. & Ghiringhelli, P. D. Evidence of recent interspecies horizontal gene transfer regarding nucleopolyhedrovirus infection of *Spodoptera frugiperda*. *BMC Genomics* **16**, (2015).
 95. Cuartas, P. *et al.* The complete sequence of the first *Spodoptera frugiperda* betabaculovirus genome: A natural multiple recombinant virus. *Viruses* **7**, 394–421 (2015).
 96. Pidre, M. L., Sabalette, K. B., Romanowski, V. & Ferrelli, M. L. Identification of an argentinean isolate of *Spodoptera frugiperda* granulovirus. *Revista Argentina de Microbiología* **51**, 381–385 (2019).
 97. Behle, R. W. & Popham, H. J. R. Laboratory and field evaluations of the efficacy of a fast-killing baculovirus isolate from *Spodoptera frugiperda*. *Journal of Invertebrate Pathology* **109**, 194–200 (2012).
 98. Vieira, C. M., Tuelher, E. S., Valicente, F. H. & Wolff, J. L. C. Characterization of a *Spodoptera frugiperda* multiple nucleopolyhedrovirus isolate that does not liquefy the integument of infected larvae. *Journal of Invertebrate Pathology* **111**, 189–192 (2012).
 99. Martínez, A.-M. *et al.* Formulation with an optical brightener does not increase probability of developing resistance to *Spodoptera frugiperda* nucleopolyhedrovirus in the laboratory. *Journal of Economic Entomology* **97**, 1202–1208 (2004).
 100. Villamizar, L., Barrera, G., Cotes, A. M. & Martínez, F. Eudragit S100 microparticles containing *Spodoptera frugiperda* nucleopolyhedrovirus: Physico-chemical characterization, photostability and in vitro virus release. *Journal of Microencapsulation* **27**, 314–324 (2009).
 101. Cuartas-Otálora, P. E., Gómez-Valderrama, J. A., Ramos, A. E., Barrera-Cubillos, G. P. & Villamizar-Rivero, L. F. Bio-insecticidal potential of nucleopolyhedrovirus and granulovirus mixtures to control the fall armyworm *Spodoptera frugiperda* (J.E. Smith, 1797) (Lepidoptera: Noctuidae). *Viruses* **11**, 684 (2019).
 102. Miller, L. K. & Dawes, K. P. Restriction endonuclease analysis for the identification of baculovirus pesticides. *Applied and Environmental Microbiology* **35**, 411–421 (1978).
-

-
103. Jehle, J. A., Fritsch, E., Huber, J. & Backhaus, H. Intra-specific and inter-specific recombination of tortricid-specific granuloviruses during co-infection in insect larvae. *Archives of Virology* **148**, 1317–1333 (2003).
104. López-Ferber, M., Simón, O., Williams, T. & Caballero, P. Defective or effective? Mutualistic interactions between virus genotypes. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **270**, 2249–2255 (2003).
105. Ladner, J. T. *et al.* Standards for sequencing viral genomes in the era of high-throughput sequencing. *mBio* **5**, (2014).
106. Merwe, M. van der *et al.* Genome analysis and genetic stability of the cryptophlebia leucotreta granulovirus (CrleGV-SA) after 15 years of commercial use as a biopesticide. *International Journal of Molecular Sciences* **18**, 2327 (2017).
107. Moldován, N. *et al.* Third-generation sequencing reveals extensive polycistronism and transcriptional overlapping in a baculovirus. *Scientific Reports* **8**, (2018).
108. Nguyen, T. T. *et al.* Comprehensive analysis of single molecule sequencing-derived complete genome and whole transcriptome of hyposidra talaca nuclear polyhedrosis virus. *Scientific Reports* **8**, (2018).
109. Renner, D. W. & Szpara, M. L. Impacts of genome-wide analyses on our understanding of human herpesvirus diversity and evolution. *Journal of Virology* **92**, (2017).
110. Chateigner, A. *et al.* Ultra deep sequencing of a baculovirus population reveals widespread genomic variations. *Viruses* **7**, 3625–3646 (2015).
111. Zwart *et al.* Identification of loci associated with enhanced virulence in spodoptera litura nucleopolyhedrovirus isolates using deep sequencing. *Viruses* **11**, 872 (2019).
112. Brito, A. F. de *et al.* The pangenome of the anticarsia gemmatalis multiple nucleopolyhedrovirus (AgMNPV). *Genome Biology and Evolution* **8**, 94–108 (2015).
-

