UNIVERSITY OF CALIFORNIA
RIVERSIDE

Transcriptome Analyses of Ascovirus Genome Expression in Lepidopteran Larvae and Host Responses

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Microbiology

by

Heba Ahmed Hamed Mohamed Zaghloul

June 2018

Dissertation Committee:
Dr. Brian A. Federici, Chairperson
Dr. Ayala L.N. Rao
Dr. Adler R. Dillman

The Dissertation of Heba Ahmed Hamed Mohamed Zaghloul is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

ACKNOWLEDGEMENT

I would like to thank my major professor Dr. Brian Federici for his guidance and support over the past five years. I have learned much from his research experience and wisdom. He constantly encouraged me during my graduate studies, research, and this dissertation. My discussions with him were always fruitful. His wide knowledge in different areas of science is impressive. I will always be proud that I have been one of his students.

Also, I would like to extend my thanks to my dissertation committee members, Dr. Ayala L. N. Rao and Dr. Adler R. Dillman. Your guidance, discussion and feedback were of great importance to my completion of this work.

Moreover, I would like to thank present and previous Federici lab members for their help and guidance with mastering different methods and techniques as I carried out my research. I especially thank Robert Hice and Peter Arensburger for their continuous support, guidance, discussions and feedback over the years that helped me develop several skills. In addition, I extend my gratitude to Dennis Bideshi, Hyun-Woo Park and Stephanie Nanneman.

I also thank the Fulbright Program, sponsored by the U.S. Department of State's Bureau of Educational and Cultural Affairs, administered by AMIDEAST in Egypt, my home country, for their financial support for my PhD at UCR from 2013-2017. Being a Fulbrighter is a great experience that I will always remember and appreciate. I hope to be a good representative for the Fulbright program in the future and inspire other Egyptian students to conduct their research and education in U.S. universities, and act as a bridge

between the two countries through education. In addition, I extend my gratitude to Mr. & Mrs. Johnson for supporting UCR graduate students with the S. Sue Johnson Endowed Graduate Fellowship Award, of which I was delighted to be the recipient during the 2016-2017 academic year. Furthermore, I would like to acknowledge the financial support for my dissertation research through the Graduate Dean's Dissertation Research Grant in 2017.

The text of chapter 2 in my dissertation, in large part, is a reprint of my second chapter that appeared in the published paper, Zaghloul HAH, Hice R, Arensburger P, Federici BA. 2017, "Transcriptome analysis of the Spodoptera frugiperda ascovirus *in vivo* provides insights into how its apoptosis inhibitors and caspase promote increased synthesis of viral vesicles and virion progeny. Journal of Virology 91:e00874-17. https://doi.org/10.1128/ JVI.00874-17." The co-authors on this paper provided important assistance and guidance for this dissertation chapter. Moreover, I also extend my acknowledgement to Dr. Yves Bigot, INRA, France, for constructive critical comments on this study.

Finally, I would like to thank all my friends in USA and Egypt for their continuous encouragement and support. Last but not least, I would like to thank all my family members especially my parents, sister and brothers. Without your unconditional love, support and encouragement, I would not be who I am today.

DEDICATION

I dedicate this work to my beloved family.

ABSTRACT OF THE DISSERTATION

Transcriptome Analyses of Ascovirus Genome Expression
in Lepidopteran Larvae and Host Responses

by

Heba Ahmed Hamed Mohamed Zaghloul

Doctor of Philosophy, Graduate Program in Microbiology
University of California, Riverside, June 2018
Dr. Brian A. Federici, Chairperson

Ascoviruses, family *Ascoviridae*, are large, enveloped ds DNA viruses that mainly attack lepidopteran larvae of the family Noctuidae causing a chronic, but fatal disease. Their cytopathology differs from other viruses by destroying the nucleus, some species using a caspase, followed by cleavage of the cell into numerous viral vesicles, where most virions are produced. Cell cleavage resembles apoptosis, but differs markedly because mitochondria are not destroyed but rescued by ascoviruses, providing energy for replication and membrane synthesis for viral vesicles and virion envelopes. Ascovirus transmission is also unique in that virions are vectored to hosts on the ovipositor of parasitic wasps.

In my research, I used molecular techniques to determine genomic expression patterns underlying ascovirus cytopathology in larvae for two ascoviruses, the Spodoptera frugiperda ascovirus (SfAV) and Trichoplusia ni ascovirus (TnAV). Specifically, through *in vivo* transcriptome studies, including RNA-Sequencing, qRT-PCR, manual genome curating and bioinformatic analysis, I determined three temporal

classes for SfAV and TnAV genes; early, late, and very late. In SfAV, three proteins that inhibit apoptosis were synthesized early, whereas the caspase was synthesized very late, which correlates with apoptotic-like events resulting in vesicle formation. I identified 15 SfAV bicistronic and tricistronic messages, and similar transcripts in TnAV, indicating multi-cistronic transcripts are common in ascoviruses, and may regulate transcription or translation. Analyses of viral vesicle transcripts from *Trichoplusia ni* hemolymph revealed much higher viral genome expression than in the fat body, epidermis and tracheal matrix, showing larval blood is important for virus reproduction. Host responses to SfAV infection in *Spodoptera frugiperda* larvae showed mitochondrial gene expression was similar to controls or upregulated slightly during vesicle formation. Interestingly, few cytoskeleton genes known to encode motor proteins were upregulated despite mitochondrial involvement in vesicle formation. Analysis of antimicrobial peptides genes showed several where highly upregulated, as much as 32-fold, for example, moricin and gloverin, whereas innate immunity genes of the Toll, melanization, and phagocytosis pathways were activated only moderately by infection. Phenoloxidase and Toll negative regulators increased 15-fold and 3-fold, respectively. Therefore, conserving mitochondria while balancing immunity gene inducers and repressors may explain the chronic nature of ascovirus diseases.

TABLE OF CONTENTS

CHAPTER I.

Literature Review and Research Objectives (Ascoviruses)

CHAPTER II.
Transcriptome Analysis of the Spodoptera frugiperda Ascovirus *In Vivo* Provides Insights into How Its Apoptosis Inhibitors and Caspase Promote Increased Synthesis of Viral Vesicles and Virion Progeny

CHAPTER III.

Mitochondrial and Innate Immunity transcriptomes from *Spodoptera frugiperda* larvae infected with the Spodoptera frugiperda Ascovirus

CHAPTER IV.

Trichoplusia ni Ascovirus: Transcriptome Analysis of Hemolymph Viral Vesicles for a Virus with a Broad Tissue Tropism

CHAPTER V.

Summary and Future Perspectives

LIST OF FIGURES

CHAPTER I.
Literature Review and Research Objectives (Ascoviruses)

CHAPTER II.
Transcriptome Analysis of the Spodoptera frugiperda Ascovirus *In Vivo* Provides Insights into How Its Apoptosis Inhibitors and Caspase Promote Increased Synthesis of Viral Vesicles and Virion Progeny

CHAPTER III.
Mitochondrial and Innate Immunity Transcriptomes of *Spodoptera frugiperda* Larvae Infected with the Spodoptera frugiperda Ascovirus

CHAPTER IV.
Trichoplusia ni Ascovirus: Transcriptome Analysis of Hemolymph Viral Vesicles for a Virus with a Broad Tissue Tropism

LIST OF TABLES

CHAPTER II.
Transcriptome Analysis of the Spodoptera frugiperda Ascovirus *In Vivo* Provides Insights into How Its Apoptosis Inhibitors and Caspase Promote Increased Synthesis of Viral Vesicles and Virion Progeny

CHAPTER III.
Mitochondrial and Innate Immunity Transcriptomes of *Spodoptera frugiperda* Larvae Infected with the Spodoptera frugiperda Ascovirus

CHAPTER IV.
Trichoplusia ni Ascovirus: Transcriptome Analysis of Hemolymph Viral Vesicles for a Virus with a Broad Tissue Tropism

Chapter I: Literature Review and Research Objectives (Ascoviruses)

1.1 Introduction

The family *Ascoviridae* was erected over a decade ago to accommodate a number of large double-stranded DNA viruses pathogenic to larvae and pupae of many lepidopteran species, including pests of agriculture. Ascovirus virions are unique among all known viruses with regard to their ultrastructure in that they are large (400 x 150 nm), enveloped, generally bean-shaped, and are vectored to their lepidopteran hosts by female endoparasitic wasps during oviposition. However, by far the most unusual property of ascoviruses is the cytopathology they cause following infection - invariably they destroy the nucleus and redirect the cell's molecular biology, inducing it to cleave into numerous highly infectious virion-containing vesicles. These anucleate vesicles also appear to function as virus mini-factories, and accumulate at a high concentration (~10^8/ml) in the insect's hemolymph (blood), to which they impart a characteristic opaque milky-white color. It has been previously demonstrated that vesicle formation is initially induced by apoptosis, apparently assisted by a virus-coded executioner caspase, as observed in the type species, *Spodoptera frugiperda* ascovirus 1a (SfAV). The presence of a functional caspase gene in SfAV is also a unique occurrence among viruses. Although induction of programmed cell death is not uncommon in virus infection, ascoviruses have evolved mechanisms to rescue, manipulate and modify the virion vesicles with a marked altered membrane structure. The evolutionary advantage of ascovirus vesicle formation is obscure although it likely plays a role in virus transmission, especially as, unlike the fate of typical apoptotic bodies, they are not degraded by the host's innate immune system.

Indeed, these virion mini-factories can circulate stably in the insect's hemolymph for weeks after synthesis, which arguably sustains viable virions for dissemination in larval populations by parasitic wasps. In this introductory chapter, after reviewing their history, I review the main characteristics of ascoviruses that differentiate them compared to other vertebrate and invertebrate viruses with respect to their classification, virion and genome structure, cytopathology, wasp associations, transmission, immune evasion and significance. Finally, I highlight the objectives of my research.

1.2 Ascovirus historical background

Ascoviruses were discovered in the mid-1970's during a baculovirus epizootic in a population of the clover cutworm, *Scottogramma trifolii* (1). During dissection, hemolymph of several larvae exhibited an unusual milky white appearance. Examination of the blood using light and electron microscopy showed it to be packed with numerous vesicles averaging 8 μm in diameter. These vesicles contained many large, enveloped bacilliform particles that did not resemble any known virus or other pathogen. Attempts to transmit this virus by feeding it to larvae were unsuccessful. However, subsequent studies showed that the virus was highly infectious when injected into larvae (2). This permitted a variety of studies that resulted in characterization of the virions (3), description of the disease caused by what are now recognized as several different viral species (4-7), genome sequencing of six species, including the type species, *Spodoptera frugiperda* ascovirus 1a (SfAV), and official recognition of a new family of viruses, *Ascoviridae* (8). Most significantly, ascoviruses were shown to induce a novel

cytopathology in which they appeared to rescue and modify developing apoptotic bodies and convert these into virion-containing vesicles, often referred to as viral vesicles or vesicles, a pathogenic process not described previously in cellular biology (2). Since these viral vesicles are considered the most unique cytopathological feature of an ascovirus infection and can be used for disease diagnosis, the Greek word "asco" meaning "sac" was selected as the root of the family name, Ascoviridae.

Further studies of this virus group showed that the virus injection step is carried out in nature by female endoparasitic wasps, mainly species belonging to the families Braconidae and Ichnuemonidae. Based on field and laboratory studies, most ascovirus infections are limited to lepidopteran larvae of the family Noctuidae. There is one exception, however, an ascovirus that attacks the pupal stage of the leak moth, *Acrolepiopsis assectella*, a member of the family Yponomeutidae (5, 6). In addition, experimental infection trials have shown that species of the lepidopteran families Crambidae and Plutellidae are susceptible to infection by HvAV-3e, indicating there is a high probability that ascoviruses have a host broader range than realized currently (9). Moreover, because species of noctuids and parasitic wasps occur globally, the geographical distribution of these viruses can be considered worldwide. They have already been reported from the United States, Europe, Indonesia, China, Australia and recently in Japan (10, 11). The signs of ascovirus disease are not very obvious under field conditions, and therefore they are likely to be much more common in nature than realized, and may in fact be quite common, probably more common than the well known baculoviruses.

1.3 Ascovirus classification

Ascoviruses are members of family Ascoviridae, a member of the Nucleo-Cytoplasmic Large DNA Viruses (NCLDV), which together with five other families, the Poxviridae, Mimiviridae, Iridoviridae, Asfarviridae and Phycodnaviridae, compose the virus order Megavirales (12). These families are characterized by viruses that produce large complex virions with genomes ranging from 100 Kb to approximately 1.2-Megabase (13), and which replicate primarily in the cytoplasm of infected cells, but with some nuclear involvement (14-16).

According to the most recent ICTV reports, the family Ascoviridae consists of two genera *Ascovirus* and *Toursvirus* (www.ictvonline.org). The genus *Ascovirus* consists of accepted three species, *Spodoptera frugiperda* ascovirus-1a (SfAV-1a), the ascovirus type species, with only 2 variants SfAV-1b and SfAV-1C, *Trichoplusia ni* ascovirus-2a (TnAV-2a) and TnAV-2b, and *Heliothis virescens* ascovirus-3a (HvAV-3a), which is characterized by having a wider geographical distribution and many variants, HvAV-3b, HvAV-3c, HvAV-3d, HvAV-3e, HvAV-3f, HvAV-3j and HvAV-3g (previously, *Spodoptera exigua* ascovirus). In addition, there are four tentative members, *Helicoverpa armigera* ascovirus-7a (HaAV-7a), *Helicoverpa punctigera* ascovirus-8a (HpAV-8a), *Spodoptera exigua* ascovirus-9a (SeAV-9a) and *Trichoplusia ni* ascovirus-6a (previously, TnAV-2c) (8, 17). To date, only eight genomes of the above isolates have been sequenced, namely, SfAV-1a, TnAV-6a, TnAV-6b, HvAV-3e, HvAV-3g, HvAV-3f, HvAV-3j and DpAV-4a (11, 18-24). The genus *Toursvirus* includes only one species, *Diadromus pulchellus* ascovirus-4a (DpAV-4a).

DpAV-4a was previously a species in the genus *Ascovirus*, but due to significant differences in its biology compared to other ascoviruses it is now placed in a separate genus. Specifically, DpAV-4a replicates in both the pupal stage of its lepidopteran host and in adults of its wasp vector. Most virus replication occurs in the lepidopteran pupa. Moreover, the DpAV genome occurs in wasp nuclei as a free genome copy, indicating this virus' ability to be vertically transmitted by its wasp vector. Thus, as the DpAV genome is not integrated into the wasp genome, its biology is very different from that of the polydnaviruses, where the viral structural proteins, enzymes, and immune peptides are integrated into the vector wasps' genome (5, 25, 26).

1.4 Ascovirus virion structure

Examination of the milky hemolymph of larvae infected with ascoviruses using phase contrast microscopy reveals that the color and opacity resulted from the accumulation of large numbers (~10^8/ml) of highly refractive densely packed spherical vesicles. Subsequent study of these using electron microscopy shows these are filled with a combination of virions in various stages of assembly, mature virions, viral inclusion bodies containing virions, and various cell organelles such as mitochondria, smooth and rough endoplasmic reticulum, nucleoli, ribosomes, and fragments of the nuclear membrane. The virions are large, enveloped, and bacilliform to allantoid (bean-like) in shape (**Fig. 1.1**), and average 200-400 nm in length by 130 nm in diameter with complex symmetry (3, 8). They exhibit a distinctive reticulate surface pattern in negatively stained preparations, and are composed of an inner particle surrounded by a membranous

envelope. The inner particle assembles first and consists of a DNA/protein core surrounded by a lipid layer that bears a distinct layer of protein subunits on its surface. After formation, the inner particle is enveloped by membranes produced in the viral vesicle matrix, a mixture of the infected cell's nuclear membrane fragments, nucleoplasm and cytoplasm. The virion, therefore, appears to contain two lipid bilayers, one internal in the inner particle, the other being part of the outer membrane envelope. As in iridoviruses, the major capsid protein is located on the outer surface of the virion envelope (27), and is partly responsible for the reticulate appearance of the virion surface.

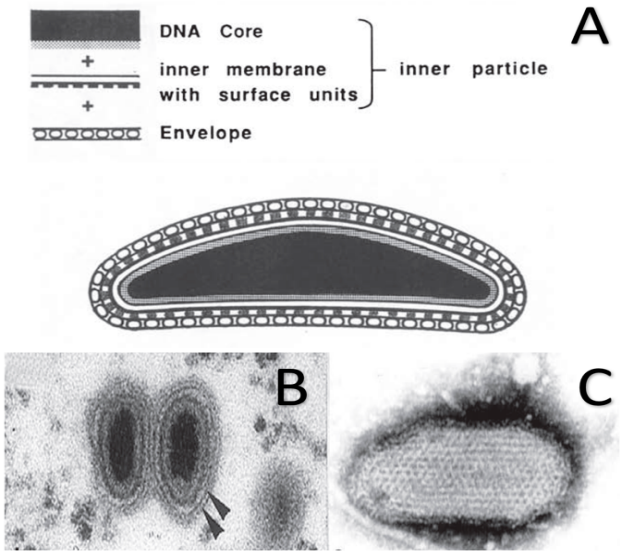


FIG 1.1 Ascovirus virion structure. **(A)** Diagrammatic representation of ascovirus different layers based on EM examination of ultrathin virion sections. **(B)** EM examination of ultrathin virion sections showing both the inner and the outer envelope of ascoviruses (arrow heads). **(C)** Net-like appearance of ascoviruse virion surface illustrated in a negative stain EM examination of purified virions (10, 28).

1.5 Ascovirus virion structural proteins

Ascovirus virions are comprised of at least 12 polypeptides ranging in mass from 10-200 kDa (3). In SfAV-1a, SDS-PAGE analysis of purified virions indicates that they consist of at least 21 virus-coded structural proteins. The nano-liquid chromatography tandem MS of SfAV-1a virion proteins illustrates that ORF041 and ORF048 are the most abundant structural proteins, identified as, respectively, the major capsid protein (MCP) and a highly basic protein (P64) (18, 29, 30). Other major proteins identified include an inhibitor of apoptosis, a yabby-like transcription factor, a myristylated membrane protein-like, a S1/P1 nuclease, a helicase, a Ser/Thr kinase, ATPase, an oxidase and CTD phosphatase. The other proteins detected have no known specific function. For TnAV-6a (TAV-2c; previously) only 8 proteins were detected on a 12% SDS-PAGE and of those identified using a high-accuracy MS only 7 matched predicted ORFs in the TnAV-6a genome. The MCP (ORF153) is the major TnAV-2c virion protein, in agreement with SfAV-1a analysis, and both a helicase and a Ser/Thr kinase were detected as well (19).

In silico analysis searching for common ascovirus virion proteins detected 11 of the 21 SfAV-1a virion's proteins homologues in DpAV4a, HvAV3e and TnAV-6a, suggesting that these proteins are part of their virions as well (21).

1.6 Ascovirus genome organization

Ascovirus genomes consist of a single circular ds DNA that ranges in size from 100-200 kbp and codes for from 117 to 180 ORFs depending on the virus species (18-24). Sequencing of the SfAV-1a genome (156,922 bp, 49.2% G+C), the ascovirus type

species, identified 123 putative ORFs, including many enzymes and proteins that could be associated with the novel ascovirus cytopathology induced in lepidopteran hosts during infection. In general, the ascovirus genes with assigned functions can be categorized into four main classes, namely, genes associated with (1) nucleotide-metabolism, (2) apoptosis, (3) lipid-metabolism and (4) baculovirus *bro* ORFs (functions not known). The majority of ascovirus genes, however, have no assigned function, especially in comparison to other large DNA viruses (18-20). Due to the novelty of caspase and caspase-like genes encoded by these viruses, much interest has been directed toward the role these genes play in apoptosis and vesicle formation (10, 31, 32). Other ascovirus or host genes that may play a role in this highly organized cytopathology remain to be identified.

The largest ascovirus genome size (199,721 bp, 45.9% G+C) is reported in one of the HvAV variants (HvAV-3g). This variant was originally referred to as SeAV-5a because it was first isolated from *Spodoptera exigua* larvae. It codes for 194 predicted ORFs, and shares high homology with the HvAV-3e variant (186,262 bp, 45.8 G+C %) isolated in Australia. Only 6 ORFs are unique to HvAV-3g. Another HvAV variant, HvAV-3f, from Australia has the second largest ascovirus genome (198,157 bp), coding for 190 ORFs with 46% G+C content (20, 22). The genome has two major homologous regions and 29 bro ORFs. Of the 190 ORFs, nine are unique to HvAV-3f. Two, ORF63 and ORF311, show significant similarity to *Agrotis segetum nucleopolyhedrovirus* (NC_007921) and *Choristoneura biennisentomo poxvirus* 'L' (NC_021248), respectively, genes not known to occur in any of the other ascoviruses. The function of

these genes is not known in any virus. In addition, three ORFs encoded by HvAV-3f (ORF17, ORF21 and ORF176) are known to have homologous proteins in Lepidopteran insects. Their role in virus infection is unclear (22, 23, 33, 34).

The TnAV-2c genome (174,059 bp) is characterized by its lower G+C content (35.4%) in comparison to other ascoviruses. The genome encodes 165 predicted ORFs with 12 ORFs being unique to this variant (19).

The DpAV-4a genome consists of 119,343 bp (G+C 49.66%), with 119 predicted ORFs. Of the 119 genes, DpAV-4a shares 34 homologous genes with other ascovirus species, and 63 genes with invertebrate Chilo iridescent virus, CIV. Therefore, DpAV-4a is considered more related to iridescent viruses than to ascoviruses (21). In addition, although ascovirus genomes are known to be methylated, DpAV-4a has the highest degree of methylation among animal DNA viruses (35). Therefore, taken together the differences in genome and virion structure, DpAV-4a, is now classified in the new genus *Toursvirus* of family Ascoviridae.

1.7 Gross pathology

Within 24-48 hours of ascovirus infection, growth and development of infected larvae are severely inhibited. Feeding slows down markedly by 48 hours, and as a result weight gain and development are very retarded. There is also difficulty in molting, and larvae that do molt typically cannot shed the cuticle. Infected larvae can remain in this chronic state for weeks, with some living for as long as six weeks until they die (36). As the infection progresses, the hemolymph changes color from being a translucent green to an opaque

milky white due to the accumulation of viral vesicles. The white color of the hemolymph, which can begin as soon as 48 hr after infection, is the classical sign of ascovirus disease, and can be used in the field to differentiate this disease from all others. However, insect dissection is generally needed to identify the color change, and it is for this reason that ascovirus infections are typically overlooked in the field (10).

1.8 Histopathology and tissue tropism

With respect to tissue tropism and histopathology, ascovirus species fall into three different types. In the first type, represented by the type species, SfAV-1a, viral replication is restricted primarily to the fat body, a tissue metabolically similar to mammalian liver, and which is completely destroyed by 12 days post-infection (4). The second type, represented by TnAV-2a and HvAV-3a, has a broader tissue tropism, with the principal tissues infected being the fat body, epidermis, and tracheal matrix. However, in these tissues the pattern of infection is uneven, with clusters of infected cells occurring randomly among uninfected healthy cells throughout each tissue. As a result, these tissues are rarely, if ever, completely destroyed (4, 7). In the third type, DpAV, which is more distantly related to the other ascoviruses, the primary tissues infected are the midgut epithelium, fat body and epidermis (5, 6).

1.9 Ascovirus cytopathology

Regardless of the differences in tissue tropism among ascovirus species, the cytopathology **(Fig. 1.2 & 1.3)** is essentially the same. Initially, it parallels apoptosis, but

differs from programmed cell death as viral replication and pathogenesis proceeds (2, 4). After infection, the nucleus hypertrophies as chromatin moves to the periphery. Subsequently, the nuclear membrane invaginates, ruptures and fragments as the cell undergoes extensive hypertrophy, enlarging to as much as 5-10 times as normal. Virions in various stages of assembly begin to appear as the nucleus fragments, which in cell culture occurs ~12 hours after infection. Simultaneously, the plasmalemma begins to invaginate along planes that partition the cell into viral vesicles. Interestingly, mitochondria align along these cleavage planes, where they appear to be involved in synthesis of new membranes that become the limiting membranes of the developing viral vesicles. After cell cleavage is complete, viral vesicles, typically about 20-30 per cell *in vivo*, separate from one another as virogenesis continues.

As the infection spreads from cell to cell, the basement membrane of the tissue in infected areas degenerates and fragments, liberating viral vesicles into the hemolymph, where they accumulate and circulate. These anucleate vesicles apparently function as mini-factories, unique in virus biology, as they continue to produce virions until the vesicles are almost completely filled with inclusions composed of virions. It has been shown *in vitro* that the apoptotic program is implemented and completed well within 9 hr following infection of the cell (37). *In vivo* the virion-containing vesicles are modified significantly as they develop and grow, but do not elicit an obvious innate immune response. Thus, in contrast to normal apoptotic bodies, which are ultimately eliminated by innate immune processes, modified ascovirus-virion vesicles appear unusually stable,

circulate in the blood, and continue to circulate in the blood for prolonged periods in morbid larvae, as long as six weeks post-infection.

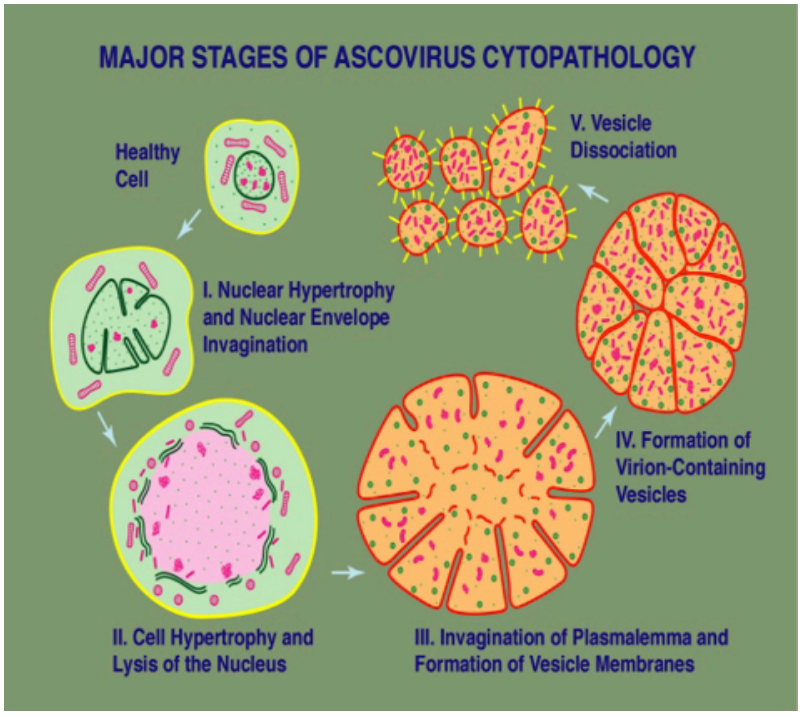


FIG 1.2 Diagrammatic representation of the major stages in ascovirus infection and vesicle formation (10).

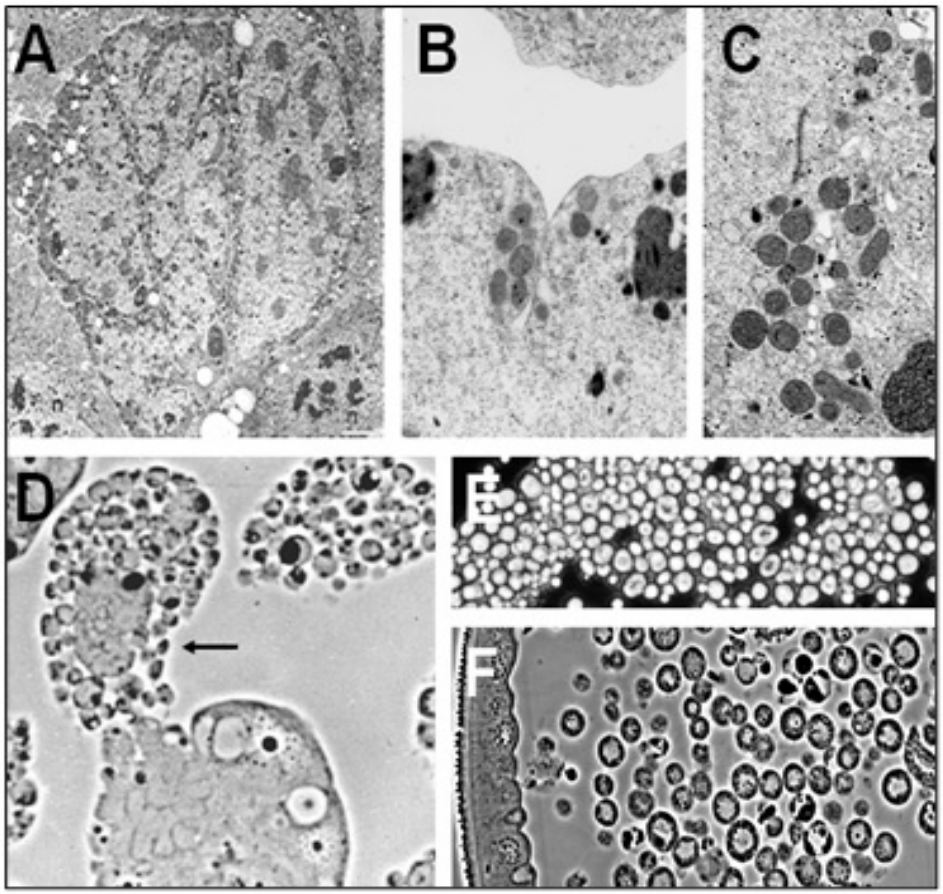


FIG 1.3 Formation of ascovirus viral vesicles. **(A)** Fat body cell in which the nucleus has fragmented and viral vesicles are beginning to form. Planes can be seen developing through the cell. The dense C-shaped areas are regions where mitochondria have aligned, apparently to assist synthesis of viral vesicle membranes. Cells at this stage of cleavage have very few virions. **(B)** Initiation of a cleavage furrow inward from the plasmalemma. Note the adjacent cluster of mitochondria. **(C)** A cleavage plane within the cell. The clear areas between the mitochondria are small lipid vesicles that will coalesce to form two lipid membranes that will become the outer membrane of two viral vesicles. **(D)** Lobe of fat body in which the cells are in advanced stages of viral vesicles formation. The cell at the top (arrow) has already formed many vesicles. These have not been liberated into the blood yet as they are retained by the basal lamina. Cleavage inward is still in progress, as indicated by the gray central area with the c-shaped arrangements of developing vesicle membranes. Just beneath this cell, a cell in an earlier stage of development, similar to that shown in A, can be seen. **(E)** Wet mount preparation of the blood of an SfAV-infected larva at 7 days post-infection. Average vesicle size is about 8 µm. **(F)** Section though the body cavity (hemocoel) of an infected larva illustrating viral vesicles circulating in the blood. **A-C** are electron micrographs and **D-F** are phase contrast micrographs of fixed tissues in plastic sections (10, 28).

1.10 Ascovirus transmission

An unusual characteristic of ascoviruses (family *Ascoviridae*) compared to other insect viruses such as baculoviruses (family *Baculoviridae*) and cytoplasmic polyhedrosis viruses (family *Reoviridae*) is that they are very poorly infectious *per os*. Instead, ascoviruses are transmitted to their lepidopteran hosts by female parasitoid wasps when they lay their eggs during oviposition. In general, under natural conditions transmission begins when a female wasp lays eggs in an ascovirus-infected larva. Virions and virion-containing vesicles circulating in the hemolymph stick to the female's ovipositor, and then subsequently she infects healthy host when laying eggs in the next host. Along with the eggs, virions and viral vesicles adhering to the ovipositor are transmitted to the new host. The female wasp transmits virions and vesicles during each oviposition event through the probing process while searching for the right host in which to lay eggs. This leads to the infection of host cells and viral replication in permissive hosts. As noted above, in the type species, the *Spodoptera frugiperda ascovirus* (SfAV-1a) the infection is restricted primarily to the fat body, whereas in other ascoviruses such as the *Trichoplusia ni ascovirus* (TnAV-2a) and *Heliothis virescens ascovirus* (HvAV-3a) the fat body, epidermis, and tracheal matrix are the main tissues infected. In all ascovirus infections none of the wasp larvae that emerge from the oviposited eggs survive in their caterpillar hosts. They either die of malnutrition or possible are killed by unknown factors produced by the virus.

The third type of ascovirus, the *Diadromus pulchellus ascovirus*, differs markedly from the typical ascoviruses described above in that it replicates at a low level in *D.*

pulchellus, i.e., its wasp host, with no significant pathology, but is vectored to this wasp's host, the pupal stage of the leak moth, *Acrolepiopsis assectella*, in which after infection the virus undergoes extensive replication, killing the pupa. In the pupal host, the wasp larvae that emerge from the eggs survive and thrive, feeding on host tissues including those in which the virus has replicated. Due to this distinctly different ascovirus biology, the *D. pulchellus ascovirus* (DpAV-4a), as noted above, has recently been reclassified in a new genus, *Toursvirus*.

Based on the different wasp-virus-host pathologies, the association between the ascovirus and its wasp vector can range from being non-pathogenic to mutualistic. The DpAV association can be explained as a mutualistic one, in which the wasp plays a role in slowing the virus replication rate and hence host cell lysis is delayed until the wasp eggs hatch. This sort of association is supported by experiments in which artificially DpAV-injected lepidopteran host cell lysis occurs very quickly preventing the parasitoid development. On the other hand, the DpAV prevents encapsulation of wasp eggs by the host innate immune system through interfering with the monophenoloxidase pathway required for melanin deposition. Interestingly, in nature the *Diadromus pulchellus* wasps usually vector DpAV along with another RNA virus known as the *Diadromus pulchellus* idnoreovirus 1 (DpRV1), which is detected in all the *Diadromus pulchellus* adult populations. This could be the main reason for the retarded rate of DpAV replication in nature and reveals the complexity of this system (6, 38-40). Interestingly, the DpAV can develop a pathogenic association and not a mutualistic one when being vectored by other non-*Diadromus* wasp species, for example, females *Itoplectis tunetana*. The pathogenic

effect is caused by the virus rapid replication in these species hindering the wasp egg development (5). Thus, DpAV is now considered a new species in the family *Ascoviridae*, with many traits that differ from classical ascoviruses.

In the case of the typical pathogenic associations found in most ascoviruses, meaning those occurring in SfAV, TnAV and HvAV and their braconid or ichneumonid vectors, the relationship is typically very detrimental to the wasp population since these viruses destroy the host tissues and interfere severely with the development of wasp larvae. Thus, the wasp basically functions as a mechanical tool used by the virus to spread it to host populations.

1.10.1 Transmission in nature

In nature, ascoviruses are more common in insecticide-free fields where their parasitoid wasp vectors can survive and transmit the virus (which, interestingly, suggests parasitic wasps are an indicator of pesticide absence). Therefore, the role of ascoviruses in controlling lepidopteran pest populations in nature cannot be underestimated, especially with its characteristic chronic infection that provides a continuous supply of virions to be picked up and spread by the wasp (41). However, the competence of ascoviruses for biological control of economically important insects, for example, *Spodoptera frugiperda*, *Spodoptera exigua*, *Heliothis virescens* and *Trichoplusia ni* larvae that serve as target hosts for many ascoviruses is still considered to be limited. This is because the infection is not mainly *per os*, as in the case of the majority of insect viruses, and thus virus injection into the hemocoel of the larval body is needed to increase the chances of

virus transmission. On the other hand, intentional introduction of wasp vectors to act synergistically with ascoviruses may improve their transmission rates. For example, it has been reported that *Microplitis similis* (Hymenoptera: Braconidae) can increase transmission rates of HvAV-3h in the *S. exigua* larvae to levels as high as >80% (42).

1.10.2 Experimental transmission in the laboratory

Ascovirus infectivity has been studied in several species of noctuid larvae using both *per os* (feeding) and injection techniques. In general, *per os* infections resulted in mortality rates of less than 15%, whereas by injection greater than 90% of the larvae developed ascovirus disease. For example, when 10 third-instar *Spodoptera frugiperda* larvae were pierced with a minute pin sequentially that had been dipped in hemolymph from an infected larva and dried for 24 hours, more than 90% of these developed ascovirus disease (36). Similarly, in a recent study, the oivipositor of *Microplitis similis* contaminated with HvAV-3h was infectious after 4.1 +/- 1.4 days (42). Therefore, inoculation is usually the method of choice to propagate ascovirus infections in the laboratory. In general, infection of early larval instars is more successful than advanced larval stages (43).

1.11 Ascovirus strategies for evading host innate immunity and host responses

One of the initial requirements for any pathogen to be successful is to evade host innate and adaptive immunity. This is usually accomplished after a long-term “arms race” between the pathogen and its host. Since viruses are obligate parasites, they are under

intensive selection pressure to evolve strategies for survival that overcome host immune responses. In this section, I highlight known and possible strategies ascoviruses have evolved to survive and replicate.

Most ascoviruses cause chronic infections in their insect hosts that are ultimately fatal. As noted above, healthy larvae typically feed, develop and advance to pupation within two weeks of hatching, whereas infected larvae live in an arrested state of development for as long as six weeks and then die (36, 43, 44). The long survival period can be explained in the case of TnAV and HvAV by the limited infection of different important host tissues, especially the epidermis, tracheal matrix, and fat body, or in SfAV, for example, by primarily infecting only the latter tissue (36). Increasing the life span of the infected host through a chronic infection and developmental retardation to avoid alarming the host and inducing an acute immune response, is advantageous to the virus because it increases the period for acquisition of infectious virions by parasitic wasps for transmission to new host larvae.

With respect to the acquisition of viruses and virus-like particles by parasitic wasps, several different strategies have evolved for avoiding or abrogating host immune responses. For example, it is known that some wasps inject virion-like particles along with their eggs during oviposition. In some wasp species, such as *Venturia canescens*, family Ichneumonidae, these particles cover the egg surface masking the egg from host immunity via molecular mimicry in which the particle surface mimics the basement membrane of the caterpillar host, *Ephestia kuhniella*, family Pyralidae (45, 4). These particles resemble virions in structure, but contain no DNA. In case of the so-called

polydnaviruses (family Polydnaviridae), however, the particles, which constitute a major component of the oviposition fluid, contain DNA that is inserted into caterpillar host cells to actively suppress the innate immune response, as well as hemocyte melanization and encapsulation by the caterpillar hosts (26, 47, 48). In the polydnaviruses, the particles of each virus species vary considerably in size, and each particle contains only one or a few genes, and these are mainly wasp genes, not viral genes *per se*. Instead, the so-called viral genome in the particles, consists of wasp circular ds DNA molecules, that code for proteins directly involved in immunosuppression (i.e., multipartite circular ds DNA molecules of different sizes enclosed within a protein particle derived from the wasp genome). The main difference between virion-like particles and viruses is the lack of genome replication and hence no progeny result in caterpillar cells. Interestingly, some of these particles resemble ascovirus virions in the structure which imply their evolution from the virus itself followed by gene gain and loss events. For example, the loss of the genome replication ability, since these particles DNA content is only transduced in the lepidopteran host but no particle DNA replication is reported (49).

On the molecular level, the large DNA genome size and the number of conserved proteins provide ascoviruses with many abilities (8, 50). For example, most large DNA viruses in the NCLDV (Nucleo-Cytoplasmic Large DNA Viruses) group, which includes the ascoviruses, encode their own DNA polymerase, RNA polymerase and transcription factors, implying a more independent life style cycle in comparison to the typical host-dependent nature of viruses (51). Moreover, ascoviruses are characterized by the presence of some lepidoptern host homologus genes in their large genomes (22, 23, 33,

34). The roles of these genes in the virus replication cycle are not know, but interestingly one of these genes found in both SfAV-1a and HvAV species, shares homology with the *Drosophila* Diedel protein (*die* gene). Diedel is a 12 kDa protein recently identified to function as a cytokine that down-regulates the immune deficiency (IMD) pathway in flies (52). Homologs of *die* occur in two other dsDNA insect viruses namely, baculoviruses and entomopoxviruses. According to the Lamiable et al. (52), *die* is defined as the first insect "virokine," which may well function in host immune system suppression through down-regulation of the IMD pathway.

