UC Riverside
UC Riverside Previously Published Works

Title

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

Permalink

https://escholarship.org/uc/item/97t8188t

Journal

Journal of Virology, 91(23)

ISSN

0022-538X

Authors

Zaghloul, Heba AH
Hice, Robert
Arensburger, Peter
et al.

Publication Date

2017-12-01

DOI

10.1128/jvi.00874-17

Peer reviewed

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

Heba A. H. Zaghloul,[a,d] Robert Hice,[a] Peter Arensburger,[c] Brian A. Federici[a,b]

Interdepartmental Graduate Program in Microbiology and Institute for Integrative Genome Biology[a] and Department of Entomology,[b] University of California, Riverside, Riverside, California, USA; California State Polytechnic University, Pomona, Department of Biological Sciences, Pomona, California, USA[c]; Department of Botany and Microbiology, Faculty of Science, Alexandria University, Alexandria, Egypt[d]

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, we 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. We 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.

IMPORTANCE Unlike all other DNA viruses, ascoviruses code for an executioner caspase, apparently involved in a novel cytopathology in which viral replication induces nuclear lysis followed by cell cleavage, yielding numerous large anucleate viral vesicles that continue to produce virions. Our transcriptome analysis of genome expression *in vivo* by the Spodoptera frugiperda ascovirus shows that inhibitors of apoptosis are expressed first, enabling viral replication to proceed, after which the SfAV-1a caspase is synthesized, leading to viral vesicle synthesis and subsequent extensive production of progeny virions. Moreover, we detected numerous bicistronic and tricistronic mRNA messages in the ascovirus transcriptome, implying that ascoviruses use other noncanonical translational mechanisms, such as internal ribosome entry sites (IRESs). These results provide the first insights into the molecular biology of a unique coordinated gene expression pattern in which cell architecture is markedly modified, more than in any other known eukaryotic virus, to promote viral reproduction and transmission.

Address correspondence to Brian A. Federici, brian.federici@ucr.edu.

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 reprogramed 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, we 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, we selected SfAV-1a, which has a genome of ~156,922 bp, with G+C 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, our 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 we report the results of our initial study, in which we determined the temporal core gene expression of SfAV-1a *in vivo*. Three temporal classes were defined: early, late, and very late. Among other results, we 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.

RESULTS

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 postinfection (hpi), then 0.06% at 12 hpi, 1.71% at 24 hpi, 1.87% at 48 hpi, 3.64% at the 4th day postinfection, and 3.76% at the 7th day postinfection. 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 (12, 13), and our edits of the published sequence corrected a number of the reported ORFs (Table 1).

In this analysis, we focused our attention on two SfAV-1a gene sets. The first consisted of 44 core genes, present in more than one ascovirus species, identified in our analysis by BLASTP (NCBI) (Table 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 (14, 15). 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

TABLE 1 Comparison of the SfAV-1a original and new positions for core and TMD-containing ORFs identified in the SfAV-1a transcriptome[a]

ORF	Ascovirus species in which ORF is conserved	No. of TMDs	Corresponding putative function (BLASTP)	ORF new position[b]	ORF old position[b]	Transcript position[b,c]
ORFs for 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		RNase III	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
ORFs for 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
ORFs for apoptosis-associated proteins						
ORF114	HvAV, TnAV, DpAV		Cathepsin B	144165 > 145391	144146 > 145507	144073 > 146526
ORF073	HvAV, TnAV		Caspase/interleukin-1 beta-converting enzyme (ICE) homologs	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-containing E3 ubiquitin ligase)	30115 > 30807	30100 > 30792	27865 > 30906
ORFs for virion structural proteins						
ORF061	HvAV, TnAV, DpAV		Erv/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

TABLE 1 (Continued)

| ORF	Ascovirus species in which ORF is conserved	No. of TMDs	Corresponding putative function (BLASTP)	ORF new position[b]	ORF old position[b]	Transcript position[b,c]
ORF091	HvAV, TnAV, DpAV		High-mobility-group protein/yabby-like protein	121630 < 121998	121613 < 121981	121482 < 122090
ORF109	HvAV, TnAV, DpAV		Haloacid dehalogenase-like hydrolase/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
ORFS for 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 metalloproteinase	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-semialdehyde dehydrogenase	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
ORFS for TMD-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 metalloproteinase	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

(Continued on next page)

TABLE 1 (Continued)

ORF	Ascovirus species in which ORF is conserved	No. of TMDs	Corresponding putative function (BLASTP)	ORF new position[b]	ORF old position[b]	Transcript position[b,c]
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

[a]Conservation in other ascovirus species is illustrated for core genes. The number of TMDs was determined using the TMHMM server v.2.0. Old positions for ORFs conserved among ascoviruses are from reference 11.