-
113. Wennmann, J., Radtke, P., Eberle, K., Alletti, G. G. & Jehle, J. Deciphering single nucleotide polymorphisms and evolutionary trends in isolates of the cydia pomonella granulovirus. *Viruses* **9**, 227 (2017).
114. Brito, A. F. *et al.* Genome-wide diversity in temporal and regional populations of the betabaculovirus erinnyis ello granulovirus (ErelGV). *BMC Genomics* **19**, (2018).
115. Ishihama, Y. *et al.* Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Molecular & Cellular Proteomics* **4**, 1265–1272 (2005).
116. Sciocco-Cap, A., Parola, A. D., Goldberg, A. V., Ghiringhelli, P. D. & Romanowski, V. Characterization of a granulovirus isolated from epinotia aporema wals. (Lepidoptera: Tortricidae) larvae. *Applied and Environmental Microbiology* **67**, 3702–3706 (2001).
117. Ferrelli, M. L. *et al.* Genome of epinotia aporema granulovirus (EpapGV), a polyorganotropic fast killing betabaculovirus with a novel thymidylate kinase gene. *BMC Genomics* **13**, 548 (2012).
118. Sánchez, N. E. & Pereyra, P. C. Neotropical soybean budborer, crocidosema aporema (walsingham, 1914) (lepidoptera: tortricidae). in *Encyclopedia of entomology* (ed. Capinera, J. L.) 2587–2588 (Springer Netherlands, 2008). doi:10.1007/978-1-4020-6359-6_2186.
119. Shapiro, M. Effect of two granulosis viruses on the activity of the gypsy moth (lepidoptera: Lymantriidae) nuclear polyhedrosis virus. *Journal of Economic Entomology* **93**, 1633–1637 (2000).
120. Miele, S. A. B., Garavaglia, M. J., Belaich, M. N. & Ghiringhelli, P. D. Baculovirus: Molecular insights on their diversity and conservation. *International Journal of Evolutionary Biology* **2011**, 1–15 (2011).
121. Harrison, R. L. *et al.* The complete genome sequence of a second distinct betabaculovirus from the true armyworm, mythimna unipuncta. *PLOS ONE* **12**, e0170510 (2017).
-

-
122. Greco, T. M., Diner, B. A. & Cristea, I. M. The impact of mass spectrometry-based proteomics on fundamental discoveries in virology. *Annual Review of Virology* **1**, 581–604 (2014).
123. Nesvizhskii, A. I. Proteogenomics: Concepts, applications and computational strategies. *Nature Methods* **11**, 1114–1125 (2014).
124. Arike, L. & Peil, L. Spectral counting label-free proteomics. in *Shotgun proteomics* 213–222 (Springer New York, 2014). doi:10.1007/978-1-4939-0685-7_14.
125. Rauniyar, N. & Yates, J. R. Isobaric labeling-based relative quantification in shotgun proteomics. *Journal of Proteome Research* **13**, 5293–5309 (2014).
126. Chen, X., Wei, S., Ji, Y., Guo, X. & Yang, F. Quantitative proteomics using SILAC: Principles, applications, and developments. *PROTEOMICS* **15**, 3175–3192 (2015).
127. Tan, Y., Bideshi, D. K., Johnson, J. J., Bigot, Y. & Federici, B. A. Proteomic analysis of the spodoptera frugiperda ascovirus 1a virion reveals 21 proteins. *Journal of General Virology* **90**, 359–365 (2009).
128. Vidick, S. *et al.* Proteomic characterization of murid herpesvirus 4 extracellular virions. *PLoS ONE* **8**, e83842 (2013).
129. İnce, İ. A., Boeren, S. A., Oers, M. M. van, Vervoort, J. J. M. & Vlak, J. M. Proteomic analysis of chilo iridescent virus. *Virology* **405**, 253–258 (2010).
130. Bézier, A., Harichaux, G., Musset, K., Labas, V. & Herniou, E. A. Qualitative proteomic analysis of tipula oleracea nudivirus occlusion bodies. *Journal of General Virology* **98**, 284–295 (2017).
131. Doellinger, J., Schaade, L. & Nitsche, A. Comparison of the cowpox virus and vaccinia virus mature virion proteome: Analysis of the species- and strain-specific proteome. *PLOS ONE* **10**, e0141527 (2015).