On the other hand, most viral miRNAs are only known in DNA viruses, not in RNA viruses. The reason for this is probably that DNA viruses have access to the host cell nucleus where the pri-miRNA processing steps take place prior to pre-miRNA formation and export to the cytoplasm. The herpesvirus (family Herpesviridae) genomes provide a well-known example of miRNAs encoded by DNA viruses. Indeed, this virus group encodes approximately 90% of those identified. It appears that in herpesviruses these miRNAs have several key roles that include evading the host immune response, elongation of infected cell longevity, and finally inducing lysis (53). In addition to the herpesviruses, miRNAs encoded by other virus families have been identified in polyomaviruses, adenoviruses, as well as in insect baculoviruses and ascoviruses (54). In case of *Heliothis virescence* ascovirus (HvAV-3e), it has been shown that HvAV-miR-1, identified in the gene encoding for the major capsid protein (MCP), down-regulates the viral DNA polymerase and inhibits the virus replication (55). However, this is the only miRNA identified among the various ascoviruses species to date, the probable reason

being that these viruses, especially their miRNAs, have received very little study. Because so little is known about miRNAs in ascoviruses, determining whether others than HvAV-miR-1 occur is worthy, and if they do occur, determination of their role in virus replication, including whether they play any role in circumventing – like herpesviruses - host immunity.

Ultimately, ascovirus infection, like any other group of viruses, is expected to induce an antiviral immune response such as RNA interference in their insect hosts. Although dsRNA formation is known to be the main trigger of the RNAi pathway in different hosts, including insects, DNA viruses can activate this immune system as well. For example, the RNAi pathway is activated in *Helicoverpa armigera* larvae upon infection with *Helicoverpa armigera* single nucleopolyhedrovirus (56). A fundamental question for DNA viruses is the source of the dsRNA or Dicer substrate. There are several possibilities for dsRNA formation during an RNA virus infection, such as possessing of a dsRNA genome or formation of dsRNA intermediates during genome replication cycle of -/+ sense single strand RNA viruses. However, in case of DNA viruses bi-directional transcription is considered the major source of dsRNA formation, and has been reported in several studies performed on both ss/ds DNA insect viruses. For example, in invertebrate iridescent virus 6 (IIV-6), the vsiRNAs were mapped equally to both virus DNA strands (57-60). It is worth mentioning that in ascoviruses a conserved RNaseIII endonuclease enzyme is reported in HvAV-3e to silence the host RNAi response. This endonuclease was found to be essential for virus replication, which implies that ascoviruses may be exposed to host RNAi immune response, but have

developed strategies to silence this effect (32). Overall, the antiviral immune response of insect hosts in case of ascoviruses has not been well studied and only one cellular miRNA, Hz-miR24, has been found during HvAV-3e infection *in vitro*, where it is associated with HvAV-3e DdRP and DdRPβ transcript down-regulation (61).

Finally, a recent study of *Spodoptera exigua* transcriptome after infection with HvAV-3h confirms the involvement of different innate immunity pathways in host response. For example, the Toll and JAK/STAT pathways. In addition, the most enriched pathways in Kyoto Encyclopedia of Genes and Genomes (KEGG) annotation involved metabolism, cytoskeleton, apoptosis, cell cycle, p53 signaling pathway and oocyte meiosis (61). Future studies should answer questions about how the insect immune system deals with ascovirus infections. Moreover, such studies might lead to basic knowledge about Dicer's substrate during DNA virus infection.

1.12 Significance of ascoviruses

Research on ascoviruses has the potential to make important contributions to two disciplines. First, in molecular cell biology, the functional executioner caspase that ascoviruses encode is a unique feature compared to all known vertebrate and invertebrate viruses. Therefore, understanding the novel cytopathology characteristic of ascoviruses may well lead to identifying novel cell biology pathways in which this caspase plays a key role, especially by defining the underlying mechanisms that result in the formation of virion-containing vesicles. This is a metabolically highly active developmental process compared to the degenerative processes that yield typical apoptotic bodies formed in the

canonical apoptosis pathway. Although induction of apoptosis is not uncommon in virus infection, ascoviruses have evolved mechanisms to apparently rescue, manipulate and modify these vesicles for dissemination in larval populations. Unlike the fate of typical apoptotic bodies, ascovirus virion vesicles are stable and not degraded by the host's innate immune system and circulate in the hemolymph for weeks after synthesis in what sounds as a re-programming step.

Second, with respect to applied sciences, ascoviruses are insecticidal viruses that target mainly lepidopteran insects, including many important agriculture pests. Therefore, they have the potential to be used in the future in the insecticidal market. For example, microbial-based control agents in 2007/2008 constituted a market of about $396 million (63). However, for ascoviruses to compete in the insecticidal market, more research is needed to optimize their use in the field under real conditions and identify new strategies of virus transmission, which is considered the main limitation for ascoviruses success. The majority of insecticidal viruses used for insect pest control are transmitted by feeding, so for ascoviruses to be economically successful this trait would have to be improved significantly.

1.13 Research objectives

As noted in the above literature review, there are now more than 40 papers dealing with various aspects of ascoviruses, including description of species and their differences, genome sequences, virion structure proteins, enzymes expressed during replication, and details of their biology such as tissue tropism, histopathology, and cytopathology.

However, virtually nothing is known about the molecular basis of ascovirus vesicle formation, the main distinguishing feature of this family. In large part, the lack of a suitable cell line to study vesicle formation is the main reason for this important lack of knowledge. Although many noctuid cell lines are permissive for ascovirus replication, none shows the typical formation of ascovirus vesicles that occurs *in vivo* (64). To help fill this knowledge gap, the main focus of my dissertation was to undertake a series of studies based on a comparison of the *in vivo* transcriptomes of two different ascovirus species, SfAV and TnAV. I have chosen to research these different species due to their different tissue tropisms, wherein SfAV replicates almost exclusively in the fat body essentially eliminating this tissue, whereas TnAV replicates in the same tissue as well as the epidermis and tracheal matrix, but never destroys more than 50% of the cells in each by ten days post-infection. Therefore, the specific objectives of my dissertation were the following:

1.13.1 Determine the SfAV-1a temporal gene expression profile for conserved core genes and those for putative membrane-associated proteins

Vesicle formation is the most prominent unique feature of ascoviruses induced upon infection of lepidopteran larvae regardless of the viral species. Therefore, studying the SfAV-1a (the ascovirus type species) transcriptome during infection of its lepidopteran host should provide valuable information about the temporal expression and the level of expression of the most important genes. This will help me to determine the patterns of expression for ascovirus core genes, those conserved in among most ascovirus species,

and will provide insights to the synthesis of viral membrane-associated proteins. The latter should enable me to identify proteins involved in the novel cytopathology that leads ultimately to viral vesicle formation.

1.13.2 Determine the SfAV-1a-infected host response especially for the mitochondrial, cytoskeleton and innate immunity genes

The changes induced in the ascovirus-infected cell cytoskeleton are remarkable and suggest the contribution of mitochondria in vesicle formation. For example, the change in mitochondria shape, size and distribution is well observed during vesicle membranes formation. Therefore, studying the host mitochondrial and cytoskeleton encoded genes expression pattern during infection may demonstrate if the mitochondria is involved at all in this process. On the other hand, the vesicles long-term of circulation in the infected host hemolymph without inducing an obvious immune response raise the question about if the vesicles formation induce any of the innate immune responses after formation. Therefore, studying the expression of innate immunity pathways for instance, humoral and cellular factors may contribute to our knowledge about the host immune response towards the vesicles.

1.13.3 Comparison of the vesicle-specific transcription patterns in blood versus the fat body, epidermis and tracheal matrix during TnAV host infection

Ascovius vesicles are known to accumulate ultimately in the insect hemolymph turning it from translucent green to milky white, a unique gross pathology that distinguishes this

disease from all others. Therefore, comparison of the viral transcription pattern from vesicles that occurs in blood compared to other host tissues, in this case using *Trichoplusia ni* as the host, may provide valuable information about unique viral genes that characterize the vesicle fraction versus the other host tissues.

1.14 References

1. Federici BA. 1978. Baculovirus epizootics in a larval population of the clover cutworm, *Scotogramma trifolii*, in Southern California. Environ Entomol 7:423-427.

2. Federici BA. 1983. Enveloped double stranded DNA insect virus with novel structure and cytopathology. Proc Natl Acad Sci USA 80:7664-7668.

3. Federici BA, Vlak JM, Hamm, JJ. 1990. Comparison of virion structure, protein composition, and genomic DNA of three ascovirus isolates. J Gen Virol 71:1661-1668.

4. Federici BA, Govindarajan R. 1990. Comparative histology of three ascovirus isolates in larval noctuids. J Invertbr Pathol 56:300-311.

5. Bigot Y, Rabouille A, Doury G, Sizaret PY, Delbost F, Hamelin MH, Periquet G. 1997a. Biological and molecular features of the relationships between Diadromus pulchellus ascovirus, a parasitoid hymenopteran wasp (*Diadromus pulchellus*) and its lepidopteran host, *Acrolepiopsis assectella.* J Gen Virol 78:1149-1163.

6. Bigot Y, Rabouille A, Sizaret PY, Hamelin MH, Periquet G. 1997b. Particle and genomic characterization of a new member of Ascoviridae, Diadromus pulchellus ascovirus. J Gen Virol 78:1139-1147.

7. Cheng XW, Carner CR, Arif BM. 2000. A new ascovirus from *Spodoptera exigua* and its relatedness to the isolate from *Spodoptera frugiperda.* J Gen Virol 81:3083-3092.

8. Bigot Y, Asgari S, Bideshi DK, Cheng X, Federici BA, Renault S. 2011. Family *Ascoviridae*, p 147-152. *In* King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (ed),Viral Taxonomy, IX Report of the International Committee on the Taxonomy of Viruses, 3rd edn. London: Elsevier–Academic Press.

9. Furlong MJ, Asgari S. 2010. Effects of an ascovirus (HvAV-3e) on diamondback moth, *Plutella xylostella*, and evidence for virus transmission by a larval parasitoid. J Invertebr Pathol 103:89-95.

10. Federici BA, Bideshi, DK, Tan Y, Spears T, Bigot Y. 2009. Ascoviruses: superb manipulators of apoptosis for viral replication and transmission. Curr Top Microbiol Immunol 328:171-196.

11. Arai E, Ishii K, Ishii H, Sagawa S, Makiyama N, Mizutani T, Omatsu T, Katayama Y, Kunimi Y, Inoue MN, Nakai M. 2018. An ascovirus isolated from *Spodoptera litura* (Noctuidae: Lepidoptera) transmitted by the generalist endoparasitoid *Meteorus pulchricornis* (Braconidae: Hymenoptera). J Gen Virol doi: 10.1099/jgv.0.001035.

12. Colson P, De Lamballerie X, Yutin N, Asgari S, Bigot Y, Bideshi DK, Cheng XW, Federici BA, Van Etten JL, Koonin EV, La Scola B, Raoult D. 2013. "Megavirales," a proposed neworder for eukaryotic nucleocytoplasmic large DNA viruses. Arch Virol 158:2517–2521.

13. Raoult D, Audic S, Robert C, Abergel C, Renesto P, Ogata H, La Scola B, Suzan M, Claverie J-M. 2004. The 1.2-Megabase Genome Sequence of Mimivirus. Science 306:1344-1350.

14. Boyer M, Yutin N, Pagnier I, Barrassi L, Fournous G, Espinosa L, Robert C, Azza S, Sun S, Rossmann MG, Suzan-Monti M, La Scola B, Koonin EV, Raoult D. 2009. Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms. Proc Natl Acad Sci USA 106:21848–21853.

15. Thomas V, Bertelli C, Collyn, F, Casson N, Telenti A, Goesmann A, Greub G. 2011. Lausannevirus, a giant amoebal virus encoding histone doublets. Environ Microbiol 13:1454–1466.

16. Yutin N, Koonin EV. 2012. Hidden evolutionary complexity of nucleo-cytoplasmic large DNA viruses of eukaryotes. Virol J 9:161.

17. Bigot Y. 2011.Genus Ascovirus. p 73-78. *In* Tidona C, Darai, G. (ed), Springer Index of Viruses, 2nd edn. Heidelberg: Springer.

18. Bideshi DK, Dematti MV, Rouleux-Bonnin F, Stasiak K, Tan Y, Bigot S, Bigot Y, Federici BA. 2006. Genomic sequence of the *Spodoptera frugiperda ascovirus 1a*, an enveloped double-stranded DNA insect virus that manipulates apoptosis for viral reproduction. J Virol 80:11791-11805.

19. Wang L, Xue J, Seaborn CP, Arif BM, Cheng XW. 2006. Sequence and organization of the *Trichoplusia ni* ascovirus 2c (Ascoviridae) genome. Virology 354:167–177.

20. Asgari S, Davis J, Wood D, Wilson P, McGrath A. 2007. Sequence and organization of the *Heliothis virescens* ascovirus genome. J Gen Virol 88:1120–1132.

21. Bigot Y, Renault S, Nicolas J, Moundras C, Demattei MV, Samain S, Bideshi DK., Federici BA. 2009. Symbiotic virus at the evolutionary intersection of three types of large DNA viruses; iridoviruses, ascoviruses, and ichnoviruses. PLoS One 4:e6397.

22. Huang GH, Wang YS, Wang X, Garretson TA, Dai LY, Zhang CX, Cheng XW. 2012. Genomic sequence of *Heliothis virescens ascovirus 3g* isolated from *Spodoptera exigua*. J. Virol. 86:12467–12468.

23. Wei YL, Hu J, Li SJ, Chen ZS, Cheng XW, Huang GH. 2014. Genome sequence and organization analysis of *Heliothis virescens* ascovirus 3f isolated from a *Helicoverpa zea* larva. J Invertebr Pathol 122:40–43.

24. Liu YY, Xian WF, Xue J, Wei YL, Cheng XW, Wang X. 2018. Complete genome sequence of a renamed isolate, Trichoplusia ni ascovirus 6b, from the United States. Genome Announc 6:e00148-18.

25. Renault S, Petit A, Be´ne´det F, Bigot S, Bigot Y. 2002. Effects of the *Diadromus pulchellus* ascovirus, DpAV4, on the hemocytic encapsulation response and capsule melanization of the leek-moth pupa, *Acrolepiopsis assectella*. J Insect Physiol 48:297–302.

26. Strand MR, Burke GR. 2014. Polydnaviruses: Nature's genetic engineers. Annu Rev Virol 1:333–354.

27. Stasiak K, Demattei MV, Federici BA, Bigot Y. 2000. Phylogenetic position of the *Diadromus pulchellus* ascovirus DNA polymerase among viruses with large double-stranded DNA genomes. J Gen Virol 81:3059-3072.

28. Federici, B. A., J. J. Hamm and E. L. Styer. 1991. Ascoviruses. Chapter 11, pp. 339-349. In "An Atlas of Invertebrate Viruses" (J. Adams and G. Bonami, editors). CRC Press, Boca Raton, Florida.

29. Tan Y, Bideshi DK, Johnson JJ, Bigot Y, Federici, BA. 2009a. Proteomic analysis of the *Spodoptera frugiperda* ascovirus 1a virion reveals 21 proteins. J Gen Virol 90:359-365.

30. Tan Y, Spears T, Bideshi DK, Johnson JJ, Hice RH, Bigot Y, Federici BA. 2009b. P64, a novel major virion DNA-binding protein potentially involved in condensing the *Spodoptera frugiperda ascovirus 1a* genome. J Virol 83:2708-2714.

31. Hussain M, Asgari S. 2008. Inhibition of apoptosis by *Heliothis virescens* ascovirus (HvAV-3e): Characterization of orf28 with structural similarity to inhibitor of apoptosis proteins. Apoptosis 13:1417–1426.

32. Hussain M, Abraham AM, Asgari S. 2010. An Ascovirus-encoded RNase III autoregulates its expression and suppresses RNA interference-mediated gene silencing. J Virol 84:3624–3630.

33. Jakubowska AK, Peters SA, Ziemnicka J, Vlak JM, van Oers MM. 2006. Genome sequence of an enhancin gene-rich nucleopolyhedrovirus (NPV) from *Agrotis segetum*: collinearity with *Spodoptera exigua* multiple NPV. J Gen Virol 87:537–55.

34. Thézé J, Takatsuka, J, Li Z. Gallais J, Doucet D, Arif B, Nakai M, Herniou EA. 2013. New insights into the evolution of Entomopoxvirinae from the complete genome sequences of four entomopoxviruses infecting *Adoxophyes honmai*, *Choristoneura biennis*, *Choristoneura rosaceana*, and *Mythimna separate*. J Virol 87:7992–8003.

35. Bigot Y, Stasiak K, Rouleux-Bonnin F, Federici BA. 2000. Characterization of repetitive DNA regions and methylated DNA in ascovirus genomes. J Gen Virol 81:3073–3082.

36. Govindarajan R, Federici, BA. 1990. Ascovirus infectivity and effects of infection on the growth and development of noctuid larvae. J Invertebr Pathol 56:291–299.

37. Bideshi DK, Tan Y, Bigot Y, Federici BA. 2005. A viral caspase contributes to modified apoptosis for virus transmission. Genes Dev 19:1416-1421.

38. Stasiak K, Renault S, Federici BA, Bigot Y. 2005. Characteristics of pathogenic and mutualistic relationships of ascoviruses in field populations of parasitoid wasps. J Insect Physiol 51:103–115.

39. Renault S, Stasiak K, Federici B, Bigot Y. 2005. Commensal and mutualistic relationships of reoviruses with their parasitoid wasp hosts. J Insect Physiol 51: 137–148.

40. Roossinck MJ. 2011. The good viruses: viral mutualistic symbioses. Nature Rev Microbiol 9:99–108.

41. Cheng XW, Wan XF, Xue JL, Moore RC. 2007. Ascovirus and its evolution. Virologica Sinica 22:137–147.

42. Li SJ, Hopkins RJ, Zhao YP, Zhang YX, Hu J, Chen XY, Xu Z, Huang GH. 2016. Imperfection works: survival, transmission and persistence in the system of Heliothis virescens ascovirus 3h (HvAV-3h), *Microplitis similis* and *Spodoptera exigua*. Sci Rep 6:21296.

43. Hu J, Wang X, Zhang Y, Zheng Y, Zhou S, Huang GH. 2016. Characterization and Growing Development of *Spodoptera exigua* (Lepidoptera: Noctuidae) Larvae Infected by Heliothis virescens ascovirus 3h (HvAV-3h*)*. J Econ Entomol. 109:2020-6.

44. Li SJ, Wang X, Zhou ZS, Zhu J, Hu J, Zhao YP, Zhou GW, Huang GH. 2013. A comparison of growth and development of three major agricultural insect pests infected with Heliothis virescens ascovirus 3h (HvAV-3h). PLoS ONE 8:e85704.

45. Rotheram S. 1967. Immune surface of eggs of a parasitic insect. Nature 214: 700.

46. Salt G. 1968. The resistance of insect parasitoids to the defense reactions of their hosts. Biol Rev 43:200–232.

47. Stoltz DB, Krell P, Summers MD, Vinson SB. 1984. Polydnaviridae – a proposed family of insect viruses with segmented, double-stranded, circular DNA genomes. Intervirology 21:1–4.

48. Vinson SB. 1990. How parasitoids deal with the immune system of their host: an overview. Arch Insect Biochem Physiol 13:2–27.

49. Federici BA, Bigot Y. 2010. In: Pontarotti, P (ed). Evolutionary Biology - Concepts, Molecular and Morphological Evolution. Springer, Berlin Heidelberg, pp. 229–248.

50. Xue JL, Cheng XW. 2011. Comparative analysis of a highly variable region of the genomes between *Spodoptera frugiperda* ascovirus 1d (SfAV-1d) and SfAV-1a. J Gen Virol 92:2797–2802.

51. Moss B. 2007. Poxviridae: The viruses and their replication, p. 2905–2946. *In* Knipe DM, Howley PM(ed), Fields virology, 5th ed, vol1. Lippincott Williams & Wilkins, Philadelphia.

52. Lamiable O, Kellenberger C, Kemp C, Troxler L, Pelte N, Boutros M, Marques JT, Daeffler L, Hoffmann JA, Roussel A, Imler JL. 2016. Cytokine Diedel and a viral homologue suppress the IMD pathway in *Drosophila*. Proc Natl Acad Sci 113:698–703.

53. Kincaid RP, Sullivan CS. 2012. Virus-encoded microRNAs: an overview and a look to the future. PLoS Pathog 8:e1003018.

54. Grundhoff A, Sullivan CS. 2011. Virus-encoded microRNAs. Virology 411:325–343.

55. Hussain M, Taft RJ, Asgari S. 2008. An insect virus-encoded microRNA regulates viral replication. J Virol 82:9164–9170.

56. Jayachandran B, Hussain M, Asgari S. 2012. RNA interference as a cellular defense mechanism against the DNA virus baculovirus. J Virol 86:13729–13734.

57. Weber F, Wagner V, Rasmussen SB, Hartmann R, Paludan SR. 2006. Double-stranded RNA is produced by positive-strand RNA viruses and DNA viruses but not in detectable amounts by negative-strand RNA viruses. J Virol 80:5059–5064.

58. Bronkhorst AW, van Cleef KW, Vodovar N, Ince IA, Blanc H, Vlak JM, Saleh MC, van Rij RP. 2012. The DNA virus Invertebrate iridescent virus 6 is a target of the *Drosophila* RNAi machinery. Proc Natl Acad Sci USA 109:E3604–E3613.

59. deFaria IJ, Olmo RP, Silva EG, Marques JT. 2013. dsRNA sensing during viral infection: lessons from plants, worms, insects, and mammals. J. Interferon Cytokine Res 33:239-253.

60. Kemp C, Mueller S, Goto A, Barbier V, Paro S, Bonnay F, Dostert C, Troxler L, Hetru C, Meignin C, Pfeffer SJ, Hoffmann A, Imler, JL. 2013. Broad RNA interference-mediated antiviral immunity and virus-specific inducible responses in Drosophila. J Immunol 190:650–658.

61. Hussain M, Asgari S. 2010. Functional analysis of a cellular microRNA in insect host-ascovirus interaction. J Virol 84:612-620.

62. Yu H, Li ZQ, He L, Ou-Yang YY, Li N, Huang GH. 2018. Response analysis of host Spodoptera exigua larvae to infection by Heliothis virescens ascovirus 3h (HvAV-3h) via transcriptome. Scientific Reports 8:5367.

63. CPL Business Consultants. 2010. The 2010 Worldwide Biopesticides Market Summary, Volume 1, June 5, 2010 - Pub ID: BGEQ2703518. MarketResearch.com, Rockville, MD, USA.

64. Asgari S. 2006. Replication of *Heliothis virescens* ascovirus in insect cell lines. Arch Virol 151**:**1689-1699.

Chapter II: Transcriptome Analysis of the Spodoptera frugiperda Ascovirus *In Vivo* Provides Insights into How Its Apoptosis Inhibitors and Caspase Promote Increased Synthesis of Viral Vesicles and Virion Progeny

2.1 Abstract

Ascoviruses are double-stranded DNA (dsDNA) viruses that attack caterpillars and differ from all other viruses by inducing nuclear lysis followed by cleavage of host cells into numerous anucleate vesicles in which virus replication continues as these grow in the blood. Ascoviruses are also unusual in that most encode a caspase or caspase-like proteins. A robust cell line to study the novel molecular biology of ascovirus replication *in vitro* is lacking. Therefore, I used strand-specific transcriptome sequencing (RNA-Seq) to study transcription *in vivo* in third instars of *Spodoptera frugiperda* infected with the type species, *Spodoptera frugiperda ascovirus 1a* (SfAV-1a), sampling transcripts at different time points after infection. I targeted transcription of two types of SfAV-1a genes; first, 44 core genes that occur in several ascovirus species, and second, 26 genes predicted *in silico* to have metabolic functions likely involved in synthesizing viral vesicle membranes. Gene cluster analysis showed differences in temporal expression of SfAV-1a genes, enabling their assignment to three temporal classes: early, late, and very late. Inhibitors of apoptosis (IAP-like proteins; ORF016, ORF025, and ORF074) were expressed early, whereas its caspase (ORF073) was expressed very late, which correlated with apoptotic events leading to viral vesicle formation. Expression analysis revealed that a Diedel gene homolog (ORF121), the only known "virokine," was highly expressed, implying that this ascovirus protein helps evade innate host immunity. Lastly, single

nucleotide resolution of RNA-Seq data revealed 15 bicistronic and tricistronic messages along the genome, an unusual occurrence for large dsDNA viruses.

2.2 Introduction

Ascoviruses are insect-pathogenic double-stranded DNA (dsDNA) viruses (family *Ascoviridae*) that attack the larval and pupal stages of lepidopterans, causing a long-term but ultimately fatal disease (1–5). They are characterized by a combination of unique features that easily distinguish them from other insect viruses, including large virion size (400 by 150 nm), reticulate virion envelope structure, and especially the formation of virion-containing vesicles derived by cleavage of infected host cells. This process resembles apoptosis but appears to differ in that what appear to be developing apoptotic bodies continue to grow and produce virions, eventually accumulating in host blood and turning it a milky white color (6). These viruses are also unique in that most have a caspase or caspase-like genes, and in at least one case, an executioner caspase was confirmed to play a role in the changes in cell architecture that lead to the development of virion-containing vesicles (7).

The unique ability of ascoviruses to apparently manipulate the apoptotic pathway, leading to the formation of virion-containing vesicles, i.e., viral vesicles, is a complex process with a high degree of cellular reorganization that possibly involves both cellular and viral genes (8, 9). Although the role of the executioner caspase may be significant in the initiation of apoptosis, other viral and host proteins that participate in the viral vesicle formation remain to be identified. Identification of these genes and how they are

temporally coordinated will contribute to our knowledge of the underlying molecular biology of this novel cytopathology in which infected cells are reprogrammed to produce the viral vesicles that enable virogenesis to continue as these circulate in the blood.

Unfortunately, understanding the molecular basis of ascovirus cytopathology is hindered by the lack of a suitable cell line in which to study viral gene expression and the interaction with host cell genes. Although previous studies show the permissiveness of noctuid insect cell lines for ascovirus propagation, viral replication varies significantly among different cell lines (10). More importantly, vesicle formation, the primary distinguishing feature of ascoviruses, is very rare and delayed in all cell lines infected with HvAV-3e (10). Moreover, expression of Spodoptera frugiperda ascovirus (SfAV) core genes in Sf9 cells was very limited and aberrant (7), yielding few virions compared to the large quantities produced *in vivo* (8).

To begin to address the knowledge gaps regarding the unique molecular biology of how ascoviruses manipulate cell architecture to generate viral vesicles, I used third-instar larvae of *Spodoptera frugiperda* infected with the ascovirus type species, *Spodoptera frugiperda ascovirus 1a* (SfAV-1a), coupled with transcriptome sequencing (RNA-Seq) technology to gain basic information about the expression of core and other viral genes. SfAV-1a is known to target mainly the insects' fat body for infection, but other ascovirus species (Heliothis virescens ascovirus [HvAV] and Trichoplusia ni ascovirus [TnAV]) have a broad tissue tropism: they infect the fat body, epidermis, and tracheal matrix (8).

Regardless of the ascovirus species and their differences in tissue tropisms, all ascoviruses are known to induce the same cytopathology, and this has been correlated with the core gene set that is shared by different ascoviruses (8, 9). Ascovirus genomes are circular and range in size from 100 to 200 kbp. For the present study, I selected SfAV-1a, which has a genome of 156,922 bp, with GC ratio of 49.2%, and codes for 123 putative proteins (11). The functions of many of these proteins fall within five main classes, namely, proteins associated with nucleotide metabolism, apoptosis, and lipid metabolism, virion structural proteins, and host interaction proteins. However, unlike baculoviruses (family *Baculoviridae*), the most widely studied dsDNA viruses that infect caterpillars, and for which the function of many proteins are known, the functions of most ascovirus proteins remain unknown. Some ascovirus proteins are related to those of baculoviruses and the related nudiviruses, but phylogenetic analysis of ascovirus DNA polymerases and major capsid proteins (MCPs) indicates that ascoviruses are more closely related to iridescent viruses and phycodnaviruses than to baculoviruses (8, 11). Thus, my SfAV-1a transcriptome study was aimed at defining the expression patterns of this virus as a prelude to defining the functions of the proteins involved in its unique reprogramming of infected cells to produce progeny virions and impede innate host immunity.

Here I report the results of my initial study, in which I determined the temporal core gene expression of SfAV-1a *in vivo*. Three temporal classes were defined: early, late, and very late. Among other results, I found that apoptosis inhibitors (ORF016, ORF025, and ORF074) were expressed early, whereas the caspase (ORF073) was

expressed very late, demonstrating coordination in this ascovirus in which virus replication was enabled first, followed much later by apoptosis and viral vesicle formation. In addition, exploration of the RNA interference pathway through expression of RNase III (ORF022) and high levels of a Diedel homolog (ORF121), a virokine, in the SfAV-1a transcriptome suggest that the function of these proteins is to overcome the host's innate immune response. Finally, the detection of 15 bi-/tricistronic mRNA messages in the SfAV-1a transcriptome is very unusual for a large DNA virus, showing that this ascovirus and likely others have noncanonical translational mechanisms.

2.3 Materials and methods

2.3.1 Virus strain and host infection

I used a wild-type strain Sf82–126 isolate of Spodoptera frugiperda ascovirus 1a (SfAV-1a) (12), the type species for the family *Ascoviridae*. This strain was used to infect 3rd-instar caterpillars of *Spodoptera frugiperda*, and all transcriptome data were derived from this developmental stage. Larvae were reared on artificial medium (Benzon Research, Carlisle, PA) and kept at room temperature (22°C). An unusual feature of ascoviruses is that feeding larvae virions or viral vesicles results in low infection levels. To mimic natural infection by parasitic wasps, a minutin pin was contaminated with SfAV-1a virions by dipping it in a suspension of viral vesicles (1 X 10^8/ml), after which the pin was inserted into the larva's abdomen.

Progression of the infection was monitored daily by examining hemolymph color changes and using phase-contrast microscopy to follow viral vesicle accumulation in this tissue.

2.3.2 Total RNA isolation

One milliliter of TRIzol reagent (Invitrogen, Life Technologies) per sample was used for isolation of total RNA from 7 time points, starting with uninfected (healthy) larvae and then sampling at 6, 12, 24, 48, 96, and 168 h post-infection (hpi). Total RNA was isolated from 3 larvae at each time point. The entire body of each larva was homogenized separately by mechanical trituration in TRIzol, followed by the addition of chloroform and separation of the RNA in the aqueous phase. The RNA was precipitated with 100% isopropanol and washed twice in 75% ethanol. RNA quantity and quality were determined using a Thermo Scientific NanoDrop ND-2000c spectrophotometer.

2.3.3 Purification and DNase treatment of RNA

An RNA Clean and Concentrator TM-5 kit (ZYMO RESEARCH) accompanied by an in-column DNase digestion using an RNase-Free DNase set (Qiagen) was used for concentrating the RNA and to eliminate contaminating DNA, respectively. After treatment, the RNA quantity was determined as described above.

2.3.4 Enrichment of mRNA, RNA-Seq library preparation, and sequencing

The NEBNext poly(A) mRNA magnetic isolation module kit (New England BioLabs) was used to enrich the mRNA from the DNase-treated total RNA samples. The poly(A) RNA was then eluted from the oligo $d(T)_{25}$ attached to paramagnetic beads in 15 µl of the first cDNA strand synthesis reaction buffer combined with random primer mix (2X), followed by 10 min of heating at 94°C. The RNA-Seq libraries were then prepared from the isolated mRNA following the manufacturer's instructions of the NEBNext Ultra Directional RNA library preparation kit (New England BioLabs) for Illumina for two biological replicates at each time point. After library preparation, the quality of the pooled libraries was determined using an Agilent 2100 bioanalyzer and the pooled libraries were sequenced using the HiSeq2500 sequencer (Illumina) in the UCR Core Facility in the Institute for Integrative Genome Biology.

2.3.5 SfAV-1a ORFs reannotation based on the RNA-Seq data

Initial mapping of the RNA-Seq reads with the published SfAV-1a genome (11) indicated the presence of a number of mismatches between the data and the reported sequence of SfAV-1a. Thus, I used the RNA-Seq data to reannotate the original genome sequence of SfAV-1a prior to proceeding with the transcriptome analysis. I initially edited the SfAV-1a genome available through the NCBI database (accession number NC_008361.1) only at those positions where 100% of the RNA-Seq reads agreed on a change to the published genome, followed by separation of plus- and minus-strand reads to facilitate reannotation

of genes in overlapping transcript regions and for accurate determination of transcript start and stop positions. The transcript start and stop positions were manually curated. The ORF-Finder (NCBI) was used for ORF predictions.

Two types of SfAV-1a genes were specifically targeted for further analysis. The first consisted of 44 core genes conserved among two or more ascovirus species, many of them containing conserved domains and motifs identified by BLASTP (**Table 1**). Two proteins conserved in both SfAV and HvAV and not in TnAV or DpAV were added to this set for either their high expression or their putative function that could be associated with SfAV-1a cytopathology. The second set consists of 26 SfAV-1a proteins identified through *in silico* analyses as having at least one or as many as seven transmembrane-spanning motifs using the TMHMM server v.2.0 (13, 14) (**Table 1**).

2.3.6 Bioinformatic analysis, RPKMs, and read mapping

RNA-Seq libraries were first filtered for low-quality reads using the FASTX toolkit (Hannon Lab [http://hannonlab.cshl.edu/fastx_toolkit/index .html]); adapter sequences were removed using Trimmomatic (15), and ribosomal contaminants were identified and removed (16). These filtered libraries were mapped to the SfAV-1a genome sequence deposited in the NCBI database (accession number NC_008361.1) using bowtie2 (17). Following reannotation of the SfAV-1a genome (see above), libraries were remapped to the reannotated genome using bowtie2. Transcript expression was estimated using reads per kilobase per million mapped reads (RPKMs). Because RNA-Seq reads were

sequenced using a directional protocol, only reads representing reads in the 5'-to-3' direction of each annotated transcript were used for RPKM calculations.

2.3.7 Promoter motif signal detection

Genomic sequences 100 bp upstream of the transcription start site of all the ORFs with a defined 5' end in the 2 gene sets (core and TMD genes) were examined to identify consensus promoter motifs for different stages of infection (early, late, and very late) using the MEME suite (18). In addition, the sequences 100 bp upstream of the transcription start site for each bi-/tricistronic message were screened using the MEME suite for identification of conserved consensus motifs in their promoters. I excluded the promoter sequence for any gene (core or TMD) that is part of a bi-/tricistronic message in analysis of monocistronic early, late, and very late genes.

2.3.8 RT-qPCR validation of the RNA-Seq data

To validate the RNA-Seq data, three viral genes were selected randomly for validation of their RPKM expression levels using RT-qPCR at 3 time points (24, 48, and 96 hpi). One microgram of the same total RNA samples as used in the RNA-Seq library preparation was reverse transcribed from each sample using a Maxima first-strand cDNA synthesis kit for RT-qPCR (Thermo Scientific). Real-time reactions were carried out using the iQ SYBR green Supermix (Bio-Rad) including 5 µl of 1:25 cDNA in each 25-µl reaction volume. The PCR was performed in a CFX Connect real-time PCR detection system

(Bio-Rad) using the following program: 95°C for 3 min, 40 cycles of 90°C for 10 s and 60°C for 30 s, followed by 72°C for 30 s. Melting curves were generated for each reaction at the end of the PCR run to confirm the specificity of the PCR. The final concentration used from each primer was 0.5 M. Primers were designed using the IDT real-time PCR design tool (**Table 3**). Standard curves were generated for each gene to calculate primer efficiency. The host L8 ribosomal gene was used for normalization since it showed stability in all the selected time points (19). Two technical and three biological replicates were tested for each sample. The third biological replicate total RNA samples were included in this test. The 24 hpi RNA was used as the reference. Each set of reactions included a no-template control and negative RT-PCR control to confirm absence of DNA contamination. The relative quantification was calculated based on the Pfaffl equation (20).

2.3.9 Validation of the bicistronic and tricistronic messages

Three tricistronic and two bicistronic messages out of the 15 identified SfAV-1a bi-/tricistronic mRNAs (**Table 2.2**) were further confirmed by using specific primer sets (**Table 2.3**) designed to amplify these messages from the 24 hpi DNase-treated total RNA sample after reverse transcriptase PCR using the following program: 94°C for 4 min followed by 30 cycles of 94°C for 20 s and annealing at 58 to 60°C for 20 s and extension at 72°C for 3 min, with a final extension at 72°C for 5 min. The Maxima first-strand cDNA synthesis kit for RT-qPCR (Thermo Scientific) was used for the reverse transcription reaction, and the PCR PreMix tubes (Bioneer) were used for the specific

primer amplification reaction. The PCR product size of each reaction and its negative RT control was confirmed on a 1% agarose gel using Tris-borate-EDTA (TBE) running buffer (Fisher Scientific). All the PCR products were sequenced using Sanger sequencing.

2.3.10 IRESPred for IRES prediction in intergenic regions

I used the IRESPred web server, available at http://bioinfo.net.in/IRESPred/, for screening of putative internal ribosome entry site (IRES) sequences in the intergenic regions identified in the bicistronic and tricistronic messages (**Table 2**). IRESPred is designed to predict both cellular and viral IRES structures. However, I was not able to test intergenic regions less than 15 bp in size because of program restrictions (21).

2.3.11 Accession number(s)

All the RNA-Seq data obtained in this study have been deposited in the NCBI Gene Expression Omnibus (GEO) and can be accessed through the GEO series accession number GSE98382.

2.4 Results

2.4.1 Editing and reannotation of SfAV-1a genes based on RNA-Seq data

The SfAV-1a replication progress and host permissiveness were illustrated by gradual increase in the number of viral reads mapping to the viral genome (in comparison to total reads, including those of the host), with only 0.007% at 6 h post-infection (hpi), then

0.06% at 12 hpi, 1.71% at 24 hpi, 1.87% at 48 hpi, 3.64% at the 4th day pi, and 3.76% at the 7th day pi. By 24 hpi, the viral reads were covering most of the SfAV-1a genome except the inverted-repeat areas, which had no transcription (11). Mapping of SfAV-1a reads to the genome revealed numerous mismatches with the reported genome sequence, including in reported open reading frames (ORFs). Therefore, editing these sequence differences in all positions that showed 100% mismatch by manual curation was an essential prerequisite to determine whether the ORFs were altered and to perform accurate expression pattern analysis. In general, RNA sequence technology is known to significantly improve genome reannotation (22, 23), and my edits of the published sequence corrected a number of the reported ORFs (**Table 2.1**). In this analysis, I focused on two SfAV-1a gene sets. The first consisted of 44 core genes, present in more than one ascovirus species, identified in my analysis by BLASTP (NCBI) (**Table 2.1**). Many of these genes have been reported in previous studies on the evolutionary relationships of ascoviruses (2, 8, 9). The functions of genes in this set are distributed among five classes, namely, genes associated with nucleotide metabolism, lipid metabolism, apoptosis, structural proteins, or host interaction proteins. Two SfAV genes, which appear to be conserved only in SfAV and HvAV, were added to this set because they were highly expressed or due to putative functions associated with SfAV-1a cytopathology. The second set consisted of 26 proteins, representing the entire transmembrane domain (TMD)-containing proteins identified after the genome reannotation using the TMHMM server v.2.0 (13, 14). The number of TMDs ranged from 1 to 7. Many of the TMD ORFs had no known function, and as a group they are unique to ascoviruses. Thus, these were

selected because of their possible association with the outer cell membrane modifications, making them good candidates for participation in viral vesicle formation. Eight of the 26 genes were also represented in the conserved SfAV-1a gene set, including genes for enzymes capable of modifying membrane lipids, more specifically, a patatin-like phospholipase, a PlsC phosphate acyltransferase, a fatty acid elongase, and an esterase/lipase. These enzymes are known to modify such things as fatty acid chain length and the degree of lipid saturation (9, 11). A comparison of the newly annotated SfAV-1a two gene sets with the original genome map positions is provided in **Table 2.1**.