[b]">" refers to the forward orientation and "<" to the reverse orientation.

[c]Undefined 3' or 5' ends (undetermined) were due to overlaps with the previous or next gene.
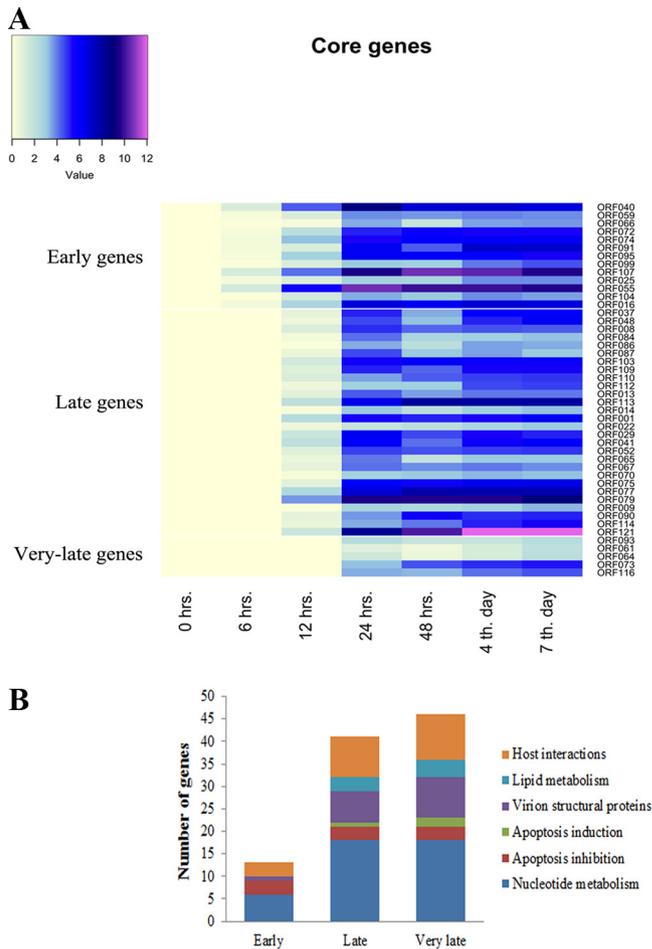
FIG 1 Cluster analysis of the SfAV-1a core gene temporal transcription pattern identified *in vivo* during infection of the 3rd-instar larvae of *Spodoptera frugiperda*. (A) Heat map representation of SfAV-1a expression of each core gene. Core genes are common genes identified in many ascovirus species. 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.

annotated SfAV-1a two gene sets with the original genome map positions is provided in Table 1.

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. 1 and 2). The early gene cluster consisted of 13 core and 4 TMD genes expressed beginning at 6 h postinfection. 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. 1 and 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