-
132. Perera, O., Green, T. B., Stevens, S. M., White, S. & Becnel, J. J. Proteins associated with culex nigripalpus nucleopolyhedrovirus occluded virions. *Journal of Virology* **81**, 4585–4590 (2007).
133. Zhu, Z. *et al.* Genome sequence and analysis of buzura suppressaria nucleopolyhedrovirus: A group II alphabaculovirus. *PLoS ONE* **9**, e86450 (2014).
134. Escasa, S. R., Harrison, R. L., Mowery, J. D., Bauchan, G. R. & Cory, J. S. The complete genome sequence of an alphabaculovirus from spodoptera exempta, an agricultural pest of major economic significance in africa. *PLOS ONE* **14**, e0209937 (2019).
135. Salvador, R. *et al.* Analysis of EpapGV gp37 gene reveals a close relationship between granulovirus and entomopoxvirus. *Virus Genes* **45**, 610–613 (2012).
136. Liu, X. *et al.* Granulovirus GP37 facilitated ODVs cross insect peritrophic membranes and fuse with epithelia. *Toxins* **11**, 145 (2019).
137. Braunagel, S. C., He, H., Ramamurthy, P. & Summers, M. D. Transcription, translation, and cellular localization of ThreeAutographa californicaNuclear polyhedrosis virus structural proteins: ODV-E18, ODV-E35, and ODV-EC27. *Virology* **222**, 100–114 (1996).
138. McCarthy, C. B. & Theilmann, D. A. AcMNPV ac143 (odv-e18) is essential for mediating budded virus production and is the 30th baculovirus core gene. *Virology* **375**, 277–291 (2008).
139. Camacho, C. *et al.* BLAST+: Architecture and applications. *BMC Bioinformatics* **10**, 421 (2009).
140. Eddy, S. R. Accelerated profile HMM searches. *PLoS Computational Biology* **7**, e1002195 (2011).
141. Li, Y. The mitotic spindle is required for loading of the DASH complex onto the kinetochore. *Genes & Development* **16**, 183–197 (2002).
142. Westermann, S. *et al.* Formation of a dynamic kinetochore- microtubule interface through assembly of the Dam1 ring complex. *Molecular Cell* **17**, 277–290 (2005).
-

-
143. Shi, A. *et al.* Autographa californica nucleopolyhedrovirus ac75 is required for the nuclear egress of nucleocapsids and intranuclear microvesicle formation. *Journal of Virology* JVI.01509–17 (2017) doi:10.1128/jvi.01509-17.
144. Yuan, M. *et al.* Identification of autographa californica nucleopolyhedrovirus ac93 as a core gene and its requirement for intranuclear microvesicle formation and nuclear egress of nucleocapsids. *Journal of Virology* **85**, 11664–11674 (2011).
145. Tao, X. Y. *et al.* The autographa californica multiple nucleopolyhedrovirus ORF78 is essential for budded virus production and general occlusion body formation. *Journal of Virology* **87**, 8441–8450 (2013).
146. Dong, F., Wang, J., Deng, R. & Wang, X. Autographa californica multiple nucleopolyhedrovirus gene ac81 is required for nucleocapsid envelopment. *Virus Research* **221**, 47–57 (2016).
147. McCarthy, C. B., Dai, X., Donly, C. & Theilmann, D. A. Autographa californica multiple nucleopolyhedrovirus ac142, a core gene that is essential for BV production and ODV envelopment. *Virology* **372**, 325–339 (2008).
148. Chen, L. *et al.* Autographa californica multiple nucleopolyhedrovirus odv-e25 (Ac94) is required for budded virus infectivity and occlusion-derived virus formation. *Archives of Virology* **157**, 617–625 (2012).
149. Alfonso, V. *et al.* AcMNPV core gene ac109 is required for budded virion transport to the nucleus and for occlusion of viral progeny. *PLoS ONE* **7**, e46146 (2012).
150. Wu, W. & Passarelli, A. L. Autographa californica multiple nucleopolyhedrovirus Ac92 (ORF92, P33) is required for budded virus production and multiply enveloped occlusion-derived virus formation. *Journal of Virology* **84**, 12351–12361 (2010).
151. Yuan, M. *et al.* A highly conserved baculovirus gene p48 (ac103) is essential for BV production and ODV envelopment. *Virology* **379**, 87–96 (2008).
-