2.4.2 Cluster analysis of SfAV-1a expressed core and TMD genes

Cluster analysis of the SfAV-1a core and TMD genes was performed based on calculations of the reads per kilobase per million mapped reads (RPKMs) to determine the transcription dynamics for each gene. Clusters were categorized into three expression classes, namely, early, late, and very late (**Fig. 2.1 and 2.2**). The early gene cluster consisted of 13 core and 4 TMD genes expressed beginning at 6 hpi. All early genes continued to be expressed during later time points and were not downregulated. The functions of early core genes mainly dealt with nucleotide metabolism (ORF040, ORF059, ORF066, ORF095, ORF099, and ORF104), inhibition of apoptosis (ORF016, ORF025, and ORF074), and host interaction (ORF055, ORF072, and ORF107). Only one structural protein was detected in this class (yabby-like transcription factor [corresponding to ORF091]). The four early TMD genes were ORF002, ORF007, ORF039, and ORF085. ORF002 corresponds to a virion structural protein, but its specific

function is unknown. None of the four lipid-modifying enzymes were detected in this early stage (**Fig. 2.1 and 2.2**).

The majority of the core (i.e., 28) and TMD (i.e., 16) genes were found to start transcription during the late stage of infection, at 12 hpi. The putative functions of the two gene sets centered mainly on DNA and RNA metabolism (**Fig. 2.1 and 2.2**). In addition, six structural proteins were transcribed in this stage, including the two major SfAV-1a structural proteins, namely, the major capsid protein (ORF041) and the DNA-condensing P64 protein (ORF048 [24]). The late gene cluster was characterized by expression of three core genes coding for lipid-modifying enzymes that have TMDs, namely, esterase (ORF013), fatty acid elongase (ORF087), and phosphate acyltransferase (ORF112). The fourth lipid patatin-like enzyme (ORF093) was detected very late in the infection. Thus, expression of lipid-modifying enzymes late and very late was consistent with the synthesis of viral vesicles.

In addition, five core and six TMD genes were expressed very late in the SfAV-1a infection cycle (expression starting at 24 hpi). Two of the core genes were identified previously as SfAV-1a virion structural proteins, namely, the Ervl/Alr family protein (ORF061) and the serine/threonine protein kinase (ORF064) (24). Both the caspase (ORF073) and putative thioredoxin-like proteins (ORF116) were detected in this cluster, where they shared similar transcription patterns for the different time points post-infection. The remainder of the very late SfAV-1a genes identified were TMD genes with unknown functions (ORF012, ORF032, ORF045, ORF062, and ORF122), with the exception of ORF93, which was identified as corresponding to a fourth lipid-modifying

enzyme (patatin-like phospholipase). Four SfAV-1a genes (ORF121, ORF107, ORF032, and ORF055) were highly expressed (>1,000 RPKMs) at certain stages of the infection, especially compared to many of the genes (<100 RPKMs) in this analysis. Interestingly, ORF121 is a late gene identified as a homolog of the *Drosophila* Diedel protein gene (*die* gene) (25). In addition, ORF107 is an early gene sharing homology with the trifunctional transcriptional regulator/ proline dehydrogenase/L-glutamate gamma-semialdehyde dehydrogenase in *Wiggles worthia glossinidia* (11). The other highly expressed ORFs (ORF032 and ORF055) represent good candidates for future studies because they are highly expressed according to my analysis and have no known function. The ORF055 product shares homology with the 254L protein in IIV-6 and is conserved in all ascovirus species, while ORF032 corresponds to a TMD protein conserved in HvAV and TnAV. Other highly expressed SfAV-1a genes (>100 RPKMs) that have interesting putative functions included genes for a BRO-like protein (ORF079), a transposase (ORF077), a putative VLTF2-like late transcription factor/Zn finger DNA binding protein (ORF113), an inhibitor of apoptosis (IAP)-like protein (ORF016), the virion structural protein yabby-like transcription factor (ORF091), and a thymidine kinase (ORF040). Also, five highly expressed (>100 RPKMs) TMD genes include those for 2 hypothetical proteins (ORF007 and ORF060), a virion structural myristylated membrane protein (ORF054), and a protein (ORF097) with a conserved RING domain (11).

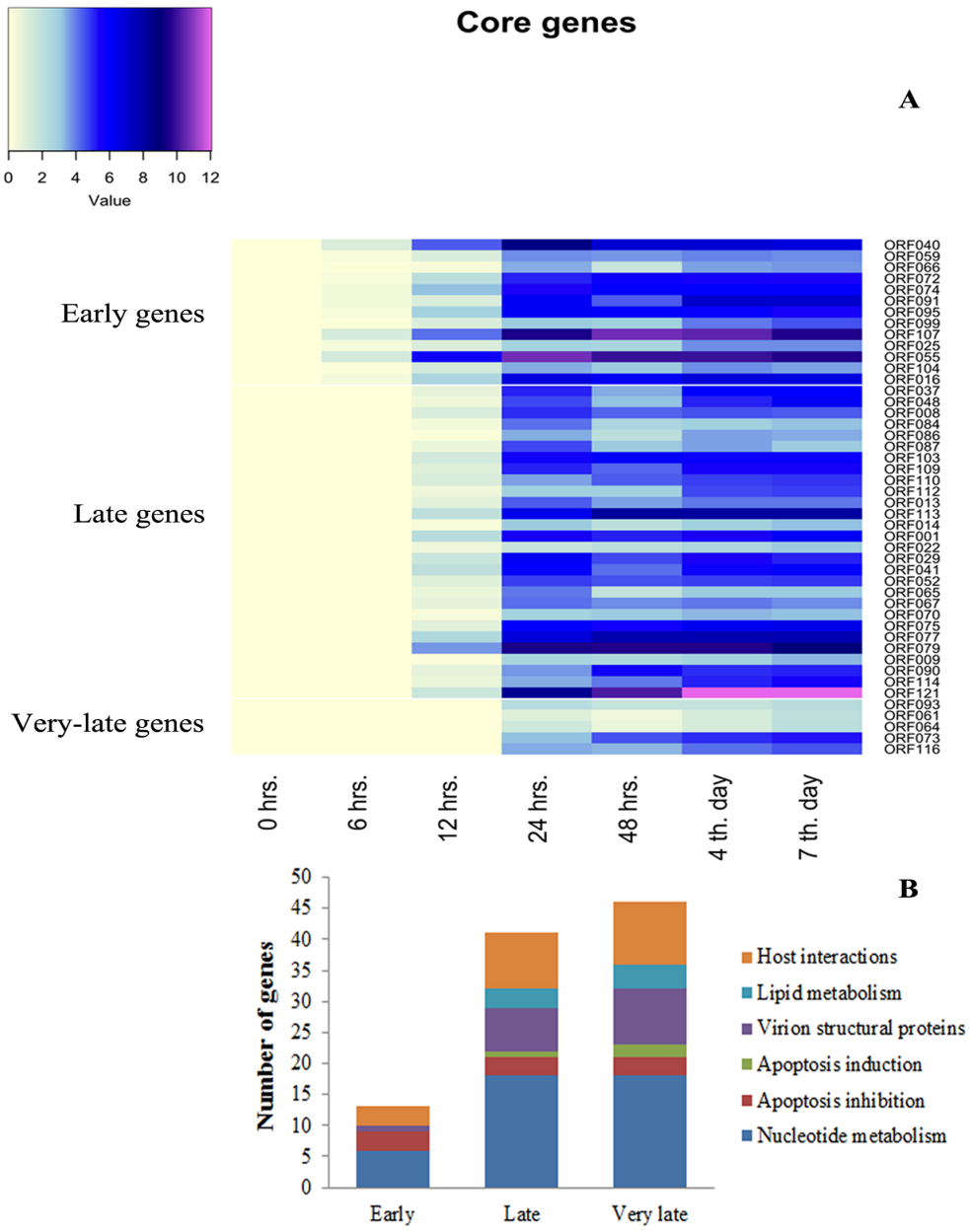


FIG 2.1 Cluster analysis of the SfAV-1a core gene temporal transcription pattern identified *in vivo* during infection of the 3rd-instar larvae of *Spodoptera frugiperda*. (**A**) Heat map representation of SfAV-1a expression of each core gene. Core genes are common genes identified in many ascovirus species. The temporal clusters are indicated on the map as early, late, and very late and are separated by white horizontal lines. Heat map colors represent the difference in transcript expression (measured in $\log_2$ RPKMs) from average replicate expression levels at 0 h. (**B**) Distribution of putative function of genes in each gene cluster.

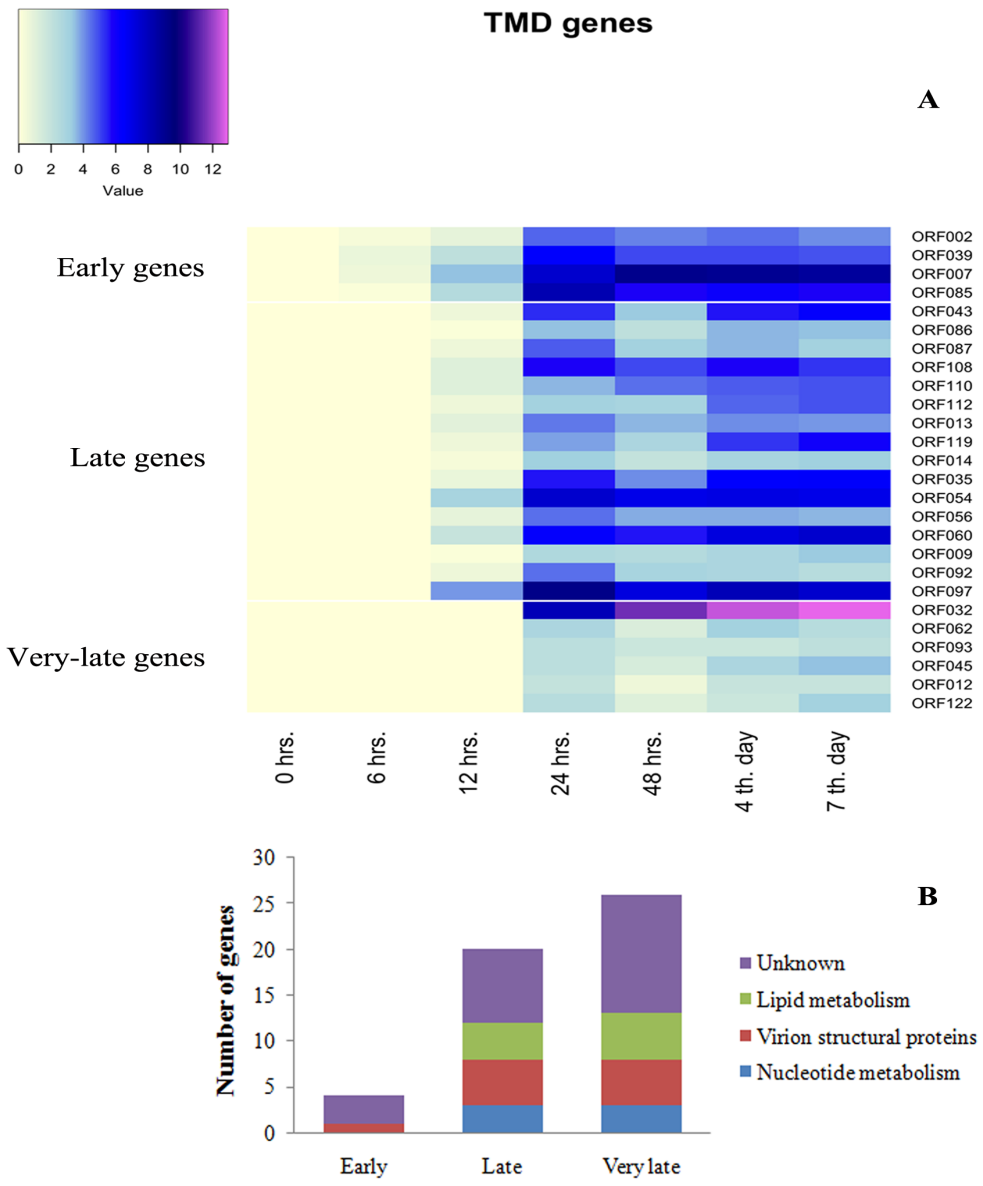


FIG 2.2 Cluster analysis of the SfAV-1a transmembrane domain (TMD)-containing genes temporal transcription pattern identified *in vivo* during infection of 3rd-instar larvae of *Spodoptera frugiperda*. (**A**) Heat map representation of SfAV-1a expression of each TMD gene. TMD genes were identified using TMHMM server v.2.0 (13, 14). The temporal clusters are indicated on the map as early, late, and very late and are separated by white horizontal lines. Heat map colors represent the difference in transcript expression (measured in $\log_2$ RPKMs) from average replicate expression levels at 0 h. (**B**) Distribution of putative function of genes in each gene cluster.

2.4.3 Identification of SfAV-1a bicistronic and tricistronic messages

A total of nine putative bicistronic and six tricistronic mRNA transcripts were detected in the SfAV-1a transcriptome, with two or three ORFs per transcript, respectively. A description of these bicistronic and tricistronic transcripts, with the position, putative function corresponding to the ORFs, intergenic regions, and poly(A) signal positions near the 3' end of the message, are listed in **Table 2.2** and described schematically in **Fig. 2.3**. Interestingly, many of the messages were missing a poly(A) signal near the 3' end, with RNA secondary structures, in these cases, observed in some messages. Involvement of stem-loop structure in transcription termination has been reported before for the major capsid protein of Spodoptera exigua ascovirus 5a (SeAV-5a) (26).

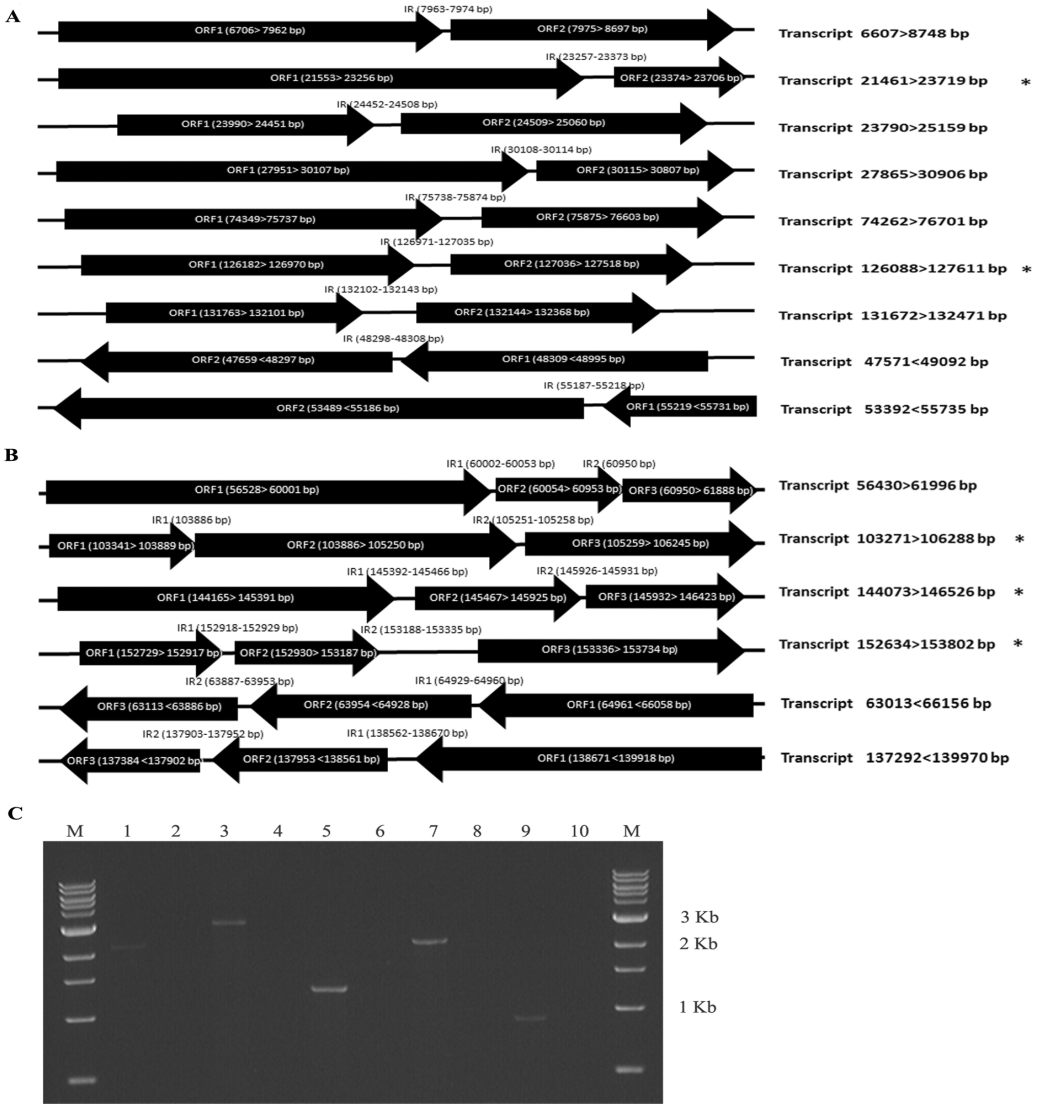


FIG 2.3 Description and validation of SfAV-1a bicistronic (**A**) and tricistronic (**B**) messages identified. IR, intergenic region. ">" refers to forward orientation and "<" refers to reverse orientation. (**C**) RT-PCR validation of mRNA message size for two bicistronic and three tricistronic messages (marked with asterisk) amplified from 24 hpi DNase-treated total RNA isolated from SfAV-1a-infected 3rd-instar larvae using specific primers set for each message (described in **Table 2.3**). Lanes 1, 3, 5, 7, and 9 represent the PCR product. Lanes 2, 4, 6, 8, and 10 contain negative RT-PCR controls, in which reverse transcriptase was omitted to confirm the absence of DNA contamination in the RNA samples. Lane M shows 1-kb DNA markers (New England BioLabs).

2.4.4 Validation of the SfAV-1a bicistronic and tricistronic messages

As noted above, I edited the SfAV-1a genome based on mapped sequence reads in different positions where only 100% mismatches were considered real. Thus, I was interested in further validating the bicistronic and tricistronic message sequences to confirm that these were not artifacts of genome editing. Therefore, five messages identified encoding apoptosis inhibitors or immune evasion proteins were selected for further validation by reverse transcriptase PCR (RT-PCR) followed by using specific primer sets for each mRNA message (**Tables 2.2 and 2.3**). Sequences of the amplified PCR products were then confirmed by Sanger sequencing. A 100% sequence homology was detected in all, and message size was confirmed by gel electrophoresis. No DNA contamination was detected in the negative RT-PCR controls (**Fig. 2.3**).

2.4.5 Validation of the RNA-Seq data by RT-qPCR analysis

I used RT-quantitative PCR (RT-qPCR) to validate the RNA-Seq data. Three SfAV-1a genes (ORF016, ORF032, and ORF037) were selected randomly for comparison of their transcription patterns at three time points (24, 48, and 96 hpi). The RT-qPCR data for the three genes were in agreement and consistent with the RNA-Seq data as illustrated in **Fig. 2.4.**

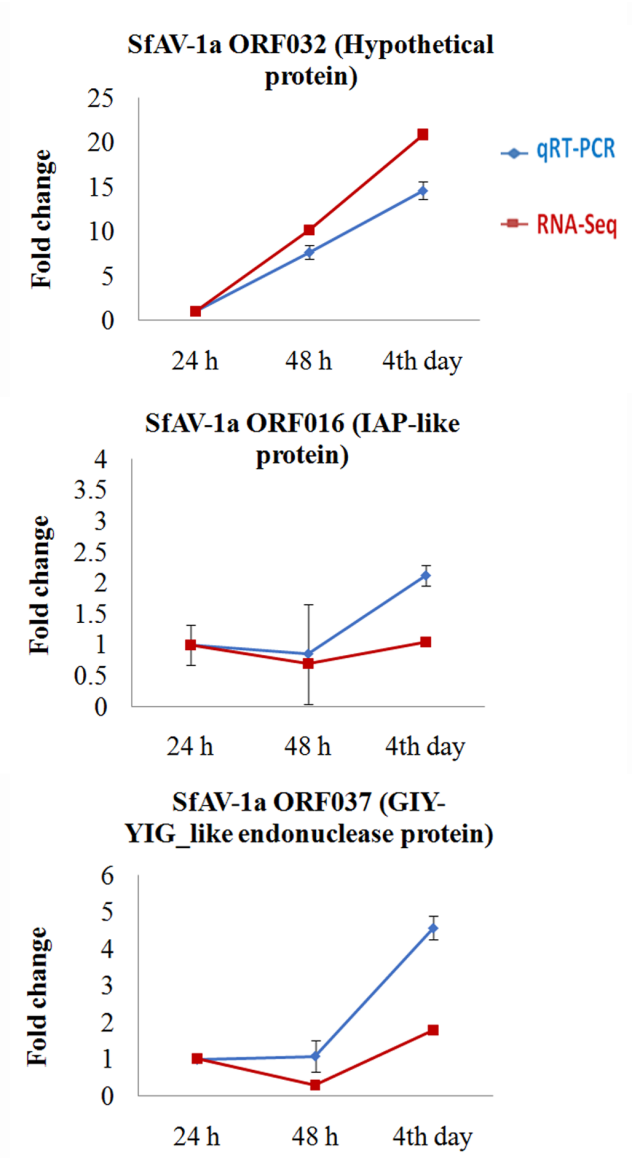


FIG 2.4 Comparison of the RNA-Seq profile with the data obtained from RT-qPCR for three viral genes (ORF032, ORF016, and ORF037). The host L8 ribosomal gene was used for normalization of the qRT-PCR data (19). Two technical and three biological replicates were tested for each sample. The 24 hpi sample was used as the reference for fold change calculations at indicated infection time points. The error bars represent the standard deviations between the three biological replicates at each time point.

2.4.6 SfAV-1a promoter consensus sequences

MEME analysis revealed conservation of different consensus motifs in the promoter sequences of several SfAV-1a core and TMD genes representing different temporal gene clusters (**Fig. 2.5**). Moreover, screening of the promoter sequences of the bi-/tricistronic messages identified a consensus motif (**Fig. 2.5**). I was able to detect the previously identified late consensus sequences TATA box-like motif (TAATTAAA) and ATTTGATCTT in two late genes, namely, ORF22 (RNase III) and ORF41 (MCP), respectively (26, 27). The TATA box like motif is located 16 bp before the ATG start codon, and the ATTTGATCTT consensus sequence is 28 bp before the ORF start codon. The two motifs are shared between ascovirus MCPs and IIV-6 (late genes). Previous molecular phylogenetic studies provide evidence that the ascoviruses evolved from iridoviruses (28). However, I did not find these motifs in other genes, i.e., in the SfAV-1a core and TMD genes I studied.

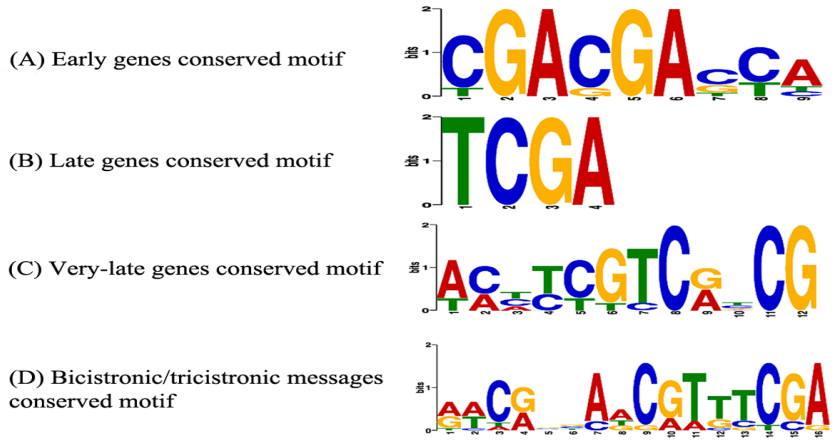


FIG 2.5 Conserved promoter motifs in regions 100 bp upstream of the transcription start site for certain SfAV-1a early, late, and very late core genes (**A to C**), including late transmembrane domain-containing genes, and bicistronic/tricistronic messages (**D**) using the MEME motif discovery suite (18).

2.4.7 Prediction of putative IRES sequences using IRESPred

Putative internal ribosome entry site (IRES) sequences were identified using the IRESPred server (http:// bioinfo.net.in/IRESPred/) in the intergenic regions of some bi-/tricistronic messages (**Table 2.2**), implying the possible use of noncanonical protein translational mechanisms such as IRESs by SfAV-1a during translation of some of these messages.

2.5. Discussion

A potential problem frequently encountered with RNA-Seq *in vivo* model systems is a low percentage of pathogen transcripts, which makes a full transcriptome study difficult because of the comparatively high levels of host RNAs (29). However, in this *in vivo* insect larval system I was able to detect strand-specific RNA sequences from most of the viral genes, even those expressed at a low level, making this system appropriate for transcriptome analysis. This is a good indication of the sensitivity of RNA-Seq technology for detection of SfAV-1a RNAs in this system and the possibility of using it in future studies for this virus family, especially given existing limitations in the *in vitro* systems (10). Moreover, *in vivo* RNA-Seq provides actual data on which viral genes are essential during host infection, thus leading to a better understanding of how these viruses establish infections in their caterpillar hosts. Several recent studies have confirmed significant differences between the *in vitro* and *in vivo* systems in which expression patterns between the same genes may vary as much as a 100-fold (30, 31). I focused the

transcription analysis on the core genes, i.e., those conserved in many ascovirus species, and transmembrane domain (TMD)-containing genes, since the latter represent good candidates to be either directly or indirectly associated with ascovirus vesicle formation, the distinguishing feature of ascovirus cytopathology. However, the single-base resolution of the virus transcripts in my analysis revealed the need for SfAV-1a genome reannotation to match the transcription pattern of the virus. Therefore, the sequence and position for genes included in this study were revised first to accurately estimate transcription levels (**Table 2.1**).

The temporal cluster analysis revealed a possible division of SfAV-1a gene transcription into three groups, early, late, and very late, a transcription pattern similar to that in other large DNA viruses, such as baculoviruses (32). In the early gene cluster expressed at 6 hpi, three inhibitor of apoptosis (IAP)-like proteins (ORF016, ORF025, and ORF074) with a RING domain containing an E3 ubiquitin ligase near the C terminus was detected in each. IAPs are known to block the apoptosis pathway through caspase degradation (33). Early expression of apoptosis inhibitors was detected in a baculovirus temporal expression analysis, in which the antiapoptosis protein p35 was detected after 6 hpi (32). The early transcription of apoptosis inhibitors can serve as an essential step in success of viral replication since both baculoviruses and ascoviruses are known to induce cell death. Therefore, early expression of proteins such as p35 or IAPs inhibiting or halting apoptosis induction may serve as a general mechanism for infection in DNA viruses. The delay in synthesis of the SfAV-1a executioner caspase (ORF073) and the early IAP-like protein expression indicate coordination between apoptosis inhibition and

the induction at specific stages of infection by this virus that ultimately enhances virus replication (**Fig. 2.6**). The caspase activity of this virus was studied previously *in vitro*, where its association with apoptosis induction was confirmed, the synthesis of this enzyme being initiated 9 hpi (7). Differences between *in vitro* and *in vivo* viral gene expression have been reported for other viruses, where the host immune system barriers and cell-specific factors are associated with these differences (34). In addition, detection of both a caspase (ORF073) and putative thioredoxin-like protein (ORF116) in the very late stage, where they shared similar transcription patterns at different time points post-infection, is reasonable. Thioredoxin is a small protein that plays a role in regulation of cell redox state. Maintenance of the cell redox state is essential for cell death induction through apoptosis rather than necrosis (35-37). Alternatively, thioredoxin may have an antiapoptotic activity through its reactive oxygen species (ROS) scavenging ability (37, 38). Moreover, detection of a protein with proline dehydrogenase activity (ORF107) highly expressed (>1,000 RPKMs) may indicate association of other apoptosis inducers in ascovirus cytopathology coordinated with caspase activity. Proline dehydrogenase can induce intrinsic apoptosis by catalyzing formation of reactive oxygen species (39).

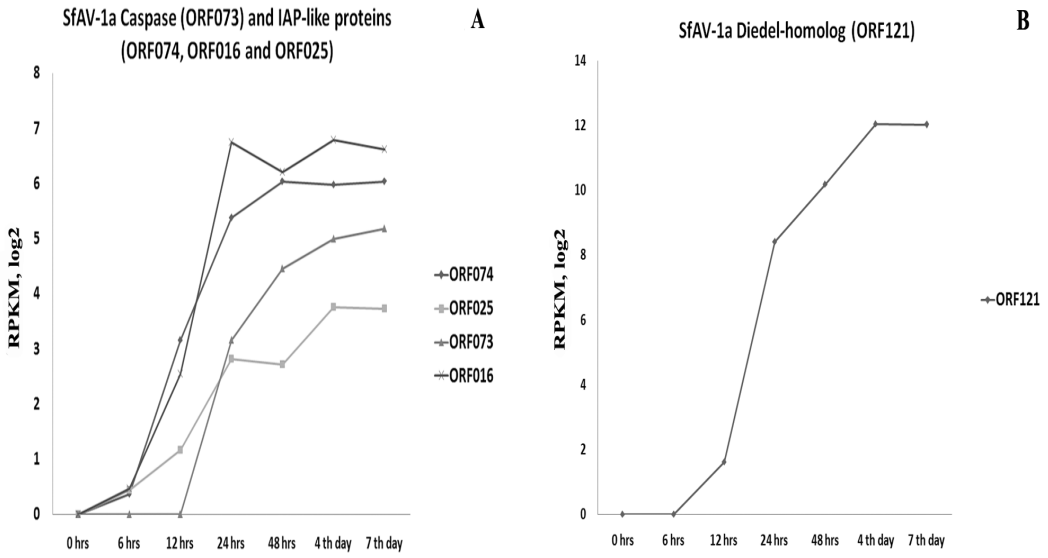


FIG 2.6 (**A**) SfAV-1a caspase (ORF073) very late expression (at 24 hpi) versus early expression (at 6 hpi) of putative inhibitor of apoptosis (IAP-like proteins) ORFs (ORF074, ORF016, and ORF025). (**B**) SfAV-1a Diedel homolog (ORF121) expression pattern *in vivo*.

It is important to note here that although the data for the SfAV-1a caspase suggest that it plays an important role in the changes in cell architecture that lead to viral vesicle formation, there is little support that the caspase-like genes that occur in other ascoviruses, such as HvAV-3a and TnAV-2a, yield functional caspases. Moreover, Diadromus pulchellus toursvirus (previously, Diadromus pulchellus ascovirus 4a; DpAV-4a), which recently has been assigned to a new ascovirus genus, *Toursvirus* (https://talk.ictvonline.org/), bears no genes resembling a caspase gene, yet this virus assembles viral vesicles similar to those formed in other ascoviruses (12). In addition, the caspase-like protein synthesized by HvAV-3e does not act as a caspase; instead, RNA interference (RNAi) gene silencing illustrates that this protein is essential for replication (40). Thus, these caspase-like genes are not orthologs of the SfAV-1a caspase, but their

maintenance in HvAV-3a and TnAV-2a suggests that they play an important yet largely unknown role in ascovirus replication and changes in cell architecture.

The SfAV-1a transcription of early genes continued at later time points, which is similar to early gene expression in both frog iridovirus 3 (FV3) and Chilo iridescent virus (CIV) (41, 42), implying that complete degradation of early messages is not essential for late expression, as reported for poxviruses (43). Moreover, transcription of viral structural genes usually occurs in late stages of infection, when virion assembly takes place. However, ORF091 shares homology with yabby-like transcription factor, which corresponds to an early gene, based on my analysis. Therefore, it is possibly associated with activation of a set of the virus genes during late stages of infection. The early synthesis of transcription factors that function as components of virion particles has been reported for herpes simplex virus, where the vp16 transcription factor detaches from the virion and is directed to the nucleus to initiate infection. Adenoviruses are also known to use the same mechanism to initiate a cascade of regulated viral gene transcription, a common character of DNA virus transcription (33).

Focusing on the TMD gene expression highlights two points. First, the late and very late expression of the four lipid-modifying enzymes (ORF013, ORF087, ORF093, and ORF112) when vesicles start to accumulate implies that these proteins are associated with vesicle membrane lipid modification or stabilization after apoptosis induction by the caspase, rather than being prerequisite building blocks for construction of the vesicle membrane. Second, the few TMD genes with unknown functions (ORF012, ORF032, ORF045, ORF062, and ORF122) detected in the very late stage represent good

candidates for *in vivo* protein localization during viral infection to check their role in vesicle membrane formation and integrity.

Comparison of the RNA-Seq profiles with data obtained from RT-qPCR for three viral genes reveals good correlation between the two methods and thus supports RPKM calculations in my study. Moreover, this reveals that the *Spodoptera frugiperda* L8 housekeeping gene (19) can be used as a reference gene in ascovirus transcription studies, since it shows great stability after the infection. Overall, the RT-qPCR is both a sensitive and reliable method for technical validation of RNA-Seq data (44).

Detection of nine putative bicistronic and six tricistronic mRNA transcripts in the SfAV-1a transcriptome ranging in size from 800 to 5,567 bp was surprising given that the majority of large DNA viruses are not known to form polycistronic messages; monocistronic messages are the dominant species. Therefore, I selected five messages for further validation using RT-PCR to rule out any possible sequence artifact after genome editing based on the RNA-Seq reads. Both the message sizes and the Sanger sequences of the RT-PCR products are in complete agreement with the RNA-Seq data. Searching for this phenomenon in other DNA viruses, the presence of longer RNA transcripts that equal two or three times the predicted size of the ORFs was found in Chilo iridescent virus, the virus most closely related to ascoviruses (45). Moreover, the same phenomenon was reported for other large DNA viruses, such as herpes simplex virus 1 (HSV-1) and vaccinia virus (46, 47). Later, HSV was found to use the noncanonical internal ribosome entry site (IRES) strategy in translation (48). Although the recruitment of the ribosome by IRES secondary structures is known for many RNA viruses, it has only rarely been

reported for DNA viruses, including herpes simplex virus 1, Kaposi's sarcoma-associated herpesvirus (KSHV), Epstein-Barr virus (EBV), and white spot syndrome virus (WSSV) (48–51). Therefore, given that ascoviruses destroy the nucleus prior to viral vesicle formation, and that eukaryotic cells favor the switch to an IRES mechanism during apoptosis (52, 53), the ability of ascoviruses to use the IRES mechanism may be advantageous to ensure that replication and translation proceed efficiently. Alternatively, because SfAV-1a expresses a FLAP-like endonuclease protein (ORF066) which shares sequence homology with the virion host shutoff proteins found in herpesviruses (54, 55), it is possible that the use of noncanonical translation mechanisms may be advantageous to the virus through suppressing the host cell canonical cap-dependent translation while keeping a high efficiency for viral translation, especially for important genes associated with apoptosis inhibition and immune evasion (**Table 2.2**). However, due to the limitation of a tightly synchronized infection in the *in vivo* system, determination of host RNA decline is difficult to rule out. Therefore, evaluation of this hypothesis may be possible in the future by focusing on a specific infected host tissue, such as the fat body in case of SfAV-1a, to detect this decline.

Overall, although I identified *in silico* several putative IRES sequences in the intergenic regions (**Table 2.2**) for the bicistronic and tricistronic messages using the IRESPred server (21), experimental validation is needed to determine whether these are functional IRES sequences and to rule out the absence of any cryptic promoter activity or ribosome reinitiation.

A future experiment that can be used for validation is to clone these sequences in a bicistronic reporter system in which the proteins are easily detected, for example, firefly and *Renilla* luciferases (56).

Interestingly, two of the tricistronic messages encoded proteins recently identified to function in host immune system suppression, namely, RNase III (ORF022) and a Diedel homolog (ORF121). RNase III represents a conserved protein among 4 different ascoviruses species, SfAV, TnAV, HvAV, and DpAV. RNase III enzymes are known to function in processing different RNAs and gene silencing. In a recent study of HvAV-3e RNase III, differential gene expression in two different tissue cultures demonstrated the important role of this gene in both virus replication and RNAi silencing specifically by targeting small interfering RNA (siRNA) molecule degradation. The RNase III transcript was detected at 16 hpi and peaked at 48 h, followed by a significant decrease in the gene expression level (57). The late expression of the RNase III gene is in agreement with what I found, as the first detection of ORF022 was at 12 hpi. However, the expression increased gradually until day 7 pi. Suppressors of host RNAi have been reported for other large DNA viruses, for example, iridoviruses and baculoviruses (58, 59). The evolutionary conservation of an RNAi suppressor in all ascovirus species implies exposure of ascoviruses to the host RNAi machinery, and probably the dsRNA, the pathway inducer signal, originates from the bidirectional transcription, the main source of dsRNA in the case of DNA viruses (60–63). This is in agreement with what I found with SfAV-1a transcripts, which overlap in many positions along the genome, potentially generating dsRNA substrates.

In addition, high expression (>1,000 RPKMs) of the Diedel protein in the SfAV-1a transcriptome may add a third host innate immune system suppression strategy explored by ascoviruses besides apoptosis and RNAi suppression. Diedel is a 12-kDa protein recently identified to function as a cytokine that downregulates the immune deficiency (IMD) pathway in flies and prolongs the life span of the fly under viral infection (25, 64). Homologs of *die* occur in two other families of dsDNA insect viruses, namely, baculoviruses and entomopoxviruses, and in the venom of two wasp species (*Leptopilina heterotoma* and *Leptopilina boulardi*). According to Lamiable et al., *die* is the first insect "virokine" gene, which may well function in host immune system suppression through downregulation of the IMD pathway. The protein is found to be induced by viral and not bacterial infection in *Drosophila*, in which it attains significantly high levels (460-fold) at 24 hpi, especially during infection by enveloped viruses (Sindbis virus [SINV] and vesicular stomatitis virus [VSV]), with gradual decline at later time points post-infection, which is in agreement with my data (64) (**Fig. 2.6**). Moreover, the *die* mutation was found to reduce the life span of the fly, especially under SINV infection, in comparison to controls. This reduction correlated with high expression of the immune deficiency pathway-regulated genes when the *die* gene was missing. The SfAV-1a Diedel homolog protein (ORF121) expression rescued the reduced life span phenotype of *die* mutant flies, at least partially. Therefore, it is expected that ascovirus Die homologs conserved in SfAV-1a and HvAV were acquired during evolution, from either their lepidopteran or wasp host, and used during lepidopteran host infection to either modulate the IMD pathway by functioning as a virokine or prolong the host life span by

functioning as a negative regulator of the IMD pathway. Prolonging the host's life span increases the virus replication period and the chance of virions or vesicles being picked up by endoparasitic wasp ovipositors in nature rather than leading to a dead end in the lepidopteran host, since the infection is not known to vertically transfer.

In conclusion, further transcriptional studies of ascoviruses will improve our knowledge about the molecular basis of their unique cell biology. RNA-Seq *in vivo* is a useful experimental approach to identify new candidate genes associated with pathogenesis that probably would be either undetected or overlooked in an *in vitro* system.

2.6. References

1. Govindarajan R, Federici BA. 1990. Ascovirus infectivity and effects of infection on the growth and development of noctuid larvae. J Invertebr Pathol 56:291–299.

2. Bigot Y, Renault S, Nicolas J, Moundras C, Demattei MV, Samain S, Bideshi DK, Federici BA. 2009. Symbiotic virus at the evolutionary inter-section of three types of large DNA viruses; iridoviruses, ascoviruses, and ichnoviruses. PLoS One 4:e6397.

3. Asgari S, Bideshi DK, Bigot Y, Federici BA, Cheng X. 2017. ICTV Report Consortium. ICTV virus taxonomy profiles: Ascoviridae. J Gen Virol 98: 4 –5.