A

TMD genes

B
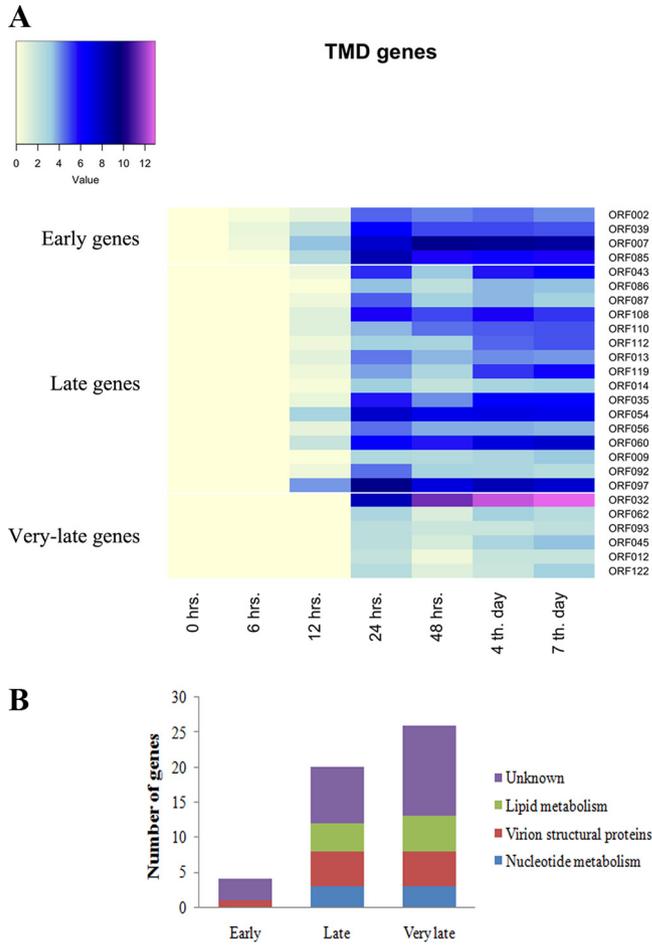
FIG 2 Cluster analysis of the SfAV-1a transmembrane domain (TMD)-containing genes temporal transcription pattern identified *in vivo* during infection of 3rd-instar larvae of *Spodoptera frugiperda*. (A) Heat map representation of SfAV-1a expression of each TMD gene. TMD genes were identified using TMHMM server v.2.0 (14, 15). 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.

two gene sets centered mainly on DNA and RNA metabolism (Fig. 1 and 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 [16]). 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) (16). 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 postinfection. 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 our analysis. Interestingly, ORF121 is a late gene identified as a homolog of the *Drosophila* Diedel protein gene (*die* gene) (17). In addition, ORF107 is an early gene sharing homology with the trifunctional transcriptional regulator/ proline dehydrogenase/L-glutamate gamma-semialdehyde dehydrogenase in *Wigglesworthia glossinidia* (11).

The other highly expressed ORFs (ORF032 and ORF055) represent good candidates for future studies because they are highly expressed according to our 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).

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 and described schematically in Fig. 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) (18).

Validation of the SfAV-1a bicistronic and tricistronic messages. As noted above, we edited the SfAV-1a genome based on mapped sequence reads in different positions where only 100% mismatches were considered real. Thus, we were 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 and 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. 3).

Validation of the RNA-Seq data by RT-qPCR analysis. We 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. 4.

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. 5). Moreover, screening of the promoter sequences of the bi-/tricistronic messages identified a consensus motif (Fig. 5). We were 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 (18, 19). 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

TABLE 2 Description of the bicistronic and tricistronic mRNA messages identified in the SfAV-1a transcriptome infecting 3rd-instar *Spodoptera frugiperda* larvae[a]

Transcript position	Size (bp)	Bicistronic/tricistronic	ORF position	Corresponding putative function	IR position	Putative IRES in IR (>15 bp)	Poly(A) signal position near the 3' message end
6607 > 8748	2,142	Bicistronic	ORF1: 6706 > 7962 ORF2: 7975 > 8697	ORF1: hypothetical protein ORF2: hypothetical protein	7963–7974	NA	No poly(A)
21461 > 23719	2,259*	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) (bp 23695)
23790 > 25159	1,370	Bicistronic	ORF1: 23990 > 24451 ORF2:24509 > 25060	ORF1: Hypothetical protein ORF2: hypothetical protein	24452–24508	Putative IRES	No poly(A)
27865 > 30906	3,042	Bicistronic	ORF1: 27951 > 30107 ORF2: 30115 > 30807	ORF1: RNase III ORF2: IAP-like (E3 ubiquitin-protein ligase)	30108–30114	NA	No poly(A)
56430 > 61996	5,567	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 bp 60950	Putative IRES NA	Poly(A) (bp 61888)
74262 > 76701	2,440	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	3,018*	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 bp 103886 Intergenic 2: 105251–105258	NA NA	Poly(A) (bp 106262)
126088 > 127611	1,524*	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) (bp 127526)
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	2,454*	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	1,169*	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) (bp 153750)
47571 < 49092	1,522	Bicistronic	ORF1: 48309 < 48995 ORF2: 47659 < 48297	ORF1: hypothetical protein ORF2: hypothetical protein	48298–48308	NA	No poly(A)

(Continued on next page)