-
152. Leroy, B., Gillet, L., Vanderplassen, A. & Wattiez, R. Structural proteomics of herpesviruses. *Viruses* **8**, 50 (2016).
153. Escasa, S. R., Lauzon, H. A. M., Mathur, A. C., Krell, P. J. & Arif, B. M. Sequence analysis of the choristoneura occidentalis granulovirus genome. *Journal of General Virology* **87**, 1917–1933 (2006).
154. Yamamoto, T. & Tanada, Y. Physicochemical properties and location of capsule components, in particular the synergistic factor, in the occlusion body of a granulosis virus of the armyworm, *pseudaletia unipuncta*. *Virology* **107**, 434–440 (1980).
155. Nan, H., Chen, H., Tuite, M. F. & Xu, X. A viral expression factor behaves as a prion. *Nature Communications* **10**, (2019).
156. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **72**, 248–254 (1976).
157. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403–410 (1990).
158. Fernandez-Escamilla, A.-M., Rousseau, F., Schymkowitz, J. & Serrano, L. Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nature Biotechnology* **22**, 1302–1306 (2004).
159. Stokstad, E. New crop pest takes africa at lightning speed. *Science* **356**, 473–474 (2017).
160. Silver, A. Caterpillar’s devastating march across china spurs hunt for native predator. *Nature* **570**, 286–287 (2019).
161. Bernardi, D. *et al.* Cross-resistance between Cry1 proteins in fall armyworm (*spodoptera frugiperda*) may affect the durability of current pyramided bt maize hybrids in brazil. *PLOS ONE* **10**, e0140130 (2015).
-

-
162. Harrison, R. D. *et al.* Agro-ecological options for fall armyworm (*spodoptera frugiperda* JE smith) management: Providing low-cost, smallholder friendly solutions to an invasive pest. *Journal of Environmental Management* **243**, 318–330 (2019).
163. Bentivenha, J. P. F. *et al.* Baseline susceptibility of *spodoptera frugiperda* (lepidoptera: Noctuidae) to SfMNPV and evaluation of cross-resistance to major insecticides and bt proteins. *Journal of Economic Entomology* **112**, 91–98 (2018).
164. Wolff, J. L. C., Valicente, F. H., Martins, R., Castro Oliveira, J. V. de & Andrade Zanotto, P. M. de. 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 (2008).
165. Simón, O. *et al.* Sequence comparison between three geographically distinct *spodoptera frugiperda* multiple nucleopolyhedrovirus isolates: Detecting positively selected genes. *Journal of Invertebrate Pathology* **107**, 33–42 (2011).
166. Mixtures of complete and pif1- and pif2- deficient genotypes are required for increased potency of an insect nucleopolyhedrovirus. *Journal of Virology* **83**, 5127–5136 (2009).
167. Simón, O., Williams, T., López-Ferber, M. & Caballero, P. Deletion of egt is responsible for the fast-killing phenotype of natural deletion genotypes in a *spodoptera frugiperda* multiple nucleopolyhedrovirus population. *Journal of Invertebrate Pathology* **111**, 260–263 (2012).
168. Lei, C., Yang, J., Wang, J., Hu, J. & Sun, X. Molecular and biological characterization of *spodoptera frugiperda* multiple nucleopolyhedrovirus field isolate and genotypes from china. *Insects* **11**, 777 (2020).
169. Obbard, D. J. & Dudas, G. The genetics of host–virus coevolution in invertebrates. *Current Opinion in Virology* **8**, 73–78 (2014).
170. Holmes, E. C. What does virus evolution tell us about virus origins? *Journal of Virology* **85**, 5247–5251 (2011).
-