4. Hu J, Wang X, Zhang Y, Zheng Y, Zhou S, Huang GH. 2016. Characterization and growing development of *Spodoptera exigua* (Lepidoptera: Noctuidae) larvae infected by *Heliothis virescens* ascovirus 3h (HvAV-3h). J Econ Entomol 109:2020–2026.

5. Li SJ, Hopkins RJ, Zhao YP, Zhang YX, Hu J, Chen XY, Xu Z, Huang GH. 2016. Imperfection works: survival, transmission and persistence in the system of Heliothis virescens ascovirus 3h (HvAV-3h), *Microplitis similis* and *Spodoptera exigua*. Sci Rep 6:21296.

6. Federici BA. 1983. Enveloped double stranded DNA insect virus with novel structure and cytopathology. Proc Natl Acad Sci U S A 80: 7664–7668.

7. Bideshi DK, Tan Y, Bigot Y, Federici BA. 2005. A viral caspase contributes to modified apoptosis for virus transmission. Genes Dev 19:1416–1421.

8. Federici BA, Bideshi DK, Tan Y, Spears T, Bigot Y. 2009. Ascoviruses: superb manipulators of apoptosis for viral replication and transmission. Curr Top Microbiol Immunol 328:171–196.

9. Bideshi DK, Bigot Y, Federici BA, Spears T. 2010. Ascoviruses, p 3–34. *In* Asgari S, Johnson KN (ed), Insect virology. Caister Academic Press, Norfolk, United Kingdom.

10. Asgari S. 2006. Replication of *Heliothis virescens* ascovirus in insect cell lines. Arch Virol 151:1689 –1699.

11. Bideshi DK, Dematti MV, Rouleux-Bonnin F, Stasiak K, Tan Y, Bigot S, Bigot Y, Federici BA. 2006. Genomic sequence of *Spodoptera frugiperda ascovirus 1a*, an enveloped, double-stranded DNA insect virus that manipulates apoptosis for viral reproduction. J Virol 80:11791–11805.

12. Bigot Y. 2011. Genus *Ascovirus*, p 73–78. *In* Tidona C, Darai G (ed), Springer index of viruses, 2nd ed. Springer, Heidelberg, Germany.

13. Krogh A, Larsson B, von Heijne G, Sonnhammer ELL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305:567–580.

14. Sonnhammer ELL, von Heijne G, Krogh A. 1998. A hidden Markov model for predicting transmembrane helices in protein sequences, p 175–182. *In* Glasgow J, Littlejohn T, Major F, Lathrop R, Sankoff D, Sensen C (ed), Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology. AAAI Press, Menlo Park, CA.

15. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for illumine sequence data. Bioinformatics 30:2114 –2120.

16. Kopylova E, Noe L, Touzet H. 2012. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics 28:3211–3217.

17. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie2. Nat Methods 9:357–359.

18. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. 2009. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res 37:W202–W208.

19. Legeai F, Gimenez S, Duvic B, Escoubas JM, Grenet ASG, Blanc F, Cousserans F, Seninet I, Bretaudeau A, Mutuel D, Girard PA, Monsempes C, Magdelenat G, Hilliou F, Feyereisen R, Ogliastro M, Volkoff AN, Jacquin-Joly E, d'Alençon E, Nègre N, Fournier P. 2014. Establishment and analysis of a reference transcriptome for *Spodoptera frugiperda*. BMC Genomics 15:704.

20. Pfaffl MW. 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 29:e45.

21. Kolekar P, Pataskar A, Kulkarni-Kale U, Pal J, Kulkarni A. 2016. IRESPred: web server for prediction of cellular and viral internal ribosome entry site (IRES). Sci Rep 6:27436.

22. Wang Z, Gerstein M, Snyder M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10:57– 63.

23. Tran VDT, De Coi N, Feuermann M, Schmid-Siegert E, Ba˘gut¸E-T, Mignon B, Waridel P, Peter C, Pradervand S, Pagni M, Monod M. 2016. RNA sequencing based genome reannotation of the dermatophyte Arthroderma benhamiae and characterization of its secretome and whole gene expression profile during infection. mSystems 1:e00036-16.

24. Tan Y, Bideshi DK, Johnson JJ, Bigot Y, Federici BA. 2009. Proteomic analysis of the *Spodoptera frugiperda* ascovirus 1a virion reveals 21 proteins. J Gen Virol 90:359 –365.

25. Coste F, Kemp C, Bobezeau V, Hetru C, Kellenberger C, Imler J-L, Roussel A. 2012. Crystal structure of Diedel, a marker of the immune response of *Drosophila melanogaster*. PLoS One 7:e33416.

26. Salem TZ, Turney CM, Wang L, Xue J, Wan XF, Cheng XW. 2008. Transcriptional analysis of a major capsid protein gene from Spodoptera exigua ascovirus 5a. Arch Virol 153:149–162.

27. Oliveira GP, Andrade AC, Rodrigues RA, Arantes TS, Boratto PV, Silva LK, Dornas FP, Trindade GS, Drumond BP, La Scola B, Kroon EG, Abrahao JS. 2017. Promoter motifs in NCLDVs: an evolutionary perspective. Viruses 9:16.

28. Stasiak K, Renault S, Demattei M, Bigot Y, Federici BA. 2003. Evidence for the evolution of ascoviruses from iridoviruses. J Gen Virol 84:2999 –3009.

29. Amorim-Vaz S, Tran VDT, Pradervand S, Pagni M, Coste AT, Sanglard D. 2015. RNA enrichment method for quantitative transcriptional analysis of pathogens *in vivo* applied to the fungus *Candida albicans*. mBio 6:e00942-15.

30. Driver AM, Penagaricano F, Huang Ahmad WKR, Hackbart KS, Wiltbank MC, Khatib H. 2012. RNA-Seq analysis uncovers transcriptomic variations between morphologically similar *in vivo* and *in vitro*-derived bovine blastocysts. BMC Genomics 13:118.

31. LoVerso PR, Wachter CM, Cui F. 2015. Cross-species transcriptomic comparison of *in vitro* and *in vivo* mammalian neural cells. Bioinform Biol Insights 9:153–164.

32. Chen YR, Zhong S, Fei Z, Hashimoto Y, Xiang JZ, Zhang S, Blissard GW. 2013. The transcriptome of the *baculovirus Autographa californica* multiple nucleopolyhedrovirus in *Trichoplusia ni* cells. J Virol 87:6391–6405.

33. Rohrmann GF. 2013. Baculovirus molecular biology, 3rd ed, p 1–341. National Center for Biotechnology Information, Bethesda, MD.

34. Parravicini C, Chandran B, Corbellino M, Berti E, Paulli M, Moore PS, Chang Y. 2000. Differential viral protein expression in Kaposi's sarcoma-associated herpesvirus-infected diseases: Kaposi's sarcoma, primary effusion lymphoma and multicentric Castleman's disease. Am J Pathol 156:743–749.

35. Ueda S, Nakamura H, Masutani H, Sasada T, Yonehara S, Takabayashi A, Yamaoka Y, Yodoi J. 1998. Redox regulation of caspase-3(-like) protease activity: regulatory roles of thioredoxin and cytochrome c. J Immunol 161:6689 – 6695.

36. Ueda S, Masutani H, Nakamura H, Tanaka T, Ueno M, Yodoi J. 2002. Redox control of cell death. Antioxid Redox Signal 4:405–414.

37. Masutani H, Ueda S, Yodoi J. 2005. The thioredoxin system in retroviral infection and apoptosis. Cell Death Differ 12:991–998.

38. Haendeler J, Hoffmann J, Tischler V, Berk BC, Zeiher AM, Dimmeler S. 2002. Redox regulatory and anti-apoptotic functions of thioredoxin depend on S-nitrosylation at cysteine 69. Nat Cell Biol 4:743–749.

39. Natarajan SK, Becker DF. 2012. Role of apoptosis-inducing factor, proline dehydrogenase, and NADPH oxidase in apoptosis and oxidative stress. Cell Health Cytoskelet 4:11–27.

40. Asgari S. 2007. A caspase-like gene from the *Heliothis virescens* ascovirus (HvAV3e) is not involved in apoptosis but is essential for virus replication. Virus Res 128:99 –105.

41. Chinchar VG, Yu W. 1992. Metabolism of host and viral mRNAs in frog virus 3-infected cells. Virology 186:435– 443.

42. D'Costa SM, Yao H, Bilimoria SL. 2001. Transcription and temporal cascade in *Chilo* iridescent virus infected cells. Arch Virol 146:2165–2178.

43. Cooper JA, Moss B. 1979. *In vitro* translation of immediate-early, early, and late classes of RNA from vaccinia virus infected cells. Virology 96:368 –380.

44. Yang M, Lin X, Rowe A, Rognes T, Eide L, Bjøra˚s M. 2015. Transcriptome analysis of human OXR1 depleted cells reveals its role in regulating the p53 signaling pathway. Sci Rep 5:17409.

45. D'Costa SM, Yao HJ, Bilimoria SL. 2004. Transcriptional mapping in Chilo iridescent virus infections. Arch Virol 149:723–742.

46. Anderson K, Frink R, Devi G, Gaylord B, Costa R, Wagner E. 1981. Detailed characterization of the mRNA mapping in the HindIII fragment K region of the HSV-1genome. J Virol 37:1011–1027.

47. Golini F, Kates JR. 1984. Transcriptional and translational analysis of a strongly expressed early region of the vaccinia virus genome. J Virol 49:459 – 470.

48. Griffiths A, Coen DM. 2005. An unusual internal ribosome entry site in the herpes simplex virus thymidine kinase gene. Proc Natl Acad Sci U S A 102:9667–9672.

49. Low W, Harries M, Ye H, Du MQ, Boshoff C, Collins M. 2001. Internal ribosome entry site regulates translation of Kaposi's sarcoma-associated herpesvirus FLICE inhibitory protein. J Virol 75:2938–2945.

50. Isaksson A, Berggren M, Ricksten A. 2003. Epstein-Barr virus U leader exon contains an internal ribosome entry site. Oncogene 22:572–581.

51. Kang ST, Wang HC, Yang YT, Kou GH, Lo CF. 2013. The DNA virus white spot syndrome virus uses an internal ribosome entry site for translation of the highly expressed nonstructural protein ICP35. J Virol 87: 13263–13278.

52. Spriggs KA, Bushell M, Mitchell SA, Willis AE. 2005. Internal ribosome entry segment-mediated translation during apoptosis: the role of IRES-transacting factors. Cell Death Differ 12:585–591.

53. Spriggs KA, Stoneley M, Bushell M, Willis AE. 2008. Re-programming of translation following cell stress allows IRES-mediated translation to predominate. Biol Cell 100:27–38.

54. Lin HW, Chang YY, Wong ML, Lin JW, Chang TJ. 2004. Function analysis of virion host shutoff protein of pseudorabies virus. Virology 324: 412–418.

55. Smiley JR. 2004. Herpes simplex virus virion host shutoff protein: immune evasion mediated by viral RNase? J Virol 78:1063–1068.

56. Thompson SR. 2012. So you want to know if your message has an IRES? Wiley Interdiscip Rev RNA 3:697–705.

57. Hussain M, Abraham AM, Asgari S. 2010. An ascovirus-encoded RNase III autoregulates its expression and suppresses RNA interference-mediated gene silencing. J Virol 84:3624–3630.

58. Bronkhorst AW, van Rij RP. 2014. The long and short of antiviral defense: small RNA-based immunity in insects. Curr Opin Virol 7:19–28.

59. Mehrabadi M, Hussain M, Matindoost L, Asgari S. 2015. The baculovirus antiapoptotic p35 protein functions as an inhibitor of the host RNA interference antiviral response. J Virol 89:8182–8192.

60. Weber F, Wagner V, Rasmussen SB, Hartmann R, Paludan SR. 2006. Double-stranded RNA is produced by positive-strand RNA viruses and DNA viruses but not in detectable amounts by negative-strand RNA viruses. J Virol 80:5059 –5064.

61. Bronkhorst AW, van Cleef KW, Vodovar N, Ince IA, Blanc H, Vlak JM, Saleh MC, van Rij RP. 2012. The DNA virus Invertebrate iridescent virus 6 is a target of the *Drosophila* RNAi machinery. Proc Natl Acad Sci U S A 109:E3604–E3613.

62. de Faria IJ, Olmo RP, Silva EG, Marques JT. 2013. dsRNA sensing during viral infection: lessons from plants, worms, insects, and mammals. J Interferon Cytokine Res 33:239 –253.

63. Kemp C, Mueller S, Goto A, Barbier V, Paro S, Bonnay F, Dostert C, Troxler L, Hetru C, Meignin C, Pfeffer SJ, Hoffmann A, Imler JL. 2013. Broad RNA interference-mediated antiviral immunity and virus-specific inducible responses in *Drosophila*. J Immunol 190:650 – 658.

64. Lamiable O, Kellenberger C, Kemp C, Troxler L, Pelte N, Boutros M, Marques JT, Daeffler L, Hoffmann JA, Roussel A, Imler JL. 2016. Cytokine Diedel and a viral homologue suppress the IMD pathway in *Drosophila*. Proc Natl Acad Sci U S A 113:698–703.

2.7 Tables:

Table 2.1: Comparison of the SfAV-1a original and new positions for core and transmembrane-domain (TMD) containing ORFs identified in the SfAV-1a transcriptome. Conservation in other ascovirus species is illustrated for core genes. The number of TMDs is determined using the TMHMM server v.2.0 (13, 14).

ORF No.	Conservation in ascovirus species	Putative function (BLASTP)	ORF new position	ORF old position (11)	Transcript position*
		Nucleotide metabolism proteins			
ORF001	HvAV, TnAV, DpAV	DNA polymerase	1>3315 (same position)	1>3315	1>3381
ORF008	HvAV, TnAV, DpAV	DNA-directed RNA polymerase largest subunit, N terminal domain	9261<11918	9257<11914	9194<12018
ORF052	HvAV, TnAV, DpAV	DNA-directed RNA polymerase subunit B	56528>60001	56523>59996	56430>61975
ORF067	HvAV, TnAV, DpAV	DNA-directed RNA polymerase largest subunit, C terminal domain	74349>75737	74345>75733	74262>76701
ORF022	HvAV, TnAV, DpAV	RNaseIII	27951>30107	27937>29190	27865>30906
ORF037	TnAV, DpAV	GIY_YIG-like endonuclease	43972<44310	43964<44302	43881<44317
ORF066	HvAV, TnAV	FLAP-like endonuclease	72899<73900	72826<73896	72856<73998
ORF059	HvAV, TnAV, DpAV	DNA repair exonuclease	64961<66058	64954<66051	63013<66156
ORF103	HvAV, TnAV, DpAV	ATPase involved in DNA repair	131096<133753	131663<133735	131002<133850
ORF110	HvAV, TnAV, DpAV	ATPase	138671<139918	138652<139824	137292<139970
ORF099	HvAV, TnAV, DpAV	Primase	127723<130203	127708<130188	127639<130290
ORF040	HvAV, TnAV, DpAV	Thymidine kinase	45529<46161	45521<46153	45430<46260
ORF113	HvAV, TnAV, DpAV	VLTF2-like late transcription factor/Zn-finger DNA binding protein	143502<144080	143483<144061	143414<144183
ORF095	HvAV, TnAV	Helicase-2	124740<126242	125523<126227	124703<126345
ORF086	HvAV, DpAV	Uvr/REP helicase	111581<112243	111570<112232	111527<112324
ORF090	HvAV, TnAV, DpAV	Putative lipopolysaccharide-modifying enzyme/Tyrosine protein kinase	118763>121543	118746>121526	118664>121642
ORF104	HvAV, TnAV, DpAV	Serine-threonine kinase	133910>135844	133892>135223	133821>135894
ORF029	HvAV, TnAV, DpAV	Poxvirus Late Transcription Factor VLTF3 like	34954>36141	34939>36126	34864>36241
		Lipid-metabolism proteins			
ORF013	HvAV, DpAV	Esterase/lipase	17498<19462	17679<19457	17448<19555
ORF087	HvAV, TnAV, DpAV	Fatty acid elongase	113621<114469	113610<114359	113522<114520
ORF093	HvAV, TnAV, DpAV	Patatin-like phospholipase	122925<123854	122909<123838	122875<123922
ORF112	HvAV, TnAV, DpAV	Phosphate acyltransferase	142131<143198	142112<143179	142038<143298
		Apoptosis-associated proteins			
ORF114	HvAV, TnAV, DpAV	CathepsinB	144165>145391	144146>145507	144073>146526
ORF073	HvAV, TnAV	Caspase/interleukin-1 beta converting enzyme (ICE) homologues	94552>95418	94542>95408	94457>95519
ORF074	HvAV, TnAV	IAP-like (RING-finger, IAP, E3 ubiquitin ligase)	95482<96234	95472<96224	95384<96337
ORF016	HvAV, TnAV, DpAV	IAP-like (RING-finger-containing E3 ubiquitin ligase)	23374>23706	23363>23695	21461>23719
ORF025	HvAV, TnAV, DpAV	IAP-like (RING-finger- E3 ubiquitin ligase)	30115>30807	30100>30792	27865>30906
		Virion structural proteins			
ORF061	HvAV, TnAV, DpAV	Ervl/Alr family protein	66958>67602	66951>67595	66916>67695
ORF064	HvAV, TnAV, DpAV	Serine/threonine protein kinase	68862>71315	68856>71309	68778>71401
ORF084	HvAV, TnAV, DpAV	Dynein-Like β chain	107285<111376	107829<110975	107234<111427
ORF091	HvAV, TnAV, DpAV	High mobility group protein/Yabby-like protein	121630<121998	121613<121981	121482<122090
ORF109	HvAV, TnAV, DpAV	Haloacid dehalogenase-like hydrolases /CTD-like phosphatase	137953<138561	137934<138542	137292<139970
ORF041	HvAV, TnAV, DpAV	Major capsid protein	46200>47585	46192>47577	46098>47623
ORF048	HvAV, TnAV, DpAV	DNA condensation, P64	53489<55186	53484<55181	53392<55735
ORF075	HvAV, TnAV	S1/P1 nuclease	96314>97099	96304>97089	96217>98183
ORF009	HvAV, TnAV, DpAV	DEAD-like helicase of the SNF2 family	12213>15221	12281>15217	12161> undefined 3' end

		Host interaction proteins			
ORF121	HvAV	Diedel-like	153336>153734	153318>153725	152634>153802
ORF077	HvAV	Transposase	98516>100336	98506>100326	98419> undefined 3' end
ORF014	HvAV, TnAV, DpAV	Zinc-dependent metaloproteinase	19762<21168	19998<21158	19710<21242
ORF116	HvAV, TnAV	Thioredoxin-Like protein	145932>146423	145914>146405	144073>146526
ORF065	HvAV, TnAV, DpAV	Immediate early protein ICP-46	71417>72556	71571>72551	Undefined 5' end >72651
ORF055	HvAV, TnAV, DpAV	Hypothetical protein/254L IIV-6	60950>61888	60944>61882	56430>61975
ORF107	HvAV, TnAV	Trifunctional transcriptional regulator/ proline dehydrogenase/ L-glutamate gamma-	136065<136718	136277<136699	135992<136821
ORF079	HvAV, TnAV, DpAV	BRO-like protein	101506>102597	101496>102587	101410>102697
ORF072	HvAV, TnAV	BRO-like protein	93898<94434	93888<94424	93904<94536
ORF070	HvAV, TnAV	BRO-like protein	78074>79393	78564>79385	77989>79470
ORF No.	**Number of TMDs**	**Putative function (BLASTP)**	**ORF new Position**	**ORF old position**	**Transcript position**
		Transmembrane-domain containing proteins			
ORF002	1	Hypothetical protein	4222<4602	4222<4602	4121<4697
ORF043	1	Hypothetical protein	48309<48995	48301<48987	47571<49092
ORF009	1	DEAD-like helicase of the SNF2 family	12213>15221	12281>15217	12161> undefined 3' end
ORF035	3	lipid membrane protein	42606>44069	42596>43321	42504>44168
ORF054	1	Putative myristylated membrane protein	60054>60953	60049>60885	56430>61975
ORF060	2	Hypothetical protein	66301>66588	66345>66581	66250>66666
ORF013	1	Esterase/lipase	17498<19462	17679<19457	17448<19555
ORF087	7	Fatty acid elongase	113621<114469	113610<114359	113522<114520
ORF093	3	Patatin-like phospholipase	122925<123854	122909<123838	122875<123922
ORF112	2	Phosphate acyltransferase	142131<143198	142112<143179	142038<143298
ORF110	1	ATPase	138671<139918	138652<139824	137292<139970
ORF086	2	Uvr/REP helicase	111581<112243	111570<112232	111527<112324
ORF014	1	Zinc-dependent metaloproteinase	19762<21168	19998<21158	19710<21242
ORF032	1	Hypothetical protein	37541<38494	37527<38480	37464<38589
ORF062	2	Hypothetical protein	68072<68284	68065<68277	67976<68349
ORF108	1	Hypothetical protein	137384<137902	137365<137883	137292<139970
ORF119	1	Hypothetical protein	152191<152730	152174<152713	152119<152833
ORF007	1	Hypothetical protein	7975>8697	7972>8634	6607>8748
ORF012	1	Hypothetical protein	16637>17389	16634>17386	16555>17472
ORF039	2	Hypothetical protein	44995>45570	44987>45562	44900>45670
ORF045	1	Hypothetical protein	50952>51833	50943>51734	50893>51872
ORF056	2	Hypothetical protein	61971>63197	61965>62996	61925>63249
ORF085	1	Hypothetical protein	111140>111529	111129>111518	111055>111631
ORF092	2	Hypothetical protein	122215>122871	122199>122855	122221>122913
ORF097	2	RING finger domain	127036>127518	127021>127503	126088>127611
ORF122	1	Hypothetical protein	154860>155696	154876>155682	154788>155793

* Undefined 3' or 5' ends (undetermined) is due to overlapping with the next or previous gene, (>) refers for forward orientation and (<) refers for reverse orientation.

Table 2.2: Primers used in this study.

Transcript position /Gene position	Primers (5' -> 3')	Annealing temperature (◦ C)	Application	Expected size & primer efficiency (E)
ORF032 (37541<38494)	For: atggatcggaacgcctattg Rev: ggccagtatcgatgtaccatg	60	RT-qPCR	142 bp (E: 91.6%)
ORF037 (43972<44310)	For: tggctgataagaacggttgg Rev: cgacctgaacaatgcatcttg	60	RT-qPCR	112 bp (E: 90.7%)
ORF016 (23374>23706)	For: aagtcctcggtcgtatttgtg Rev: ctatcctagccctacacaatgg	60	RT-qPCR	103 bp (E: 90.9%)
rep_c559 - ribosomal protein L8 (37)	For: gtgattcgtgctcagcgtaa Rev: accagggtcatggatgatgt	60	RT-qPCR	(E: 90.2%)
21461>23719	For_mesg 1: ggtgaaacgaacgaaacattc Rev_mesg 1: ctatcctagccctacacaatg	58	RT-PCR	2116 bp
103271>106288	For_mesg 3: ggtatagtggtgaggttgtac Rev_mesg 3: cgtagttctcgttgagattg	58	RT-PCR	2777 bp
126088>127611	For_mesg 4: ccacgatgactgatccattg Rev_mesg 4: gcaatcctctcttgtacacg	60	RT-PCR	1265 bp
144073>146526	For_mesg 5: ggatcgtcgattgtcatacc Rev_mesg 5: gtcgtcgttccaatcgtgaa	60	RT-PCR	2115 bp
152634>153802	For_mesg 6: cgaatcggtcaatatctgttc Rev_mesg 6: ttctccaatgcgtctccatc	60	RT-PCR	898 bp

Table 2.3: Description of the bicistronic and tricistronic mRNA messages identified in the SfAV-1a transcriptome infecting 3rd instar *Spodoptera frugiperda* larvae.

Transcript position	Size (bp)	Bicistronic /tricistronic	ORFs position	ORFs putative function	Intergenic region(s) (IR) position	Putative IRES in IR (>15 bp)	Poly-A signal position near the 3' message end
6607>8748	2142	bicistronic	ORF1: 6706>7962 ORF2: 7975>8697	ORF1: Hypothetical protein ORF2: Hypothetical protein	7963-7974	(NA)	No poly-A
21461>23719	2259*	bicistronic	ORF1:21553>23256 ORF2: 23374>23706	ORF1: Zinc-RING finger domain (E3 ubiquitin-protein ligase) ORF2: IAP-like (E3 ubiquitin-protein ligase)	23257-23373	Putative IRES	Poly-A (23695 bp)
23790>25159	1370	bicistronic	ORF1: 23990>24451 ORF2:24509>25060	ORF1: Hypothetical protein ORF2: Hypothetical protein	24452-24508	Putative IRES	No poly-A
27865>30906	3042	bicistronic	ORF1: 27951>30107 ORF2: 30115>30807	ORF1: RNaseIII ORF2: IAP-like (E3 ubiquitin-protein ligase)	30108-30114	(NA)	No poly-A
56430>61996	5567	tricistronic	ORF1: 56528>60001 ORF2: 60054>60953 ORF3: 60950>61888	ORF1: DNA-directed RNA polymerase II ORF2: Putative myristylated membrane protein ORF3: Hypothetical protein/254L IIV-6 (Domain of unknown function)	Intergenic 1: 60002-60053 Intergenic 2: Overlapping at 60950 bp	Putative IRES (NA)	Poly-A (61888 bp)
74262>76701	2440	bicistronic	ORF1: 74349>75737 ORF2: 75875>76603	ORF1: DNA-directed RNA polymerase largest subunit, C terminal domain ORF2: Hypothetical protein	75738-75874	Putative IRES	No poly-A
103271>106288	3018*	tricistronic	ORF1: 103341>103889 ORF2: 103886>105250 ORF3: 105259>106245	ORF1: RING-finger-containing E3 ubiquitin ligase ORF2: Serine/Threonine protein kinases/predicted eukaryotic translation initiation factor 2-alpha kinase 1-like ORF3: Hypothetical protein	Intergenic 1: overlapping at 103886 bp Intergenic 2: 105251-105258	(NA) (NA)	Poly-A (106262 bp)
126088>127611	1524*	bicistronic	ORF1: 126182>126970 ORF2: 127036>127518	ORF1: Hypothetical protein ORF2: RING-finger-containing E3 ubiquitin ligase (Predicted breast cancer susceptibility type 1 protein)	126971-127035	No IRES	Poly-A (127526 bp)
131672>132471	800	bicistronic	ORF1: 131763>132101 ORF2: 132144>132368	ORF1: No homology ORF2: No homology	132102-132143	No IRES	No poly-A
144073>146526	2454*	tricistronic	ORF1: 144165>145391 ORF2: 145467>145925 ORF3: 145932>146423	ORF1: Cathepsin B ORF2: No homology ORF3: Thiol reductase thioredoxin	Intergenic 1: 145392-145466 Intergenic 2: 145926-145931	Putative IRES (NA)	No poly-A
152634>153802	1169*	tricistronic	ORF1: 152729>152917 ORF2: 152930>153187 ORF3: 153336>153734	ORF1: Hypothetical protein ORF2: autotransporter domain-containing protein like ORF3: Diedel-like protein	Intergenic 1: 152918-152929 Intergenic 2: 153188-153335	(NA) Putative IRES	Poly-A (153750 bp)
47571<49092	1522	bicistronic	ORF1: 48309<48995 ORF2: 47659<48297	ORF1: Hypothetical protein ORF2: Hypothetical protein	48298-48308	NA	No poly-A
53392<55735	2344	bicistronic	ORF1: 55219<55731 ORF2: 53489<55186	ORF1: chromosome segregation protein SMC ORF2: 2-cysteine adaptor domain, P64 protein	55187-55218	No IRES	No poly-A
63013<66156	3144	tricistronic	ORF1: 64961<66058 ORF2: 63954<64928 ORF3: 63113<63886	ORF1: DNA repair exonuclease SbcCD nuclease subunit ORF2: Hypothetical protein ORF3: Hypothetical protein	Intergenic 1: 64929-64960 Intergenic 2: 63887-63953	Putative IRES No IRES	No poly-A
137292<139970	2679	tricistronic	ORF1: 138671<139918 ORF2: 137953<138561 ORF3: 137384<137902	ORF1: ATPase ORF2: Putative haloacid dehalogenase-like ORF3: Hypothetical protein	Intergenic 1: 138562-138670 Intergenic 2: 137903-137952	Putative IRES No IRES	No poly-A

*Message size and sequence were further validated by RT-PCR and Sanger sequencing of the PCR product. (>) refers for forward orientation and (<) refers for reverse orientation. Intergenic Region (IR). Internal Ribosome Entry Site (IRES). Intergenic regions less than 15 bp in size are referred to as (NA).

Chapter III: Mitochondrial and Innate Immunity Transcriptomes from *Spodoptera frugiperda* Larvae Infected with the Spodoptera frugiperda ascovirus

3.1 Abstract

Ascoviruses are large DNA viruses transmitted by parasitic wasps and typically infect lepidopteran larvae. They cause a chronic but ultimately fatal disease during which remarkable changes in cellular architecture partition the cell into numerous vesicles in which most virus replication occurs. These vesicles accumulate in larval hemolymph turning it milky white for weeks before the larvae die. Previous studies show these vesicles arise from changes in cell morphology resembling apoptosis, yet differ in advanced stages after nuclear lysis in that mitochondria are not degraded but instead are modified by the virus, changing in size, shape, and motility as they appear to participate in synthesizing membranes that delimit the viral vesicles. Moreover, ascovirus infection does not provoke an obvious innate immune response. Thus, in the present study I used *in vivo* RNA-sequencing to determine whether infection by the Spodoptera frugiperda ascovirus (SfAV-1a) modified expression of the host mitochondrial genome, as well as cytoskeleton genes known to control organelle movement. Additionally, I examined changes in expression of innate immunity genes. Here I show that while expression of many mitochondrial genes was similar to uninfected controls, others were upregulated, especially ATP synthase subunits 6 and 8 during vesicle formation, indicating the importance of conserving these organelles for virus replication. Of 106 genes that code for cytoskeleton proteins potentially involved in organelle movement, most showed a

minor change (less than 2-fold), while 32 were downregulated. Only three gene orthologs of kinesin and another for dynein, major motor proteins involved in mitochondrial movement, were upregulated (more than 2-fold). Regarding innate immunity, moderate increases occurred in expression of many Toll, melanization and phagocytosis genes and their negative regulators. Genes for the antimicrobial peptides moricin and gloverin, however, were upregulated more than 32-fold, and for lebocin-1, lebocin-2 and Hdd23-like proteins, which are also antimicrobial, were upregulated more than 15-fold, as was a negative regulator prophenoloxidase gene expression. SfAV-1a destroys most fat body cells, so expression of these innate immunity genes apparently occurs in cells that remain viable in this tissue at seven days post-infection, and possibly in tissues such as the epidermis and tracheal matrix.

3.2 Introduction

Ascoviruses are viruses of insects characterized by large enveloped virions with ds DNA genomes ranging from 150 – 180 kbp, and primarily infect larvae of the lepidopteran family Noctuidae. These viruses have several unique characteristics with respect to their biology and pathology compared to other insect viruses. For example, transmission through feeding is possible but much more typically they are transmitted mechanically on the ovipositor of parasitic wasps when probing their hosts for oviposition. Infection results in a chronic disease that significantly retards development, the larval stage often lasting four to six weeks prior to death, compared to two weeks for normal growth resulting in pupation. Additionally, a few days after infection these viruses produce

millions of spherical virion-containing vesicles, 2-15 μm in diameter, that accumulate in the hemolymph (blood) turning it from translucent green to milky white (1-3).

While these characteristics easily differentiate ascoviruses from all other insect viruses, their most unique feature is the remarkable combination of changes in cell architecture that result in the formation of virion-containing vesicles. Referred to hereafter as viral vesicles, the early stages of their morphogenesis resemble apoptosis, with invagination of the nuclear membrane followed by lysis of the nucleus (1-6). In fact, the ascovirus type species, Spodoptera frugiperda ascovirus-1a (SfAV-1a), encodes an executioner caspase synthesized *in vitro* nine hours post-infection (4) and observed *in vivo* 24 hours post-infection (7). However, rather than resulting in death, the cell undergoes significant hypertrophy, typically enlarging *in vivo* more than ten times the size of healthy cells. Simultaneously, the mitochondria change in shape, divide, and are moved, aligning along cleavage planes throughout the enlarged cell where new double layers of lipid membranes are formed. The latter become large fragments of the delimiting membrane of viral vesicles. The formation of viral vesicles is completed when these large vesicle membrane fragments formed along cell cleavage planes fuse with membrane derived by invagination of the plasmalemma, a process in which the mitochondria also appear to participate (1, 3).

The *Spodoptera frugiperda* mitochondrial genome consists of 15365 bp (NCBI KM362176.1; described in 8) and has the typical gene order for mitochondrial genomes of noctuid larvae (8, 9), consisting of genes coding for 13 proteins, 22 tRNAs and 2 rRNAs. The accumulation of dense aggregations of mitochondria during ascovirus

infection at sites of internal membrane synthesis and plasmalemma invagination, along with several lipids metabolizing enzymes encoded by all ascoviruses (10), provides evidence these host cell organelles are reorganized to provide energy for virus replication and vesicle formation. Thus, one objective of the present study was to investigate this possibility by examining the host mitochondrial genome expression along with that of cytoskeletal genes possibly involved in moving mitochondria to sites of membrane synthesis in fat body tissue infected with SfAV-1a.

In addition, very little is known about host responses at the molecular level to the chronic disease caused by ascoviruses, especially given the circulation of foreign viral vesicles in the hemolymph for weeks, as was as the almost complete destruction of the larval fat body tissue during SfAV-1a infection (2). This raises the question of whether the host's innate immunity pathways such as Toll, Imd, melanization and phagocytosis respond to ascovirus infection. Thus, my second objective was to use the *in vivo* larval RNA-Sequencing system to determine whether these innate immune pathways responded in *Spodoptera frugiperda* 3rd-instars infected with SfAV-1a. Previous evidence for any innate host response to ascovirus infection is limited to conservation of an RNAseIII endonuclease in all ascovirus species associated with virus evasion of host RNAi machinery (11), and a recent transcriptomic study of HvAV-3h during infection in *Spodoptera exigua* larvae (12). The latter study showed activation of certain immunity gene pathways such as Toll-like receptor signaling and JAK/STAT.

Here we show that during the first week of SfAV-1a infection, genes of the *S. frugiperda* mitochondrial genome were expressed at levels comparable to the healthy

larvae or slightly higher during vesicle formation, especially ATP synthase subunits. Cytoskeletal gene expression varied from being stable to downregulated, with only a few genes for mitochondrial motor proteins, specifically, dynein and kinesin, upregulated. Regarding host innate immunity, SfAV-1a infection triggered upregulation of important antimicrobial genes, and as the infection advanced, and for several key humoral and cellular genes of Toll, melanization, and phagocytosis pathways. However, I also detected simultaneous increases of certain negative regulators of these pathways, suggesting this might moderate the host immune response prolonging the larval life span, thereby extending ascovirus transmission.

3.3 Materials and methods

3.3.1 Virus infection and disease development in infected larvae

The Spodoptera frugiperda ascovirus 1a (SfAV-1a), strain Sf82-126 (13), was used to inoculate *Spodoptera frugiperda* 3rd-instar larvae. Larvae were grown on artificial diet (Benzon Research, Carlisle, PA) under ambient laboratory conditions (22°C). To mimic infection by parasitic wasps, virus inoculation was performed manually using a minutin pin dipped in a suspension of viral vesicles ($1x10^8$/ml), which was then inserted briefly through the posterior abdominal cuticle into the fat body. Disease development was confirmed by examining a small drop of hemolymph to ensure it was opaque white, and then by viewing the drop by phase contrast microscopy for the presence of high concentrations of viral vesicles.

3.3.2 Extraction of total RNA from healthy and infected larvae

Total RNA was isolated from healthy and SfAV-infected *S. frugiperda* larvae using the TRIzol (Invitrogen, Life Technologies) following the protocol provided by the manufacturer. RNA was isolated from 50-100 mg of tissue by mechanical homogenization of the larval body in 1 ml TRIzol. RNA samples were isolated at 0 hours (uninfected control larvae), and for SfAV-infected larvae at 6, 12, 24, 48 h, and 4 and 7 days post-infection (pi). Total RNA was extracted from three biological replicates at each time point. Two replicates were included in the Transcripts Per Million (TPM) calculations (described below).

3.3.3 DNase-treatment and RNA purification

To purify and remove any DNA contamination in the isolated RNA, instructions for the RNA Clean and Concentrator TM-5 kit manufacturer (ZYMO RESEARCH) were followed combined with the RNase-Free DNase set (Qiagen). RNA quantity and quality was determined using a Thermo Scientific NanoDrop ND-2000c spectrophotometer.

3.3.4 Enrichment of mRNA, RNA-Seq library preparation, and sequencing

RNA-Seq library preparation and sequencing was previously described in detail (7). Briefly, mRNA was enriched using the NEBNext poly(A) mRNA magnetic isolation module kit (New England BioLabs), followed by library preparation using NEBNext Ultra Directional RNA library preparation kit (New England BioLabs) for Illumina sequencing.

For sequencing the HiSeq2500 Illumina sequencer was used (Noll Core Facility at the UC Riverside Institute for Integrative Genome Biology).

3.3.5 Bioinformatic analysis and TPM calculations

For quantification of *S. frugiperda* mitochondrial genes expression I used the published *S. frugiperda* mitochondrial genome (8) (available through NCBI, KM362176.1) for transcript locations. In this reference genome, two transcripts are annotated at multiple loci but with identical transcript names: tRNA-Leu and tRNA-Ser. For clarity, I refer to these transcripts as tRNA-Leu-1 (located on the reference mitochondrial genome at positions 2992-3059), tRNA-Leu-2 (positions 12714-12783), tRNA-Ser-1 (positions 6164-6233) and tRNA-Ser-2 (positions 11690-11756). For quantification of innate immunity and cytoskeleton genes, RNA-Seq reads were mapped to the recently published *Spodoptera frugiperda* corn variant genome (14). The assembled genome can be accessed through the European Bioinformatics Institute (EMBL-EBI) (accession number: PRJEB13110). The innate immunity and cytoskeleton gene ID's listed in this study can be accessed through the SfruDB Information system http://bipaa.genouest.org/is/lepidodb/spodoptera_frugiperda/. For innate immunity genes, I included representatives (97 genes) for every innate immunity pathway identified in *Spodoptera frugiperda* transcriptome (Sf_TR2012b) and listed in Legeai et al., 2014 (15). Thus, genes were included for Toll, Imd, JAK/STAT, JNK, phenoloxidase system (PO), extra cellular signal transduction and cytokines, antimicrobial peptides (AMPs), transmembrane receptors, Peptidoglycan Recognition Protein (PGRP) and the Gram-

negative bacteria-binding protein (GNBP). For cytoskeleton genes, I included 106 genes in the analysis. These included representatives for actin, actin-related proteins, tubulin, dynein, kinesin, lamin, filamin, myosin and profilin. Transcript expression estimates for these genes were estimated as Transcripts Per Million (TPM) using the RSEM program (16).

3.3.6 Determination of *Spodoptera frugiperda* genes upregulated and downregulated post-infection

After TPM quantification for each *S. frugiperda* mitochondrial, cytoskeletal or innate immunity gene, the genes were separated into two classes, upregulated or downregulated, based on the increase or decrease in expression compared to the 0 h control. **Upregulated genes** were defined as those for which transcription increased 2-fold or more post-infection, and had a basal transcription level equal to or greater than 1 RPKM at 0 h or any time post-infection. Similarly, **downregulated genes** were those for which transcription decreased by greater than 2-fold, and had a basal transcription level equal to or greater than 1 RPKM at 0 h or any time post-infection.

3.3.7 Electron Microscopy

The transmission EM imaging was done following the protocols described in (1, 17). Briefly, for tissue fixation 3% glutaraldehyde and 1% OsO_4 buffers were used. Followed by, tissue embedding in Epon-Araldite. Ultrathin section were cut on a Sorvall Ultrathin Microtome and examined in a Philips electron microscope.