TABLE 2 (Continued)

Transcript position	Size (bp)	Bicistronic/ tricistronic	ORF position	Corresponding putative function	IR position	Putative IRES in IR (>15 bp)	Poly(A) signal position near the 3′ message end
53392 < 55735	2,344	Bicistronic	ORF1: 55219 < 55731 ORF2: 53489 < 55186	ORF1: chromosome segregation protein SMC ORF2: 2-cysteine adapter domain, P64 protein	55187–55218	No IRES	No poly(A)
63013 < 66156	3,144	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	2,679	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)

^aAsterisks indicate that message size and sequence were further validated by RT-PCR and Sanger sequencing of the PCR product. ">" refers to forward orientation and "<" refers to reverse orientation. IR, intergenic region; IRES, internal ribosome entry site; NA, intergenic regions less than 15 bp.
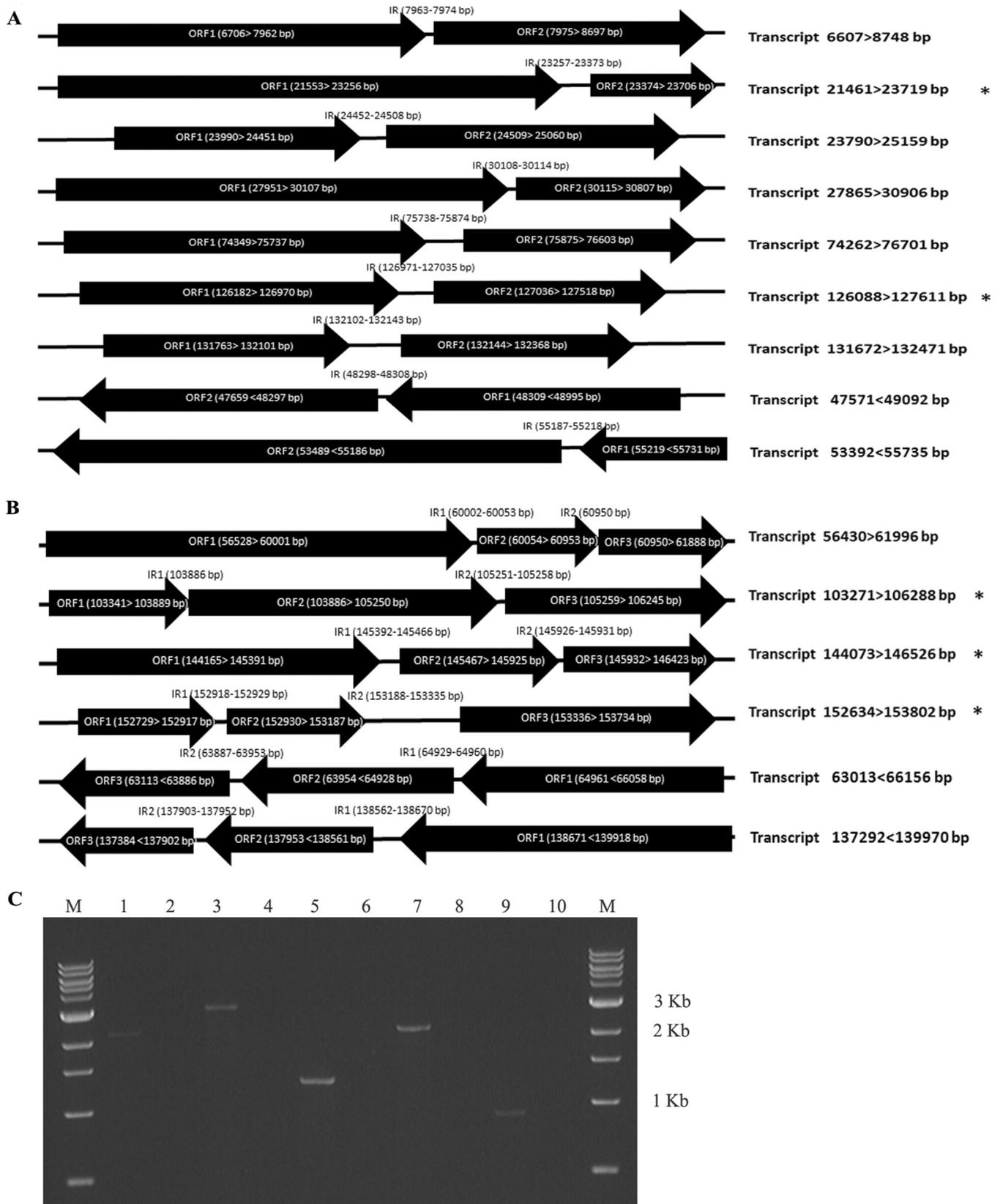
FIG 3 Description and validation of SfAV-1a bicistronic (A) and tricistronic (B) messages identified. IR, intergenic region. ">" refers to forward orientation and "<" refers to reverse orientation. (C) RT-PCR validation of mRNA message size for two bicistronic and three tricistronic messages (marked 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 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).
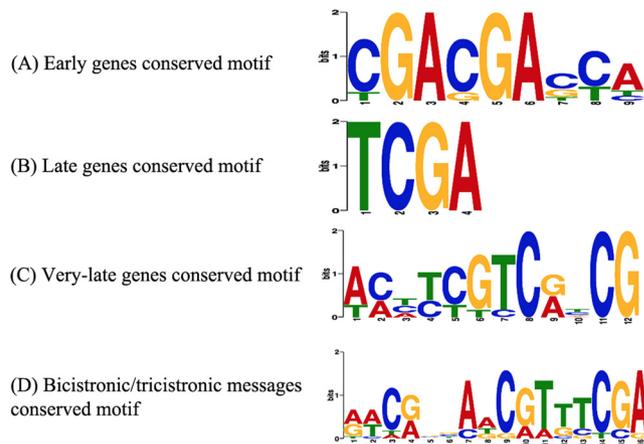
TABLE 3 Primers used in this study