-
171. Hill, T. & Unckless, R. L. Baculovirus molecular evolution via gene turnover and recurrent positive selection of key genes. *Journal of Virology* **91**, (2017).
172. Berretta, M. F., Rios, M. L. & Cap, A. S. de. Characterization of a nuclear polyhedrosis virus of *spodoptera frugiperda* from argentina. *Journal of Invertebrate Pathology* **71**, 280–282 (1998).
173. Loiseau, V. *et al.* Wide spectrum and high frequency of genomic structural variation, including transposable elements, in large double-stranded DNA viruses. *Virus Evolution* **6**, (2020).
174. Serrano, A. *et al.* Analogous population structures for two alphabaculoviruses highlight a functional role for deletion mutants. *Applied and Environmental Microbiology* **79**, 1118–1125 (2012).
175. Gilbert, C. *et al.* Population genomics supports baculoviruses as vectors of horizontal transfer of insect transposons. *Nature Communications* **5**, (2014).
176. Simoń, O. *et al.* Sf29 gene of *spodoptera frugiperda* multiple nucleopolyhedrovirus is a viral factor that determines the number of virions in occlusion bodies. *Journal of Virology* **82**, 7897–7904 (2008).
177. Rawlings, N. D., Waller, M., Barrett, A. J. & Bateman, A. MEROPS: The database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Research* **42**, D503–D509 (2013).
178. Geoghegan, J. L. & Holmes, E. C. The phylogenomics of evolving virus virulence. *Nature Reviews Genetics* **19**, 756–769 (2018).
179. Beperet, I., Simoń, O., Williams, T., López-Ferber, M. & Caballero, P. The ‘11K’ gene family members sf68, sf95 and sf138 modulate transmissibility and insecticidal properties of *spodoptera frugiperda* multiple nucleopolyhedrovirus. *Journal of Invertebrate Pathology* **127**, 101–109 (2015).
-

-
180. Daugherty, M. D. & Malik, H. S. Rules of engagement: Molecular insights from host-virus arms races. *Annual Review of Genetics* **46**, 677–700 (2012).
181. Wang, W., Zhao, H. & Han, G.-Z. Host-virus arms races drive elevated adaptive evolution in viral receptors. *Journal of Virology* **94**, (2020).
182. Rezelj, V. V., Levi, L. I. & Vignuzzi, M. The defective component of viral populations. *Current Opinion in Virology* **33**, 74–80 (2018).
183. Leeks, A., Segredo-Otero, E. A., Sanjuán, R. & West, S. A. Beneficial coinfection can promote within-host viral diversity. *Virus Evolution* **4**, (2018).
184. Thézé, J., Takatsuka, J., Nakai, M., Arif, B. & Herniou, E. Gene acquisition convergence between entomopoxviruses and baculoviruses. *Viruses* **7**, 1960–1974 (2015).
185. Rodrigues, D. T. *et al.* Characterization of a novel alphabaculovirus isolated from the southern armyworm, *spodoptera eridania* (cramer, 1782) (lepidoptera: Noctuidae) and the evolution of odv-e66, a bacterium-acquired baculoviral chondroitinase gene. *Genomics* **112**, 3903–3914 (2020).
186. Hill, T. & Unckless, R. L. Recurrent evolution of two competing haplotypes in an insect DNA virus. *BioRxiv* (2020) doi:10.1101/2020.05.14.096024.
187. Köster, J. & Rahmann, S. Snakemake—a scalable bioinformatics workflow engine. *Bioinformatics* **34**, 3600–3600 (2018).
188. Chen, S., Zhou, Y., Chen, Y. & Gu, J. Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).
189. Li, D., Liu, C.-M., Luo, R., Sadakane, K. & Lam, T.-W. MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de bruijn graph. *Bioinformatics* **31**, 1674–1676 (2015).
190. Sayers, E. W. *et al.* Database resources of the national center for biotechnology information. *Nucleic Acids Research* **39**, D38–D51 (2010).
-