3.3.8 Accession numbers for *Spodoptera frugiperda* mitochondrial and innate immunity gene RNA-Seq data

The *Spodoptera frugiperda* RNA-Seq data and TPM values for mitochondrial, cytoskeletal and innate immunity genes obtained in this study can be accessed through the NCBI Gene Expression Omnibus (GEO) series accession numbers (GSE114901).

3.4 Results

3.4.1 Host mitochondrial genome transcriptome pattern

Ascoviruses have significant and unusual effects on mitochondrial shape, size and distribution within infected cells regardless of the tissue, especially where mitochondria aggregate along cellular cleavage planes as viral vesicle delimiting membranes form (3). These observations imply mitochondria provide energy for this *de novo* membrane synthesis. The transcriptome data support these observations. For example, Both ATP synthase subunits (i.e. [ATP6 and ATP8] or [mitochondria complex V]) were upregulated, especially during the vesicle formation period from 6 to 48 hpi. The upregulation level ranged from 1.05 to 3.43-fold change (**Table 3.1**). At days 4 and 7 post-infection (pi) transcription for both subunits was downregulated slightly (decreases ranged from -1.02 to -1.58-fold). In addition, the increases were typically lower than two-fold (ranged from 1.10 to 2.22-fold) for transcripts of NADH-dehydrogenase subunits 1, 2, 3, 4L and 6 at all time points post-infection, and similarly for subunit 5 at all time

points except for day 4 and 7 pi, by which time the expression began to decrease slightly (**Table 1, Fig. 3.1**). NADH subunit 4 showed a slight upregulation (1.52-fold) at 24 hpi and then decreased at all time points by (-1.05 to -1.33-fold). Therefore, the expression of ATP synthase and NADH-dehydrogenase genes in infected larvae is best described as varying between stability and upregulation during viral vesicle formation. In contrast, mitochondrial 12S and 16S ribosomal RNA transcripts (rRNA) were upregulated late in the infection by greater than 2-fold at 48 hours and day 7 pi, respectively (**Table 3.1, Fig. 3.1**), implying active mitochondrial translation after vesicle formation.

With respect to cytochrome-c-oxidase subunits, COII and COIII transcript levels increased, but by less than 2-fold, at all times after infection compared to control levels. Alternatively, COI transcription decreased by -1.06 to -1.21-fold at almost all time points pi. Moreover, the *S. frugiperda* mitochondrial cytochrome b (complex III) transcripts decreased by -1.10-fold to -1.54-fold. Finally, most tRNA expression patterns were undetected or fluctuated at the time points tested (**Fig. 3.1**).

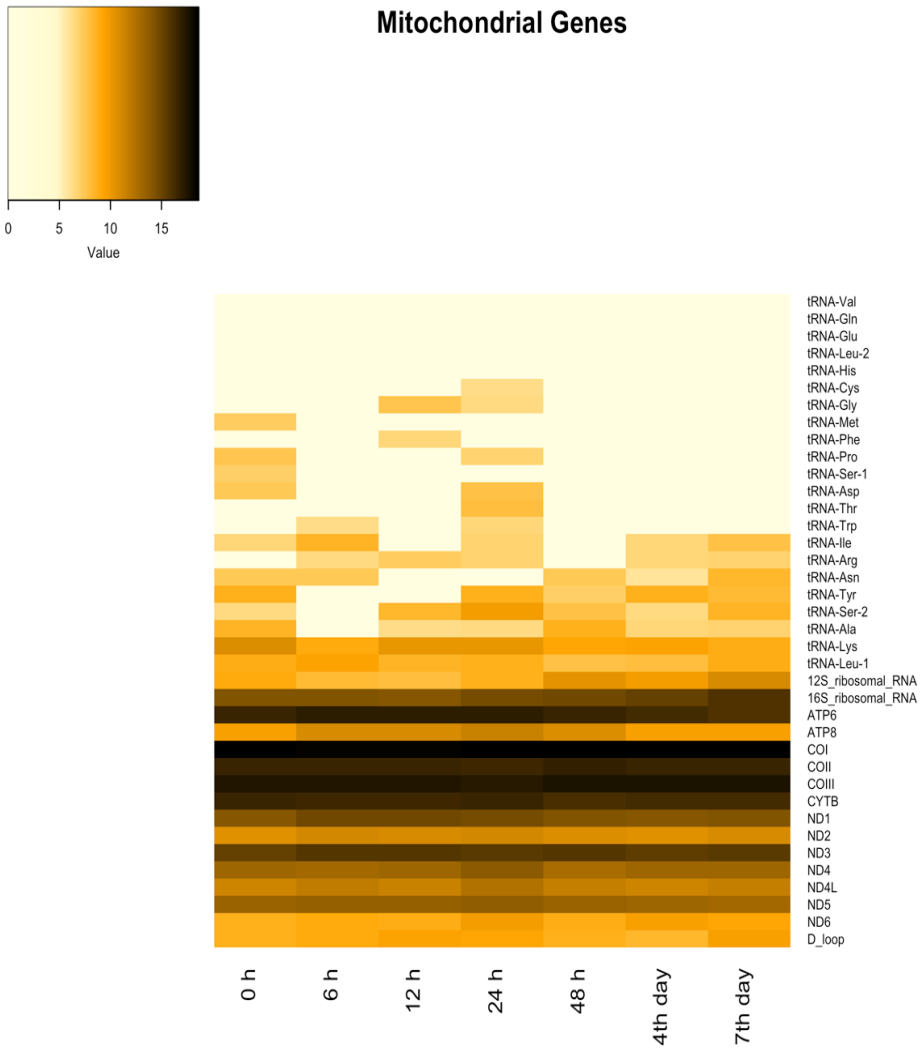


FIG 3.1 Heat map representation of mitochondrial genome expression in *Spodoptera frugiperda* 3rd-instar larvae infected with the Spodoptera frugiperda ascovirus 1a. Mitochondrial gene expression is for the 0h uninfected control compared with time points from 6h to 7 days post-infection. The heat map color scale represents the $\log^2$ TPM value of replicate expression level averages at each time. The *Spodoptera frugiperda* mitochondrial genome used for mapping RNA-Seq reads is available at NCBI KM362176.1 (8). ND1-6 refers to NADH dehydrogenase subunits and COI-III refers to cytochrome c oxidase subunits.

3.4.2 Cytoskeleton gene transcription pattern

Out of 106 cytoskeleton genes involved in movement, including genes coding for actin, actin-related proteins, tubulin, dynein, kinesin, lamin, filamin, myosin and profilin, 42

were upregulated or downregulated by 2-fold or more at certain time point post-infection compared to the control (**Fig 3.2**). Most of these, 32, were downregulated at all time points (**Table 3.2**). The only genes upregulated (by more than 2-fold change) were those coding for a kinesin light chain, two kinesin-like proteins (KIF19 and KIF23), a dynein light chain roadblock-type 2, a myosin-VIIa, and a tubulin gamma-2 chain-like (**Table 3.2; Fig 3.2**).

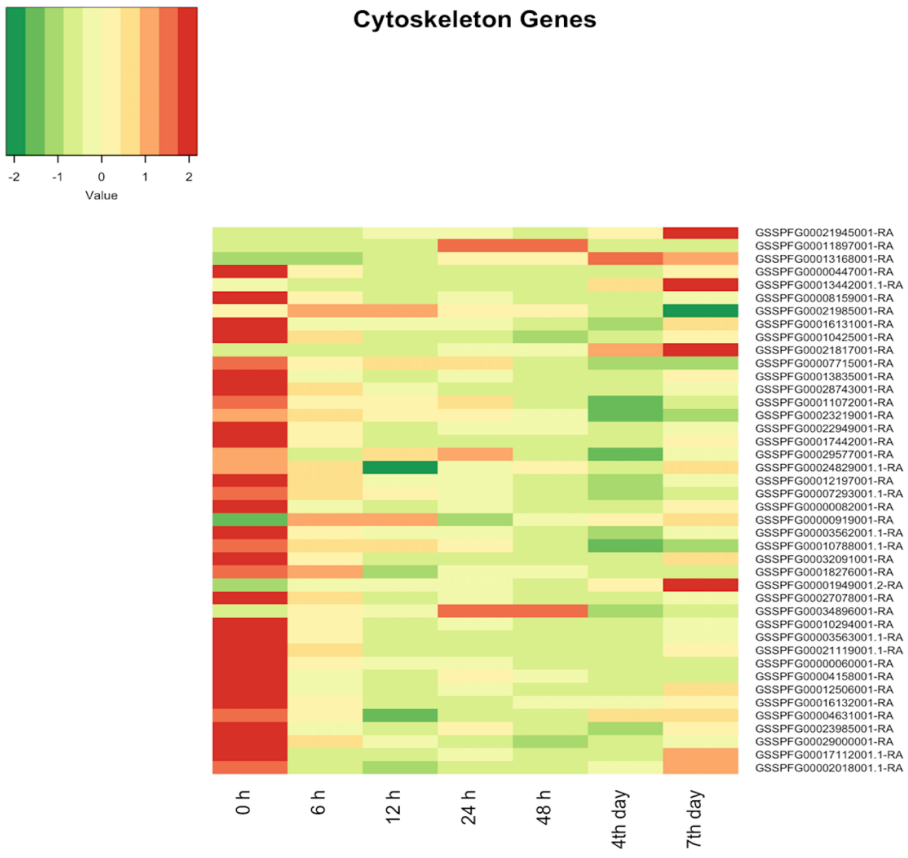


FIG 3.2 Heat map representation of cytoskeleton genes expression in *Spodoptera frugiperda* 3rd-instar larvae infected with the Spodoptera frugiperda ascovirus 1a. Expression patterns are for genes changing by 2-fold or more at any time point post-infection. The heat map color scale represents the z-score normalization for Transcripts Per Million (TPM) of replicate expression level averages at each time point. The *Spodoptera frugiperda* cytoskeleton gene identifications can be accessed through: the SfruDB Information system http://bipaa.genouest. org/is/lepidodb/spodoptera_frugiperda/ (14).

3.4.3 Innate immunity genes transcription pattern

I quantified expression of the primary innate immunity genes identified in the *S. frugiperda* transcriptome and genome (14, 15), and characterized the results as local, humoral, or cellular, based on the structural characterization of the *Drosophila* immune system (18). As discussed below, major innate immunity genes were found to be upregulated or downregulated during ascovirus infection (**Table 3.3, Fig. 3.3**).

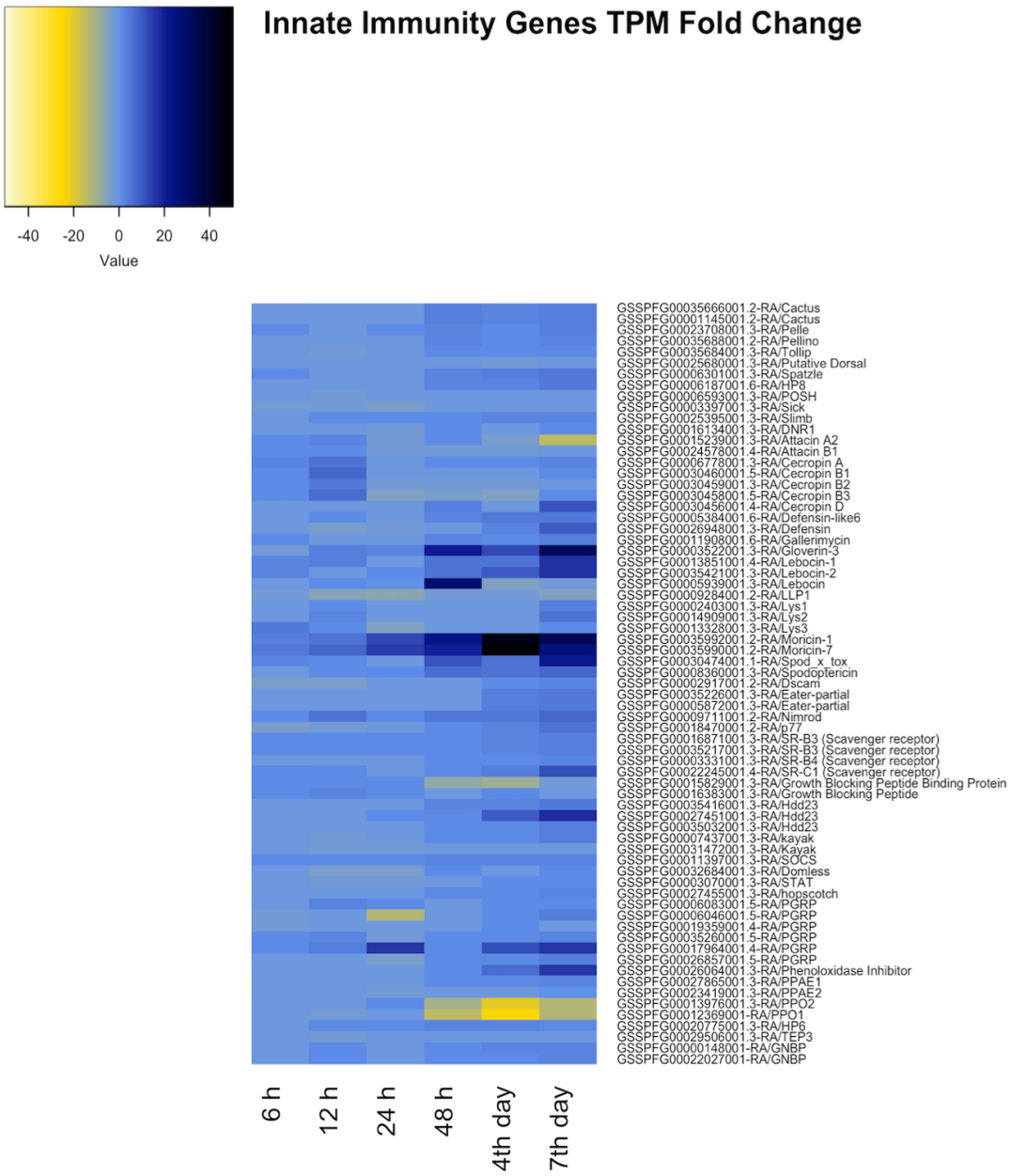


FIG 3.3 Heat map representation of innate immunity gene Transcripts Per Million (TPM) changes in *Spodoptera frugiperda* 3rd-instar larvae infected with the Spodoptera frugiperda ascovirus 1a. The level of upregulation or downregulation was calculated by comparing the gene expression level at the above time points (6h - 7 days post-infection) to the 0h uninfected controls. The heat map color scale shown in the upper left panel represents the TPM fold change of upregulated and downregulated genes. The *Spodoptera frugiperda* innate immunity gene identifications used can be accessed through: the SfruDB Information system http://bipaa.genouest. org/is/lepidodb/spodoptera_frugiperda/ (14).

3.4.3.1 Local immunity - Antimicrobial peptides (AMPs)

The local immune response is the first line of defense against the foreign bodies, and is provoked by entry of microbial pathogens. Activation of this response leads to the production of many antimicrobial peptides (AMPs), limiting the spread of pathogens from wounds. During SfAV-1a infection I found that attacin A2, cecropins (A, B1, B2 and B3), gloverin-3, lebocin (1 and 2), lys (2 and 3), moricin (1 and 7) and spod_x_tox genes were upregulated greater than 2-fold by 6 and/or 12 hpi (**Fig 3.3, Table 3.2**). In addition, cecropin C (GSSPFG00030457001.5-RA) increased from 0 $\log^2$ RPKM at 0 h to more than 1 $\log^2$ RPKM at 6 h, and even higher by the 4th and 7th day pi by (1.85, 5.53 and 9.33 log2 RPKM, respectively). Some of these AMPs continued their expression at similar levels at later time points (i.e. from 24 hpi to 7th day pi). However, gloverin-3, lebocin-1, lebocin-2 and spod_x_tox achieved markedly higher levels at later stages, increasing on day 7 pi by 32.56-fold, 17.02-fold, 16.88-fold and 21.82-fold; respectively. Only moricin-1 and moricin-7 maintained an expression level higher than 2-fold, at sometimes markedly higher, at all times post-infection (ranged from 3.51 to 45.55-fold and 4.37 to 50.28-fold; respectively), indicating their importance to *S. frugiperda* immunity and survival. In addition, other AMPs were upregulated by more than 2-fold later in the infection (from 48 hpi to 7th day pi) either consistently, for example defensin-like 6 and spodoptericin, or only at certain times, for example, cecropin D, defensin, gallerimycin and lys1 (**Table 3.3**). On the other hand, several AMPs were downregulated by more than 2-fold at certain time points during the infection including attacin (A2 and B1), cecropin (B2 and B3), defensin, gloverin-3 (only at 6 hpi), lebocin, lys3. Only,

lysozyme LLP1 was downregulated by more than 2-fold at all the time points post-infection (ranging from -2.10 to -6.76-fold).

3.4.3.2 Humoral Immunity - Immune signaling pathways

Toll pathway

Most of the *S. frugiperda* Toll pathway genes were upregulated in infected larvae by 2-fold or more, especially by 48 hpi, specifically, Pelle, Pellino, Spätzle, HP8 (Spätzle-processing enzyme) and two Cactus genes (**Table 3.3, Fig. 3.4**). Cactus is a negative regulator of the Toll pathway (19). Interestingly, both Cactus genes were upregulated starting from day 2 and remained upregulated by 2-fold or more until day 7 pi (**Fig. 3.4**). The increase in Cactus ranged from 2.55 - 3.49-fold. Tollip was increased beginning day 2 but by less than 2-fold. Finally, only a putative dorsal gene was downregulated at all the time points, ranged from -1.54 to -2.12-fold.

Immune Deficiency (Imd) and Jun N-terminal kinase (JNK) pathways

Although I noticed alterations in expression of most Imd genes, the level of up or down regulation was less than 2-fold. Out of seventeen *S. frugiperda* Imd pathway genes and negative regulators included in the analysis, only four Imd negative regulators showed a change of 2-fold or slightly more at any time after infection. These were POSH, Slimb, DNR1 and Sick (**Table 3.3**). In the JNK pathway only two gene homologs (Fos [Kayak]) changed by 2-fold or more. Kayak (GSSPFG00031472001.3-RA) was downregulated at all time points, while kayak (GSSPFG00007437001.3-RA) was upregulated over 2-fold late in the infection (**Fig. 3.3**).

Janus kinase/signal transducers and activators of transcription (JAK/STAT) pathway

The SOCS (a negative JAK/STAT regulator) and hopscotch were upregulated by 2-fold or more, at 48 hpi and 4th day, respectively. STAT and Domless were downregulated early during infection, however, thereafter were upregulated by 1.78-fold and 1.61-fold at day 7th pi, respectively demonstrating activation of some JAK/STAT genes (**Table 3.3, Fig. 3.3**).

Melanization

Both prophenoloxidase subunit 1, PPO1 and prophenoloxidase subunit 2, PPO2 genes and the prophenoloxidase-activating enzyme (PPAE2) gene were downregulated at almost all time points, and markedly by 4 days pi (**Table 3.3**). Alternatively, prophenoloxidase-activating enzyme (PPAE1) and Phenoloxidase Inhibitor (POI) were upregulated, respectively, at 48 h by 2.07-fold, and by about 16-fold at day 7 pi, (**Fig. 3.4**).

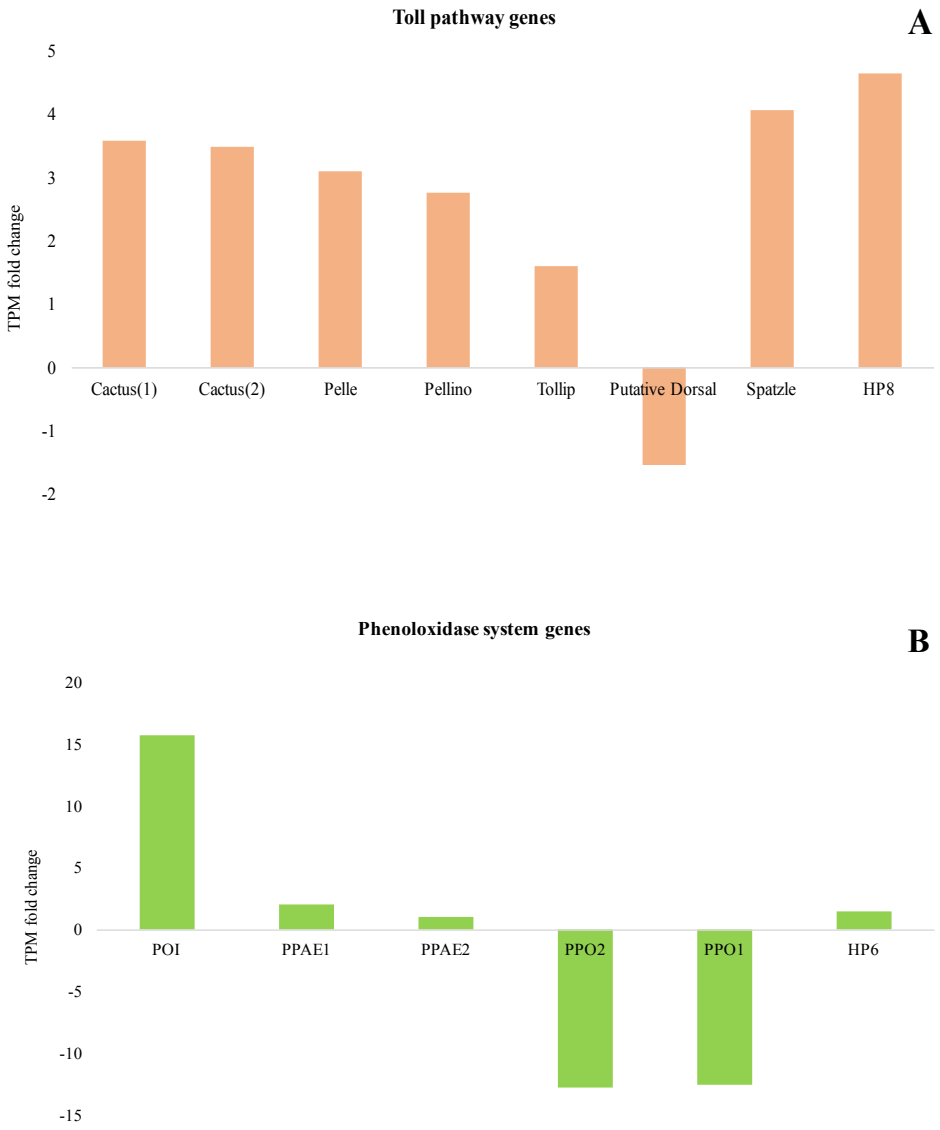


FIG 3.4 Upregulated and downregulated innate immunity genes of 3rd-instar *Spodoptera frugiperda* infected with Spodoptera frugiperda ascovirus 1a. **A)** Toll pathway and **B)** phenoloxidase system (PO) and melanization cascade associated gene Transcripts Per Million (TPM) levels at 7 days post-infection. Expression level changes were calculated by comparing levels at day 7 post-infection with the 0h control (uninfected) expression level. Cactus and Phenoloxidase Inhibitor (POI) are negative regulators of Toll pathway and PO system, respectively. Cactus(1) refers to GSSPFG00035666001.2-RA and Cactus(2) refers to GSSPFG00001145001.2-RA gene identifications for the *Spodoptera frugiperda* genome (14).

3.4.3.3 Cellular immunity

The electron micrograph and RNA-Seq results provide support for the activation of the cellular immunity. The marked reduction in prophenoloxidase transcripts and some AMPs, probably due to the loss of healthy fat body tissue, may also result in greater cellular immunity. However, the only cellular immunity I observed were increases in the phagocytosis pathway.

Phagocytosis

Many Pattern recognition receptors (PRRs) associated with phagocytosis were upregulated by SfAV-1a infection. The Nimrod and two Scavenger receptors (SR-B3) were upregulated at all time points post-infection. A DSCAM, two Eater genes and Scavenger receptors (SR-B4 and SR-C1) were also upregulated, but late in infection (**Table 3.3, Fig. 3.5**). In addition, six Peptidoglycan Recognition Proteins (PGRPs) were upregulated late during infection, whereas the TEP3 (Thioester containing protein) was downregulated at all time points (**Table 3.3**).

Finally, the Growth Blocking Peptide Binding Protein (GBP-BP) gene was downregulated by -2.79 at 48 hpi and by -9.65-fold at 7 days pi. Simultaneously, the Growth Blocking Peptide (GBP) gene showed a slight increase at 4^{th} day pi by 1.37-fold. Moreover, P77 or the GBP adapter was upregulated starting from 48 hpi to 7^{th} day (by 1.02 to 5.47-fold). GBP is a multifunctional insect cytokine associated with stimulating plasmatocyte spreading, oenocytoid lysis, and growth regulation, whereas GBP-BP is a negative regulator of GBP in that after binding together the GBP-GBP-BP complex is

cleared from the blood (20), thereby suggesting that the downregulation of GBP-BP preserves plasmatocyte spreading and oenocyte lysis.

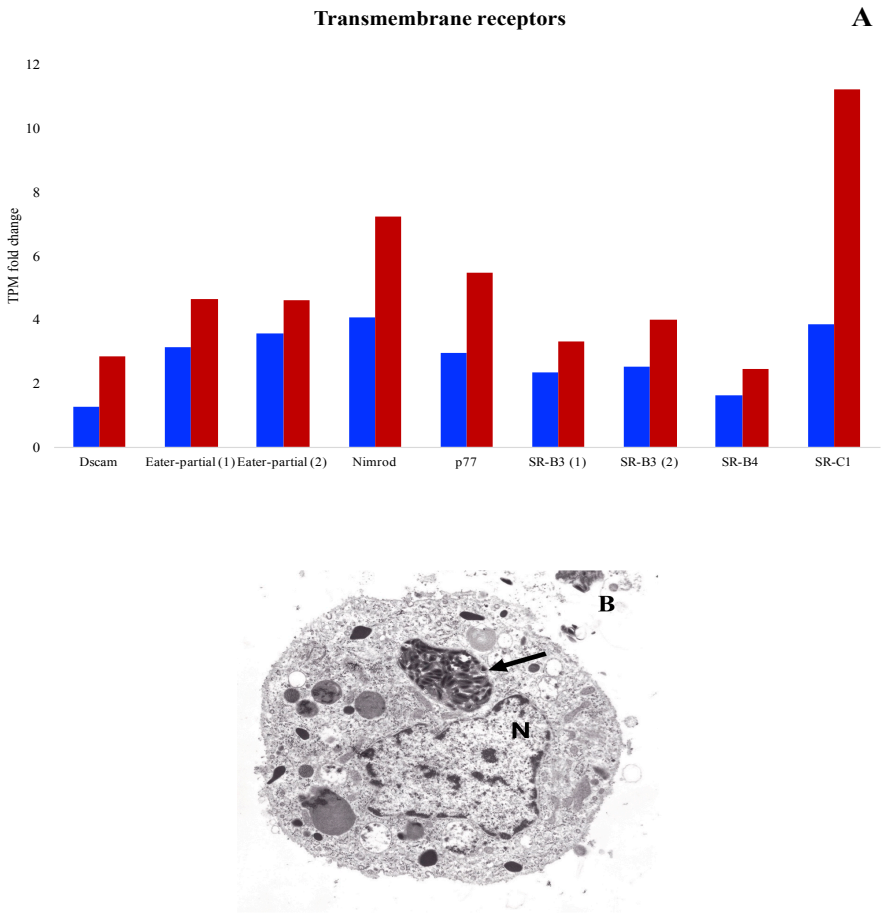


FIG 3.5 Upregulation of *Spodoptera frugiperda* 3rd-instar larvae transmembrane receptor genes after infection with Spodoptera frugiperda ascovirus-1a. These transmembrane receptor proteins are known to be associated with phagocytosis. **A)** Transmembrane receptors gene Transcripts Per Million (TPM) fold changes at 4th day (Blue bars) and 7th day (Red bars) post-infection. Eater-partial (1) refers to GSSPFG00035226001.3-RA and Eater-partial (2) refers to GSSPFG00005872001.3-RA gene identifications for the *Spodoptera frugiperda* genome (14). SR-B3 (1) refers to GSSPFG00016871001.3-RA and SR-B3 (2) refers to GSSPFG00035217001.3-RA gene identification in the same genome. **B)** Hemocyte with phagocytosized ascovirus vesicle containing virions (arrow). Note that the nucleus (N) is intact.

3.4.3.4 Genes associated with immunity

The innate immunity gene, Hdd23-like GSSPFG00027451001.3-RA and two homologs (GSSPFG00035416001.3-RA and GSSPFG00035032001.3-RA) were upregulated, starting from 48 hpi. By day 7 pi their expression level increased by 17.99, 4.03 and 3.35-fold, respectively (**Table 3.3, Fig. 3.6**).

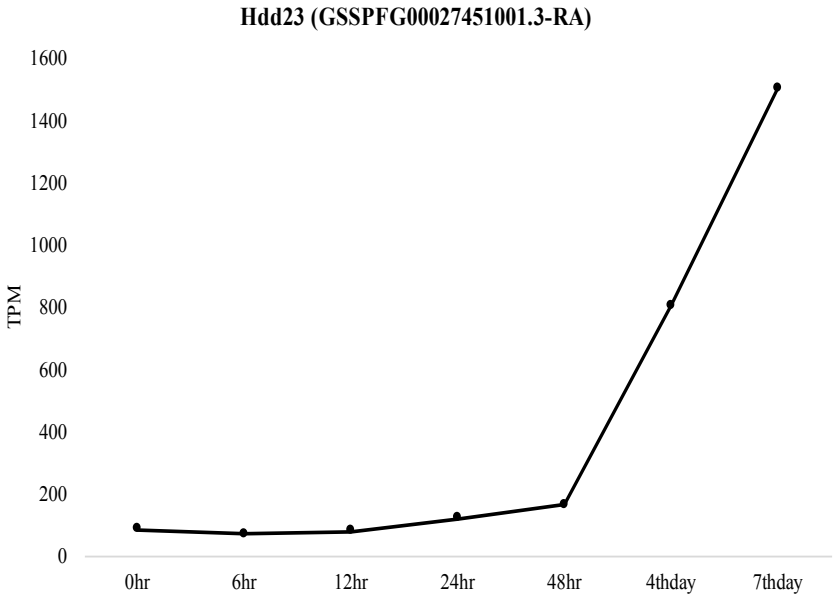


FIG 3.6 Expression of the *Spodoptera frugiperda* Hdd23 gene from 0h to 7 days post-infection with Spodoptera frugiperda ascovirus 1a. The high expression level is shown in Transcripts Per Million (TPM).

3.5 Discussion

Studies of ascovirus cytopathology, especially ultrastructural investigations of the remarkable changes in cell architecture that result in the formation of viral vesicles, suggest a complex interaction between these viruses and their lepidopteran hosts. The changes in mitochondrial shape and their aggregation at sites where the delimiting membranes of the viral vesicles form, for example, suggest these organelles are producing ATP required for synthesis of these membranes (1, 21). Moreover, previous studies suggest a combination of viral gene alteration of proteins known to be components of innate immune responses, as well as possible ascovirus encoded evasion of the host innate immune system. The former is supported by ascovirus synthesis of the Diedel protein, an important antimicrobial innate immunity protein, and the latter by the persistence of viral vesicles for four to six weeks or more while circulating in the hemolymph, especially in *S. frugiperda* larvae infected with SfAV-1a, despite almost complete destruction of the fat body tissue during the first two weeks of infection. In the present study, therefore, I extended the previous *in vivo* investigations of the SfAV-1a transcriptome in infected *S. frugiperda* 3rd-instars (7) by focusing on host mitochondrial, cytoskeleton and innate immunity gene expression patterns. As reported and discussed below, I found stability for most mitochondrial genes to upregulation during vesicle formation and upregulation of some mitochondria motor genes and key innate immune pathway genes.

I detected *Spodoptera frugiperda* 3rd-instar larvae mitochondrial transcriptome gene expression variation from being stable to increase by levels typically less than 2-

fold, after infection with Spodoptera frugiperda ascovirus-1a (SfAV-1a). Interestingly, the transcripts upregulation temporally correlated with vesicle formation stage, implying the conservation of this organelle after the nucleus destruction rather than being destroyed as occurs in apoptosis. Thus, the data support previous ultrastructural studies that predicted mitochondria involvement in vesicle membrane formation (1, 3). Overall, there are very few studies that examine expression of the mitochondrial genome in other insects infected with viruses, or even vertebrate animals or cell lines infected with mammalian viruses, so it is not possible to make direct comparisons. There is, however, some evidence in at least one case where mitochondria play a role in human cells infected with cytomegalovirus. This virus typically causes a chronic infection, but when treated with chloramphenicol it significantly inhibited mitochondrial translation and virus replication indicating the mitochondria directly enhanced viral replication (22). In other cases, for example where viruses replicate quickly and have short life cycles, the role of mitochondria in viral replication appears unimportant, and mitochondrial DNA can be rapidly degraded. In addition, in vertebrate cells such as the Vero cell line infected with African swine fever virus, or chick embryo fibroblasts infected with Frog virus 3, microtubules mediate the movement of mitochondria to be close to the virogenic stroma (23-25), apparently to make virus replication more efficient. Alternatively, in herpes simplex virus type 1 (HSV-1) mitochondrial DNA is degraded by a virus-encoded mitochondrial-targeted nuclease (26, 27). The chronic nature of ascovirus infection and the use of ascovirus vesicles for virus replication and virion assembly explains the need

for maintaining mitochondrial function for many days as the infection spreads throughout the fat body tissue.

The changes in mitochondrial shape and size during ascovirus infection may be partially explained by the RNA-Seq data as an outcome of expression level alterations of complex I (NADH dehydrogenases), complex III (cytochrome b), complex IV (cytochrome c oxidases) and complex V (ATP synthases), given that these subunits are embedded in the inner mitochondrial membrane. However, these changes alone most likely do not account for the radical changes in cell architecture characteristic of ascovirus infection. Furthermore, it is unlikely that mitochondrial genes encoded in the nuclear genome are involved in these changes because the nucleus and nuclear DNA is degraded early during cell infection. Thus, it is more likely these marked changes in the size, shape, and movement of mitochondria are due to several viral genes that contain putative sequences targeting the encoded proteins to mitochondria, identified in the previous study of the SfAV-1a transcriptome (7).

While the mitochondria changes are marked in ascovirus infections, based on biochemical studies it is known that many mammalian viral proteins manipulate these organelles (reviewed in 23, 28). For example, viral proteins can affect regulation of mitochondrial membrane potential (29), alter the intracellular distribution of mitochondria (24), "hijack" mitochondrial proteins enabling their genomes to enter the nucleus (30), and deplete mitochondrial DNA (31). It is also possible that other complex insect viruses like baculoviruses and entomopoxviruses directly manipulate the mitochondrial transcriptome, but none has been identified. In general, more transcriptome

studies are needed to determine which viruses are capable of manipulating mitochondria for virus replication and the mechanisms underlying these virus-directed changes. Such studies could well be of health importance to humans because it is now known that several neurological diseases of aging are due to mutants in mitochondrial genes encoded in either the mitochondrial or nuclear genome (32-34).

On the other hand, the downregulation of cytoskeleton genes by 2-fold or more was surprising given the drastic changes induced in the infected cell to transform it into 20-30 or more viral vesicles. However, I noticed that the few upregulated genes (**Table 3.2**) included those for three kinesins and a dynein-related protein. Kinesin and dynein are main motor proteins involved in moving mitochondria along axons. Moreover, a virus structural gene known as a Dynein-like b chain (ORF084) that is conserved in TnAV and HvAV species, was identified recently as a late SfAV-1a gene for which expression begins by 12 hpi and peaks at 24 hpi (4.14 $\log^2$ RPKM) (7). Specifically, kinesin is associated with antegrade and dynein with retrograde movement. For example, in *Drosophila*, cytoplasmic dynein was associated with retrograde mitochondrial transport. In addition, I found a myosin gene (myosin-VIIa) was upregulated. Myosins are associated with mitochondrial movement, however, only for short distances (35, 36).

Previous studies of lepidopteran larvae infected with ascoviruses reported the accumulation of high concentrations of viral vesicles in the hemolymph, which turns this tissue milky white. These studies also showed viral vesicles continued to be produced and circulate in the hemolymph for weeks, thereby favoring transmission of the virus by parasitic wasps (1, 2). Other than the change in color, however, there were no obvious

signs of an innate immune response such as melanization of the vesicles or phagocytosis by hemocytes as they circulate in the hemolymph. Moreover, the long-term survival of infected larvae for transmission might favor an immune response to prevent them from succumbing to infections by other pathogens. This raises the question; does the larval host not recognize the vesicles as foreign. Based on a recent transcriptomic study for *Spodoptera exigua* larvae infected with Heliothis virescens ascovirus-3h (HvAV-3h), the Toll-like receptor signaling and JAK/STAT pathways along with other genes involved with immunity were upregulated upon infection (12).

The insect innate immune system is classified as being local, humoral, or cellular, based on the structural characterization of the *Drosophila* immune system (18). Interestingly, the local immunity represented in antimicrobial peptides such as attacin, cecropin, moricin, defensin, gloverin, gallerimycin and lebocin, all known to be effective in combination against a wide range of bacterial, fungal, and viral pathogens that infect lepidopteran larvae (37, 41) were upregulated either consistently or at different time points post-infection. Thus, they can be understood as a systemic antimicrobial defense induced by wounding of the insect cuticle (42, 43), which our infection process may have triggered, as it mimicked the type of wound made by a wasp ovipositor. Therefore, the induction of these antimicrobial peptides might prevent opportunistic pathogens from competing for host resources during infection. With respect to lysozymes, upregulation of Lys1 and Lys2, and downregulation of lysozyme-like LLP1 appears contradictory. However, these enzymes are known to be associated with different roles other than peptidoglycan attack, for example, melanization (44).

Induction of the Toll pathway is initiated by interactions between pathogen-associated molecular patterns (PAMPs) and host pattern recognition receptors (PRRs). Gram-positive bacterial and fungal cell wall components, for example, are the main inducers of the Toll cascade (45). Whereas it has not been determined how the NF-κB-dependent pathways (Toll and Imd) are induced by a virus infection, these pathways have nevertheless been shown to limit insect virus infections. Specifically, the Toll pathway and cellular immunity limit Drosophila X virus (46), and the Imd pathway and cellular immunity, also in *Drosophila*, limit Cricket Paralysis Virus (47). Alternatively, it has been shown that certain Toll-Dorsal pathway genes, specifically, *pelle*, *tube* and *cactus*, are involved in development in *Drosophila* dorsal-ventral patterning, a difference that distinguishes the invertebrate Toll pathway from that of mammals (48-51). Therefore, the upregulation of Toll genes after SfAV-1a infection may be both an immune response and to modify larval development. Of interest to the latter, development is very retarded during ascovirus disease compared to that of healthy larvae (3). Moreover, I detected the upregulation of some JAK/STAT associated genes (Domless, STAT, SOCS and hopscotch) especially late in the infection starting between 48 hpi and 4th day and lasted till day 7 pi (**Table 3.3**). Interestingly, the Unpaired cytokine (Upd3) or the JAK/STAT pathway activator was not detected in the *S. frugiperda* transcriptome (15). Similarly, the *Drosophila* extracellular cytokine unpaired3 (Upd3) ortholog was missing in *Manduca sexta* and *Tribolium castaneum* (52, 53), implying the involvement of an undefined JAK/STAT pathway activating cytokine that replaces Upd3 in these insects. The insect JAK/STAT is like the Toll pathway in terms of being associated with immunity and

development (51, 53). The upregulation of Toll and JAK/STAT genes after SfAV-1a infection agrees with the *S. exigua* response to HvAV-3h (12).