Transcript position/gene position	Primers (5′→ 3′)	Annealing temp (°C)	Application	Expected size (bp) (primer efficiency)
ORF032 (37541 < 38494)	For: ATGGATCGGAACGCCTATTG Rev: GGCCAGTATCGATGTACCATG	60	RT-qPCR	142 (91.6%)
ORF037 (43972 < 44310)	For: TGGCTGATAAGAACGGTTGG Rev: CGACCTGAACAATGCATCTTG	60	RT-qPCR	112 (90.7%)
ORF016 (23374 > 23706)	For: AAGTCCTCGGTCGTATTTGTG Rev: CTATCCTAGCCCTACACAATGG	60	RT-qPCR	103 (90.9%)
rep_c559-ribosomal protein L8 (37)	For: GTGATTCGTGCTCAGCGTAA Rev: ACCAGGGTCATGGATGATGT	60	RT-qPCR	90.2%
21461 > 23719	For_mesg 1: GGTGAAACGAACGAAACATTC Rev_mesg 1: CTATCCTAGCCCTACACAATG	58	RT-PCR	2,116
103271 > 106288	For_mesg 3: GGTATAGTGGTGAGGTTGTAC Rev_mesg 3: CGTAGTTCTCGTTGAGATTG	58	RT-PCR	2,777
126088 > 127611	For_mesg 4: CCACGATGACTGATCCATTG Rev_mesg 4: GCAATCCTCTCTTGTACACG	60	RT-PCR	1,265
144073 > 146526	For_mesg 5: GGATCGTCGATTGTCATACC Rev_mesg 5: GTCGTCGTTCCAATCGTGAA	60	RT-PCR	2,115 bp
152634 > 153802	For_mesg 6: CGAATCGGTCAATATCTGTTC Rev_mesg 6: TTCTCCAATGCGTCTCCATC	60	RT-PCR	898