-
191. Jones, P. *et al.* InterProScan 5: Genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240 (2014).
192. Li, H. & Durbin, R. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
193. Li, H. *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
194. Wilm, A. *et al.* LoFreq: A sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. *Nucleic Acids Research* **40**, 11189–11201 (2012).
195. Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly* **6**, 80–92 (2012).
196. Rausch, T. *et al.* DELLY: Structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics* **28**, i333–i339 (2012).
197. Layer, R. M., Chiang, C., Quinlan, A. R. & Hall, I. M. LUMPY: A probabilistic framework for structural variant discovery. *Genome Biology* **15**, R84 (2014).
198. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution* **30**, 772–780 (2013).
199. Page, A. J. *et al.* SNP-sites: Rapid efficient extraction of SNPs from multi-FASTA alignments. *Microbial Genomics* **2**, (2016).
200. Minh, B. Q. *et al.* IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution* **37**, 1530–1534 (2020).
201. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., Haeseler, A. von & Jermin, L. S. ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods* **14**, 587–589 (2017).
-

-
202. Hoang, D. T., Chernomor, O., Haeseler, A. von, Minh, B. Q. & Vinh, L. S. UFBoot2: Improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution* **35**, 518–522 (2017).
203. Eaton, D. A. R. Toytree: A minimalist tree visualization and manipulation library for python. *Methods in Ecology and Evolution* **11**, 187–191 (2019).
204. Pond, S. L. K. *et al.* HyPhy 2.5—a customizable platform for evolutionary hypothesis testing using phylogenies. *Molecular Biology and Evolution* **37**, 295–299 (2019).
205. Pond, S. L. K., Posada, D., Gravenor, M. B., Woelk, C. H. & Frost, S. D. W. Automated phylogenetic detection of recombination using a genetic algorithm. *Molecular Biology and Evolution* **23**, 1891–1901 (2006).
206. Pond, S. L. K. & Frost, S. D. W. Not so different after all: A comparison of methods for detecting amino acid sites under selection. *Molecular Biology and Evolution* **22**, 1208–1222 (2005).
207. Murrell, B. *et al.* Detecting individual sites subject to episodic diversifying selection. *PLoS Genetics* **8**, e1002764 (2012).
208. Smith, M. D. *et al.* Less is more: An adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Molecular Biology and Evolution* **32**, 1342–1353 (2015).
209. Weaver, S. *et al.* Datamonkey 2.0: A modern web application for characterizing selective and other evolutionary processes. *Molecular Biology and Evolution* **35**, 773–777 (2018).
210. Mészáros, B., Erdős, G. & Dosztányi, Z. IUPred2A: Context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Research* **46**, W329–W337 (2018).
211. Larsson, A. AliView: A fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* **30**, 3276–3278 (2014).
-

-
212. Beperet, I. *et al.* Superinfection exclusion in alphabaculovirus infections is concomitant with actin reorganization. *Journal of Virology* **88**, 3548–3556 (2014).
213. Nguyen, Q., TB.Tran, T., Chan, L. C., Nielsen, L. K. & Reid, S. In vitro production of baculoviruses: Identifying host and virus genes associated with high productivity. *Applied Microbiology and Biotechnology* **100**, 9239–9253 (2016).
214. Summers, M. D. Milestones leading to the genetic engineering of baculoviruses as expression vector systems and viral pesticides. in *Advances in virus research* 3–73 (Elsevier, 2006). doi:10.1016/s0065-3527(06)68001-9.
215. Kost, T. A. & Kemp, C. W. Fundamentals of baculovirus expression and applications. in *Advanced technologies for protein complex production and characterization* 187–197 (Springer International Publishing, 2016). doi:10.1007/978-3-319-27216-0_12.
216. Vanarsdall, A. L., Okano, K. & Rohrmann, G. F. Characterization of a baculovirus with a deletion of vlf-1. *Virology* **326**, 191–201 (2004).
217. Marek, M., Romier, C., Galibert, L., Merten, O.-W. & Oers, M. M. van. Baculovirus VP1054 is an acquired cellular PUR , a nucleic acid-binding protein specific for GGN repeats. *Journal of Virology* **87**, 8465–8480 (2013).
218. Alfonso, V., Amalfi, S., López, M. G. & Taboga, O. Effects of deletion of the ac109 gene of autographa californica nucleopolyhedrovirus on interactions with mammalian cells. *Archives of Virology* **162**, 835–840 (2016).
219. Airene, K. J. *et al.* Baculovirus: An insect-derived vector for diverse gene transfer applications. *Molecular Therapy* **21**, 739–749 (2013).
220. Oers, M. M. van, Pijlman, G. P. & Vlak, J. M. Thirty years of baculovirus–insect cell protein expression: From dark horse to mainstream technology. *Journal of General Virology* **96**, 6–23 (2015).
-