The low variation in expression levels of Imd genes is not surprising because there is evidence that expression of this pathway can be deleterious to insects. For example, when the Imd pathway in *Drosophila* was downregulated, i.e., suppressed by a small fly protein known as Diedel, flies lived longer. When the *die* gene was mutated, the life span of the flies was reduced by about 10 days, which is significantly different from wild type flies that typically live for 70 days (54). Interestingly, the Imd-associated genes that varied (increase or decrease) with 2-fold or more were four negative regulators (POSH, Sick, Slimb, DNR1) (**Table 3.3**). Of relevance here is that SfAV-1a synthesizes high levels of a diedel-homolog (ORF121) during infection and thus its function may be to suppress the Imd pathway. If so, this would help explain the longevity of larvae infected with ascoviruses as well as the low expression levels of Imd pathway genes. Alternatively, the expression level of Imd genes may be tightly controlled by mechanisms not yet identified during virus infections. Overall, the upregulation of the Toll pathway negative regulator of Cactus in combination with the highly expressed ascovirus virokine diedel (ORF121), which apparently acts as an Imd suppressor (7), may ultimately help explain at the molecular level how SfAV-1a controls expression of two of the most important signaling innate immunity pathways.

In lepidopteran insects, the oenocytoids are the hemocytes associated with the production of melanization cascade enzymes. The antiviral and antibacterial effect of the melanin cascade associated compounds has been demonstrated in many recent studies

(55-57). Therefore, the phenoloxidase pathway gene alterations indicate the activation of the melanization cascade against SfAV-1a, however, the high upregulation of the host phenoloxidase inhibitor (~16-fold by day 7^{th} pi) may decrease the effectiveness of this pathway against this virus. The dedication of a specific viral protein to inhibit the PO is also reported for the viral protein WSSV453 produced by the white spot syndrome virus, which inhibits the melanization cascade in the shrimp, *Penaeus monodon*, by interacting with and preventing activation of PmproPPAE2, pro PO-activating enzyme 2 (57). Whether ascovirus genomes possess a melanization inhibitor is not known, but ascoviruses have large DNA genomes and synthesize many proteins with unknown functions.

The SfAV-1a tissue tropism is primarily restricted to the *S. frugiperda* fat body. This tissue is completely destroyed by day 12 pi, with dense accumulations of viral vesicles (10^8/ml) in the hemolymph by 7 days pi (21). Although the fat body, with functions similar to the mammalian liver, is not essential for larval survival, it does regulate many aspects of insect growth, development and metamorphosis (58, 59). Very importantly, it also has primary responsibility for innate immunity responses (60). Interestingly, I detected many PRRs associated with phagocytosis upregulation by SfAV-1a infection for example, Nimrod, Scavenger receptors, DSCAM and Eater. For example, many PRRs are associated with phagocytic events (61-63) against bacteria, plasmodia, (64) and recently viruses. For instance, recently the class A (65) and C (66) scavenger receptors role in virus phagocytosis was confirmed. The class C receptor in the shrimp *Marsupenaeus japonicus* is efficient in inducing hemocyte phagocytosis of white spot

syndrome virus thereby preventing systemic infection (66). Phagocytosis is a cellular immune response that takes place by either circulating or immobile hemocytes. In lepidopterans, the plasmatocytes are responsible for this process (57, 67). Each phagocyte can internalize and degrade numerous invading pathogens (68). Therefore, given that the fat body is almost completely destroyed by SfAV-1a infection by day 7, it is expected that humoral immunity would be significantly reduced as the infection increases. Indeed, detection of a phagocytosis event to one of the ascovirus vesicles by host blood cells using EM and upregulation of phagocytosis-associated receptors support the induction of cellular immunity.

Finally, I detected the high upregulation of an Hdd23 gene or an immune associated protein by ~18-fold at day 7 pi. The Hdd23 gene was identified in the fall webworm, *Hyphantria cunea*, as an early acute-phase immunity associated protein upregulated continuously after a bacterial challenge (69). However, a protein that shares sequence similarity with Hdd23 (known as stress-responsive peptide, SRP) was characterized recently in *Helicoverpa armigera* and found to be associated with prophenoloxidase and nodulation activation (70). The high upregulation of Hdd23-like proteins (**Fig. 4.5**) shows that this is in response to SfAV-1a infection, even though they apparently are largely ineffective against this virus. Thus, as in the case of the antimicrobial peptides, given the response to bacterial challenge in *H. cunea* and activation of nodule formation in *H. armigera*, the function of these proteins may be to inhibit bacterial infections during the prolonged larval stage caused by ascovirus infection.

In conclusion, the evolutionary success of the chronic characteristic of ascovirus disease in their lepidopteran larval hosts may reflect their unique ability to conserve mitochondrial functions while keeping the balance between the different innate immunity cascade inducers and inhibitor/negative regulators that are expressed simultaneously. The transcriptome data shows that some of these processes are controlled by viral genes, such as the preservation of mitochondrial functions, whereas other are likely direct innate immune responses due directly to cuticular damage due to oviposition attempts. Studies of many vertebrate viruses over the past decade, especially of mammalian viruses, have identified many proteins involved in manipulating mitochondria and inducing innate immune responses. Ascoviruses have large genomes and the function of many genes remain unknown. Thus, it is possible that several of these if not more are involved in the direct manipulation of mitochondria and innate immunity, suggesting a wealth of new information may be derived by transcriptome studies of additional ascoviruses and other large DNA insect viruses.

3.6 References

1. Federici BA. 1983. Enveloped double stranded DNA insect virus with novel structure and cytopathology. Proc Natl Acad Sci U S A 80: 7664–7668.

2. Govindarajan R, Federici BA. 1990. Ascovirus infectivity and effects of infection on the growth and development of noctuid larvae. J Invertebr Pathol 56:291–299.

3. Federici BA, Bideshi DK, Tan Y, Spears T, Bigot Y. 2009. Ascoviruses: superb manipulators of apoptosis for viral replication and transmission. Curr Top Microbiol Immunol 328:171–196.

4. Bideshi DK, Tan Y, Bigot Y, Federici BA. 2005. A viral caspase contributes to modified apoptosis for virus transmission. Genes Dev 19:1416–1421.

5. Asgari S. 2007. A caspase-like gene from the *Heliothis virescens* ascovirus (HvAV3e) is not involved in apoptosis but is essential for virus replication. Virus Res 128:99 –105.

6. Bigot Y, Asgari S, Bideshi DK, Cheng X, Federici BA, Renault S. 2011. Family *Ascoviridae*, p 147-152. *In* King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (ed), Viral Taxonomy, IX Report of the International Committee on the Taxonomy of Viruses, 3rd edn. London: Elsevier–Academic Press.

7. Zaghloul HAH, Hice R, Arensburger P, Federici BA. 2017. Transcriptome analysis of the Spodoptera frugiperda ascovirus *in vivo* provides insights into how its apoptosis inhibitors and caspase promote increased synthesis of viral vesicles and virion progeny. J Virol 91:e00874-17.

8. Liu QN, Chai XY, Ge BM, Zhou CL, Tang BP. 2016. The complete mitochondrial genome of fall armyworm *Spodoptera frugiperda* (Lepidoptera: Noctuidae). Genes Genom 38:205–216.

9. Babbucci M, Basso A, Scupola A, Patarnello T, Negrisolo E. 2014. Is it an ant or a butterfly? Convergent evolution in the mitochondrial gene order of Hymenoptera and Lepidoptera. Genome Biol Evol 6:3326–3343.

10. Bideshi DK, Bigot Y, Federici BA, Spears T. 2010. Ascoviruses, p 3–34. *In* Asgari S, Johnson KN (ed), Insect virology. Caister Academic Press, Norfolk, United Kingdom.

11. Hussain M, Abraham AM, Asgari S. 2010. An ascovirus-encoded RNase III autoregulates its expression and suppresses RNA interference- mediated gene silencing. J Virol 84:3624–3630.

12. Yu H, Li Z-Q, He L, Ou-Yang Y-Y, Li N, Huang G-H. 2018. Response analysis of host Spodoptera exigua larvae to infection by Heliothis virescens ascovirus 3h (HvAV-3h) via transcriptome. Scientific Reports 8:5367.

13. Bigot Y. 2011. Genus *Ascovirus*, p 73–78. *In* Tidona C, Darai G (ed), Springer index of viruses, 2nd ed. Springer, Heidelberg, Germany.

14. Gouin, A. et al. 2017. Two genomes of highly polyphagous lepidopteran pests (*Spodoptera frugiperda*, Noctuidae) with different host-plant ranges. Sci Rep 7:11816.

15. Legeai F, Gimenez S, Duvic B, Escoubas JM, Grenet ASG, Blanc F, Cous- serans F, Seninet I, Bretaudeau A, Mutuel D, Girard PA, Monsempes C, Magdelenat G, Hilliou F, Feyereisen R, Ogliastro M, Volkoff AN, Jacquin- Joly E, d'Alençon E, Nègre N, Fournier P. 2014. Establishment and analysis of a reference transcriptome for *Spodoptera frugiperda*. BMC Genomics 15:704.

16. Li B, Dewey CN. 2011. "RSEM: Accurate Transcript Quantification from RNA-Seq Data with or without a Reference Genome." BMC Bioinformatics 12: 323.

17. Federici, BA. 1982. A new type of insect pathogen in larvae of the clover cutworm, *Scotogramma trifolii*. J lnvertebr Pathol 40:41-54.

18. Valanne S, Kallio J, Kleino A, Rämet M. 2012. Large-scale RNAi screens add both clarity and complexity to *Drosophila* NF-κB signaling. Dev Comp Immunol 37: 9–18.

19. Lindsay SA, Wasserman SA. 2014. Conventional and non-conventional *Drosophila* toll signaling. Dev Comp Immunol 42:16e24.

20. Matsumoto Y, Oda Y, Uryu M, Hayakawa Y. 2003. Insect cytokine growth-blocking peptide triggers a termination system of cellular immunity by inducing its binding protein. J Biol Chem 278:38579–38585.

21. Federici BA, Govindarajan R. 1990. Comparative histology of three ascovirus isolates in larval noctuids. J Invertbr Pathol 56: 300-311.

22. Karniely S, Weekes MP, Antrobus R, Rorbach J, van Haute L, Umrania Y, Smith DL, Stanton RJ, Minczuk M, Lehner PJ, Sinclair JH. 2016. Human cytomegalovirus infection upregulates the mitochondrial transcription and translation machineries. MBio 7:e00029.

23. Anand, SK, Tikoo SK. 2013. Viruses as modulators of mitochondrial functions. Adv Virol 2013. 738794.

24. Rojo G, Chamorro M, Salas ML, Vinuela E, Cuezva JM, Salas J. 1998. Migration of mitochondria to viral assembly sites in African Swine Fever Virus-infected cells. J Virol 72:7583–7588.

25. Kelly DC. 1975. Frog virus 3 replication: electron microscope observations on the sequence of infection in chick embryo fibroblasts. J Gen Virol 26:71–86.

26. Claus C, Liebert UG. 2014. A renewed focus on the interplay between viruses and mitochondrial metabolism. Arch Virol 159:1267–1277.

27. Duguay BA, Saffran HA, Ponomarev A, Duley SA, Eaton HE, Smiley JR. 2014. Elimination of mitochondrial DNA is not required for herpes simplex virus 1 replication. J Virol 88:2967–2976.

28. Khan M, Syed GH, Kim SJ, Siddiqui A. 2015. Mitochondrial dynamics and viral infections: A close nexus. Biochim Biophys Acta 1853:2822–33.

29. Nudson WA, Rovnak J, Buechner M, Quackenbush SL. 2003. Walleye dermal sarcoma virus orf C is targeted to the mitochondria. J Gen Virol 84:375-81.

30. Matthews DA, Russell WC. 1998. Adenovirus core protein V interacts with p32--a protein which is associated with both the mitochondria and the nucleus. J Gen Virol 79:1677-85.

31. Saffran HA, Pare JM, Corcoran JA, Weller SK, Smiley JR. 2007. Herpes simplex virus eliminates host mitochondrial DNA. EMBO Rep 8:188-93.

32. Smith DR. 2013. RNA-Seq data: a goldmine for organelle research. Brief Funct Genom 12: 454–456.

33. Hodgkinson A, Idaghdour Y, Gbeha E, Grenier JC, Hip-Ki E, Bruat V, Goulet JP, de Malliard T, Awadalla P. 2014. High-resolution genomic analysis of human mitochondrial RNA sequence variation. Science 344:413–415.

34. Tian Y, Smith DR. 2016. Recovering complete mitochondrial genome sequences from RNA-Seq: a case study of *Polytomella* non-photosynthetic green algae. Mol Phylogenet Evol 98:57–62.

35. Pilling AD, Horiuchi D, Lively CM, Saxton WM. 2006. Kinesin-1 and Dynein are the primary motors for fast transport of mitochondria in *Drosophila* motor axons. Mol Biol Cell 17:2057–2068.

36. Wu M, Kalyanasundaram A, Zhu J. 2013. Structural and biomechanical basis of mitochondrial movement in eukaryotic cells. Int J Nanomedicine 8:4033–4042.

37. Hara S, Yamakawa M. 1995. Moricin, a novel type of antibacterial peptide isolated from the silkworm, *Bombyx mori*. J Biol Chem 270:29923–29927.

38. Brown SE, Howard A, Kasprzak AB, Gordon KH, East PD. 2008. The discovery and analysis of a diverged family of novel antifungal moricin-like peptides in the wax moth *Galleria mellonella*. Insect Biochem Mol Biol 38:201–212.

39. Hultmark D, Engstrom A, Andersson K, Steiner H, Bennich H, Boman HG. 1983. Insect immunity. Attacins, a family of antibacterial proteins from *Hyalophora cecropia*. EMBO J 2:571–576.

40. Yi H-Y, Chowdhury M, Huang Y-D, Yu X-Q. 2014. Insect Antimicrobial Peptides and Their Applications. Appl Microbiol Biotechnol 98:5807-5822.

41. Badapanda C, Chikara SK. 2016. Lepidopteran Antimicrobial Peptides (AMPs): Overview, Regulation, Modes of Action, and Therapeutic Potentials of Insect-Derived AMPs. In: Raman C, Goldsmith M, Agunbiade T (eds) Short Views on Insect Genomics and Proteomics. Entomology in Focus, vol 4. Springer, Cham.

42. Brey PT, Lee WJ, Yamakawa M, Koizumi Y, Perrot S, François M, Ashida M. 1993. Role of the integument in insect immunity: epicuticular abrasion and induction of cecropin synthesis in cuticular epithelial cells. Proc Natl Acad Sci USA 90:6275-6279.

43. Ferrandon D, Jung AC, Criqui M, Lemaitre B, Uttenweiler- Joseph S, Michaut L, Reichhart J, Hoffmann JA. 1998. A drosomycin- GFP reporter transgene reveals a local immune response in *Drosophila* that is not dependent on the Toll pathway. EMBO J 17:1217-1227.

44. Rao XJ, Ling E, Yu XQ. 2010. The role of lysozyme in the prophenoloxidase activation system of *Manduca sexta*: an in vitro approach. Dev Comp Immunol 34:264e271.

45. Tanji T, Hu X, Weber AN, Ip YT. 2007. Toll and IMD Pathways Synergistically Activate an Innate Immune Response in *Drosophila melanogaster*. Mol Cell Biol 27:4578-88.

46. Zambon RA, Nandakumar M, Vakharia VN, Wu LP. 2005. The Toll pathway is important for an antiviral response in *Drosophila*. Proc Natl Acad Sci USA 102:7257–7262.

47. Costa A, Jan E, Sarnow P, Schneider DS. 2009. The Imd pathway is involved in antiviral immune responses in *Drosophila*. PLoS ONE 4:e7436.

48. Lemaitre B, Nicolas E, Michaut L, Reichhart JM, Hoffmann JA. 1996. The dorsoventral regulatory gene cassette *spatzle/Toll/cactus* controls the potent antifungal response in *Drosophila* adults. Cell 86:973–983.

49. Belvin MP, Anderson KV. 1996. A conserved signaling pathway: the *Drosophila* toll-dorsal pathway. Annu Rev Cell Dev Biol 12:393–416.

50. Hultmark D. 2003. *Drosophila* immunity: paths and patterns. Curr Opin Immunol 15:12–19.

51. Xu J, Cherry S. 2014. Viruses and antiviral immunity in *Drosophila*. Dev Comp Immunol 42:67e84.

52. Zou Z, Evans J, Lu Z, Zhao P, Williams M, Sumathipara N, Hetru C, Hultmark D, Jiang H. 2007. Comparative genome analysis of the *Tribolium* immune system. Genome Biol 8:R177.

53. Cao X, He Y, Hu Y, Wang Y, Chen YR, Bryant B, Clem RJ, Schwartz LM, Blissard G, Jiang H. 2015.The immune signaling pathways of *Manduca sexta*. Insect Biochem. Mol Biol 62:64-74.

54. Lamiable O, Kellenberger C, Kemp C, Troxler L, Pelte N, Boutros M, Marques JT, Daeffler L, Hoffmann JA, Roussel A, Imler JL. 2016. Cytokine Diedel and a viral homologue suppress the IMD pathway in *Drosophila*. Proc Natl Acad Sci U S A 113:698–703.

55. Zhao P, Lu Z, Strand MR, Jiang H. 2011. Antiviral, anti-parasitic, and cytotoxic effects of 5,6-dihydroxyindole (DHI), a reactive compound generated by phenoloxidase during insect immune response. Insect Biochem Mol Biol 41:645–652.

56. Charoensapsri W, Amparyup P, Suriyachan C, Tassanakajon A. 2014. Melanization reaction products of shrimp display antimicrobial properties against their major bacterial and fungal pathogens. Dev Comp Immunol 47:150–159.

57. Sutthangkul J, Amparyup P, Eum JH, Strand MR, Tassanakajon A. 2017. Anti-melanization mechanism of the white spot syndrome viral protein, WSSV453, via interaction with shrimp proPO-activating enzyme, PmproPPAE2. J Gen Virol 98:769-778.

58. Mirth CK, Riddiford LM. 2007. Size assessment and growth control: how adult size is determined in insects. BioEssays 29:344–55.

59. Hoshizaki DK. 2005. Fat-cell development. In Complete Molecular Insect Science, ed. LI Gilbert, K Iatrou, S Gill, 2:315–45. Berlin: Elsevier.

60. Ferrandon D, Imler JL, Hetru C, Hoffmann JA. 2007. The *Drosophila* systemic immune response: sensing and signalling during bacterial and fungal infections. Nat Rev Immunol 7:862–74.

61. Mukhopadhyay S, Gordon S. 2004. The role of scavenger receptors in pathogen recognition and innate immunity. Immunobiol 209:39–49.

62. Kurucz E, Markus R, Zsamboki J, Folkl-Medzihradszky K, Darula Z, Vilmos P, Udvardy A, Krausz I, Lukacsovich T, Gateff E, Zettervall CJ, Hultmark D, Ando I. 2007. Nimrod, a putative phagocytosis receptor with EGF repeats in *Drosophila* plasmatocytes. Curr Biol 17:649e654.

63. Ramet M, Manfruelli P, Pearson A, Mathey-Prevot B, Ezekowitz RA. 2002. Functional genomic analysis of phagocytosis and identification of a *Drosophila* receptor for *E. coli*. Nature 416:644e648.

64. Dong Y, Taylor HE, Dimopoulos G. 2006. AgDscam, a hypervariable immunoglobulin domain-containing receptor of the *Anopheles gambiae* innate immune system. PLoS Biol 4:e229.

65. Dansako H, Yamane D, Welsch C, McGivern DR, Hu F, Kato N, Lemon SM. 2013. Class A scavenger receptor 1 (MSR1) restricts hepatitis C virus replication by mediating toll-like receptor 3 recognition of viral RNAs produced in neighboring cells. PLoS Pathog 9:e1003345.

66. Yang M, Shi X, Yang H, Sun J, Xu L, Wang X, Zhao X, Wang J. 2016. Scavenger receptor C mediates phagocytosis of white spot syndrome virus and restricts virus proliferation in shrimp. PLoS Pathog 12:e1006127.

67. Honti V, Csordas G, Kurucz E, Markus R, Ando I. 2014. The cell-mediated immunity of *Drosophila melanogaster*: hemocyte lineages, immune compartments, microanatomy and regulation. Dev Comp Immunol. 42:47e56.

68. Oliver, JD, Dusty Loy J, Parikh G, Bartholomay L. 2011. Comparative analysis of hemocyte phagocytosis between six species of arthropods as measured by flow cytometry. J Invertebr Pathol 108:126e130.

69. Shin SW, Park SS, Park DS, Kim MG, Kim SC, Brey PT, Park HY. 1998. Isolation and characterization of immune-related genes from the fall webworm, *Hyphantria cunea*, using PCR-based differential display and subtractive cloning. Insect Biochem Mol Biol 28:827–837.

70. Qiao C, Li J, Wei XH, Wang, JL, Wang YF, Liu XS. 2014. SRP gene is required for *Helicoverpa armigera* prophenoloxidase activation and nodulation response. Dev Comp Immunol 44:94-99.

3.7 Tables:

TABLE 3.1: Changes in *Spodoptera frugiperda* 3rd-instar larval mitochondrial gene expression levels at different time points post-infection with Spodoptera frugiperda ascovirus-1a.

[a]Mitochondrial genes	6 h	12 h	24 h	48 h	4th day	7th day
12S_ribosomal_RNA	[b]-1.75	-2.00	-1.23	2.78	1.81	3.93
16S_ribosomal_RNA	1.06	-1.05	1.31	1.59	1.93	3.21
ATP6	1.36	1.44	1.32	1.05	-1.20	-1.58
ATP8	2.45	2.32	3.43	2.16	-1.02	-1.09
COI	-1.19	-1.21	-1.10	-1.07	1.00	-1.06
COII	1.07	1.07	-1.13	1.13	1.02	1.04
COIII	1.04	1.00	-1.11	1.11	1.08	1.10
CYTB	-1.16	-1.10	-1.02	-1.54	-1.24	-1.25
ND1	1.70	1.81	1.53	1.23	1.10	1.15
ND2	1.50	1.38	1.58	1.17	1.10	1.29
ND3	1.40	1.38	1.16	1.31	1.10	1.30
ND4	-1.14	-1.08	1.52	-1.33	-1.05	-1.05
ND4L	1.62	1.13	2.22	1.40	1.10	1.39
ND5	1.22	1.23	1.38	1.05	-1.15	-1.20
ND6	1.30	1.15	2.13	1.18	1.95	1.53
D_loop	1.33	1.60	1.44	1.03	-1.27	1.86

[a]The *Spodoptera frugiperda* mitochondrial genome used in this study is available through (NCBI KM362176.1 (described in 8). tRNAs are not listed in this table.

[b]Red and blue fonts refer to upregulated and downregulated genes, respectively. ND1-6 refers to NADH dehydrogenase subunits and COI-III refers to cytochrome c oxidase subunits.

TABLE 3.2: Changes in *Spodoptera frugiperda* 3rd-instar larval cytoskeleton gene levels at different time points post-infection with Spodoptera frugiperda ascovirus 1a. Only genes that changed by 2-fold or more are included in this table.

Description	Gene ID*	6hr	12 h	24h	48h	4th	7th
actin-binding LIM protein 2	GSSPFG00027078001-RA	-1.53	-2.44	-2.36	-2.95	-2.89	-2.04
actin-binding protein anillin-like	GSSPFG00004631001-RA	-1.32	-2.02	-1.67	-1.71	-1.19	-1.18
putative actin binding protein, partial	GSSPFG00004158001-RA	-1.97	-2.38	-1.74	-2.16	-2.75	-2.33
dystonin/microtubule-actin cross-linking factor 1-like	GSSPFG00017442001-RA	-2.44	-4.81	-4.60	-6.22	-5.75	-2.49
dystonin/microtubule-actin cross-linking factor 1-like	GSSPFG00032091001-RA	-2.03	-4.08	-3.97	-4.98	-4.68	-1.65
dynactin subunit 1	GSSPFG00000447001-RA	-1.53	-2.30	-2.29	-2.20	-2.18	-1.62
dynein assembly factor with WDR repeat domains 1	GSSPFG00024829001.1-RA	-1.15	-4.15	-1.46	-1.30	-1.85	-1.18
dynein heavy chain, cytoplasmic	GSSPFG00017112001.1-RA	-2.38	-2.29	-1.94	-2.42	-2.65	-1.24
dynein heavy chain, cytoplasmic	GSSPFG00002018001.1-RA	-1.93	-2.31	-1.87	-1.93	-1.77	-1.10
dynein light chain roadblock-type 2	GSSPFG00034896001-RA	1.83	1.34	2.92	2.98	-1.32	1.07
kinesin light chain	GSSPFG00000919001-RA	4.14	4.05	1.23	2.82	3.01	3.65
kinesin-like protein KIF13B	GSSPFG00007715001-RA	-1.37	-1.14	-1.25	-1.96	-2.35	-2.16
kinesin-like protein KIF13B	GSSPFG00010294001-RA	-1.98	-2.98	-2.27	-2.73	-3.40	-2.33
kinesin-like protein KIF18A	GSSPFG00021945001-RA	-1.10	1.02	1.10	-1.06	1.23	2.00
kinesin-like protein KIF19	GSSPFG00011897001-RA	1.62	2.62	108.85**	112.69**	7.54	2.62
kinesin-like protein KIF21B	GSSPFG00028743001-RA	-1.85	-3.53	-3.91	-6.17	-5.63	-2.82
kinesin-like protein KIF23	GSSPFG00001949001.2-RA	1.36	1.43	1.29	1.25	1.57	2.48
kinesin-like protein KIF3A	GSSPFG00000060001-RA	-1.50	-1.80	-1.85	-2.10	-2.33	-2.01
kinesin-like protein Klp10A	GSSPFG00012197001-RA	-1.38	-1.74	-1.68	-1.93	-2.14	-1.75
kinesin-like protein Klp68D	GSSPFG00010425001-RA	-1.49	-2.29	-2.56	-2.82	-2.24	-1.63
kinesin-like protein unc-104	GSSPFG00016131001-RA	-2.23	-2.64	-2.46	-3.67	-4.59	-1.64
kinesin-like protein unc-104	GSSPFG00003563001.1-RA	-2.27	-4.63	-4.60	-5.21	-4.42	-2.93
kinesin-like protein unc-104	GSSPFG00016132001-RA	-2.55	-6.79	-6.69	-3.06	-3.63	-2.44

kinesin-like protein unc-104	GSSPFG00003562001.1-RA	-1.77	-2.62	-2.38	-2.74	-5.35	-2.30
chromosome-associated kinesin KIF4	GSSPFG00013442001.1-RA	-1.19	-1.26	-1.06	-1.13	1.53	2.18
myosin heavy chain, muscle	GSSPFG00021119001.1-RA	-1.62	-3.23	-2.90	-3.16	-3.87	-1.96
myosin heavy chain, non-muscle	GSSPFG00012506001-RA	-1.71	-1.98	-1.68	-1.96	-2.12	-1.28
myosin heavy chain, non-muscle	GSSPFG00023985001-RA	-1.84	-1.99	-1.63	-2.05	-2.31	-1.47
myosin-I heavy chain	GSSPFG00013835001-RA	-2.21	-3.78	-2.50	-2.99	-3.75	-1.90
myosin-I heavy chain	GSSPFG00029577001-RA	-2.48	-1.06	1.06	-2.07	-4.00	-1.64
myosin-IA	GSSPFG00023219001-RA	-1.08	-1.21	-1.20	-1.31	-2.04	-1.61
myosin-IIIb-like	GSSPFG00007293001.1-RA	-1.36	-1.61	-1.92	-2.75	-4.23	-2.81
myosin-VIIa	GSSPFG00013168001-RA	-1.20	1.28	1.99	2.24	3.24	2.90
tubulin alpha-1 chain-like	GSSPFG00018276001-RA	-1.09	-7.76	-2.79	-2.86	-3.39	-4.01
tubulin gamma-2 chain-like	GSSPFG00021817001-RA	-1.02	1.08	1.22	1.17	1.88	2.37
tubulin polyglutamylase TTLL4-like	GSSPFG00010788001.1-RA	-1.16	-1.20	-1.25	-1.56	-2.17	-1.78
tubulin polyglutamylase TTLL5	GSSPFG00011072001-RA	-1.51	-1.61	-1.22	-2.15	-4.75	-2.30
tubulin-folding cofactor B	GSSPFG00021985001-RA	1.50	1.47	-1.01	1.07	-1.49	-31.49
laminin subunit alpha	GSSPFG00008159001-RA	-1.91	-3.25	-2.74	-2.92	-3.34	-2.17
Titin-homolog	GSSPFG00022949001-RA	-2.20	-3.63	-2.39	-2.53	-4.53	-2.84
filamin-A	GSSPFG00029000001-RA	-1.45	-2.50	-3.36	-4.59	-4.13	-2.17
dystonin	GSSPFG00000082001-RA	-2.27	-3.59	-2.62	-3.83	-3.56	-1.94

* Gene ID's used in this study can be accessed through: the SfruDB Information system http://bipaa.genouest.org/is/lepidodb/spodoptera_frugiperda/ (14).

** The values in this case were numerically high because the expression level for the gene changed from 0.065 TPM at 0 h into 7.075 and 7.32 TPM at 24 and 48 hpi, respectively.

*** Red and blue font refers to upregulated and downregulated genes, respectively.

TABLE 3.3: Changes in *Spodoptera frugiperda* 3rd instar larval innate immunity gene levels at different time points post-infection with Spodoptera frugiperda ascovirus-1a.

Innate immunity gene ID/Description*	6 h	12 h	24 h	48 h	4th day	7th day
Toll pathway						
GSSPFG00035666001.2-RA/Cactus	-1.29	-1.79	-1.40	3.41	2.55	3.58
GSSPFG00001145001.2-RA/Cactus	-1.18	-1.72	-1.44	3.27	2.60	3.49
GSSPFG00023708001.3-RA/Pelle	1.15	-1.03	1.06	2.63	1.97	3.09
GSSPFG00035688001.2-RA/Pellino	-1.18	-1.71	-1.41	2.21	1.76	2.75
GSSPFG00035684001.3-RA/Tollip	-1.69	-2.04	-1.73	1.38	1.35	1.59
GSSPFG00025680001.3-RA/Putative Dorsal	-1.55	-1.70	-1.85	-1.57	-2.12	-1.54
GSSPFG00006301001.3-RA/ Spätzle	1.14	-1.31	-1.04	2.79	3.70	4.06
GSSPFG00006187001.6-RA/HP8 (Spätzle-processing enzyme)	-1.35	-1.03	-1.29	2.13	2.29	4.63
IMD negative regulators						
GSSPFG00006593001.3-RA/POSH	-1.87	-2.17	-1.57	-1.86	-1.19	-1.16
GSSPFG00003397001.3-RA/Sick	-2.18	-2.67	-3.02	-1.36	-1.54	-1.08
GSSPFG00025395001.3-RA/Slimb	-1.04	1.20	1.57	1.60	2.10	2.42
GSSPFG00016134001.3-RA/DNR1	-1.58	-1.72	-2.09	1.01	-1.21	1.21
Antimicrobial Peptides						
GSSPFG00015239001.3-RA/Attacin A2	1.71	2.41	-2.24	1.07	-3.15	-14.53
GSSPFG00024578001.4-RA/Attacin B1	1.15	1.89	-2.94	-2.91	-2.58	-1.67
GSSPFG00006778001.3-RA/Cecropin A	2.03	5.53	-1.31	1.86	1.06	2.79
GSSPFG00030460001.5-RA/Cecropin B1	1.84	7.32	-1.82	-1.83	-1.77	1.66
GSSPFG00030459001.3-RA/Cecropin B2	1.16	4.46	-2.75	-2.91	-2.43	-1.04
GSSPFG00030458001.5-RA/Cecropin B3	1.60	6.87	-4.56	-3.71	-4.91	1.32
GSSPFG00030456001.4-RA/Cecropin D	-1.42	-1.15	-1.31	3.05	-1.03	10.94
GSSPFG00005384001.6-RA/Defensin-like6	-1.22	1.54	-1.05	2.21	4.07	4.85
GSSPFG00026948001.3-RA/Defensin	-1.35	-3.08	-2.02	-1.46	2.73	9.25
GSSPFG00011908001.6-RA/Gallerimycin	1.73	-1.54	-1.20	2.74	1.27	3.75

GSSPFG00003522001.3-RA/Gloverin-3	-2.62	3.89	2.06	20.65	13.01	32.56
GSSPFG00013851001.4-RA/Lebocin-1	2.16	3.39	-1.56	5.84	5.34	17.02
GSSPFG00035421001.3-RA/Lebocin-2	2.47	-1.36	1.34	5.37	9.71	16.88
GSSPFG00005939001.3-RA/Lebocin	-1.18	1.53	1.00	28.23	-4.20	-2.34
GSSPFG00009284001.2-RA/LLP1	-2.10	-5.65	-6.76	-2.81	-2.61	-4.46
GSSPFG00002403001.3-RA/Lys1	-1.12	1.94	-1.92	-1.60	-1.11	3.10
GSSPFG00014909001.3-RA/Lys2	-1.07	2.23	-1.34	-1.31	-1.43	5.04
GSSPFG00013328001.3-RA/Lys3	4.43	1.24	-4.82	-1.49	-1.13	1.30
GSSPFG00035992001.2-RA/Moricin-1	3.51	5.18	12.85	23.58	45.55	33.98
GSSPFG00035990001.2-RA/Moricin-7	4.37	7.12	14.44	19.33	50.28	25.25
GSSPFG00030474001.1-RA/Spod_x_tox	2.36	1.64	-1.75	10.59	5.48	21.82
GSSPFG00008360001.3-RA/Spodoptericin	-1.02	1.09	1.73	6.67	5.08	7.72
Transmembrane receptors						
GSSPFG00002917001.2-RA/Dscam	-3.26	-3.28	-1.74	-1.74	1.26	2.86
GSSPFG00035226001.3-RA/Eater-partial	-1.64	-1.80	-1.93	-1.64	3.15	4.64
GSSPFG00005872001.3-RA/Eater-partial	-1.43	-1.60	-1.83	-1.46	3.57	4.62
GSSPFG00009711001.2-RA/Nimrod	1.33	5.08	1.43	4.87	4.06	7.24
GSSPFG00018470001.2-RA/p77	-3.63	-2.75	-1.51	1.02	2.96	5.47
GSSPFG00016871001.3-RA/SR-B3	1.62	1.33	1.24	1.41	2.35	3.33
GSSPFG00035217001.3-RA/SR-B3	1.61	1.19	1.11	1.22	2.51	3.98
GSSPFG00003331001.3-RA/SR-B4	-1.27	-1.40	-1.07	1.95	1.64	2.47
GSSPFG00022245001.4-RA/SR-C1	1.46	1.13	-1.25	1.94	3.84	11.21
Extra cellular signal transduction and cytokines						
GSSPFG00015829001.3-RA/Growth Blocking Peptide Binding Protein	1.41	1.50	1.22	-8.03	-9.65	-2.79
GSSPFG00016383001.3-RA/Growth Blocking Peptide	1.57	2.73	1.14	-1.09	1.37	-1.09
GSSPFG00035416001.3-RA/Hdd23	-1.86	-1.19	-1.36	2.51	2.92	4.03
GSSPFG00027451001.3-RA/Hdd23	-1.19	-1.08	1.45	1.96	9.63	17.99
GSSPFG00035032001.3-RA/Hdd23	-1.64	-1.83	-1.81	1.71	2.01	3.35
JNK pathway						

GSSPFG00007437001.3-RA/kayak	-1.86	-2.02	-1.59	1.31	1.72	3.19
GSSPFG00031472001.3-RA/Kayak	-1.52	-2.16	-2.23	-1.18	-1.37	-1.08
JAK/STAT pathway						
GSSPFG00032684001.3-RA/Domless	-1.66	-3.43	-3.26	1.03	-1.71	1.61
GSSPFG00003070001.3-RA/STAT	-1.66	-2.53	-2.65	-1.01	1.03	1.78
GSSPFG00011397001.3-RA/SOCS	1.05	1.29	1.21	2.85	2.11	2.25
GSSPFG00027455001.3-RA/hopscotch	-1.91	-1.61	-1.55	1.45	1.97	2.84
Peptidoglycan recognition proteins (PGRP)						
GSSPFG00006083001.5-RA/PGRP	-1.36	2.11	1.29	-1.55	1.71	1.79
GSSPFG00006046001.5-RA/PGRP	-2.77	-1.49	-12.14	-1.08	1.75	3.39
GSSPFG00019359001.4-RA/PGRP	-2.62	-1.48	-1.48	-1.25	1.66	-1.09
GSSPFG00035260001.5-RA/PGRP	1.28	2.30	-2.36	1.09	1.47	2.81
GSSPFG00017964001.4-RA/PGRP	1.60	3.93	16.03	1.67	11.83	15.97
GSSPFG00026857001.5-RA/PGRP	-1.59	-1.06	-3.98	1.24	1.79	3.28
Phenoloxidase system						
GSSPFG00026064001.3-RA/Phenoloxidase Inhibitor	-1.06	-1.18	-1.81	1.55	6.11	15.75
GSSPFG00027865001.3-RA/PPAE1	-1.07	-1.10	-1.33	1.68	1.88	2.07
GSSPFG00023419001.3-RA/PPAE2	-1.54	-1.95	-2.55	-1.44	-1.34	1.00
GSSPFG00013976001.3-RA/PPO2	-1.01	-1.17	1.02	-10.24	-20.83	-12.74
GSSPFG00012369001-RA/PPO1	-1.79	-2.10	-1.24	-14.11	-25.07	-12.48
GSSPFG00020775001.3-RA/HP6	-1.01	1.27	1.26	2.16	2.37	1.49
Thioester containing protein						
GSSPFG00029506001.3-RA/TEP3	-1.71	-2.23	-1.83	-1.47	-1.43	-1.24
Gram-negative bacteria binding proteins (GNBP)						
GSSPFG00000148001-RA/GNBP	-1.02	2.00	-1.36	1.58	2.12	2.17
GSSPFG00022027001-RA/GNBP	-1.04	1.70	-1.18	1.34	1.93	2.18

*The *Spodoptera frugiperda* innate immunity genes used in this study are listed in 14 and 15. Only genes that were upregulated or downregulated by more than 2-fold at any time-point post-infection are listed. Gene identifications can be accessed through: the SfruDB Information system http://bipaa.genouest. org/is/lepidodb/spodoptera_frugiperda/.

**Red and blue font refers to upregulated and downregulated genes, respectively.

Chapter IV: Trichoplusia ni Ascovirus: Transcriptome Analysis of Hemolymph Viral Vesicles for a Virus with a Broad Tissue Tropism

4.1 Abstract

Ascoviruses are large double-stranded DNA viruses that primarily attack larvae of the lepidopteran family Noctuidae. They produce large, enveloped virions, 130 x 300 - 400 nm in length, and are transmitted on the ovipositor of parasitic wasps as they lay their eggs. The most unique feature of ascoviruses, however, is the changes in cell architecture they control during replication. After infecting a cell, the nucleus is lysed in a process resembling apoptosis, after which the cell enlarges as much as ten-fold. During this process, the mitochondria are not destroyed but rather participate in cleaving the cell into numerous viral vesicles, from 20 – 30 or more, which accumulate in the hemolymph where most progeny virions are formed. In the present study, I used high-throughput genome sequencing to determine the core genes of the Trichoplusia ni ascovirus, which has a broad tissue tropism, and their pattern of expression during the first seven days of viral replication using RNA-sequencing. Specifically, I analyzed the mRNA isolated from hemolymph over the first seven days post-infection (pi), during which this tissue became packed with viral vesicles, and compared expression primarily of the viral core genes to those of other major larval tissues, specifically the fat body, epidermis, and tracheal matrix (FET) tissues. At 48 hpi, only 26 genes were expressed in the later tissues, at a level equal to or more than 5 $\log^2$ RPKM, demonstrating infection was limited. In the hemolymph viral vesicles, however, 48 genes were expressed at a similar level at the same time point. This increase in viral transcripts from hemolymph viral vesicles indicates a high level of viral replication early

during ascovirus disease. Two viral lipid-modifying enzymes, a fatty acid elongase and a patatin-like phospholipase, were highly expressed in vesicles at 48 hpi and 7 days pi, 10.15 $\log^2$ and 8.13 $\log^2$ RPKM, respectively, implying their importance for assembly of virion progeny, and possibly the increase in vesicle size. In addition, I detected bidirectional transcription and five common bicistronic and tricistronic mRNA messages in all tissues, indicting conservation of these phenomena in ascoviruses, and possible regulatory role in transcription or translation.