between ascovirus MCPs and IIV-6 (late genes). Previous molecular phylogenetic studies provide evidence that the ascoviruses evolved from iridoviruses (20). However, we did not find these motifs in other genes, i.e., in the SfAV-1a core and TMD genes we studied.
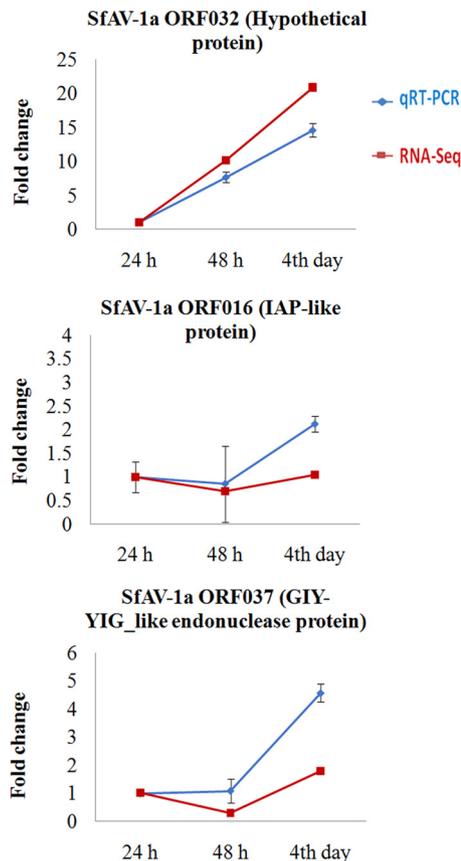
FIG 4 Comparison of the RNA-Seq profile with the data obtained from RT-qPCR for three viral genes (ORF032, ORF016, and ORF037). The host L8 ribosomal gene was used for normalization of the qRT-PCR data (37). Two technical and three biological replicates were tested for each sample. The 24-hpi sample 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.

(A) Early genes conserved motif

(B) Late genes conserved motif

(C) Very-late genes conserved motif

(D) Bicistronic/tricistronic messages conserved motif

FIG 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 (63).

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), implying the possible use of noncanonical protein translational mechanisms such as IRESs by SfAV-1a during translation of some of these messages.

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 (21). However, in our *in vivo* insect larval system we were 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 (22, 23). We focused our 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 our 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 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 (24). 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 (25). Early expression of apoptosis inhibitors was detected in a baculovirus temporal expression analysis, in which the antiapoptosis protein p35 was detected after 6 hpi (24). The early transcription of apoptosis inhibitors can serve as an essential step
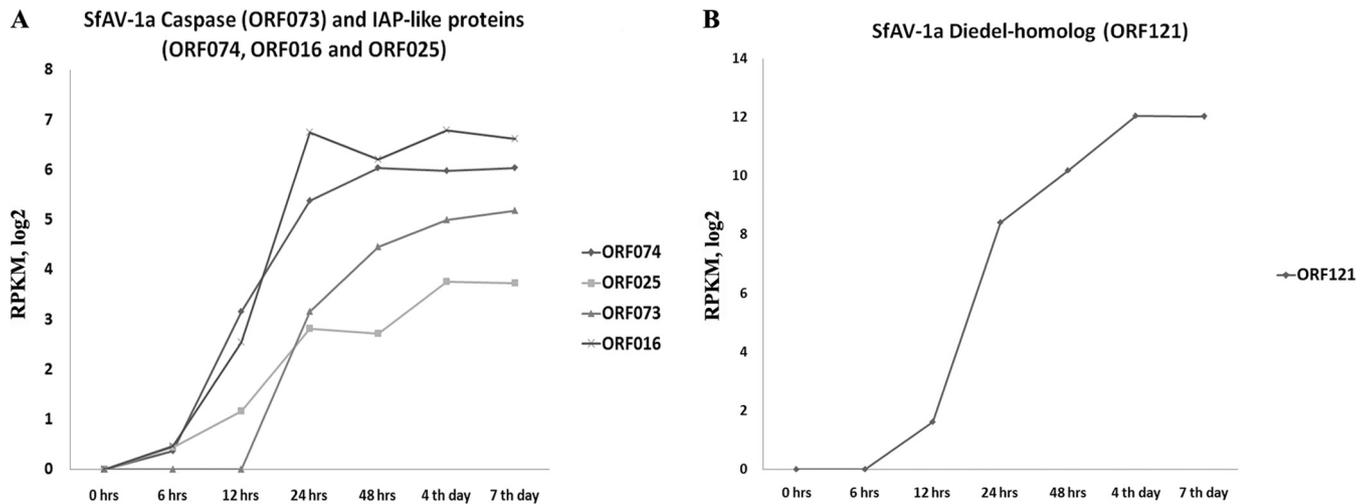
FIG 6 (A) SfAV-1a caspase (ORF073) very late expression (at 24 hpi) versus early expression (at 6 hpi) of putative inhibitor of apoptosis (IAP-like proteins) ORFs (ORF074, ORF016, and ORF025). (B) SfAV-1a Diedel homolog (ORF121) expression pattern *in vivo*.