-
221. Thompson, C. M., Montes, J., Aucoin, M. G. & Kamen, A. A. Recombinant protein production in large-scale agitated bioreactors using the baculovirus expression vector system. in *Methods in molecular biology* 241–261 (Springer New York, 2016). doi:10.1007/978-1-4939-3043-2_11.
222. López, M. G., Diez, M., Alfonso, V. & Taboga, O. Biotechnological applications of occlusion bodies of baculoviruses. *Applied Microbiology and Biotechnology* **102**, 6765–6774 (2018).
223. Granados, R. R., Derksen, A. C. G. & Dwyer, K. G. Replication of the trichoplusia ni granulosis and nuclear polyhedrosis viruses in cell cultures. *Virology* **152**, 472–476 (1986).
224. Léry, X., Giannotti, J., Taha, A., Ravalec, M. & Abol-Ela, S. Multiplication of a granulosis virus isolated from the potato tuber moth in a new established cell line of phthorimaea operculella. *In Vitro Cellular & Developmental Biology - Animal* **33**, 640–646 (1997).
225. Naser, W. L., Miltenburger, H. G., Harvey, J. P., Huber, J. & Huger, A. M. In vitro replication of the cydia pomonella(codling moth) granulosis virus. *FEMS Microbiology Letters* **24**, 117–121 (1984).
226. Dwyer, K. G., Webb, S. E., Shelton, A. M. & Granados, R. R. Establishment of cell lines from pieris rapae embryos: Characterization and susceptibility to baculoviruses. *Journal of Invertebrate Pathology* **52**, 268–274 (1988).
227. Winstanley, D. & Crook, N. E. Replication of cydia pomonella granulosis virus in cell cultures. *Journal of General Virology* **74**, 1599–1609 (1993).
228. Hu, Y. & Li, L.-L. The p35 and ie1 of autographa californica multiple nucleopolyhedrovirus could rescue late gene expression of plutella xylostella granulovirus in nonpermissive cell lines. *Virus Genes* **48**, 343–355 (2013).
229. Hershberger, P. A., LaCount, D. J. & Friesen, P. D. The apoptotic suppressor P35 is required early during baculovirus replication and is targeted to the cytosol of infected cells. *Journal of Virology* **68**, 3467–3477 (1994).
-

-
230. Yin, F. *et al.* A functional f analogue of autographa californica nucleopolyhedrovirus GP64 from the agrotis segetum granulovirus. *Journal of Virology* **82**, 8922–8926 (2008).
231. Ardisson-Araújo, D. M. P., Melo, F. L., Clem, R. J., Wolff, J. L. C. & Ribeiro, B. M. A betabaculovirus-encoded gp64 homolog codes for a functional envelope fusion protein. *Journal of Virology* **90**, 1668–1672 (2015).
232. Biedma, M. E. Desarrollo de un sistema de recombinación para su aplicación al estudio de determinantes genéticos que afectan la infectividad de baculovirus de interés agronómico. (Universidad Nacional de La Plata, 2009). doi:10.35537/10915/90015.
233. Eberle, K. E., Wennmann, J. T., Kleespies, R. G. & Jehle, J. A. Basic techniques in insect virology. in *Manual of techniques in invertebrate pathology* 15–74 (Elsevier, 2012). doi:10.1016/b978-0-12-386899-2.00002-6.
234. Ogay, I. D. *et al.* Transfection of insect cell lines using polyethylenimine. *Cytotechnology* **51**, 89–98 (2006).