4.2 Introduction

The family *Ascoviridae* consists of a small family of ds DNA viruses that primarily attack larvae of the lepidopteran family Noctuidae. These viruses produce large, enveloped virions (150 x 400 nm), with genomes ranging from 120 – 180 kbp. Depending on the species the virions vary from being bacilliform to bean-shaped, and have a distinctive reticulate surface structure when visualized in negatively stained electron microscope preparations. Ascoviruses have two unusual characteristics compared to most other large DNA viruses. First, they are poorly infectious by feeding, typically being transmitted in nature on the ovipositor of parasitic wasps when these probe host larvae for laying eggs (1-7). Second, viral replication results in unique changes in cellular architecture marked by nuclear lysis, enormous cell hypertrophy, followed by cleavage of the cell into clusters of 20 or more anucleate viral vesicles in which most virion progeny are produced. Ultrastructural and more recent transcriptome studies indicate that ascoviruses manipulate host cell mitochondria to provide the energy required for synthesizing the membranes that

delimit viral vesicles (8, 9). After formation, these vesicles accumulate in the larval hemolymph, turning it milky white, and continue to assemble progeny virions. The name "ascoviruses" is derived from this characteristic of forming vesicles, i.e., "asco" in Greek meaning "sac".

The *Trichoplusia ni ascovirus* (TnAV-2) is one of four ascovirus species recognized. The name is derived from the name of host from which the first isolate was obtained. The type species, the *Spodoptera frugiperda ascovirus* (SfAV-1a), primarily attacks the fat body tissue and appears to only replicate in species of the genus *Spodoptera*. TnAV-2, however, has a wide host range including species belonging to at least several genera. Moreover, this virus has a broad tissue tropism, the major tissues attacked being the epidermis, tracheal matrix, and fat body. While many cells are infected in each, virus replication rarely destroys more than fifty percent of these tissues ten days after infection (3, 4, 10). Of interest for the current study is that the TnAV genome is 174,059 bp, encoding 165 ORFs (11), about 11% larger than the SfAV genome, which is 156,922 bp and encodes 123 ORFs (12). The difference in the genome size and encoding capacity raises the question of whether the additional proteins of TnAV are responsible for the wider tissue tropism of this virus, since both viruses have the same cytopathology.

In a recent study of SfAV-1a transcriptome, I showed that the 44 core genes, i.e., those common in the two ascovirus genera, *Ascovirus* and *Toursvirus*, could be assigned to three temporal classes during expression; early, late, and very late (Chapter 2). Because no suitable cell system exists for studying transcription *in vitro* (13), I determined the transcription profile through *in vivo* infection of *Spodoptera frugiperda*

3rd-instar caterpillars. In the present study, I used the same approach to analyze the transcriptome of a variant of TnAV-6a (previously TnAV-2c; 5, 11). Referred to hereafter as TnAV-6a1. Early experiments showed that this ascovirus initiated infection of host tissues quickly, and numerous virion-containing vesicles, i.e., viral vesicles, began accumulating in the hemolymph within 48 hours of infection, whereas most cells in the infected tissues appeared to remain healthy, consistent with previous studies of TnAV-2 histopathology (3). Thus, in the present study, I compared the transcriptome of TnAV-6a1 viral vesicles circulating in hemolymph, where virus replication and the generation of progeny virions was more extensive compared to that apparent in the above major host tissues. I targeted ascovirus genus core genes (i.e. those shared between the members of ascovirus genus namely, SfAV, TnAV and HvAV) and a few ORFs that occur only in HvAV or are unique to the TnAV species.

My comparison illustrated the following: **First**, the infection in the host **F**at body, **E**pidermis and **T**racheal matrix (FET) tissues during 6 hpi through 48 hpi can be described as partial or limited. **Second**, the enormous increase in transcripts in the hemolymph 48 hours after infection demonstrates these are critical to virus replication and synthesis of progeny virions, in essence serving as anucleate minicells for replication. **Third**, the early and high expression of a TnAV caspase-2-like gene in all tissues indicates it is most likely not associated with apoptosis induction, which occurs in the ascovirus type species, SfAV-1a, but is nevertheless essential for replication, as reported in HvAV (14). **Fourth**, lipid-modifying enzymes are highly expressed in vesicles, suggesting these play an important role in vesicle growth and virion synthesis. **Finally**,

the detection of bicistronic and tricistronic messages and bidirectional transcripts in the TnAV transcriptome, reported previously for certain SfAV-1a transcripts, shows these atypical mRNAs occur in other ascovirus species, and may be important in regulating transcription and translation.

4.3 Materials and methods

4.3.1 TnAV ascovirus and host infection

For all infections, I used a minor variant of the Trichoplusia ni ascovirus 6a (referred to hereafter as TnAV-6a1) and larvae of the cabbage looper, *Trichoplusia ni*. TnAV-6a was previously known as TnAV-2c (5, 11). For studies of the fat body, epidermis and tracheal matrix, early 3rd-instars were infected with TnAV-6a1 and tissues were harvested from infected larvae at 6, 12, 24 and 48 hours post-infection (hpi), when larvae were late third or early fourth instars. The same tissues were collected from control *T. ni* larvae at 0, 24 and 48 h. Viral vesicles liberated from infected tissues begin to appear in the hemolymph 24 – 48 hpi, and increase in concentration to 10^7 or more per ml by 7 days pi, far outnumbering hemocytes. Thus, to analyze transcripts from viral vesicles, hemolymph was collected from infected larvae at 6, 12, 24, 48 and 168 (7 days) hpi. For uninfected *T. ni* control larvae, hemolymph was harvested at 0 h. All transcriptomic data were derived from these time points post-infection. Infected and control larvae were reared on artificial diet (Benzon Research, Carlisle, PA) and kept at room temperature (22^oC). To infect larvae, a minutin pin was contaminated with TnAV-6a1 virions by dipping it into a suspension of viral vesicles (10^8/vesicles/ml). Control larvae were pierced with a pin

dipped in phosphate buffered saline (PBS). The pins were inserted into larval abdomen, which mimicked parasitic wasp infection in nature. The progression of infection was monitored daily by the color of the hemolymph, which changed from translucent light green to opaque white as viral vesicles accumulated in the hemolymph. In addition, the infected hemolymph was examined by phase microscopy to insure the vesicles were characteristic in size and shape as the disease progressed.

4.3.2 Isolation of RNA from hemolymph

Total RNA was isolated from each infected tissue and control at six time points, at 0, 6, 12, 24, 48, 168 hpi. For each hemolymph sample and control, 0.9 ml TRIzol reagent (Thermo Fisher Scientific) was combined with 0.1 ml of blood collected from the larva and held on ice and quickly mixed by vortexing for 30 seconds. Chloroform was added for RNA separation in the aqueous phase followed by RNA precipitation with 100% isopropanol. Finally, precipitated RNA was washed twice in 75% ethanol. RNA quality and quantity monitored using a Thermo Scientific NanoDrop ND-2000c Spectrophotometer.

4.3.3 Isolation of RNA from fat body, tracheal matrix and epidermis

Total RNA was collected from two larval replicates for each of five time points prior to and after infection for both infected and control larvae. For infected larvae, the time points were 6, 12, 24, and 48 hpi, and for the BPS controls, 0, 24 and 48 h. Each 3rd-instar larva was dissected on a wax plate surface-sterilized with 95% ethanol. The head and posterior body segment were cut away with a razor, and then the alimentary canal

(foregut, midgut, and hindgut) was removed from the body using dissecting forceps. The fat body, tracheal matrix and epidermis in both infected and control larvae, were then homogenized in 1 ml TRIzol reagent per sample by mechanical trituration, followed by addition of chloroform and isopropanol as above. RNA quantity and quality were monitored using the Thermo Scientific NanoDrop ND-2000c Spectrophotometer.

4.3.4 RNA purification and DNA removal

For each sample, 5 μg was used as the starting material for purifying and concentrating the RNA using an RNA Clean and Concentrator TM-5 kit (ZYMO RESEARCH). The RNase-Free DNase Set (QIAGEN) was used for DNA removal following the in-column DNase digestion protocol. The quantity and quality of purified RNA were determined as described above. Complete removal of DNA was verified by RT-PCR reactions in which no Reverse Transcriptase was included.

4.3.5 mRNA isolation, cDNA library preparation and sequencing

After the DNase-treatment of total RNA, the mRNA was isolated from each sample using the NEB NextPoly(A) mRNA magnetic isolation module kit (New England Biolabs). The poly(A) RNA was then eluted from the oligo $d(T)_{25}$ attached to paramagnetic beads in 15 μl of the first cDNA strand synthesis reaction buffer combined with random primer mix (2x), followed by heating for 10 min at 94°C. The cDNA libraries were prepared following the instructions of the manufacturer of NEBNext® Ultra™ Directional RNA Library Prep Kit (New England Biolabs) for Illumina®. For each time point in the two

tissues under study I sequenced two biological replicates. Agilent 2100 Bioanalyzer was used to ensure the quality of the pooled libraries before sequencing using the HiSeq2500 sequencer (Illumina) in Core Facility at the UCR Institute for Integrative Genome Biology.

4.3.6 TnAV-6a1 variant genome sequencing and contig assembly

Viral DNA was isolated from the TnAV-6a1-infected *Trichoplusia ni* hemolymph following the protocol used by (15, 16). Purified viral DNA was fragmented with a Bioruptor (Diagenode, PA, USA) and was used in genomic DNA library preparation following manufacturer instructions of the NEBNext DNA Library Prep Master Mix Set for Illumina (New England BioLabs). This was followed by library quantity and quality checks using the Qubit 2.0 Fluorometer (Invitrogen, Life Technologies) and Agilent 2100 Bioanalyzer. Resulting library was sequenced using the Illumina MiSeq (2 x 300, Paired-end) at the UCR Institute for Integrative Genome Biology Core Facility. For genome assembly, I used the TnAV-6a genome available in the NCBI database (Accession number: DQ517337.1) as a reference genome. Thirty-four contig sequences were generated using Trinity software with "genome guided" assembly option (17)

4.3.7 TnAV-6a1 variant Reads Per Kilobase Per Million (RPKM) quantification in hemolymph viral vesicles compared to FET tissues

The fat body, epidermis and tracheal matrix tissues and hemolymph (vesicle fraction) RNA-Seq reads were mapped to the TnAV-6a1-variant assembled contigs by using the

bowtie2 (18). The Reads Per Kilobase Per Million mapped reads (RPKMs) were calculated for each gene using custom written scripts, only reads sequenced in the same orientation as the gene ORF were considered. The position of core genes (i.e. those shared between the members of ascovirus genus namely, SfAV, TnAV and HvAV) and TnAV and/or HvAV-species specific ORFs identified using the NCBI ORF Finder application (https://www.ncbi.nlm.nih.gov/orffinder/), in contig sequences are listed in **Table 4.2**. The ORF number of the 70 genes listed in **Table 4.1**, is derived from the TnAV-6a reference genome (Accession number: DQ517337.1).

4.3.8 RNA-Seq data validation

The qRT-PCR was used to quantify the expression of three randomly selected virus genes namely (homologs of- ORF029, ORF042 and ORF077). I used total RNA samples derived from the FET tissues (i.e. partially infected tissues) for this analysis to facilitate the identification of a host stable gene during the infection. For data normalization, I used the host eukaryotic Initiation Factor-4a gene (*eIF4*a; 19). In addition, I designed a set of primers (**Table 4.1**) using the Integrated DNA Technologies (IDT) Real Time PCR design tool, to amplify the three genes above. For efficiency estimates, standard curves were generated for each primer. Briefly, 1 µg of DNAse-treated RNA was reverse transcribed following the instruction of the Maxima First Strand cDNA Synthesis kit for RT-qPCR (Thermo Scientific). For Real-time PCR reaction, I used the Luna Universal qPCR Master Mix (New England Biolabs). One ul of ten-fold diluted cDNA in each 20 ul reaction volume. The CFX Connect TM Real-Time PCR Detection System (BIORAD)

was used following this program for the 4 genes: 95 °C for 3 min, 40 cycles of 90 °C for 10 sec and 55 °C for 30 sec followed by 72 °C for 30 sec. For reach reaction, melting curve was generated, two technical and three biological replicates were examined. Negative RT-PCR and no-template controls were included. Pfaffl equation (20) was used for quantification and 6 hpi sample was used as the reference.

4.3.9 RT-PCR validation of bicistronic and polycistronic mRNA messages and intergenic region analysis

Three bicistronic TnAV-6a1 messages with intergenic regions identified *in silico* using IRESPred web server, available at http://bioinfo.net.in/IRESPred/, to possess a putative IRES sequence (**Table 4.3**) were further validated. The RT-PCR product of a DNase-treated fat body, epidermis and tracheal matrix (FET) tissues total RNA (24 hpi sample) was used as a template for all reactions. A negative control for all RT-PCR reactions (where RT enzyme was omitted) was performed to confirm the absence of DNA contamination. For reverse transcriptase PCR, I used the Maxima First Strand cDNA Synthesis kit for RT-qPCR (Thermo Scientific), followed by PCR using specific primers designed to amplify either the junction region only between the two ORFs encoded in the same message, or amplify almost the full-size message (**Table 4.1**). For amplification, the Q5® High-Fidelity 2X Master Mix kit (New England Biolabs) was used. The PCR product size and sequence were confirmed using 1% agarose gel and Sanger sequence, respectively. The following program was used for PCR amplification: 98°C for 30 sec,

followed by 30 cycles of 98°C for 10 sec and annealing at 60-61°C for 20 sec and extension at 72°C for 30-70 sec, and a final extension at 72°C for 2 min.

4.4 Results

4.4.1 TnAV-6a1 variant genome assembly and genes annotation

Initial mapping of viral RNA-Seq reads derived from different infected insect tissues (hemolymph viral vesicles or FET tissues, the combination of fat body, epidermis, and tracheal matrix) to the published TnAV-6a (TnAV-2c previously; 11) genome using the program Bowtie2 (18), showed that our TnAV strain was a slightly different variant closely related to the published TnAV-6a strain (NCBI, accession number DQ517337.1). Our TnAV-6a variant (TnAV-6a1) high-throughput DNA genome sequencing resulted in 3,075,028 virus specific reads. The assembly of the virus using the TnAV-6a (Accession number: DQ517337.1) as a reference genome generated 34 contigs. Sixty of the ascovirus genus core ORFs were identified and compared to their homologues in TnAV-6a and TnAV-6b (21) in order to define the ORF number and expected size (**Table 4.1**). In addition, ten ORFs that occur only in HvAV or are unique to the TnAV species were identified in the assembled sequence of the contigs and added to the gene list. I refer to the genes identified in the assembled contigs by their ORF order in the TnAV-6a variant (11), **Table 4.1**. Our sequences can be accessed through (GSE114902).

4.4.2 Expression patterns of TnAV-6a1 genes in the combination of fat body, epidermis and tracheal matrix tissues compared to hemolymph viral vesicles

TnAV is known to have a broad tissue tropism. Therefore, most larval tissues present in the Fat body, Epidermis and Tracheal Matrix (FET) tissue samples were infected with TnAV-6a1. I was able to detect the expression of most genes (**Table 4.2**) by 6 hpi in FET tissues. However, the Reads Per Kilobase Per Million (RPKM) values did not exceed ≥5 $\log^2$ RPKM. For example, of the 70 genes included in my study, at 6, 12, 24 and 48 hpi, there were only 6, 6, 14 and 26 genes expressed, respectively, at the above rate (**Fig. 4.1**). These results imply that the infection start initially in FET tissues but due to either host or virus constraints the increase in virus transcripts above a specific level was limited. This agrees with previous studies of ascovirus histopathology showing TnAV infection of these tissues is limited in *Trichoplusia ni* at least during the first week of infection.

Based on ascoviruses pathogenesis, viral vesicles accumulate in the hemolymph as the basement membrane of infected tissues degenerate and rupture (1, 3). As more and more vesicles accumulate and circulate in the hemolymph it becomes milky white. In general, from 6-24 hpi the TnAV gene expression pattern in the hemolymph vesicle fraction was low, ranging from 0 to 3.24 $\log^2$ RPKM. Subsequently, beginning at 48 hpi, TnAV transcripts began to increase markedly after more and more vesicles accumulated in the hemolymph. Specifically, at 6, 12, 24 and 48 hpi there were 0, 0, 0 and 48 genes expressed at ≥5 $\log^2$ RPKM, respectively (**Fig. 4.2**). This marked increase in TnAV hemolymph viral vesicle transcripts continued through to 7th day post-infection, demonstrating ongoing viral replication in these anucleate vesicles.

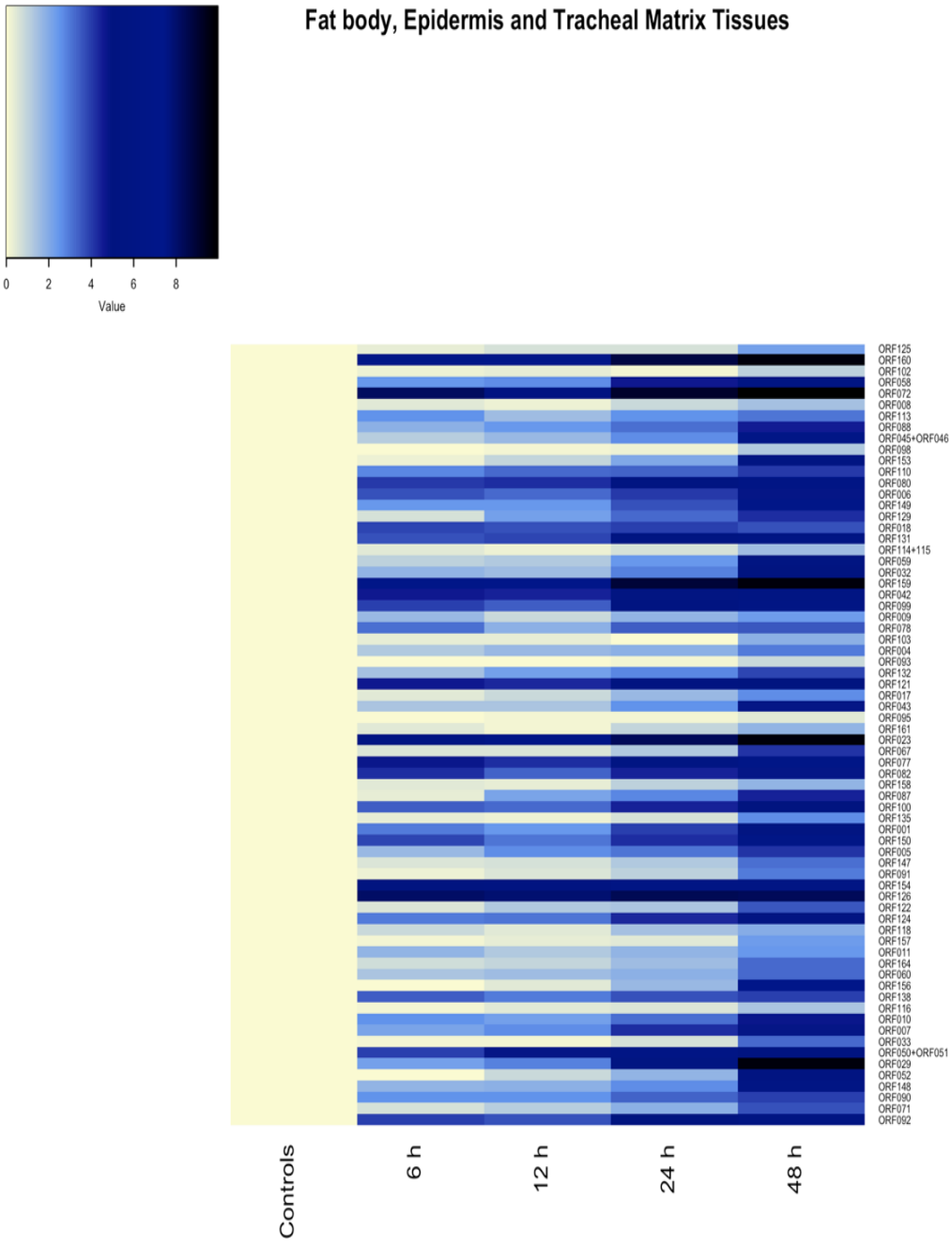


FIG 4.1 Heat-map representation of TnAV temporal expression trend for TnAV (core and species-specific) genes in Fat body, Epidermis, Tracheal matrix (FET) tissues. The color scale represents the $\log^2$ RPKM value from average replicate expression level at 0, 24 and 48 hrs. The ORFs number for our TnAV-variant strain is derived from the ORFs order in TnAV-6a genome (11). The plus sign refers to two ORFs that were found in our variant strain as a single ORF.

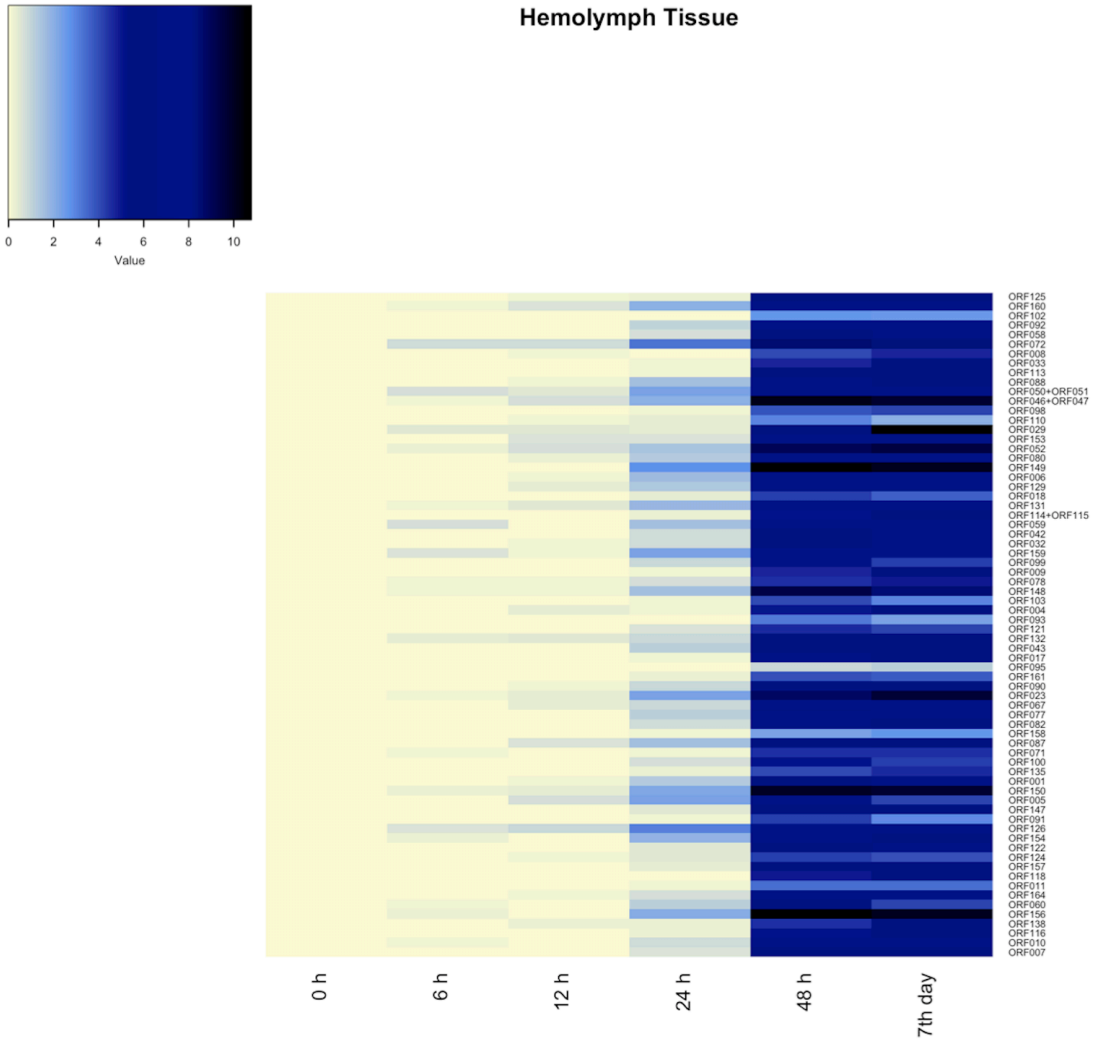


FIG 4.2 Heat-map representation of TnAV temporal expression trend for TnAV (core and species-specific) genes in hemolymph viral vesicles. The color scale represents the $\log^2$ RPKM value from average replicate expression levels at 0h. The ORF numbers for our TnAV variant is derived from the order of ORFs in the TnAV-6a genome (11). The plus sign refers to two ORFs found in our minor TnAV variant strain.

4.4.3 Highly expressed genes in Fat body, Epidermis and Tracheal matrix (FET) tissues compared to hemolymph viral vesicles

The expression patterns of FET tissues and hemolymph viral vesicles for the 70 TnAV-6a1 genes (**Fig 4.1, Fig 4.2**) studied revealed three main levels of expression; low (from less than 1 to 5 $\log^2$ RPKM), medium (from >5 to 8 $\log^2$ RPKM), or high **(**more than 8 to ~11 $\log^2$ RPKM). At 48 hpi, in FET tissues, 44 genes were low, 20 were medium and 6 were high in terms of their expression. In vesicles, at the same time point, 22 genes were low, 37 were medium and 11 were high. Therefore, the expression of viral genes in FET tissues was low to medium. Alternatively, the expression level of TnAV-6a1 genes in the hemolymph viral vesicles was medium to high.

Genes that were the most highly expressed (>8 $\log^2$ RPKM) in both sets of tissues were the TnAV-6a homologs for a caspase-2-like enzyme (ORF072), and two hypothetical proteins (ORF023 and ORF159). On the other hand, several highly expressed genes were unique in the vesicles (**Fig 4.2**). For example, two lipid-modifying enzymes namely, a fatty acid elongase combined with a protein responsible for elongation of very long chain fatty acids (ORF046+ORF047), and a patatin-like phospholipase (ORF067) were highly expressed. The plus sign in the former refers to two ORFs that occurred as a single ORF in our variant strain. Notably, in the TnAV6b variant (21) this long protein also occurs in the longer ORF. Two other lipid-modifying enzymes, phosphate acyltransferase (ORF098) and a lipase (ORF132) were also more highly expressed in the vesicles compared to the FET tissues. The lipase (ORF132) was expressed at 48 hpi in FET and vesicles by 3.87 and 5.77 $\log^2$ RPKM, respectively. The

phosphate acyltransferase (ORF098) was expressed at 48 hpi in FET and vesicles by 1.29 and 3.80, respectively. In addition, two structural proteins, a Yabby-like transcription factor (ORF059) and a Dynein-like β chain (ORF043) and three hypothetical proteins (ORF149, ORF150, ORF156) were expressed at levels >8 $\log^2$ RPKM at 48 hpi and/or 7 days pi. In FET tissues, two hypothetical proteins (ORF160 and ORF126) were highly expressed by 9.72 and 8.30 $\log^2$ RPKM, respectively, levels comparable to those in vesicles, by 7.91 and 7.95 $\log^2$ RPKM, respectively.

Interestingly, the TnAV-6a homolog (ORF072; caspase-2-like protein) was detected at 6 hpi by 0.81 and 8.24 $\log^2$ RPKM in vesicles and FET tissues, respectively (**Fig. 4.3**). Moreover, the expression level increased very markedly, especially in vesicles, after 2 days post-infection, to levels of 8.52 and 9.97 $\log^2$ RPKM, respectively, in vesicles and FET tissues.

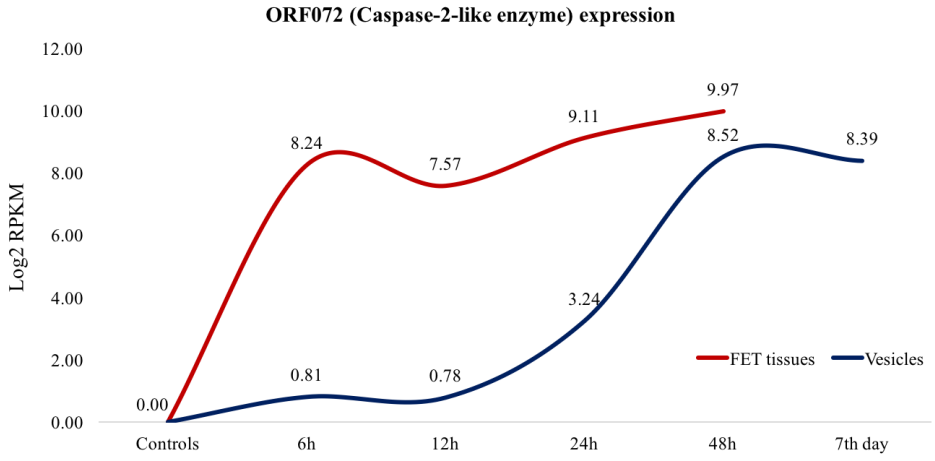


FIG 4.3 Early and high expression levels for TnAV-6a homolog ORF072, a caspase-2-like enzyme, in hemolymph viral vesicles compared to expression of the same gene in the combination of fat body, epidermis, and tracheal matrix tissues of *Trichoplusia ni.*

4.4.4 TnAV and/or HvAV species-specific genes

In addition to ascovirus core genes, several genes identified in our TnAV contigs were not conserved in SfAV (i.e., unique to TnAV or HvAV). These genes include, ORF029, ORF033, ORF050+ORF051, ORF052, ORF071, ORF090, ORF092, ORF148, ORF132 and ORF147 (**Table 4.1**). Only three of these were identified as highly expressed genes in the viral vesicles (>8 $\log^2$ RPKM at 48hrpi and/or 7 days pi), namely an aegerolysin gene (ORF029) and two hypothetical proteins (ORF052 and ORF148). In FET tissues, only the aegerolysin gene was identified as one of high expression (9.67 $\log^2$ RPKM at 48 hpi).

4.4.5 Detection of bicistronic and tricistronic mRNA messages

Alignment of viral vesicle and FET TnAV-6a1 RNA-Seq reads to contig sequences revealed these tissues shared five bicistronic and tricistronic transcripts (**Table 4.3; Fig. 4.4**). In addition, some bicistronic and tricistronic transcripts only occurred in vesicles, implying that there might be differential transcriptional regulation in the vesicles that does not occur in the tissues from which the vesicles originate. I recently identified nine bicistronic and six tricistronic messages in *Spodoptera frugiperda* 3rd-instars infected with SfAV-1a (9). The detection of these polycistronic messages in TnAV, therefore, indicates that such transcripts may be a feature of many ascovirus species.

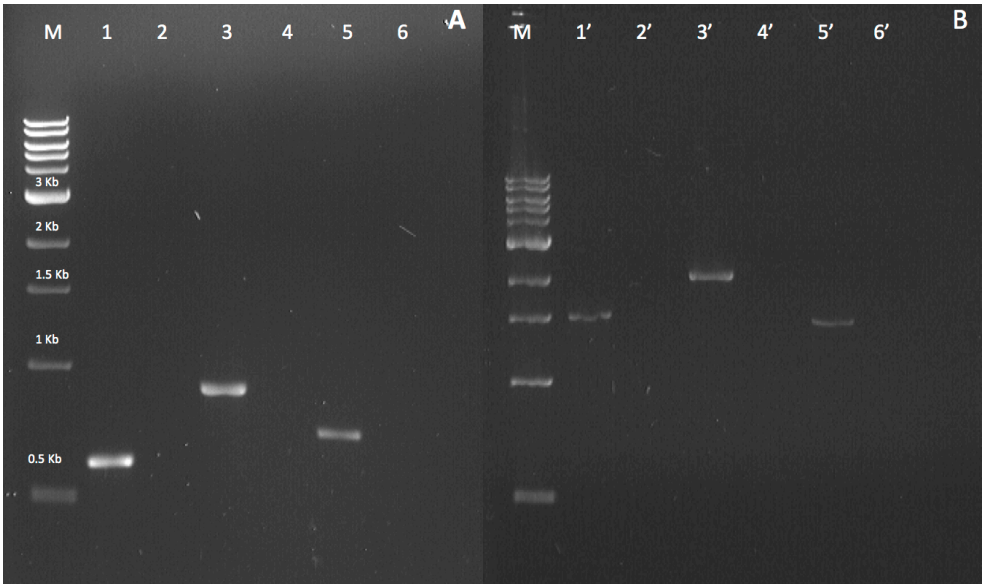
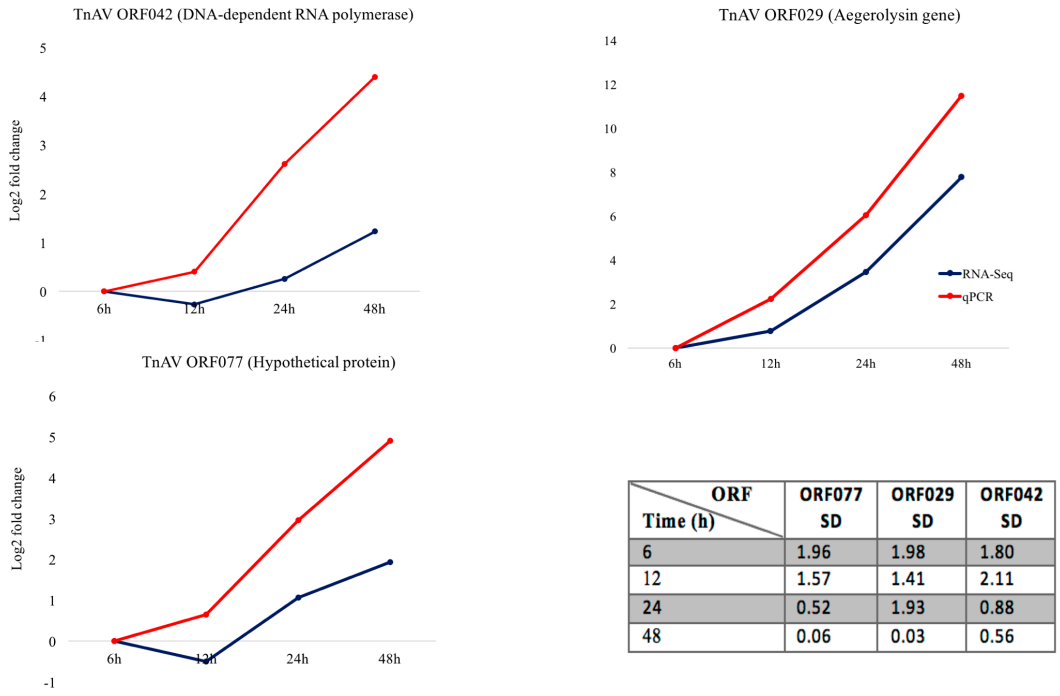


FIG 4.4 RT-PCR validation of three different bicistronic messages identified in the transcriptome of our TnAV strain. **(A)** Junction sequence amplification, and **(B)** Almost full message size amplification for the same three messages. Lane M contains a 1 Kb Marker (M). Lanes 1, 3, and 5 are amplified cDNAs for the junction regions of these bicistronic messages, while 1', 3', and 5' contain almost the full cDNA regions from these messages. All PCR products were sequenced to confirm PCR product identity.

4.4.6 RNA-Seq data validation

The RPKM value estimates for the RNA-Seq data were validated by qRT-PCR. Three TnAV-6a homologs selected randomly, specifically, the genes for aegerolysin (ORF029), DNA-dependent RNA polymerase (ORF042), and a hypothetical protein (ORF077) were quantified by qRT-PCR. The results using each technique were in agreement regarding gene expression patterns. In general, the qRT-PCR estimates were higher than the RNA-Seq, but the expression trend was similar (**Fig. 4.5**). It is possible that differences in the estimates of the mRNA levels between these methods may be due to the differences in the normalization procedure between the two technologies. Moreover, qRT-PCR provides more sensitive expression estimates on single gene level (22, 23).



ORF / Time (h)	ORF077 SD	ORF029 SD	ORF042 SD
6	1.96	1.98	1.80
12	1.57	1.41	2.11
24	0.52	1.93	0.88
48	0.06	0.03	0.56

FIG 4.5 qRT-PCR validation of the RNA-Sequencing data for three TnAV-6a homolog genes (ORF029, ORF042, ORF077). The table illustrates the standard deviation (SD) between three biological replicates examined in each time point.

4.4.7 TnAV-6a1 RNA-Seq data accession number

The TnAV-6a1 variant RNA-Seq data from the FET tissues and viral vesicles along with the RPKM values from different tissues and contig sequences were deposited in the NCBI Gene Expression Omnibus (GEO), and can be accessed through the GEO series accession number (GSE114902).

4.5 Discussion

Ascoviruses constitute a family of novel large DNA insecticidal viruses with the unique ability to reorganize the infected cell cytoskeleton. The outcome of this reorganization after nuclear lysis is the formation of numerous virion-containing vesicles cleaved from each infected which accumulate in the hemolymph as the basement membrane of infected tissues ruptures. The role of these vesicles in the ascovirus life cycle is obscure and many hypotheses have been proposed. For example, the vesicles serve as a tool for virus transmission by circulating in larval hemolymph where they can be acquired on the ovipositor of parasitic wasps during egg-laying, and then transmitted to other lepidopteran larvae. This hypothesis is supported by several studies in which it has been demonstrated that long-term contamination of the ovipositor of parasitic wasps can acquire and transmit ascovirus virions that result in infection (24-27). For example, the infectivity of HvAV-3h on *Microplitis similis* wasp ovipositor to *Spodoptera exigua* larvae lasted for 4.10 ± 1.44 days (27). On the other hand, an immune evasion role for ascovirus vesicles has been proposed because they circulate in the blood for weeks without eliciting an obvious host innate immune response (25, 28, 29). Data support both of the above hypotheses, but my *in vivo* RNA-Seq transcriptome analyses of the Trichoplusia ni ascovirus (TnAV-6a1) suggest another important role for these vesicles – they contribute significantly to the production of progeny virions by absorbing hemolymph nutrients as they circulate in this tissue for several weeks. In support of this hypothesis, in the present study I found a much higher level of expression of more genes in the viral vesicles than in the combination of fat body, epidermis, and tracheal (FET)

tissues from which these vesicles originated. Thus, producing progeny virions must be considered a very important function of viral vesicles once they enter the hemolymph.

According to my RNA-Seq analysis of *Trichoplusia ni* FET tissues after TnAV-6a1 infection, its gene expression patterns are rapid yet limited to being low or medium for most ascovirus genes, at least for the 70 genes I studied, which include most ascovirus core and a few genes found in either other TnAV strains or HvAV species. For example, I was able to detect the immediate expression of many ascovirus genes as early as 6 hpi, in FET tissues, however, RPKM values for most did not exceed ($\log^2$ RPKM = 5) at all the tested time points (**Fig. 4.1**). This demonstrates infection likely begins in these tissues, but spreads slowly during the first days of infection. Generally, TnAV-2 and HVAV-3 isolates do not spread quickly within infected FET tissues during the first 7 – 10 days, whereas the SfAV-1 isolate, which predominantly infects only the fat body, almost completely destroys this tissue within a period of 7 – 12 days (3). Alternatively, although the level of expression in the hemolymph viral vesicles was low during the first 48 hours post-infection, from then until 7 days post-infection TnAV-6a1 expression in the vesicles was high (**Fig. 4.2**), and for 49 genes, suggesting that liberation into the hemolymph, and thus direct access to nutrients there, enhanced viral gene expression for a wider variety of genes.