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. 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 (26). 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 postinfection, 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 (27–29). Alternatively, thioredoxin may have an antiapoptotic activity through its reactive oxygen species (ROS) scavenging ability (29, 30). 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 (31).

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 (32). 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 (33). 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) (34, 35), implying that complete degradation of early messages is not essential for late expression, as reported for poxviruses (36). 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 our 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 (25).

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 our study. Moreover, this reveals that the *Spodoptera frugiperda* L8 housekeeping gene (37) 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 (38).

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, we 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 (39). Moreover, the same phenomenon was reported for other large DNA viruses, such as herpes simplex virus 1 (HSV-1) and vaccinia virus (40, 41). Later, HSV was found to use the noncanonical internal ribosome entry site (IRES) strategy in translation (42). 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) (42–45). 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 (46, 47), 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 (48, 49), 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). However, due to the limitation of a tightly synchronized infection in our *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 we identified *in silico* several putative IRES sequences in the intergenic regions (Table 2) for the bicistronic and tricistronic messages using the IRESPred server (50), 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 (51).

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 h postinfection and peaked at 48 h, followed by a significant decrease in the gene expression level (52). The late expression of the RNase III gene is in agreement with what we found, as the first detection of ORF022 was at 12 hpi. However, the expression increased gradually until day 7 postinfection. Suppressors of host RNAi have been reported for other large DNA viruses, for example, iridoviruses and baculoviruses (53, 54). 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 (55–58). This is in agreement with what we 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 (17, 59). 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 postinfection, which is in agreement with our data (59) (Fig. 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 knowl-

edge 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.

MATERIALS AND METHODS

Virus strain and host infection. As in previous studies, we used the wild-type strain Sf82–126 isolate of Spodoptera frugiperda ascovirus 1a (SfAV-1a) (32), 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 × 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.

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 postinfection (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.

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.

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 (2×), 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.

SfAV-1a ORFs reannotation based on the RNA-Seq data. Initial mapping of our RNA-Seq reads with the published SfAV-1a genome (11) indicated the presence of a number of mismatches between our data and the reported sequence of SfAV-1a. Thus, we used the RNA-Seq data to reannotate the original genome sequence of SfAV-1a prior to proceeding with the transcriptome analysis. We initially edited the SfAV-1a genome available through the NCBI database (accession number NC_008361.1) only at those positions where 100% of our 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 (14, 15) (Table 1).

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 (60), and ribosomal contaminants were identified and removed (61). These filtered libraries were mapped to the SfAV-1a genome sequence deposited in the NCBI database (accession number NC_008361.1) using bowtie2 (62). 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.

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 (63). 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. We 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.

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 μM. 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 (37). 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 (64).

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) were further confirmed by using specific primer sets (Table 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.

IRESPred for IRES prediction in intergenic regions. We 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, we were not able to test intergenic regions less than 15 bp in size because of program restrictions (50).

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.

ACKNOWLEDGMENTS

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. https://doi.org/10.1016/0022-2011(90)90115-M.
2. 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. https://doi.org/10.1371/journal.pone.0006397.
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. https://doi.org/10.1093/jee/tow183.
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. https://doi.org/10.1038/srep21296.
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. https://doi.org/10.1073/pnas.80.24.7664.
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. https://doi.org/10.1101/gad.1300205.
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. https://doi.org/10.1007/s00705-006-0762-7.
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. https://doi.org/10.1128/JVI.01639-06.
12. Wang Z, Gerstein M, Snyder M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10:57–63. https://doi.org/10.1038/nrg2484.
13. Tran VDT, De Coi N, Feuermann M, Schmid-Siegert E, Băguţ 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. https://doi.org/10.1128/mSystems.00036-16.
14. 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. https://doi.org/10.1006/jmbi.2000.4315.
15. 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.

16. 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. https://doi.org/10.1099/vir.0.005934-0.

17. 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. https://doi.org/10.1371/journal.pone.0033416.

18. 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. https://doi.org/10.1007/s00705-007-1081-3.

19. 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. https://doi.org/10.3390/v9010016.

20. 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. https://doi.org/10.1099/vir.0.19290-0.

21. 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. https://doi.org/10.1128/mBio.00942-15.