TnAV-6a1 gene expression patterns in the different infected tissues agree with earlier histopathological studies carried on TnAV infected *Trichoplusia ni* fourth instar larvae that described the TnAV infection in FET tissues as sporadic and partial. For instance, the fat body infected cells were detected in each infected larva but most of the

fat body cells were intact even in advanced stages of the infection (1). Overall, the limited infection of the host tissues can be evolutionary selected for to allow the virus to achieve a long-term chronic infection rather than a severe acute infection, resulting in death in 7 – 10 days, as is typical of most baculoviruses. In SfAV-1 infections, the virus primary infects the fat body, as noted above, almost completely destroying this tissue, but infected larvae can survive for weeks, as the other tissues are not infected to any extent. Thus, evolution has apparently selected for two types of ascoviruses, both of which extend the larval life span to several weeks. One type predominantly infects most cells in a single tissue (e.g., SfAV-1a) maintaining the other tissues, whereas another type infects only a portion of the cells in a wider variety of tissues (e.g., TnAV-2 and HvAV-3 isolates). In addition, the limited effect of the virus or partial infection in the infected tissues may be favored to avoid alarming the host innate immune system.

Similarly, the proposed role for TnAV anucleate vesicles in virus replication is supported by the early ultrastructural studies explored ascovirus vesicle composition using EM. **First**, the localization of what seems to be a virogenic stroma inside these vesicles agrees with the proposed role in virus replication. **Second**, the vesicles increase in size after release from the infected tissue into the hemolymph. **Third**, vesicles originate from enlarged infected cells that become as much as 10 times or more the size of the original uninfected cell, yet each of the vesicles formed ranged in size from 5-10 μM, and were fully packed with virions (1). The combination of these results provides strong evidence that these viral vesicles use the hemolymph as a primary tissue, if not the

most important tissue in which to replicate, making these anucleate viral factories a markedly different mechanism for virus replication.

Surprisingly, I detected the early and high expression of TnAV-6a ORF072-homolog identified as a caspase-2 precursor-like protein in *Trichoplusia ni* FET tissues and vesicle fraction (**Fig. 4.3**). This expression pattern was unexpected for two reasons: First, the SfAV-1a caspase protein (ORF073) is a very-late gene (expressed by 24 hpi *in vivo*) and is associated with apoptosis induction in Spodoptera frugiperda ascovirus (SfAV-1a), the ascovirus type species (9, 12). Second, in vesicles, the caspase activity in such high level (started at 48 hpi and lasted till day 7th post-infection won't be needed and may interfere with the virus cycle completion, because vesicles are considered the last station in the virus pathogenesis pathway where they accumulate and virus assembly process is ongoing. However, given that the HvAV-3e caspase-like protein was unable to induce the apoptosis in Sf9 cells (14), due to the lack of cysteine residue in the "QACXG" catalytic domain (conserved in all functional caspase proteins and also present in SfAV-1a caspase protein), and the absence of p20 (large unit) and p10 (small unit) potential cleavage site, I can conclude that these caspase-like enzymes may be associated with other functions during the infection. In HvAV-3e the RNAi silencing of the caspase-like gene demonstrated its essential role in viral replication, which would fit more with the proposed function for the vesicles as anucleate virus factory. In general, caspases as intracellular cysteine proteases are associated with many functions other than cell killing, for example, signal transduction, cell differentiation and cytoskeleton remodeling (30).

In addition to the ORF072-homolog high expression that was common between the different tissues, some genes were expressed in high level but only in the vesicle fraction implying its direct association with the vesicle formation or virus replication. Interestingly, two of the lipid-modifying enzymes were identified as highly expressed genes only in the vesicles, these are the fatty acid elongase or elongation of very long chain fatty acids protein 1-like (ORF046-homolog) and a patatin-like phospholipase (ORF067-homolog). The association of these lipid-modifying enzymes with vesicle formation was anticipated since the presence of a cassette (4 genes) of lipid-modifying enzymes is unique to ascovirus genomes. The remaining lipid-modifying enzymes homologs namely, phosphate acyltransferase (ORF098) and lipase (ORF132) were expressed in higher levels in vesicles especially if compared to their levels in FET tissues. Therefore, I conclude that these enzymes must be essential in vesicle membrane lipid modification or virus replication. Interestingly, analysis of a lipase-like enzyme (ORF019) in HvAV-3e demonstrated that this enzyme lacks any functional lipase or esterase activity. However, RNAi knock-down of this transcript demonstrated that it was nevertheless essential for virus replication (31). Other ORFs that were highly expressed in the vesicle fraction included some structural proteins and hypothetical proteins with unknown function. The increase in virion structural proteins is expected to be directly proportional to the increase in the number of virions inside the vesicles. Characterization of ORF149, 150 and 156 hypothetical proteins may help in understanding their role in vesicle formation.

The TnAV genome size and encoding capacity difference from the type species SfAV-1a is expected to be associated with the existing difference in host range and tissue-tropism between ascovirus species. Interestingly, I detected the high expression of a TnAV unique gene identified as an aegerolysin-like protein (ORF029). Aegerolysin proteins are pore-forming proteins, characterized by their β-sheet structure and its low molecular weight (15-17 KDa). They are identified in fungi, bacteria, plants, insects and protozoa (32). This makes ascoviruses, represented by TnAV and HvAV, the only viruses that encode an aegerolysin gene homologs. The biological role of these proteins is diverse and include, antibacterial, antitumor and antiproliferative (33). Moreover, some are known to interact with lipid membranes and lipid vesicles where they induce pore formation and permeabilization, respectively (32). Whether this protein is associated with TnAV/HvAV wider tissue tropism or virulence is not examined yet.

Finally, detection of bidirectional transcription and five bicistronic and tricistronic mRNA messages in the TnAV-6a1 transcriptome indicates this phenomenon may occur in other ascovirus species. I detected 9 bicistronic and 6 tricistronic messages in SfAV transcriptome during *in vivo* infection of the *Spodoptera frugiperda* 3rd-instar larvae (9). Detection of this number of bi- and tri-cistronic messages in the ascovirus transcriptomes may imply the use of non-canonical protein translation strategies for the efficient translation of the proteins encoded by these messages, such as Internal Ribosome Entry Site (IRES) and ribosome re-initiation (34).

In conclusion, ascovirus vesicle formation following destruction of the nuclear membrane represents a unique biology and likely underlies novel cell biology pathways

of this group of viruses. Understanding ascovirus genome expression in different tissues through conducting more transcriptomic studies is a powerful tool that adds more information about their unique cytopathology at the molecular level.

4.6 References

1. Federici BA. 1983. Enveloped double stranded DNA insect virus with novel structure and cytopathology. Proc Natl Acad Sci USA 80: 7664-7668.

2. Hamm J J, Styer E L, Federici BA. 1998. Comparison of field-collected ascovirus isolates by DNA hybridization, host range, and histopathology. J Invertebr Pathol 72:138 –146.

3. Federici BA, Govindarajan R. 1990. Comparative histology of three ascovirus isolates in larval noctuids. J Invertebr Pathol 56: 300-311.

4. Federici BA, Bideshi, DK, Tan Y, Spears T, Bigot Y. 2009. Ascoviruses: superb manipulators of apoptosis for viral replication and transmission. Curr Top Microbiol Immunol 328: 171-196.

5. Bigot Y. 2011.Genus Ascovirus. p 73-78. *In* Tidona C, Darai, G. (ed), Springer Index of Viruses, 2nd edn. Heidelberg: Springer.

6. Bigot Y, Asgari S, Bideshi DK, Cheng X, Federici BA, Renault S. 2011. Family Ascoviridae, p 147-152. *In* King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (ed),Viral Taxonomy, IX Report of the International Committee on the Taxonomy of Viruses, 3rd edn.London: Elsevier–Academic Press.

7. Asgari S, Bideshi DK, Bigot Y, Federici BA, Cheng X. 2017. ICTV Report Consortium. ICTV virus taxonomy profiles: Ascoviridae. J Gen Virol 98: 4 –5.

8. Federici BA. 1982. A new type of insect pathogen in larvae of the clover cutworm, *Scotogramma trifolii*. J Inverteb Pathol 40:41–54.

9. Zaghloul HAH, Hice R, Arensburger P and Federici BA. 2017. Transcriptome analysis of the Spodoptera frugiperda ascovirus *in vivo* provides insights into how its apoptosis inhibitors and caspase promote increased synthesis of viral vesicles and virion progeny. J Virol.

10. Bideshi DK, Bigot Y, Federici BA, Spears T. 2010. Ascoviruses, p 3–34. *In* Asgari S, Johnson KN (ed), Insect virology. Caister Academic Press, Norfolk, United Kingdom.

11. Wang L, Xue J, Seaborn CP, Arif BM, Cheng XW. 2006. Sequence and organization of the *Trichoplusia ni* ascovirus 2c (Ascoviridae) genome. Virology 354: 167–177.

12. Bideshi DK, Dematti MV, Rouleux-Bonnin F, Stasiak K, Tan Y, Bigot S, Bigot Y, Federici BA. 2006. Genomic sequence of the *Spodoptera frugiperda ascovirus 1a*, an enveloped double-stranded DNA insect virus that manipulates apoptosis for viral reproduction. J Virol 80: 11791-11805.

13. Asgari S. 2006. Replication of *Heliothis virescens* ascovirus in insect cell lines. Arch Virol 151:1689 –1699.

14. Asgari S. 2007. A caspase-like gene from *Heliothis virescens* ascovirus (HvAV-3e) is not involved in apoptosis but is essential for virus replication. Virus Res 128:99–105.

15. Federici BA, Vlak JM, Hamm JJ. 1990. Comparative study of virion structure, protein composition and genomic DNA of three ascovirus isolates. J Gen Virol. 71:1661-1668.

16. Bigot Y, Rabouille A, Doury G, Sizaret PY, Delbost F, Hamelin MH, Periquet G. 1997. Biological and molecular features of the relationships between Diadromus pulchellus ascovirus, a parasitoid hymenopteran wasp (*Diadromus pulchellus*) and its lepidopteran host, *Acrolepiopsis assectella*. J Gen Virol. 78:1149-1163.

17. Grabherr MG, Brian JH, Moran Y, Joshua ZL, Dawn AT, Ido A, Xian A et al. 2011. Full-Length Transcriptome Assembly from RNA-Seq Data without a Reference Genome. Nature Biotechnol 29:644-52.

18. Langmead B, Steven SL. 2012. "Fast Gapped-Read Alignment with Bowtie 2." Nature Methods 9:357–59.

19. Simmons J, D'Souza O, Rheault M, Donly C. 2013. Multidrug resistance protein gene expression in *Trichoplusia ni* caterpillars. Insect Mol Biol. 22:62-71.

20. Pfaffl MW. 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 29:e45.

21. Liu Y-Y, Xian W-F, Xue J, Wei Y-L, Cheng X-W, Wang X. 2018. Complete genome sequence of a renamed isolate, Trichoplusia ni ascovirus 6b, from the United States. Genome Announc 6:e00148-18.

22. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat. Methods 5:621– 628.

23. Wu, L, Zhou P, Ge X, Wang LF, Baker ML, Shi Z. 2013. Deep RNA sequencing reveals complex transcriptional landscape of a bat adenovirus. J Virol 87:503–511.

24. Hamm, JJ, Pair SD, Marti OG. 1986. Incidence and host range of a new ascovirus isolated from fall armyworm, *Spodoptera frugiperda* (Lepidoptera: Noctuidae). Fla Entomol. 69:524–531.

25. Govindarajan R, Federici BA. 1990. Ascovirus infectivity and effects of infection on the growth and development of noctuid larvae. J Invertebr Pathol. 56:291–299.

26. Cheng XW, Carner CR, Arif BM. 2000. A new ascovirus from *Spodoptera exigua* and its relatedness to the isolate from *Spodoptera frugiperda.* J Gen Virol 81: 3083-3092.

27. Li SJ, Hopkins RJ, Zhao YP, Zhang YX, Hu J, Chen XY, Xu Z, Huang GH. 2016. Imperfection works: survival, transmission and persistence in the system of Heliothis virescens ascovirus 3h (HvAV-3h), *Microplitis similis* and *Spodoptera exigua*. Sci Rep 6: 21296.

28. Li SJ, Wang X, Zhou ZS, Zhu J, Hu J, Zhao YP, Zhou GW, Huang GH. 2013. A comparison of growth and development of three major agricultural insect pests infected with Heliothis virescens ascovirus 3h (HvAV-3h). PLoS ONE 8: e85704.

29. Hu J, Wang X, Zhang Y, Zheng Y, Zhou S, Huang GH. 2016. Characterization and Growing Development of *Spodoptera exigua* (Lepidoptera: Noctuidae) Larvae Infected by Heliothis virescens ascovirus 3h (HvAV-3h). J Econ Entomol. 109: 2020-6.

30. Connolly PF, Jäger R, Fearnhead HO. 2014. New roles for old enzymes: killer caspases as the engine of cell behavior changes. Front Physiol 5: 149.

31. Smede M, Hussain M, Asgari S. 2009. A lipase-like gene from *Heliothis virescens* ascovirus (HvAV-3e) is essential for virus replication and cell cleavage. Virus Genes 39: 409.

32. Butala M, Novak M, Krasevec N, Skocaj M, Veranic P, Macek P, Sepcic K. 2017. Aegerolysins: Lipid-binding proteins with versatile functions. Semin Cell Biol. 72:142-151.

33. Berne S, Lah L, Sepcić K. 2009. Aegerolysins: structure, function, and putative biological role. Protein Sci 18:694-706.

34. Griffiths A, Coen DM. 2005. An unusual internal ribosome entry site in the herpes simplex virus thymidine kinase gene. Proc Natl Acad Sci USA 102:9667–9672.

4.7 Tables:

Table 4.1: Primers used for RT-qPCR and RT-PCR.

Gene position/ Transcript position	Primers (5' -> 3')	Annealing temperature (◦ C)	RT-PCR/RT-qPCR	Expected size & primer efficiency (E)
eukaryotic Initiation Factor-4a (*eIF4*a) (19)	*eif4*a Forward: ctggatacgctgtgtgac *eif4*a Reverse: accttgcggcgagtgtt	55	RT-qPCR	(E: 91.2%)
ORF077(Contig 18: 2963>5746)	ORF77F: gagtgtttaagtgctggtgaag ORF77R: caatgacaatcgggcagtttc	55	RT-qPCR	141 bp (E: 91%)
ORF042(contig 9: 1428>3938)	ORF42F: gatagagacaaagcactagccg ORF42R: gatcccacgaaccatatcacc	55	RT-qPCR	105 bp (E: 92.5%)
ORF029(Contig 7: 448>1116)	ORF29F: catcaataccgccatacctagag ORF29R: ttaccccatgtcagaactgc	55	RT-qPCR	114 bp (E: 91%)
55-1653	FN: cgaatcctattacatttgtggt RN: catgctccgtatccatagtt	61	RT-PCR	1483 bp
1881-4003	FN: cgtggattaccatatcaacgta RN: agcaacactaccatcacatg	61	RT-PCR	1943 bp
588-2269	FN: ccacgggtagtggtaaatattc R: cgacacagtttatgagaacac	60	RT-PCR	1360 bp
55-1653	F: accaatccatgggtgataaga R: ggttcgagtcaattctgagattg	60	RT-PCR	612 bp
1881-4003	F: cactggaatttatatgtgccg R: ggacactgcatattgaatttcac	60	RT-PCR	889 bp
588-2269	F: caacaccatgttgaacatacg R: cgacacagtttatgagaacac	60	RT-PCR	713 bp

Table 4.2: Trichoplusia ni ascovirus core genes conserved in the genus *Ascovirus* identified in TnAV-6a1 variant contig sequences. Conservation of ORFs among acoviruses species in the genus *Ascovirus* is illustrated.

ORF No.	Contig No.	Start	Stop	Strand	Size (aa)	Size (aa) in TnAV-2c *	Conservation in genus ascovirus	BLASTP (Putative Function)
ORFs for nucleotide metabolism proteins								
ORF001	Contig 34	98	3343	Contig 34 (98>3343)	1081	1026	SfAV, HvAV	DNA polymerase
ORF110	Contig 20	24719	23364	Contig 20 (23364<24719)	451	453	SfAV, HvAV	DNA-directed RNA polymerase subunit b
ORF042	Contig 9	1428	3938	Contig 9 (1428>3938)	836	833	SfAV, HvAV	DNA-dependent RNA polymerase
ORF138	Contig 29	3801	322	Contig 29 (3801<322)	1159	1160	SfAV, HvAV	DNA-directed RNA polymerase II
ORF008	Contig 2	6754	8769	Contig 2 (6754>8769)	671	678	SfAV, HvAV	RNaseIII
ORF087	Contig 20	2961	3323	Contig 20 (2961>3323)	120	135	SfAV, HvAV	GIY-YIG nuclease superfamily protein
ORF121	Contig 21	8159	9307	Contig 21 (8159>9307)	382	380	SfAV, HvAV	DNA repair exonuclease SbcCD nuclease subunit
ORF077	Contig 18	2963	5746	Contig 18 (2963>5746)	927	937	SfAV, HvAV	DNA repair exonuclease SbcCD ATPase subunit
ORF095	Contig 20	11597	10659	Contig 20 (10659<11597)	312	312	SfAV, HvAV	ATPase
ORF078	Contig 18	6049	9159	Contig 18 (6049>9159)	1036	1046	SfAV, HvAV	D5 family helicase-primase
ORF154	Contig 25	10010	10675	Contig 25 (10010>10675)	221	219	SfAV, HvAV	Thymidin kinase
ORF100	Contig 20	15975	14998	Contig 20 (14998<15975)	325	320	SfAV, HvAV	Zn-finger/nucleic acid binding
ORF082	Contig 19	2350	4047	Contig 19 (2350>4047)	565	565	SfAV, HvAV	Helicase-2
ORF058	Contig 11	7456	10398	Contig 11 (7456>10398)	980	979	SfAV, HvAV	Lipopolysaccharide modifying enzyme/Putative tyrosine protein kinase
ORF088	Contig 20	3396	5156	Contig 20 (3396>5156)	586	588	SfAV, HvAV	Serine/Threonine protein kinases
ORF017	Contig 3	5643	6779	Contig 3 (5643>6779)	378	378	SfAV, HvAV	Poxvirus Late Transcription Factor VLTF3 like/DNA binding/packing protein
ORFs for lipid metabolism proteins								
ORF132	Contig 21	17762	16659	Contig 21 (16659<17762)	367	367	TnAV-6b	lipase
**ORF046 + ORF047	Contig 11	1317	553	Contig 11 (553<1317)	254	ORF046 (181 bp), ORF047 (73 bp)	SfAV, HvAV	Fatty acid elongase/elongation of very long chain fatty acids protein 1-like
ORF067	Contig 13	1395	415	Contig 13 (415<1395)	326	327	SfAV, HvAV	patatin-like phospholipase
ORF098	Contig 20	14668	13688	Contig 20 (13688<14668)	326	326	SfAV, HvAV	Phosphate acyltransferase
ORFs for apoptosis-associated proteins								
ORF102	Contig 20	17136	18668	Contig 20 (17136>18668)	510	509	SfAV, HvAV	Cathepsin B
ORF072	Contig 14	3481	2786	Contig 14 (2786<3481)	231	231	SfAV, HvAV	Caspase-like protein
ORF009	Contig 2	8771	9532	Contig 2 (8771>9532)	253	282	SfAV, HvAV	IAP-like protein
ORF006	Contig 2	2144	3964	Contig 2 (2144>3964)	606	598	SfAV, HvAV	Modified RING finger, HC subclass (C3HC5-type)
ORF032	Contig 7	4289	2805	Contig 7 (2805<4289)	494	487	SfAV, HvAV	IAP-like protein, RING-finger-containing E3 ubiquitin ligase
ORFs for virion structural proteins								
ORF118	Contig 21	6978	6298	Contig 21 (6298<6978)	226	226	SfAV, HvAV	Erv1 / Alr family protein
ORF114 + ORF115	Contig 21	4986	2542	Contig 21 (2542<4986)	814	ORF114 (334 bp), ORF115 (456 bp)	SfAV, HvAV	Serine/threonine protein kinase
ORF043	Contig 10	132	3119	Contig 10 (132>3119)	995	977	SfAV, HvAV	Dynein-like beta chain
ORF059	Contig 11	10810	10463	Contig 11 (10463<10810)	115	114	SfAV, HvAV	Yabby-like transcription factor/HMG box
ORF153	Contig 25	9945	8512	Contig 25 (9945<8512)	477	480	SfAV, HvAV	MCP
ORF135	Contig 21	19748	20527	Contig 21 (19748>20527)	259	259	SfAV, HvAV	S1/P1 nuclease
ORF147	Contig 25	43	4470	Contig 25 (43>4470)	1475	1481	TnAV-6b	Chromosome segregation protein SMC
ORF093	Contig 20	9853	9302	Contig 20 (9302<9853)	183	183	SfAV, HvAV	Haloacid dehalogenase-like hydrolase/CTD phosphatase-like
ORF161	Contig 24	4188	7112	Contig 24 (4188>7112)	974	974	SfAV, HvAV	DEAD-like helicase/SWI/SNF2 family helicase

ORFs for host interaction proteins								
ORF158	Contig 24	1867	362	Contig 24 (1867<362)	501	501	SfAV, HvAV	Zinc metalloproteinase
ORF103	Contig 20	19116	19742	Contig 20 (19116>19742)	208	212	SfAV, HvAV	Thioredoxin domain-containing protein
ORF080	Contig 19	1143	676	Contig 19 (676<1143)	155	155	SfAV, HvAV	Transcription repressor MOT2/ ubiquitin ligase
ORF129	Contig 21	15342	14455	Contig 21 (14455<15342)	295	220	SfAV, HvAV	Hypothetical protein/ Myristylated membrane-like protein
ORF099	Contig 20	15023	14766	Contig 20 (14766<15023)	85	85	SfAV, HvAV	Putative transcription elongation factor TFIIS
Hypothetical proteins conserved in ascovirus genus								
ORF126	Contig 21	14402	13503	Contig 21 (13503<14402)	299	299	SfAV, HvAV	Hypothetical protein/254L IIV-6
ORF131	Contig 21	16458	16027	Contig 21 (16027<16458)	143	140	SfAV, HvAV	Hypothetical protein
ORF113	Contig 21	2453	1101	Contig 21 (1101<2453)	450	449	SfAV, HvAV	Hypothetical protein
ORF156	Contig 25	11465	12151	Contig 25 (11465>12151)	228	225	SfAV, HvAV	Hypothetical protein
ORF164	Contig 24	9187	8474	Contig 24 (9187<8474)	237	237	SfAV, HvAV	Hypothetical protein
ORF023	Contig 3	12956	11640	Contig 3 (11640 <12956)	438	438	SfAV, HvAV	Hypothetical protein
ORF007	Contig 2	4291	6609	Contig 2 (4291>6609)	772	772	SfAV, HvAV	Hypothetical protein
ORF010	Contig 3	467	60	Contig 3 (60<467)	135	135	SfAV, HvAV	Hypothetical protein
ORF011	Contig 3	2773	536	Contig 3 (536<2773)	745	744	SfAV, HvAV	Hypothetical protein
ORF018	Contig 3	7129	7914	Contig 3 (7129>7914)	261	257	SfAV, HvAV	Hypothetical protein
ORF157	Contig 25	13719	12208	Contig 25 (13719<12208)	503	507	SfAV, HvAV	Hypothetical protein
ORF150	Contig 25	6090	6779	Contig 25 (6090>6779)	229	228	SfAV, HvAV	Hypothetical protein
ORF149	Contig 25	5393	6088	Contig 25 (5393>6088)	231	233	SfAV, HvAV	Hypothetical protein
ORF125	Contig 21	13196	11712	Contig 21 (11712<13196)	494	495	SfAV, HvAV	Hypothetical protein
ORF124	Contig 21	10981	11739	Contig 21 (10981>11739)	252	251	SfAV, HvAV	Hypothetical protein
ORF122	Contig 21	9375	10550	Contig 21 (9375>10550)	391	388	SfAV, HvAV	Hypothetical protein
ORF116	Contig 21	5551	4973	Contig 21 (4973<5551)	192	192	SfAV, HvAV	Hypothetical protein
ORF159	Contig 24	2693	1944	Contig 24 (2693<1944)	249	245	SfAV, HvAV	Hypothetical protein
ORF091	Contig 20	7745	7206	Contig 20 (7206<7745)	179	180	SfAV, HvAV	Hypothetical protein
ORF060	Contig 11	10931	11578	Contig 11 (10931>11578)	215	215	SfAV, HvAV	Hypothetical protein
ORF004	Contig 2	940	1572	Contig 2 (940>1572)	210	210	SfAV, HvAV	Hypothetical protein
ORF160	Contig 24	3716	2865	Contig 24 (3716<2865)	283	287	SfAV, HvAV	Hypothetical protein
ORF005	Contig 2	1615	2013	Contig 2 (1615>2013)	132	132	SfAV, HvAV	Hypothetical protein
Conserved in HvAV or unique to TnAV								
ORF148	Contig 25	4550	5341	Contig 25 (4550>5341)	263	266	HvAV	Hypothetical protein
ORF033	Contig 7	4651	4307	Contig 7 (4307<4651)	114	114	HvAV	MBL fold hydrolase
ORF050 + ORF051	Contig 11	2339	3091	Contig 11 (2339>3091)	250	ORF51 (81 bp), ORF52 (243 bp)	HvAV	RNA polymerase Rpb5/Cinnamyl alcohol dehydrogenases
ORF029	Contig 7	448	1116	Contig 7 (448>1116)	222	222	HvAV	Aegerolysin-like protein
ORF052	Contig 11	3222	3956	Contig 11 (3222>3956)	244	243	HvAV	Hypothetical protein
ORF090	Contig 20	6367	6849	Contig 20 (6367>6849)	160	211	TnAV	AcOrf-84 peptide [Autographa californica multiple nucleopolyhedrovirus]
ORF092	Contig 20	7812	9230	Contig 20 (7812>9230)	472	472	TnAV	Tryptophan-repeat gene family protein/RING domain in cIAP1, cIAP2
ORF071	Contig 14	1770	2639	Contig 14 (1770>2639)	289	265	TnAV	NAD(P)-dependent dehydrogenase

*The TnAV-6a (TnAV-2c, previously) (11) was used as a reference genome to refer to the ORFs order in our TnAV-6a1 variant.

**The plus sign refers to two ORFs that were found in our TnAV-6a1 variant strain as a single ORF, moreover the longer size of the ORF is supported by TnAV-6b variant (21).

Table 4.3: Trichoplusia ni ascovirus bicistronic and tricistronic messages identified in both the hemolymph viral vesicles and fat body, epidermis and tracheal matrix (FET) tissues.

Contig number	Position	ORFs	Putative function (BLASTP)	IRES** prediction	PolyA
Contig 11	3-undefined*	ORF045 ORF046+ORF047	Hypothetical protein Elongation of very long chain fatty acids protein	Only 12 bases between ORFs	No
Contig 12	55-1653	ORF062 ORF063	Shiga-like toxin/Apoptosis inducing protein PDZ-domain containing protein	Putative IRES	Yes, three poly-A signals at 3' end
Contig 19	588-2269	ORF080 ORF081	MOT2 domain-containing protein/RING finger domain/ BRCA1-like Hypothetical protein	Putative IRES	No (AT rich 3'end)
Contig 24	1881-4003	ORF159 ORF160	Hypothetical protein Hypothetical protein	Putative IRES	Yes (located at the 3' end of the 2nd ORF)
Contig 25	4529-6822	ORF148 ORF149 ORF150	Hypothetical protein Hypothetical protein Hypothetical protein	No IRES predicted between the first and 2nd ORFs. Only 1 base between 2nd and 3rd ORFs	Yes

* Undefined, due to overlapping transcripts.
**Internal Ribosome Entry Site (IRES).

Chapter V: Summary and Future Perspectives

5.1 Introduction

Ascoviruses, discovered in the mid-1970's, are a type of entomopathogen, and thus are considered natural enemies for insects, especially because they primarily attack the larvae of many lepidopteran pest species. In addition, because these viruses have a unique pattern of viral replication and cell pathology, their study holds promise for contributing to our knowledge of molecular cell biology. Among the known hosts of ascoviruses are many of the most economically important and devastating agricultural pests including the fall armyworm (*Spodoptera frugiperda*), beet armyworn (*S. exigua*), cabbage looper (*Trichoplusia ni*), corn earworm, (*Helicoverpa zea*), and tobacco budworm (*Heliothis virescens*). The isolation of ascoviruses from different continents (1, 2) supports their worldwide distribution. Moreover, because the gross pathology of ascovirus disease – retarded development – is difficult to recognize in the field, their abundance is likely underestimated. These viruses, thus, may be more common as natural enemies than the much better known baculoviruses of lepidopteran larvae. Indeed, their restricted transmission by feeding and their dependence on endoparasitic wasp vectors for transmission, limit their success as biological control agents in agricultural applications. With respect to studying their molecular biology, the absence of a suitable cell line to study their unique biology and cytopathology has hindered ascovirus research progress. Therefore, one of the main unresolved questions about ascoviruses is how they transform each infected cell *in vivo* to about 20-30 or more anucleate viral vesicles in which these viruses replicate after destruction of the nucleus.

Therefore, in my dissertation, I focused on contributing to our fundamental knowledge about the underlying molecular biology of ascovirus replication in host larvae that results in the formation of viral vesicles. To do this I used RNA-sequence technology, i.e., transcriptomic studies, to achieve the following three primary objectives: **Objective 1**, characterization and analysis of genome expression patterns and identity of core genes for the Spodoptera frugiperda ascovirus (SfAV-1a), the ascovirus type species, which has a tissue tropism primarily restricted to the fat body tissue, in *S. frugiperda*; **Objective 2**, characterization of the effects of SfAV-1a infection on *S. frugiperda* host mitochondrial and cytoskeletal gene expression and the host's innate immune response; and **Objective 3**, characterization and analysis of genome expression patterns of the Trichoplusia ni ascovirus (TnAV-6a1 variant), which has a broad tissue tropism, replicating in at least three different major tissues, the fat body, tracheal matrix, and epidermis, in *T. ni* larvae, with an emphasis on viral vesicles produced in these tissues, which I collected from the hemolymph (blood).

Here I first provide a very brief summary of my results, followed by a more detailed presentation of these in later paragraphs. Objective 1, addressed in Chapter 2, consists of the first ascovirus transcriptome analysis, conducted on the ascovirus type species or SfAV-1a. I identified the core ascovirus genes including genes with transmembrane-domains, and their expression patterns. I identified the temporal coordination between the Inhibitors of apoptosis, IAP-like proteins (ORF016, ORF025, and ORF074) expressed early in infection compared to late and very late expression of apoptosis inducing enzymes such as cathepsin B (ORF114) and caspase (ORF073),

respectively. For Objective 2, addressed in Chapter 3, I focused on *Spodoptera frugiperda* mitochondrial, cytoskeletal and innate immune gene responses to SfAV-1a infection. The results of this study demonstrated the importance of mitochondrial encoded gene conservation or activation during SfAV-1a infection, and specifically during formation of viral vesicles. My results show that several antimicrobial genes were highly expressed, and that certain cellular immunity genes were involved in host immune defense, but with a limited effect. For Objective 3, dealt with in Chapter 4, I focused on the TnAV transcriptome of the *T. ni* larval hemolymph collected during the first seven days after infection, during which this tissue became densely packed with viral vesicles, and compared the pattern of genome expression to that of other major tissues of this host at the same time. My results revealed high expression of lipid-modifying enzymes in the vesicle fraction, suggesting their importance in vesicle membrane formation. In combination, the results of these studies add considerable new knowledge at the molecular level to how ascovirus replication alters cell architecture to generate viral vesicles and increase the amount of progeny virions produced when infecting their lepidopteran larval hosts. Below I provide a more detailed summary for each research chapter of my dissertation along with my principal conclusions, followed by avenues of future research suggested by my results.

5.2 Summary of Chapter 2 results

The first objective for my study was to determine the SfAV-1a temporal gene expression profile for conserved 44 core genes and those for putative membrane-associated proteins (26 genes). I used strand-specific RNA-Sequencing to follow the SfAV-1a genes

expression pattern at different time points post-infection of 3rd-instar larvae of *Spodoptera frugiperda* followed by qRT-PCR to validate the RNA-Sequencing expression estimates. My transcriptome analysis revealed that using an *in vivo* approach can provide enough coverage for all viral and host genes, an approach that provides a partial solution for lack of a robust cell line for ascovirus studies. In addition, the transcriptome revealed the need for SfAV-1a genome reannotation based on the single-nucleotide resolution of RNA-Sequencing data that performed by manual curation. Moreover, the difference in genes expression revealed the possible clustering of virus genes into three classes based on their temporal expression pattern namely: early, late and very late, an expression pattern that mimic large DNA viruses. Most importantly, the detection of temporal coordination between the virus encoded executioner caspase enzyme (ORF073) and inhibitors of apoptosis or IAP-like proteins (ORF016, ORF025 and ORF074) correlates well with apoptotic events that lead to vesicles formation. Interestingly, the transcriptome revealed the common bidirectional transcription and occurrence of 15 bi-/tri-cistronic messages, an unusual phenomenon for a large ds DNA virus, implying the possible regulatory role for these messages on transcriptional or translational levels. Finally, I identified the marked expression of the first identified insect virokine known as Diedel protein (*die* gene). The high expression of this gene during *in vivo* infection agrees with the proposed role of this protein in *Drosophila* as an IMD pathway suppressor (3).

5.3 Summary of Chapter 3 results

In Objective 2, I followed my initial study (Chapter 2) for the virus transcriptome by studying the response of the SfAV-1a-infected host, *S. frugiperda*, especially for the mitochondrial, cytoskeleton and innate immunity genes (Chapter 3). RNA-Sequencing revealed that the *S. frugiperda* mitochondrial genome transcriptome is modified upon the SfAV-1a infection. Specifically, mitochondria encoded genes are transcribed in levels either comparable to the control or as much as two-fold during vesicle formation. This transcription pattern resembles human cytomegalovirus, during which a chronic infection is induced that differs from other viruses that induce an acute infection because the latter degrade the mitochondrial genome quickly. This pattern reveals the importance of this organelle in the vesicle formation or virus replication, and supports the previous ultrastructural studies on tissues infected with ascovirus, which demonstrated that mitochondrial shape, size and localization changed after infection (1). On the other hand, the innate immunity genes represented in humoral and cellular effectors demonstrated moderate increases after infection, implying that the vesicles induce an immune response. However, I noticed a simultaneous upregulation of negative regulators of these pathways. For example, the upregulation of the Toll pathway negative regulator, Cactus, as well as upregulation of the prophenoloxidase negative regulator. Moreover, the longevity of the ascovirus disease may be supported by the more than 32-fold increase of some antimicrobial peptides such as moricin and gloverin. This study demonstrated that even after the fat body destruction some antimicrobial peptides were effectively expressed.

Therefore, other tissues such as epidermis and tracheal matrix may be associated with this defense.

5.4 Summary of Chapter 4 results

In Objective 3 (Chapter 4)**,** I compared the vesicle-specific transcriptome patterns in blood versus the fat body, epidermis and tracheal matrix during TnAV-6a1 variant infection in *Trichoplusia ni* larvae. From some larvae, the blood was collected and from others the larval tissues were collected except from the alimentary canal. Then, the genome from our TnAV lab isolate was isolated and sequenced, followed by viral genome assembly using the TnAV-6a strain as a reference genome. Finally, RNA-Sequencing for the different tissues was performed and validated by qRT-PCR. The comparison between these tissues at 48 hpi illustrated limited infection of the Fat body, Epidermis and Tracheal matrix (FET) tissues, especially if compared to the hemolymph vesicle fraction, which is characterized by extremely high levels of virus transcription after formation. For example, at 48 hpi, 26 viral genes were expressed in FET tissues by equal to or more than 5 $\log^2$ RPKM, whereas almost double the number of genes (48 genes) were expressed by similar levels in the vesicles. Therefore, the vesicles were functioning as anuclear virus factories for replication. Moreover, the analysis of the vesicle fraction demonstrated the importance of the lipid modifying enzymes (for example, a fatty acid elongase ORF046 and a patatin-like phospholipase ORF067), as these were expressed in higher levels only in the hemolymph. Finally, the conservation of the phenomena of expressing some virus genes as a bicistronic or tricistronic messages

implies the importance of this phenomenon as a possible important characteristic of ascoviruses.

5.5 Future studies on ascoviruses

The above studies contributed new information about the molecular basis for ascovirus replication and vesicle formation. However, as is typical of research, my results point to many new questions that may represent good research avenues to explore in the future. Below I suggest three areas of research that could yield new and potentially important information about the unique molecular biology of ascoviruses,

5.5.1 Identification of functional Internal Ribosome Entry Sites (IRES) in ascoviruses

The identification of bi- and tri-cistronic messages in the dsDNA genome of SfAV-1a and TnAV-6a1, imply the possible importance of this phenomena in regulation of transcriptional or translational levels. Whether this means that ascoviruses can use non-canonical translation mechanism such as IRES has not been confirmed and thus needs further investigation and analysis. For example, the absence of the cap structure in these messages needs to be proven, and the IRES sequences identified *in silico* in my study have to be tested for their ability to efficiently translate each of the proteins encoded by the mRNA message. The use of a bicistronic reporter system has been described for firefly and *Renilla* luciferase proteins, and both transcripts were easily detected and quantified (4), and thus this type of reporter system can be used to test this multi-cistronic possibility.

5.5.2 Identification of ascovirus encoded mitochondria-targeting proteins and their role in mitochondria manipulation

Based on the analysis of the mitochondrial gene transcription during the SfAV-1a infection in the 3rd-instar larvae of *S. frugiperda*, it is expected that the mitochondria-encoded proteins participate in vesicle formation. However, which viral genes are associated with this process remains unknown. Therefore, a good starting point for answering this question is to functionally characterize the SfAV-1a genes that possess a mitochondria-targeting signal in their protein sequence because these represent good candidates for involvement in mitochondrial manipulation. Although the mitochondrial manipulation by mammalian viruses is well documented (5, 6), this research area has not been explored for any virus that attacks insects.

5.5.3 Identification and functions of ascovirus encoded miRNAs

Given that most of the identified virus-encoded miRNAs are in DNA viruses, it is likely that ascoviruses encode miRNAs in their genomes. This is supported by a study in which a single miRNA, HvAV-miR-1, was found encoded by HvAV-3 (7). Therefore, investigating the SfAV-1a genome for miRNAs and identifying their targets in the host cell may help explain at the molecular level whether this mechanism occurs in other ascovirus species. Given the complexity of ascovirus viral vesicle formation, it would not be surprising that miRNAs are involved regulation of this process, and even in regulating innate immunity protein pathways.

5.6. References

1. Federici BA, Bideshi, DK, Tan Y, Spears T, Bigot Y. 2009. Ascoviruses: superb manipulators of apoptosis for viral replication and transmission. Curr Top Microbiol Immunol 328:171-196.

2. Arai E, Ishii K, Ishii H, Sagawa S, Makiyama N, Mizutani T, Omatsu T, Katayama Y, Kunimi Y, Inoue MN, Nakai M. 2018. An ascovirus isolated from *Spodoptera litura* (Noctuidae: Lepidoptera) transmitted by the generalist endoparasitoid *Meteorus pulchricornis* (Braconidae: Hymenoptera). J Gen Virol doi: 10.1099/jgv.0.001035.

3. Lamiable O, Kellenberger C, Kemp C, Troxler L, Pelte N, Boutros M, Marques JT, Daeffler L, Hoffmann JA, Roussel A, Imler JL. 2016. Cytokine Diedel and a viral homologue suppress the IMD pathway in Drosophila. Proc Natl Acad Sci 113:698–703.

4. Thompson SR. 2012. So you want to know if your message has an IRES? Wiley Interdiscip Rev RNA 3:697–705.

5. Anand, SK, Tikoo SK. 2013. Viruses as modulators of mitochondrial functions. Adv Virol 2013. 738794.

6. Khan M, Syed GH, Kim SJ, Siddiqui A. 2015. Mitochondrial dynamics and viral infections: A close nexus. Biochim Biophys Acta 1853:2822–33.

7. Hussain M, Taft RJ, Asgari S. 2008. An insect virus-encoded microRNA regulates viral replication. J Virol 82:9164–9170.