22. 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. https://doi.org/10.1186/1471-2164-13-118.

23. 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.

24. 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. https://doi.org/10.1128/JVI.00194-13.

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

26. 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. https://doi.org/10.1016/S0002-9440(10)64940-1.

27. 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.

28. Ueda S, Masutani H, Nakamura H, Tanaka T, Ueno M, Yodoi J. 2002. Redox control of cell death. Antioxid Redox Signal 4:405–414. https://doi.org/10.1089/15230860260196209.

29. Masutani H, Ueda S, Yodoi J. 2005. The thioredoxin system in retroviral infection and apoptosis. Cell Death Differ 12:991–998. https://doi.org/10.1038/sj.cdd.4401625.

30. 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. https://doi.org/10.1038/ncb851.

31. 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.

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

33. 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. https://doi.org/10.1016/j.virusres.2007.04.020.

34. Chinchar VG, Yu W. 1992. Metabolism of host and viral mRNAs in frog virus 3-infected cells. Virology 186:435–443. https://doi.org/10.1016/0042-6822(92)90008-D.

35. D'Costa SM, Yao H, Bilimoria SL. 2001. Transcription and temporal cascade in *Chilo* iridescent virus infected cells. Arch Virol 146:2165–2178. https://doi.org/10.1007/s007050170027.

36. 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. https://doi.org/10.1016/0042-6822(79)90095-3.

37. 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. https://doi.org/10.1186/1471-2164-15-704.

38. Yang M, Lin X, Rowe A, Rognes T, Eide L, Bjørås M. 2015. Transcriptome analysis of human OXR1 depleted cells reveals its role in regulating the p53 signaling pathway. Sci Rep 5:17409. https://doi.org/10.1038/srep17409.

39. D'Costa SM, Yao HJ, Bilimoria SL. 2004. Transcriptional mapping in *Chilo* iridescent virus infections. Arch Virol 149:723–742. https://doi.org/10.1007/s00705-003-0232-4.

40. 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-1 genome. J Virol 37:1011–1027.

41. 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.

42. 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. https://doi.org/10.1073/pnas.0504132102.

43. 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. https://doi.org/10.1128/JVI.75.6.2938-2945.2001.

44. Isaksson A, Berggren M, Ricksten A. 2003. Epstein-Barr virus U leader exon contains an internal ribosome entry site. Oncogene 22:572–581. https://doi.org/10.1038/sj.onc.1206149.

45. 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. https://doi.org/10.1128/JVI.01732-13.

46. Spriggs KA, Bushell M, Mitchell SA, Willis AE. 2005. Internal ribosome entry segment-mediated translation during apoptosis: the role of IRES-trans-acting factors. Cell Death Differ 12:585–591. https://doi.org/10.1038/sj.cdd.4401642.

47. 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. https://doi.org/10.1042/BC20070098.

48. 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. https://doi.org/10.1016/j.virol.2004.04.015.

49. Smiley JR. 2004. Herpes simplex virus virion host shutoff protein: immune evasion mediated by viral RNase? J Virol 78:1063–1068. https://doi.org/10.1128/JVI.78.3.1063-1068.2004.

50. 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. https://doi.org/10.1038/srep27436.

51. Thompson SR. 2012. So you want to know if your message has an IRES? Wiley Interdiscip Rev RNA 3:697–705. https://doi.org/10.1002/wrna.1129.

52. 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. https://doi.org/10.1128/JVI.02362-09.

53. 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. https://doi.org/10.1016/j.coviro.2014.03.010.

54. 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. https://doi.org/10.1128/JVI.00802-15.

55. 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. https://doi.org/10.1128/JVI.80.10.5059-5064.2006.

56. 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. https://doi.org/10.1073/pnas.1207213109.

57. 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. https://doi.org/10.1089/jir.2013.0026.

58. 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. https://doi.org/10.4049/jimmunol.1102486.

59. 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. https://doi.org/10.1073/pnas.1516122113.

60. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for illumine sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

61. Kopylova E, Noe L, Touzet H. 2012. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics 28:3211–3217. https://doi.org/10.1093/bioinformatics/bts611.

62. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923.

63. 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. https://doi.org/10.1093/nar/gkp335.

64. Pfaffl MW. 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 29:e45. https://doi.org/10.1093/nar/29.9.e45.